Essays on quantitative risk management with an application to asset allocation problems

Von der Mercator School of Management, Fakultät für Betriebswirtschaftslehre, der

Universität Duisburg-Essen

zur Erlangung des akademischen Grades

eines Doktors der Wirtschaftswissenschaft (Dr. rer. oec.)

genehmigte Dissertation

von

Sascha Offermann

aus

Oberhausen

Referentin/Referent: Prof. Dr. Antje Mahayni
Koreferentin/Koreferent: Prof. Dr. Volker Krätschmer

Tag der mündlichen Prüfung: 12. August 2022

Acknowledgments

First, I would like to express my sincere gratitude to my supervisor Prof. Dr. Antje Mahayni. Her trust, support and especially her great expertise have led and accompanied me on my academic path. Due to her caring and loving nature, it was always a pleasure to work at her chair and she gave me a lot of freedom to develop my own research.

A special thank you goes to Prof. Dr. Volker Krätschmer and Dr. Nikolaus Schweizer. Without them, my path might have never led me into academia.

Furthermore, I would like to thank my co-authors and (former) colleagues Dr. Sven Balder, Lara Becker, Dr. Richard Biegler-König, Prof. Dr. Nicole Branger, Dr. Cathleen Kroh, Oliver Lubos, Susanne Lucassen, Prof. Dr. Judith C. Schneider, Dr. Daniel Schwake, Evamaria Speich and Katharina Stein for many great moments that I will always remember: Be it the last exhausting days before a deadline, the conferences, designing of lectures, the annual Topsim week or the daily fight for sweets in the office.

Most importantly, none of this would have been possible without my parents, Andreas Offermann and Iris Offermann, whose love and guidance are with me in everything I do in life. My achievements and success are because they believe in me every day. My deepest gratitude and love go to my wonderful girlfriend Jennifer Sowa. I am forever thankful for your unconditional love, your unfailing support and every day with you!

Danksagungen

An erster Stelle möchte ich mich bei meiner Betreuerin und Doktormutter Prof. Dr. Antje Mahayni herzlich bedanken. Ihr Vertrauen, ihre Unterstützung und vor allem ihre große Expertise haben mich auf meinem akademischen Weg geleitet und begleitet. Durch ihre fürsorgliche und liebevolle Art war es immer eine Freude, an ihrem Lehrstuhl zu arbeiten und sie gab mir viel Freiraum, um meine eigene Forschungsarbeit zu entwickeln.

Ein besonderes Dankeschön geht an Prof. Dr. Volker Krätschmer und Dr. Nikolaus Schweizer. Ohne sie hätte mich mein Weg vielleicht nie in die Wissenschaft und Forschung geführt.

Des Weiteren möchte ich mich bei meinen Co-Autoren und Kollegen Dr. Sven Balder, Lara Becker, Dr. Richard Biegler-König, Prof. Dr. Nicole Branger, Dr. Cathleen Kroh, Oliver Lubos, Susanne Lucassen, Prof. Dr. Judith C. Schneider, Dr. Daniel Schwake, Evamaria Speich und Katharina Stein für viele tolle Momente bedanken, an die ich mich immer erinnern werde: Seien es die letzten anstrengenden Tage bis zu einer Deadline, die Konferenzen, die Gestaltung von Vorlesungen, die jährliche Topsim Woche oder der tägliche Kampf um Süßigkeiten im Büro.

Am wichtigsten ist jedoch, dass all dies nicht ohne meine Eltern, Andreas Offermann und Iris Offermann, möglich gewesen wäre, deren Liebe und Unterstützung mich bei allem, was ich im Leben tue, begleiten. Mein Erfolg ist darauf zurückzuführen, dass sie jeden Tag an mich glauben. Meine tiefste Dankbarkeit und Liebe gilt meiner wundervollen Freundin Jennifer Sowa. Ich bin dir für immer dankbar für deine bedingungslose Liebe, deine unermüdliche Unterstützung und jeden Tag, den ich mit dir zusammen verbringen darf!

Table of Contents

List of Abbreviations

BS	Black-Scholes
CE	Certainty Equivalent
CM	Constant Mix
CMR	Constant Management Rule
CPT	Cumulative Prospect Theory
CRRA	Constant Relative Risk Aversion
EU	Expected Utility
FRTB	Fundamental Review of the Trading Book
GBM	Geometric Brownian Motion
GMDB	Guaranteed Minimum Death Benefit
HJB	Hamilton-Jacobi-Bellman
MCPT	Multi-Cumulative Prospect Theory
MRRG	Minimum Return Rate Guarantee
MC	Markov Chain
MV	Mean-Variance
PI	Portfolio Insurance
RM	Risk Measure
SDE	Stochastic differential equation
SFP	Shortfall Probability
SST	Swiss Solvency Test
VaR	Value at Risk
VAR	Vector Auto-regressive
VMR	Variable Management Rule

List of Tables

List of Figures

Chapter 1

General Introduction

Quantitative risk management is arguably the most fundamental step in every risk management process. It builds the theoretical foundation to quantify and analyze risks in every company. In the present thesis, each chapter contributes to the literature by dealing with aspects of fundamental problems and discussions in quantitative risk management. Special emphasis is put on three major topics: First, on the shortfall probability resp. the risk measure concepts in the context of portfolio allocation problems. Then on the impact of periodic premium payments on pricing, the shortfall probability and the optimal portfolio allocation of an insurance contract. Finally, on time-inconsistent optimal portfolio allocations that stem from uncertainty aspects because of different market states (regimes).

The usage of risk measures in quantitative finance and insurance is at least since the axiomatic foundation of Artzner et al. (1999) omnipresent. Companies use them to quantify the riskiness of an investment and even to optimize their asset allocation under some shortfall constraints. Also the regulatory discusses different risk measures and their meaningfulness to calculate capital requirements s.t. the companies and investors' are secured against a shortfall. The companies have the motivation to reduce the capital requirements s.t. the capital commitment costs do not exceed their earnings. This leads to a conflict of goals between the regulatory, that wants to protect the investors' capital and the companies that want to maximize their profit. Since the introduction of the Value at Risk to the public by J.P. Morgan in 1994 and its implementation in the first Basel accord (Basel I) to calculate the capital requirements of the banks' market risks (cf. Basel Committee on Banking Supervision (1996)), the Value at Risk has been an indispensable tool in risk management for insurance companies and banks. There have been many discussions in academia, the regulatory frameworks and amongst practitioners about the benefits and drawbacks of the VaR and possible alternative risk measures as the Expected Shortfall which has become more and more prominent since its appearance in Artzner et al. (1999). Concerning the regulatory frameworks for banks, the Basel accords, there have been many revisions since the Basel I framework: The current Basel III accord after the so-called fundamental review of the trading book (cf. Basel Committee on Banking Supervision (2013)) accounts for some of the arguments from academia and have changed the measurement of market risks from the VaR towards the Expected Shortfall. On the insurance side, the current regulatory framework is Solvency II. Here the capital requirements are still calculated with the VaR, but the discussion is ongoing and the results of a review from the year 2020 are expected soon (cf. EIOPA (2020)).

Different payment structures of the policyholder and guarantee features are other important aspects in the context of risk quantification, especially for insurance companies. In the ongoing low or even negative interest rate environment these two components become even more important.[1] The companies may no longer be able to guarantee all of the policyholder's contributions and need to reduce the guaranteed interest rate even further. Moreover, alternative guarantee concepts and innovative products are needed to compensate for these problems. This results in a more risky investment policy of the insurance company s.t. there exists the possibility that parts of the insureds' contributions may default.[2] Another important aspect for life insurance companies is the perspective of the insured's benefits: Traditionally, life insurance contracts in academia are mostly calculated and priced with a single, upfront contribution of the insured. But in reality insureds often want to split the contributions over time instead of paying an upfront premium payment. Thus, the company's room for maneuver is even more reduced and the premium splitting affects the pricing of the contracts as pointed out by Bernard et al. (2017).

Finally, uncertainty aspects are omnipresent in the companies' daily business and thus of high relevance for their risk management: Regulatory requirements may change because of new capital standards. Traditional, rather conservative assets may become more risky or even worthless because of changing environmental awareness and including sustainability aspects in the analysis of the company.[3] Even entire economic systems suddenly may change because of pandemic risks as seen in the ongoing COVID 19 pandemic. [4] This is a current and really powerful example of parameter uncertainty resulting from uncertainty in the world. The modeling of the company's assets becomes even more complicated than before because of uncertain regimes. Thus, it is reasonable to investigate the optimal investment policy under

[1] The effects of low-interest rate environment on life insurers are discussed in Berdin and Gründl (2015).

[2] The major insurance company in Germany, the Allianz, is no longer offering a full guarantee on new life insurance products (cf. Handelsblatt (2020)).

[3] This behavior that assets become worthless because of more prone environmental standards is called *stranded assets*. A discussion on this topic is e.g. given in Bos and Gupta (2019) and Caldecott (2017). A recent paper on uncertainty effects induced by climate change is Barnett et al. (2020) and the mentioned papers within this work.

[4] The paper of Ali et al. (2020) analyzes the reaction of the financial markets to the COVID-19 pandemic and presents the volatility evolution of corresponding markets.

different regime possibilities that influence the drift and the volatility of the assets. The uncertainty might be that great, that even probabilities for different regimes are not observable anymore.

We contribute to these, as mentioned, highly relevant topics in quantitative risk management in insurance and finance in the thesis. Special emphasis is put on the pricing of a typical participating life insurance product, the minimum return rate guarantee (MRRG) contract including default risk, and solving the resulting asset allocation problem including a shortfall probability constraint. Moreover, another interesting aspect arises if we not only allow for upfront premium payments in this context: we investigate the impact of a periodic payment structure combined with a tool that controls the investment fraction in the risky asset on the pricing, the optimal asset allocation and risk management. Finally, we analyze a more general asset allocation problem in finance where different possibilities for a regime are included and time-inconsistency arises because of a pre-commitment investment strategy. We derive the optimal investment decision as also study the impact of the value of information about the regime.
Methodically, the thesis contributes to the literature on model-free asset pricing, payoff modification and asset allocation in Chapter 2. Moreover, Chapter 3 addresses the literature of pricing, risk management and portfolio planning of periodic payments contracts with guarantee features. Chapter 4 contributes to parameter uncertainty, ambiguity and optimal asset allocation. Based on these objectives, the structure of the thesis is as follows:

Chapter 2 initiates the thesis by presenting the most common risk measure concepts and comparing the two risk measures Value at Risk and Expected Shortfall in detail. The cost-minimal payoff modification concept is introduced and applied to a payoff s.t. a VaR constraint is fulfilled. Finally, the main result is presented: We investigate participating life insurance contracts with minimum return rate guarantees where we include default risk in our analysis. Without default risk the insured receives the maximum of a guaranteed rate and a participation in the investment returns. With default risk, the payoff is modified by a default put implying a compound option. We represent the yearly returns of the liabilities in a model-free manner by a portfolio of plain vanilla options. In a Black and Scholes model, the optimal payoff constrained by a maximal shortfall probability can be stated in closed form. Due to the completeness of the market, it can be implemented for any equity to debt ratio. An analysis on the impact of risk measures constraints on the optimal solution of portfolio allocation problems in terms of a literature overview concludes the chapter.

Building upon the insights of the previous chapter, Chapter 3 considers MRRG contracts in a periodic premium payment setting under a terminal guarantee feature. We find a representation of the periodic account value based on just one splitting factor (i.e. two premium dates). This allows us to analyze qualitatively the effect of periodic premium payments in a stylized setup. In a two-period Black-Scholes Model, we first discuss the impact of periodic premium payments on the guarantee costs. Splitting the premium fraction instead of paying an upfront contribution leads to a more risky contract for the insurance company. By introducing a management rule that controls the investment fraction in the risky asset, the insurance company can react to the risk shift introduced by the periodic contributions of the insured. Analyzing the impact of the splitting factor and the management rule, we find quasi closed-form solutions for the guarantee costs. Moreover, we point out that the management rule can be used by the insurance company to reduce the risk of periodic contributions: The required capital for fulfilling a SFP can be dramatically be reduced if a management rule is implemented. Finally, we analyze the expected utility of the insured. We study the optimal splitting factor for a given management rule and investment fraction that maximizes the expected utility. We find that splitting the contributions has a huge impact on the optimal strategy: For an investment fraction that is greater than the Merton fraction the splitting factor is used to adapt the investment fraction to be as close as possible to the Merton one. If we additionally include a management rule, the overall optimal investment fraction even differs from the Merton solution. We complete the chapter with a literature overview by investigating the impact different guarantee features on the optimal portfolio planning.

Next, Chapter 4 contributes to the literature on optimal asset allocation in the context of parameter uncertainties. First, we discuss the impact of ambiguity and learning on the optimal investment decision by reviewing the most important literature on portfolio planning in this research area. Afterward, we consider a stylized setup of an investment decision to shed light on the impact of time-inconsistency on pre-commitment strategies. First, we use a double risk situation where the outer risk is given by a simple a-priori lottery and the inner risk situation is a regime coinciding with the classic Merton problem. While in the myopic case, the weights resemble the regime probabilities, there is a time-dependent probability reduction of the good regime, i.e. as the investment horizon increases, the worst-case regime gets more important. We also account for ambiguity about the "success" probability of this lottery. Preferences towards risk and ambiguity are modeled using the smooth ambiguity approach by Klibanoff et al. (2005) under a double power utility assumption. Again, we can separate the effects of the two risk situations as well as the ambiguity aversion. We explain why the impact of time-inconsistency gets more ambiguous since varying the ambiguity situation may also change the risk situation. At the end of the chapter, we extend the previously described model by including the possibility of one regime switch over the investment horizon and we review the portfolio planning literature that includes regime-switching and analyze its impact on the optimal portfolio allocation.

Finally, Chapter 5 summarizes the thesis and gives an outlook on further research.

Chapter 2

Risk Measure Concepts, Shortfall Probability Constraints and Optimal Design of Quantile Guarantees

The chapter discusses the concept of shortfall probability (SFP) constraints and the impact of default risk in the context of portfolio optimization. The SFP is a typical and well-used benchmark in insurance and finance to calculate the riskiness of a portfolio resp. to limit the risk of a potential shortfall. It can be classified into the so-called shortfall risk measures. In general, risk measures are used to determine the risk of a financial position resp. the required capital s.t. a position is acceptable.[5] The regulatory frameworks for banks, the Basel accords, as also the ones for the insurance companies, Solvency II and the Swiss Solvency Test (SST), require from the companies that their portfolios fulfill a SFP benchmark s.t. the portfolio has a limited shortfall percentage.[6] To see how different risk measures are connected and to get an overview of different risk measure classes it is convenient to start the chapter with a review of prominent risk measure concepts.

In general, companies are interested in minimizing their required capital resp. their costs. But they should also take their investors' interests into account s.t. they are committed to stay as long as possible with them. Therefore, it is also of interest for them to maximize their investors expected utility. Thus, it is reasonable for the company to determine the terminal payoff that maximizes the expected utility of the investor, including a terminal SFP constraint s.t. the regulatory requirements are

[5] For a detailed discussion on acceptance sets in the risk measure context we refer to Artzner et al. (1999) resp. to Artzner et al. (2009).

[6] The calculation and concrete assumption of a shortfall probability differs. For banks, the Basel regulatory prescribes different SFPs for market risks, interest rate risks and operational risks. The Solvency regulatory for insurance companies in Europe differentiates between many risk types but prescribes just one SFP on the aggregated risks. For more details, we refer to the following sections.

fulfilled. [7] The general solution for the mentioned problem with a utility function that implies constant relative risk aversion is given by Merton (1971). Modifying this payoff s.t. the SFP is fulfilled is thus of great interest for the companies. This modification mechanism is the so-called *cost-minimal payoff modification*, first introduced by Dybvig (1988a) and Dybvig (1988b). The cost-minimal modification should have a positive impact on both, the financial company and the investor: For the first one, because the shortfall restriction is fulfilled and for the second one, because she receives the optimal payoff under these circumstances. But in reality, the investor suffers from this solution as she is not protected on the bad states of the expected utility-maximizing payoff. This mitigates the idea of protection that should be fulfilled with the SFP constraint. A detailed discussion is given in the second part of this chapter.

Analyzing (life) insurance companies in particular, there is the current problem of a low-interest-rate phase, s.t. traditional products cannot serve the insureds' needs anymore. Thus, there is the need for insurance companies to model innovative contract designs. In so-called participating life insurance contracts with minimum return rate guarantees (MRRG), the insured participates with her payments in the asset side of the insurance company: The company invest the premium payments in risky resp. risk-free assets and at the end of the contract period the insured receives the maximum of a terminal guarantee and the terminal asset result. This payoff has the possibility to default if the asset strategy has performed not well and is smaller than the guaranteed rate. We provide a model-independent analysis of this payoff under the possibility of a default. Furthermore, following the argumentation from above, it is interesting to solve the resulting expected utility maximization problem under a shortfall probability constraint. Stating the resulting optimal payoff, we draw the connection to the cost-efficient payoff modification and find the impeding behavior that the optimal payoff does not secure the insured on the bad states of the world.

Following this line of arguments, there exist three main aims in this chapter:

The first aspect is to present the state-of-the-art risk measure concepts and the comparison of the different regulatory frameworks on the level of risk measures to discuss the differences and commonalities of these concepts. The second aim is to present the cost-efficient payoff modification and motivate the idea behind it to understand how the protection of the insured resp. the investor is mitigated. Here we derive first results in a toy example and contribute to the literature. The third and main aim in this chapter is to present MRRG products under default risk: We derive

[7] Portfolio optimization resp. portfolio planning itself dates back to Merton (1971). There are many possibilities for the optimization argument, e.g.we can maximize over all investment fractions that are invested in the risky asset or over all possible premium payment schemes. In Chapter 3 we analyze a expected utility maximization problem where we determine for a given investment fraction the optimal premium payment fraction. In Chapter 4 we discuss a expected utility maximization problem over the investment fraction in the risky asset in a situation under uncertainty.

model-free insights and find an application of the cost-efficient payoff modification in an optimization problem, where we maximize the expected utility of the insured under a shortfall constraint by determining the optimal payoff structure. Furthermore, we give a literature overview on optimal portfolio planning under risk measure constraints.
We proceed as follows. First we introduce and review the most important concepts and results in the theory of risk measures that are of importance for the thesis. Here we rely on the path-breaking work of Artzner et al. (1999) and the resulting papers of Acerbi (2002), Föllmer and Schied (2002) and Frittelli and Gianin (2002). We discuss the most common risk measures in practice, the VaR and ES and embed them into a broader class of risk measures, the so-called distortion and spectral risk measures. Afterward, we present the cost-minimal payoff modification mechanism by Dybvig (1988a) and Bernard et al. (2014). Using that technique we construct modified payoffs that fulfill shortfall probability requirements imposed by a regulator but impede the idea that the bad states of the world are the ones that need to be secured. Our findings can even be generalized by applying the work of Wei (2018). The main result is achieved by analyzing minimum return rate guarantee contracts under default risk. We give model-free insights of the pricing and the SFP of these insurance contracts and find in a Black-Scholes application that quantile guarantees maximize the insureds expected utility and fulfill the imposed SFP. At the end of this chapter we discuss the impact of different risk measure constraints resp. different optimization arguments on the optimal expected utility maximizing solution by giving a literature overview.

2.1 Concepts of Risk Measures

One of the most important topics in the finance and insurance industry is to account for the corresponding risks of the firms. Risky investments might lead to high losses that affect the performance of the bank resp. the insurance company and might endanger the investors' or insureds' contributions. Thus, an intact risk management should be at the heart of every company. Hereby the risks are modeled with random variables, so-called loss variables or profit & loss variables, depending on the context.[8]
Some company goals might contradict prudent risk management or the companies themselves underestimate some of their taken risks. Here regulatory authorities come into play. They want to protect with their frameworks the investors and insureds on the one hand and want to force the companies to protect themselves against high risks on the other hand. For this protection the companies have to calculate

[8] The modeling of the random variables itself is not an easy task. In practice the modeling is based on time-series samples from recent years which is prone to errors.

capital requirements to absorb a large loss. Depending on the company there are different regulatory frameworks: For banks there exists the Basel accords, currently the revised version of Basel III but there are ongoing discussions s.t. a new regulatory framework in terms of Basel IV seems not that far away. The need for capital requirements can be found in Basel Committee on Banking Supervision (2020). For the insurance side there is the Solvency II framework for European insurance companies. Switzerland as a non-European country has its own insurance regulatory, the Swiss Solvency Test (SST). The need for capital requirements in the Solvency II context is given in Chapter 3, Art. 37, §1 in the 2009 directive of the European Parliament about Solvency II (cf. EC (2009)). In the technical documentation of the SST, capital requirements are discussed in Section 2.1 (cf. Swiss Financial Market Supervisory Authority (2006)).

All different regulatory frameworks imply that the companies' risks should be condensed into one number to calculate the capital requirement based on the risks of the company. As mentioned before the risks are modeled with random variables and thus because of the randomness is it not possible to find one perfect function that can handle this task for every risk in every company equally good and reliable. At this point we start with our analysis and introduce the concept of risk measures. A risk measure is a mapping that assigns every random variable X a real number c. This number can be interpreted as the capital requirement. Of course there are many possibilities to choose such a function. There exists e.g. the expected value or the variance that exactly does the desired mapping. Thus there is the need to conceptualize risk measures and discuss desired properties that a risk measure should fulfill. On the basis of these properties we can evaluate if a risk measure should be used to calculate capital requirements for banks and insurance companies or not. Therefore, we first present the axiomatic of risk measures in the following subsection and review the most common concepts of risk measures afterward.

2.1.1 The axiomatic of risk measures

As discussed before, all information of a random variable X are condensed into one number to quantify the risk of X. This is problematic because one number is not able to reflect the whole behavior of the corresponding risk. But if we have to rely on that calculated number the corresponding functional should at least fulfill some useful and meaningful properties. The quantification and discussion of these properties will be the main aim in this section. Before that, let us start with an example that will be of importance trough out the whole thesis.

Intuitively the first idea to quantify the risk of a random variable X is to ask for the probability that X is smaller than a certain value s resp. the loss is greater than s, depending on the definition of the random variable X. Mathematically this can be described by

$$\mathbb{P}(X < s) \text{ resp. } \mathbb{P}(X > s).$$

This probability is the so-called *shortfall probability* (SFP). The next natural question that arises is how much capital c we need to add to the random variable X s.t. the SFP can be minimized to at least a value of ε.[9] This is described by

$$\mathbb{P}(X + c < s) \le \varepsilon \text{ resp. } \mathbb{P}(X + c > s) \le \varepsilon.$$

This is the so-called SFP-concept. Determining the capital c is an important task for many different research questions. It will be a major component in the analyses of the following chapters. Of course this concept has also some disadvantages when it comes to the point that a shortfall occurs: we do not know how severe the shortfall is. The SFP-concept is closely related to the risk measure Value at risk that will be discussed in details in the next section.
Another simple method for determining the risk could be to calculate the volatility of the corresponding risk X. But if X has a profit and loss distribution also the gains would be used to determine a number for the corresponding risk. This would distort the risk situation of X. Thus properties for an appropriate definition of risk measures are important and meaningful.

For the definition of risk measures we work on an atomless probability space $(\Omega, \mathcal{F}, \mathbb{P})$, i.e. there exists at least one continuously distributed random variable on that space. Furthermore, we define $L^0(\Omega, \mathcal{F}, \mathbb{P})$ as the set of all random variables in that probability space. The set of all integrable random variables is denoted with $L^1(\Omega, \mathcal{F}, \mathbb{P})$ and the set of all bounded random variables is given by $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$. Moreover, let $\mathcal{X} \subseteq L^0(\Omega, \mathcal{F}, \mathbb{P})$ be a vector space of random variables over $(\Omega, \mathcal{F}, \mathbb{P})$. We interpret a random variable $X \in \mathcal{X}$ as a loss variable, i.e. if $X_1 \ge X_2$ $\mathbb{P} - a.s.$ the risk X_1 has a higher loss than X_2 and thus it is more risky. With these assumptions we can define monetary risk measures. The definition already dates back to Artzner et al. (1999).

Definition 2.1 (Monetary Risk Measure)
Let $\mathcal{X}$ with $\mathbb{R} \subseteq \mathcal{X}$ be a $\mathbb{R}$ vector space of random variables over the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The mapping $\rho : \mathcal{X} \to \mathbb{R}$ is called a ***monetary risk measure*** *if it holds*

(i) For $X_1, X_2 \in \mathcal{X}$ with $X_1 \ge X_2$ $\mathbb{P} - a.s.$ it holds $\rho(X_1) \ge \rho(X_2)$.

(ii) ρ is cash-invariant, i.e. for $X \in \mathcal{X}$ and $c \in \mathbb{R}$ it holds $\rho(X + c) = \rho(X) + c$.

[9] In general it is not possible to reduce the SFP to zero. The capital requirement would be infinite.

The outcome of the risk measure ρ is interpreted as the minimum capital requirement. This is one of the most important regulatory requirements for banks and insurance companies in the Basel resp. the Solvency accords. Notice that the definition of a monetary risk measure is given for loss variables X which takes values greater than zero, e.g. damage claims. In the finance context loss variables are connected with negative values X. In this case a definition of a monetary risk measure R can be found e.g. in Föllmer and Schied (2016) and is given by

(i) $X_1 \geq X_2 \ \mathbb{P} - a.s. \Rightarrow R(X_1) \leq R(X_2)$

(ii) $R(X + c) = R(X) - c.$

These interpretations are convertible into each other: if ρ is a monetary risk measure as defined in Definition 2.1, then $R(X) = -\rho(X)$ or $R(X) = \rho(-X)$ defines a monetary risk measure in the sense of Föllmer and Schied (2016).
This is the most canonic definition of a risk measure. To achieve a more detailed classification we introduce in the next step the so-called convex risk measures which have been first introduced by Föllmer and Schied (2002) and Frittelli and Gianin (2002).

Definition 2.2 (Convex Risk Measure)
Let $\mathcal{X}$ with $\mathbb{R} \subseteq \mathcal{X}$ be a $\mathbb{R}$ vector space of random variables over the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. A monetary risk measure $\rho : \mathcal{X} \to \mathbb{R}$ is called a ***convex risk measure*** *if ρ is a convex mapping.*

This definition implies the following: Let ρ be a convex risk measure, X_1 and X_2 risks and $\lambda \in (0, 1)$, then it holds

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

This leads to one of the most important and desired property of risk measures as also often discussed in theory and practice, the so-called subadditivity. Before discussing this property, we define the class of coherent risk measures.

Definition 2.3 (Coherent Risk Measure)
Let $\mathcal{X}$ with $\mathbb{R} \subseteq \mathcal{X}$ be a $\mathbb{R}$ vector space of random variables over the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. A monetary risk measure $\rho : \mathcal{X} \to \mathbb{R}$ is called ***coherent risk measure*** *if it holds*

(i) ρ is a convex risk measure.

(ii) ρ is positive homogene, i.e. for every random variable $X \in \mathcal{X}$ and $\lambda > 0$ it holds $\rho(\lambda X) = \lambda\rho(X)$.

If ρ is a coherent risk measure and $X_1, X_2 \in \mathcal{X}$, the **subadditivity property** is fulfilled, i.e.

$$\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2).$$

Subadditivity can be interpreted as the property which accounts for diversification. The capital requirement for a portfolio of risks should be less or equal to the capital requirements of the risks measured separately. In the literature this property is widely accepted as one of the key properties of a risk measure. However, there also exist some critics regarding the this property: Dhaene et al. (2008) show that the subadditivity axiom can lead to an increase in the shortfall risk by a merger and also Rootzen and Klüppelberg (1999) and Kou et al. (2013) argue that the property of subadditivity is not indispensable and might be relaxed. The subadditivity discussion will be taken up again in the next section.

2.1.2 Value at Risk and Expected Shortfall as most common risk measures in theory and practice

In this subsection we want to discuss the most common risk measures that are used in theory and practice, the Value at Risk (VaR) and the Expected Shortfall (ES). Before we start with the discussion of the VaR and the ES let us take a glance at the definitions of them and embed them into the coherence axioms.

Definition 2.4 (Value at Risk)
*For $\alpha \in (0,1)$ the mapping $VaR_\alpha : \mathcal{X} \to \mathbb{R}$ defines the risk measure **Value at Risk**, where*

$$VaR_\alpha(X) := \inf\{x \in \mathbb{R} : F_X(x) \geq \alpha\} = \inf\{x \in \mathbb{R} : \mathbb{P}(X > x) \leq 1 - \alpha\} = F_X^{\leftarrow}(\alpha).$$

The function $F_X^{\leftarrow}(\alpha)$ denotes the left-inverse function of the random variable X. For further properties of $F_X^{\leftarrow}(\alpha)$ we refer to Embrechts and Hofert (2013). We can see in the definition that the VaR is a quantile of the random variable X. This is an easy to calculate value and one of the reasons why the VaR is so famous in practice. Furthermore, the connection to the SFP-concept mentioned in the beginning of this section is obvious: The VaR denotes the capital c that is required s.t. the random variable $X - c$ will not be greater than zero with probability $1 - \alpha$.

Remark 2.1
Remember that $\mathcal{X}$ is a set of loss variables. In this case $VaR_\alpha(X)$ denotes the minimum capital requirement which is needed s.t. the loss variable exceeds this value with a probability of only $1 - \alpha$. Or to state it the other way around: with a probability of α the loss variable has a smaller outcome than $VaR_\alpha(X)$.
In the context of profit variables $VaR_\alpha(X)$ denotes the profit which does not fall

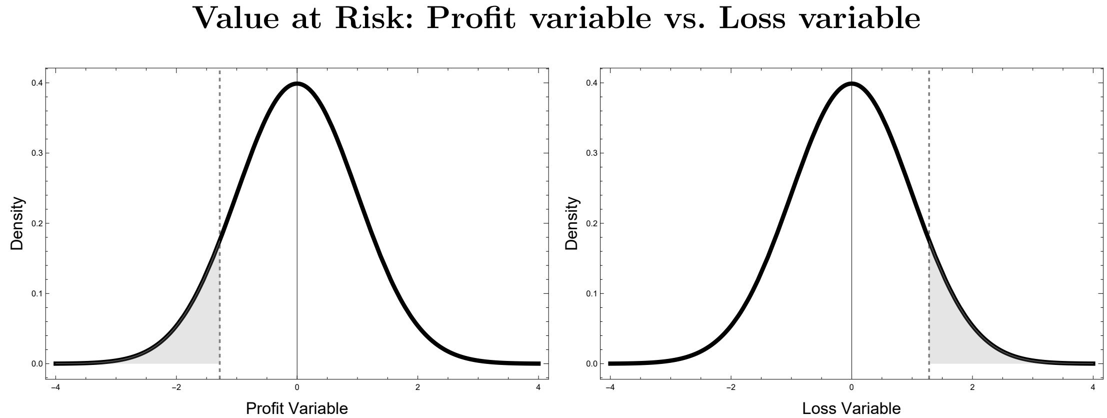


Figure 2.1: Distinction between the VaR of a loss and a profit variable. The gray dashed lines refer to the VaR with $\alpha = 0.9$ resp. $\alpha = 0.1$.

below with probability α. Thus in the context of loss variables we are interested in the 'high' quantile regions of the distribution of X, in the context of profit variables the 'low' quantile regions are of interest. Figure 2.1 shows the relation of loss and profit variables and the corresponding VaR levels.

The VaR is not a convex risk measure and thus not coherent which is already shown and discussed in Artzner et al. (1999). Especially the subadditivity property is not fulfilled. To given an example let X_1 and X_2 be i.i.d. random variables with the outcomes and corresponding probabilities p

$$X_i = \begin{cases} 0, & p = 0,99 \\ 1, & p = 0,01 \end{cases}, i = 1, 2. \tag{2.1}$$

Then $VaR_{0,99}(X_i) = 0$, but $VaR_{0,99}(X_1 + X_2) = 1$.
In contrast to this the ES is a coherent measure of risk defined in the following.

Definition 2.5 (Expected Shortfall)
Let $\mathcal{X} \subseteq L^1(\Omega, \mathcal{F}, \mathbb{P})$ be a set of loss variables and $\alpha \in (0, 1)$. The risk measure ***Expected Shortfall (ES)*** *is defined by*

$$ES_\alpha(X) := \frac{1}{1-\alpha} \int_\alpha^1 VaR_\beta(X) \mathrm{d}\beta = VaR_\alpha(X) + \frac{1}{1-\alpha} \int_{VaR_\alpha(X)}^\infty 1 - F_X(x) \mathrm{d}x,$$

where the second equation holds when X is a continuously distributed random variable.

Again notice that this is the definition for loss variables X. In the context of profit variables we refer to Föllmer and Schied (2016). Here the ES is defined by

$$ES_\alpha(X) := -\frac{1}{\alpha}\int_0^\alpha VaR_\beta(X)\mathrm{d}\beta.$$

For a discussion on the coherence properties of the ES we refer e.g. to Acerbi and Tasche (2002). Especially the subadditivity property is fulfilled s.t. the ES accounts for diversification. Moreover, the ES at level α is more sensitive to the shape of the tail of the loss distribution compared to the VaR because it evaluates the expected return on the portfolio in the worst α percent of cases. Thus the ES is in a continuous setting often defined as

$$ES_\alpha(X) = \mathbb{E}[X|X > VaR_\alpha(X)].$$

This definition goes back to Rockafellar and Uryasev (2000). Using this definition, it is easy to see that the ES prescribes a more conservative capital requirement than the VaR. These are two important properties why the ES is preferred by many academics over the VaR. An axiomatic foundation for the ES based on portfolio capital requirement calculation is presented in Wang and Zitikis (2021).

As stated before different capital requirements have to be fulfilled depending on the company. European insurance companies are settled under the regulatory framework of Solvency II established in the year 2006. The overall capital requirements are divided into submodules, e.g. market risks, life insurance risks, non-life risks etc. Every submodule in the standard approach (standard formula) is measured with a one-year VaR with a confidence level of $99,5\%$.[10] In the year 2020 the EIOPA reviewed the Solvency II framework and a revised version is planned to be published in 2022.[11] Switzerland as a non-European country has its own insurance regulatory, the so-called Swiss Solvency Test (SST). It uses a multi-period risk measure based on the ES.[12]

On the banking side there are the Basel accords as the European regulatory framework. The current version is given by Basel III. One of the most extensive revisions

[10] An overview about the Solvency II process and a critical analysis of that framework is given in Eling et al. (2007) and Doff (2008). The paper of Floreani (2013) criticizes the risk measure and capital requirement approach in Solvency II and Höring (2013) discusses the impact of the market risk on the capital allocation of insurance companies. Gatzert and Martin (2012) quantify the market and credit risk under Solvency II by comparing the standard approach with internal models. BaFin (2018) discusses the standard formula of Solvency II and highlights the problem of interest rate risks: there is no possibility to account for negative interest rates which is crucial in the current low-interest-rate environment. Thus a detailed discussion of interest rate risks is of great concern in the actual development and there are still many open research questions to answer.

[11] For more insights we refer to EIOPA (2020).

[12] Filipović and Vogelpoth (2008) discuss in detail the SST risk measure and find that it is not coherent. A general discussion of the SST can e.g. be found in Eling et al. (2008).

has been carried out in Basel Committee on Banking Supervision (2013), the so-called "Fundamental Review of the Trading Book" (FRTB).[13] Before the review, market risks have been measured with the VaR at the 99% confidence interval. After the review, market risks are now measured with the ES at level $97,5\%$, but credit risks are still calculated with the VaR. A discussion on this topic is given in Bugár and Ratting (2016) as also in Laurent et al. (2016). For an analysis of credit risks in the Basel III framework measured with the ES we refer to Osmundsen (2018). Kellner and Rösch (2016) highlight the consequences of that change for the model risk. A comparison of regulatory requirements for Solvency II and Basel III for and after the fundamental review of the trading book can be found in Laas and Siegel (2017) and Gatzert and Wesker (2012). The differences of insurance regulatory frameworks (Solvency II, Swiss Solvency Test (SST) and US-RBC) is provided by Holzmüller (2009). Comparing these regulatory frameworks, the two risk measures VaR and ES are commonly used in practice and highly relevant for banks and insurance companies. Thus we focus on these two risk measures. Later we discuss alternatives for regulatory practice.

Many research papers consider the VaR and ES from a more practical point of view by analyzing them concerning the regulatory frameworks Solvency II, SST and the Basel III accords. An early paper in this stream of literature is provided by Berkowitz and O'Brien (2002), who are the first that analyze the performance of Value-at-Risk models for large trading firms. They find modeling constraints and regulatory practices that harm the VaR calculation. Kaplanski and Levy (2007) study the efficiency of the VaR regulation in the Basel II accord and find that there is an optimal level of required eligible capital from the regulators point of view. A comparison of bank risk measures before, during and after the financial crisis is presented in O'Brien and Szerszen (2017). They find that banks' VaR is conservative if times are normal but understate risks in a period of market instability. Chang et al. (2019) compares the VaR and ES in the Basel III accord regarding market risks and find that policymakers prefer the ES using a stochastic dominance approach. A general discussion of the FRTB and the Basel 3.5 accord is presented by Embrechts et al. (2014).

Coming back to the subadditivity axiom we have seen that the VaR does not fulfill this property in general. But there exist some important exceptions where the feature is met when the random variables fulfill some distributional assumptions.[14]
Starting with Bernoulli distributed random variables, Hofert and McNeil (2015) find

[13] The banking book was completely revised in 2013. For all detailed changes in the FRTB we refer to Basel Committee on Banking Supervision (2013).

[14] In reality the distributions of random variables have to be estimated s.t. there is the potential of an incorrect model. The testing on how accurate the corresponding risk measure model has been in the last period is called backtesting. This method is highly relevant for practitioners. In the context of our two discussed risk measures the VaR has thereby better properties: Gneiting (2011) shows that the ES is not an elicitable risk measure. A formal definition of elicitability

a necessary and sufficient condition for a homogeneous portfolio of i.i.d. Bernoulli random variables s.t. the VaR is subadditive.

Proposition 2.1 (Subadditivity VaR - i.i.d. Bernoulli Variables)
Let $X_j \sim B(1, p)$ be i.i.d. Bernoulli variables with $j \in \{1, 2, \cdots, d\}$ with $d \geq 2$ and $p \in [0, 1]$. Then $VaR_\alpha(\sum_{i=1}^{d} X_j)$ is superadditive if and only if $(1-p)^d < \alpha \leq (1-p)$.

For a proof of Proposition 2.1 we refer to Hofert and McNeil (2015) (Theorem 2.1). This proposition is of importance because credit risks resp. the default of a credit can be modeled via a Bernoulli random variable. If the quantile α is large enough or the default probability p is big enough, the VaR is subadditive. In equation (2.1) we used two i.i.d. Bernoulli random variables with $p = 0,01$ and $\alpha = 0.99$. Because $\alpha = 1-p = 0.99$ we have constructed an example where the VaR is superadditive. If we choose the level $\alpha = 0.98$ or $\alpha = 0.995$, the VaR is subadditive. Even in the case of a heterogeneous portfolio of Bernoulli variables it is possible to state necessary and sufficient conditions s.t. the VaR is subadditive. This finding is given in the next proposition. The proof can be found in Hofert and McNeil (2015) (Theorem 3.1).

Proposition 2.2 (Subadditivity VaR - Heterogeneous Bernoulli Variables)

Let $X_j \sim B(1, p_j)$ be independent Bernoulli variables with $j \in \{1, 2, \cdots, d\}$ with $d \geq 2$. Furthermore, let $0 =: p_{(0)} < p_{(1)} \leq p_{(2)} \leq \cdots \leq p_{(d)} < p_{(d+1)} := 1$ denote the ordered default probabilities. Then it holds

(i) $VaR_\alpha(\sum_{i=1}^{d} X_j)$ is superadditive for all $\prod_{j=1}^{d}(1-p_j) < \alpha \leq (1-p_{(d)})$.

(ii) $VaR_\alpha(\sum_{i=1}^{d} X_j)$ is subadditive for $0 \leq \alpha \leq \prod_{j=1}^{d}(1-p_j)$ resp. $1-p_{(2)} < \alpha \leq 1$.

Again, this proposition implies that the VaR is subadditive for Bernoulli variables whenever the level α is large enough. On the one hand these are good news for practitioners because in the Basel III accords, credit risks are measured with the VaR with a confidence interval of $\alpha = 0.999$. This level should be sufficiently large s.t. the risk measure stays subadditive for many credit risks. On the other hand when we have credit risks with a very good investment grade of AAA, the default probabilities are smaller than $0,1\%$ (see e.g. in Bundesbank (2019)). In this situation the VaR remains superadditive.

is given in Lambert et al. (2008). This missing property has been used by many authors in the past as a reason to say that the ES is not backtestable. But recent papers of Du and Escanciano (2017), Moldenhauer and Pitera (2019), Dimitriadis et al. (2019) and Kratz et al. (2018) provide some techniques to backtest the ES based on quantiles. For the VaR there are many backtesting procedures: Two great overviews of backtesting procedures are the papers of Campbell (2005) and Nieto and Ruiz (2016).

To analyze a more general situation we take a look at elliptical distributions.[15] Typical examples of these distribution classes are the multivariate $t-$distribution, the logistic distribution, symmetric multivariate Laplace distributions or the multivariate normal distributions. The interesting aspect arises if we take a look at a $n-$dimensional elliptic distribution X with marginal distributions X_i s.t. $Var[X_i] < \infty$, for all $i = 1, \cdots, n$. For the set of all linear portfolios the VaR has the desired subadditivity property.

Proposition 2.3 (Subadditivity VaR - Elliptical Distributions)
Let X be a n-dimensional elliptical distribution as defined above and

$$\mathcal{P} = \left\{ Z = \sum_{i=1}^{n} \lambda_i X_i : \lambda_i \in \mathbb{R} \right\}$$

the set of all linear portfolios. For any $Z_1, Z_2 \in \mathcal{P}$ and $\frac{1}{2} \leq \alpha < 1$ it holds

$$VaR_\alpha(Z_1 + Z_2) \leq VaR_\alpha(Z_1) + VaR_\alpha(Z_2).$$

The proof of Proposition 2.3 is given in Theorem 1 of Embrechts et al. (2002).[16] Among other things because of this property elliptical distributions are popular in the context of finance and insurance. This has been first discussed by Owen and Rabinovitch (1983). But the assumption that the joint distribution of a portfolio follows an elliptical distribution might be not realistic because mostly only the marginal distributions of the risks are known.

Let us finally briefly comment on the situation where heavy-tailed distributions are involved in the calculation of capital requirements. Danielsson et al. (2006) presents approximations for downside risk measures such as the VaR and the ES for heavy-tailed risks with regular variation.[17] Under this assumption the corresponding risks have a tail similar to the Pareto distribution. With this result Danielsson shows that all downside risk measures order and measure heavy tail risks in a similar manner. Further insights and an empirical study on this topic is presented in Daníelsson et al. (2013).

[15] Elliptical distributions are a whole class of multidimensional probability distributions. They are strongly connected to the so-called spherical distributions: While spherical distributions are a generalization of the multivariate standard normal distribution $\mathcal{N}(0, I)$, elliptical distributions extend the multivariate normal $\mathcal{N}(\mu, \Sigma)$ distributions. Thus, the class of spherical distributions is a subset of all elliptical distributions. For the concrete mathematical definitions and an in-depth analysis we refer to McNeil et al. (2015) and to Fang et al. (1987).

[16] The statement of Proposition 2.3 can even be generalized to all positive-homogeneous, translation-invariant risk measures. See e.g. Theorem 8.28 in McNeil et al. (2015).

[17] A cumulative density function $f(x)$ varies regularly at minus infinity with tail index $\beta > 0$ if $\lim_{t \to \infty} \frac{f(-tx)}{f(-t)} = x^{-\beta}$ for all $x > 0$.

To ask the general question of what risk measure is an appropriate choice for regulatory and risk management there are many papers that discuss this aspect. Dowd and Blake (2006) analyze with an application to insurance companies what risk measure can come after the VaR, Chen (2018) does this analysis with a focus on the Basel accords. Moreover, Emmer et al. (2015) search for 'the best risk measure' in theory and practice. What all of these papers have in common is that they compare the VaR and ES and they all mention other interesting classes of risk measures that may be suitable alternatives: the classes of distortion and spectral risk measures. This will be the topic of our next subsection before we start with the analysis of the cost-efficient payoff modification and the MRRG contract design.

2.1.3 Classes of Risk Measures: Distortion and Spectral Risk Measures

Besides the two prominent risk measures VaR and ES, there are many possibilities to choose a risk measure. Two of the most important classes of risk measures are the so-called distortion and spectral risk measures. Especially we will see that the ES is a member of both classes, the VaR a member of the distortion risk measures family. The concept of distortion risk measures goes back to Wang et al. (1997) while spectral risk measures date back to Acerbi (2002). In this section we want to present the main findings regarding these classes of risk measures and show that they are under a simple condition equivalent.

Properties and Examples of Distortion Risk Measures

At first we shed a light on the class of distortion risk measures and start with an intuition behind these risk measures:

Let X be an integrable random variable, then the following integral representation for the expected value of X holds:

$$\mathbb{E}[X] = -\int_{-\infty}^{0} 1 - (1 - F_X(x))\mathrm{d}x + \int_{0}^{\infty} (1 - F_X(x))\mathrm{d}x,$$

where $(1 - F_X(x))$ is the so-called *tail function* of X. This representation holds because every random variable X can be represented as

$$X := X^+ - X^-,$$

where $X^+ := \max\{X, 0\}$ and $X^- := -\min\{X, 0\}$. We assume integrability of the risks, s.t. it holds $\mathbb{E}[X^+] < \infty$ and $\mathbb{E}[X^-] < \infty$ and therefore

$$\mathbb{E}_{\mathbb{P}}[X] := \mathbb{E}_{\mathbb{P}}[X^+] - \mathbb{E}_{\mathbb{P}}[X^-].$$

Moreover for $\omega \in \Omega$ it holds

$$X^-(\omega) = \int_{-X^-(\omega)}^{0} 1\mathrm{d}x = \int_{-\infty}^{0} \mathbf{1}_{\{X^-(\omega) \geq -x\}}\mathrm{d}x = \int_{-\infty}^{0} \mathbf{1}_{\{X(\omega) \leq x\}}\mathrm{d}x.$$

The last equation holds because $X^-(\omega) \geq -x$ for $x < 0$ implies that $X^+(\omega) = 0$ and thus $X(\omega) \leq x$. Conversely, if $X(\omega) \leq x < 0$, then $X^+(\omega) = 0$ and thus $X^-(\omega) \geq -x$. Analogue we can represent the variable X^+ as

$$X^+ := \int_{0}^{\infty} \mathbf{1}_{\{X(\omega) > x\}}\mathrm{d}x.$$

Using these two representations and the Theorem of Fubini, which allows to change the order of integration, it holds

$$\begin{aligned} \mathbb{E}_{\mathbb{P}}[X^+] &= \int_{0}^{\infty} \mathbb{P}(X(\omega) > x)\mathrm{d}x = \int_{0}^{\infty} 1 - F_X(x)\mathrm{d}x \text{ and} \\ \mathbb{E}_{\mathbb{P}}[X^-] &= \int_{-\infty}^{0} \mathbb{P}(X(\omega) \leq x)\mathrm{d}x = \int_{-\infty}^{0} F_X(x)\mathrm{d}x = \int_{-\infty}^{0} 1 - (1 - F_X(x))\mathrm{d}x. \end{aligned}$$

Finally, using the fact that we assume integrable random variables, the claimed result holds with $\mathbb{E}_{\mathbb{P}}[X] := \mathbb{E}_{\mathbb{P}}[X^+] - \mathbb{E}_{\mathbb{P}}[X^-]$.

We now create a distortion risk measure by distorting the expected value (more precisely the tail function) with a mapping $g : [0,1] \to [0,1]$, a so-called distortion function.

Definition 2.6 (Distortion Function)
An increasing function $g : [0,1] \to [0,1]$ with $g(0) = 0$ and $g(1) = 1$ is called a ***distortion function****. Additionally, if g is concave, g is called a* ***concave distortion function****.*

The next proposition shows us how a distortion function generates a distortion risk measure.

Definition 2.7 (Distortion Risk Measure)
Let $g : [0,1] \to [0,1]$ be a (concave) distortion function, such that the two integrals $\int_{-\infty}^{0} 1 - g(1 - F_X(x))\mathrm{d}x$ and $\int_{0}^{\infty} g(1 - F_X(x))\mathrm{d}x$ are real valued, then

$$\rho_g(X) := -\int_{-\infty}^{0} 1 - g(1 - F_X(x))\mathrm{d}x + \int_{0}^{\infty} g(1 - F_X(x))\mathrm{d}x$$

defines a positive homogeneous, law-invariant risk measure, a so-called ***(concave) distortion risk measure****.*

In the following concave distortion risk measures will be of great interest because if g is concave then $\rho_g(X)$ is a coherent and convex risk measure. In the literature there is also known an other representation for the expected value connected to the Value at risk:

$$\mathbb{E}[X] = \int_0^1 VaR_\alpha(X)\mathrm{d}\alpha. \tag{2.2}$$

With this connection we can reformulate the distortion risk measure $\rho_g(X)$ as follows:

$$\rho_g(X) = \int_0^1 VaR_\alpha(X)\mathrm{d}g(\alpha).$$

For a detailed discussion about quantiles and distortion risk measures we refer to Dhaene et al. (2012) and for a note on generalized distortion risk measures we refer to Hürlimann (2006). The following proposition gives a 1:1 connection between concave distortion risk measures, coherent risk measures, convex risk measures and the so-called **weighted AVaR** $\rho_\mu(X) := \int_0^1 ES_\alpha(X)\mu(\mathrm{d}\alpha)$, first introduced by Cherny (2006).[18]

Proposition 2.4 (Concave Distortion Risk Measures - Properties)
Let $\mathcal{X}$ be a vector space of random variables which contains a continuous random variable X^*, for $X \in \mathcal{X} \Rightarrow |X| \in \mathcal{X}$ and $\{X : X \sim B(1,t), t \in [0,1]\} \subseteq \mathcal{X}$.[19] Then the following statements are equivalent:

(i) ρ_g *is convex.*

(ii) ρ_g *is a coherent.*

(iii) g *is a concave distortion function.*

(iv) It exists a distribution μ *with* $\mu([0,1]) = 1$, *s.t.* $\rho_g(X) = \rho_\mu(X)$.

For the proof of Proposition 2.4 we refer to Theorem 4.82 and 4.88 in Föllmer and Schied (2016). In a later section we refer to the so-called weighted VaR first introduced by He et al. (2015). To eliminate confusion we use the following remark to compare the weighted VaR and weighted AVaR.

Remark 2.2 (weighted VaR vs. weighted AVaR)
As seen in Proposition 2.4 the weighted AVaR $\rho_\mu(X) = \int_0^1 ES_\alpha\mu(\mathrm{d}\alpha)$ *introduced by Cherny (2006) represents the class of concave distortion risk measures. In contrast,*

[18] Notice that in our case the terms Average Value at Risk and Expected Shortfall coincide.
[19] Choose for example $\mathcal{X}$ as the set of all essentially bounded random variables.

the weighted VaR $\rho_\mu(X) = \int_0^1 VaR_\alpha \mu(\mathrm{d}\alpha)$ *represents a much broader class of risk measures s.t. the weighted AVaR is a special case of the weighted VaR.*[20]

Let us take a look at some examples of distortion risk measures. Setting the distortion function to

$$g_1(x) := 1_{(1-\alpha,1]}(x)$$

we receive $\rho_{g_1}(X) = VaR_\alpha(X)$, i.e. the VaR is a distortion risk measure with distortion function g_1. Obviously g_1 is not concave s.t. the VaR is not coherent as discussed before. For the distortion function

$$g_2(x) := \min\left\{\frac{x}{1-\alpha}, 1\right\}$$

we receive $\rho_{g_2}(X) = ES_\alpha(X)$. Notice that g_2 is a concave distortion function and thus coherent. A third example of distortion risk measures is the so-called **Wang transformation** which can be traced back to Wang (2000). Using the concave distortion function

$$g_{wang}(t) := \Phi(\Phi^{-1}(t) - \Phi^{-1}(q)), \text{ for some } q \in \left(0, \frac{1}{2}\right],$$

where Φ denotes the distribution function of a standard normal distribution, $\rho_{g_{wang}}(X)$ is called the Wang transformation risk measure. A remark of Adam et al. (2008) gives a nice interpretation of this distortion risk measure, if the risk X follows a normal distribution $\mathcal{N}(\mu, \sigma^2)$. Under this assumption one can show that it holds

$$\rho_{g_{wang}}(X) = \mu + \sigma\Phi^{-1}(q).$$

This is exactly the Value at Risk of X to the level q, i.e. with the Wang transformation we can force a valuation of the risk X to the level of $VaR_q(X)$, whenever X is $\mathcal{N}(\mu, \sigma^2)$ distributed. The Wang transformation will be of great interest in the next section in the context of cost-minimal modifications of risk measures.

Spectral Risk Measures and their relation to Distortion Risk Measures

In the literature, spectral risk measures are widely discussed and first introduced by Acerbi (2002). Adam et al. (2008) analyzes asset allocation problems with a spectral risk measure as a constraint. In the context of regulatory capital and optimal reinsurance under Solvency II, Brandtner and Kürsten (2014) show that the ES is an adequate risk measure as long as the re-insurance component is not involved. Adding

[20] For a detailed discussion of the weighted VaR and the comparison to the weighted AVaR we refer to He et al. (2015).

this concept to the analysis, the ES gets suboptimal: it seems too restrictive s.t. no reinsurance contracts are signed at all and thus spectral risk measures become striking. In a follow-up paper Brandtner and Kürsten (2015) discuss the decision making under ES and spectral risk measures and highlight the problem of comparative risk aversion hereby and Brandtner (2018) analyze spectral risk measures in the context of background risks. Before we discuss the properties we want to give an intuition behind this class of risk measures.

Let X be an integrable random variable. As seen before the following VaR representation for the expected value holds:

$$\mathbb{E}[X] = \int_0^1 VaR_\alpha(X)\mathrm{d}\alpha.$$

Now, similar to the distortion risk measure concept, we can re-weight the VaR levels with some weight function $\Psi(t)$, a so-called **admissible risk spectrum** or **risk aversion function**, with $\int_0^1 \Psi(z)\mathrm{d}z = 1$. The following is taken from Acerbi (2002).

Definition 2.8 (Admissible Risk Spectrum)
*A function $\Psi \in L^1([0,1])$ is called **admissible risk spectrum** or **risk aversion function** if the following properties hold:*

(i) Ψ *is positive, i.e.* $\int_I \Psi(p)\mathrm{d}p \geq 0$*, for all* $I \subset [0,1]$.

(ii) Ψ *is decreasing, i.e.* $\int_{q-\varepsilon}^{q} \Phi(p)\mathrm{d}p \geq \int_q^{q+\varepsilon} \Psi(p)\mathrm{d}p$*, for all* $q \in (0,1)$ *and for all* $\varepsilon > 0$*, s.t.* $[q-\varepsilon, q+\varepsilon] \subset [0,1]$.

(iii) $\int_0^1 \Psi(p)\mathrm{d}p = 1$.

With this, a spectral risk measure can be defined.

Definition 2.9 (Spectral Risk Measure)
Let Ψ be an admissible risk spectrum, then

$$\rho_\Psi(X) := \int_0^1 VaR_\beta(X)\Psi(\beta)\mathrm{d}\beta$$

*defines a risk measure, a so-called **spectral risk measure**.*

We see the re-weighting of the expected value by multiplying the VaR at level β with the function value $\Psi(\beta)$. Because of the fact that Ψ is normalized the re-weighting is possible.

Remark 2.3 (Value at Risk)
By setting $\Psi^*(t) = 1_{\{t=\alpha\}}(t)$ we can create the $VaR_\alpha(X)$. But Ψ^* is not decreasing and thus it is not an admissible risk spectrum. Therefore, the VaR does not belong to the class of spectral risk measures.

To get an intuition for the risk aversion function Ψ we take a closer look at the Expected Shortfall: By setting

$$\Psi(t) := \frac{1}{\alpha} 1_{\{0 \leq t \leq \alpha\}}$$

(or $\Psi(t) := \frac{1}{1-\alpha} 1_{\{\alpha \leq t < 1\}}$, depending on the context) we can represent the $ES_\alpha(X)$ as a spectral risk measure. The $ES_\alpha(X)$ averages the possible outcomes in the α-right tail with equal weights. Looking at the risk aversion function we see that this is also just a weight function of the average. But here every tail quantile gets the same weight s.t. one does not allow for risk aversion. A further discussion on spectral risk measures as a generalization of the ES can e.g. be found in Tasche (2002). Generally $\Psi(t)$ gives different weights to different "p-confidence level slices" of the tail. Because of the normalization property of Ψ the weights in the average sum up to 1.The work of Dowd et al. (2008) discusses the use of different functions Ψ that account for risk aversion and their limitations. They show risk aversion functions Ψ that rebuild utility functions, e.g. if we choose $\Psi(t) = \gamma(1-t)e^{(1-\gamma)}$ for $\gamma < 1$ we receive the CRRA utility function. The function faces the problem, that for large t and a high risk aversion γ, lower losses are overweighted and high losses are underweighted. For further discussions about this topic and other risk aversion functions Ψ we refer to Dowd et al. (2008), the comment on this paper by Brandtner (2016a) and the paper of Brandtner (2016b).

Now that we have discussed the two risk measure classes separately, we close the gap between spectral risk measures and distortion risk measures. As stated before, the concave distortion risk measures are in this context highly important: Using the remark in McNeil et al. (2015) on page 287, we can write every concave distortion function because of absolute continuity in the form $g(t) = \int_0^t \Psi(u)\mathrm{d}u$. Thus every concave distortion risk measure can be represented in form of a spectral risk measure and vice versa. Thus the classes of concave distortion risk measures and spectral risk measures coincide and in general the class of spectral risk measures is a subset of the class of distortion risk measures.

2.2 Shortfall constraints and cost-minimal payoff modification

The protection of the investor's needs should be of great interest at least for the regulator if not for the bank or insurance company. In this section, we investigate some kind of natural portfolio optimization problem when it comes to this subject: maximizing the investor's expected utility under a prescribed terminal shortfall constraint. This constraint can e.g. be modeled with the VaR or with the ES. To be more general the investor receives an (optimal) payoff which is secured in form of a shortfall constraint.[21] The terminal shortfall constraint setup seems to be a good and helpful tool at first sight. But in this section we will see that this setting can impede the protection concept in times of bad market behavior because the state prices are the more expensive the worse the market. It is thus more costly to adapt the payoff on states of the world ω where the state prices are high, s.t. the payoff is modified on the cheaper state prices and thus better market conditions. The modification is thus given to better states of the world while the bad states stay unchanged and thus unprotected. This gives a first and simple introduction to the concept of cost-minimal payoff modifications:

The theory of cost-minimal payoff modifications dates back to Dybvig (1988b) and Dybvig (1988a). He analyzes state prices in discrete and continuous-time settings and calculates cost-efficient payoff modifications. Bernard et al. (2014) build upon the work of Dybvig and generalize many of his results. They find that cost-efficiency does not offer protection against a decline in an economy. Furthermore, they provide a condition that is sufficient for a cost-efficient payoff: If the random variable X_T and the state price density ξ_T are countermonotone, then X_T is cost-efficient. The concept of countermonotonicity is widely spread in the context of risk measures resp. valuation bounds for risks, the so-called Fréchet-Hoeffding bounds. Countermonotone risks have some kind of opposing behavior: When X_T has a high outcome in a state of the world ω^* then $\xi_T(\omega^*)$ has a low outcome. This is a first hint for the problem we are dealing with: The state price densities ξ_T are more expensive when the random variable X_T has a low outcome, i.e. there is no incentive to modify the optimal payoff in the bad states of the world. Bernard et al. (2015) analyze portfolio selection problems where they include some constraints on the state price densities. They find again that optimal portfolios do not fit with the aspiration of investors who seek protection. Bernard et al. (2019) even construct an algorithm to obtain numerically an investor's optimal portfolio under general preferences. Wei (2018) uses this cost-efficient modification to define the so-called risk reduction per cost (RRPC) to measure the trade-off between reducing the risk and incurring the costs. He analyzes many risk measures in the context of expected utility maximization and finds in a very general setting (the so-called weighted-VaR) that VaR and even the ES as a constraint do not insure the investor on the bad states of the distribution.

[21] For example the shortfall constraint in the Solvency II framework is given in terms of a VaR constraint with level $\alpha = 0.995$.

2.2.1 VaR constraint and cost-minimal payoff modification - A toy example

We want to motivate the concept of cost-minimal modifications with a simple example. In this toy example, we can see how the concept works and can transfer this to more sophisticated applications.
For this, let X be a discrete random variable in an arbitrage-free model setup. The discrete sample space is given by $\Omega = \{\omega_1, \dots, \omega_N\}$, the outcome by $X(\omega_1) = 1, \dots, X(\omega_N) = N$, i.e. the payoff in the state ω_1 has the smallest value, the payoff in state ω_N has the highest value. The probabilities of the states are given by

$$\mathbb{P}(\{\omega \in \Omega : X(\omega) = i\}) = \frac{1}{N}, \text{ for all } i = 1, \dots, N.$$

The distribution function $F_X(x)$ resp. the left inverse function $F_X^{\leftarrow}(\alpha)$ can be stated as

$$F_X(x) = \begin{cases} 0 & , x < 1 \\ \frac{1}{N} & , 1 \leq x < 2 \\ \vdots & \\ \frac{N-1}{N} & , N-1 \leq x < N \\ 1 & , x \geq N \end{cases}, \quad F_X^{\leftarrow}(\alpha) = \begin{cases} 1 & , 0 < \alpha \leq \frac{1}{N} \\ 2 & , \frac{1}{N} < \alpha \leq \frac{2}{N} \\ \vdots & \\ j & , \frac{j-1}{N} < \alpha \leq \frac{j}{N} \\ \vdots & \\ N & , \frac{N-1}{N} < \alpha < 1 \end{cases}.$$

The state prices $\kappa_1, \dots, \kappa_N$ are because of the no-arbitrage assumption strictly positive and because of the increasing payoff also increasing, i.e. $0 < \kappa_1 \leq \kappa_2 \leq \dots \leq \kappa_N$. Now let us take the regulator's point of view and prescribe a constraint at the VaR, i.e.

$$VaR_\alpha(X) \overset{!}{\leq} d.$$

The constraint d can be interpreted for example in the context of an insurance contract as a (constant) guarantee level G_T promised to the insured. In the context of Solvency II, the upper bound on the shortfall probability ($\alpha = 0,995$) determines the amount of equity that is needed to honor the liabilities of the insured, i.e. $\mathbb{P}(X < G_T) \leq 0.005$ or in other words: $VaR_{0.995}(X) \overset{!}{\leq} G_T$.
For a level of $\alpha = 1 - \frac{j}{N}$ with $j > d$, we find in our example that

$$VaR_{1-\frac{j}{N}}(X) = F_X^{\leftarrow}\left(\frac{j}{N}\right) = j,$$

i.e. the regulatory requirement is not fulfilled s.t. X has to be modified.
The modification is quite obvious by taking a closer look at the left inverse function of X: To achieve $F_X^{\leftarrow}(\frac{j}{N}) = d$, the outcome of X has to be reduced in the states where $X(\omega)$ is between $d+1$ and j to the value d. For higher outcomes of X than j

the reduction is not necessary because the VaR at level $\alpha = 1 - \frac{j}{N}$ is not influenced. The calculation scheme of the modified payoff $\tilde{X}$ is presented in Table 2.1, s.t. the distribution function of $\tilde{X}$ is given by

$$F_{\tilde{X}}(x) = \begin{cases} 0 & , \; x < 1 \\ \frac{1}{N} & , \; 1 \leq x < 2 \\ \vdots & \\ \frac{d-1}{N} & , \; d-1 \leq x < d \\ \frac{j}{N} & , \; d \leq x < j \\ \frac{j+1}{N} & , \; j \leq x < j+1 \\ \vdots & \\ 1 & , \; x \geq N \end{cases}.$$

ω_i	$X(\omega_i)$	$Y(\omega_i)$	κ_i	$\tilde{X}(\omega_i) = X(\omega_i) - Y(\omega_i)$
ω_1	1	0	κ_1	1
ω_2	2	0	κ_2	2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
ω_{d-1}	$d-1$	0	κ_{d-1}	$d-1$
ω_d	d	0	κ_d	d
ω_{d+1}	$d+1$	1	κ_{d+1}	d
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
ω_{j-1}	$j-1$	$j-1-d$	κ_{j-1}	d
ω_j	j	$j-d$	κ_j	d
ω_{j+1}	$j+1$	0	κ_{j+1}	$j+1$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
ω_N	N	0	κ_N	N

Table 2.1: Cost-efficient Payoff Modification Scheme

The modified payoff $\tilde{X}$ fulfills the regulators' specification, i.e.

$$VaR_{1-\frac{j}{N}}(\tilde{X}) = \inf \left\{ x \in \mathbb{R} : F_{\tilde{X}}(x) \geq \frac{j}{N} \right\} = d.$$

This is also the cheapest way to modify X, s.t. the regulatory constraint is fulfilled with the modification costs

$$\sum_{i=1}^{N} \kappa_i Y(\omega_i) = \sum_{i=1}^{j-d} \kappa_{d+i} i. \tag{2.3}$$

The procedure in the toy example can be applied to any situation where a random variable has to be modified s.t. a shortfall condition based on the VaR is fulfilled. This is in many applications of interest, e.g. in the context of maximizing the expected utility over all possible payoffs $X \in \mathcal{X}$ of an investor s.t. a shortfall constraint is fulfilled, i.e.

$$\max_{X \in \mathcal{X}} \mathbb{E}_{\mathbb{P}}[u(X)]$$
$$s.t.\ \mathbb{P}(X < d) \leq \alpha.$$

The cost-efficient modification in our example is given in the following proposition.

Proposition 2.5 (Cost-efficient Payoff modification - VaR)
Let X be a random variable and d a prescribed constraint on the Value at risk with level α, where $VaR_\alpha(X) > d$. The cost-efficient payoff modification $\tilde{X}^{VaR_\alpha}$, s.t. the VaR constraint is fulfilled, is given by

$$\tilde{X}^{VaR_\alpha}(\omega) := X(\omega) + [d - X(\omega)]\mathit{1}_{\{d_1 \leq X(\omega) \leq VaR_\alpha(X)\}}$$
$$= \begin{cases} X(\omega) & , X(\omega) < d \\ d & , d \leq X(\omega) \leq VaR_\alpha(X) \\ X(\omega) & , X(\omega) > VaR_\alpha(X) \end{cases}.$$

The modification costs are given by

$$P_D(F_X, F_\xi) - P_D(F_{\tilde{X}}, F_\xi) = \int_0^1 [F_X^{\leftarrow}(\alpha) - F_{\tilde{X}}^{\leftarrow}(\alpha)] F_\xi^{\leftarrow}(\alpha) \mathrm{d}\alpha,$$

where $P_D(F_X, F_\xi)$ denotes the distributional price of the variable X with the state price density ξ.

The representation of the modification costs in Proposition 2.5 are proven in Theorem 3 of Dybvig (1988a). Using the new representation of the modification costs we calculate these costs in our toy example. Recall that the VaR constraint is given by the level of $\alpha = 1 - \frac{j}{N} > d$ the distribution function and left inverse function of X is given by

$$F_X(x) = \begin{cases} 0 & , x < 1 \\ \frac{1}{N} & , 1 \leq x < 2 \\ \vdots & \\ \frac{N-1}{N} & , N-1 \leq x < N \\ 1 & , x \geq N \end{cases}, \quad F_X^{\leftarrow}(\alpha) = \begin{cases} 1 & , 0 < \alpha \leq \frac{1}{N} \\ 2 & , \frac{1}{N} < \alpha \leq \frac{2}{N} \\ \vdots & \\ j & , \frac{j-1}{N} < \alpha \leq \frac{j}{N} \\ \vdots & \\ N & , \frac{N-1}{N} < \alpha < 1 \end{cases}$$

and for $\tilde{X}$ given by

$$F_{\tilde{X}}(x) = \begin{cases} 0 & , \ x < 1 \\ \frac{1}{N} & , \ 1 \leq x < 2 \\ \vdots & \\ \frac{d-1}{N} & , \ d-1 \leq x < d \\ \frac{j}{N} & , \ d \leq x < j \\ \frac{j+1}{N} & , \ j \leq x < j+1 \\ \vdots & \\ 1 & , \ x \geq N \end{cases}, \quad F_{\tilde{X}}^{\leftarrow}(\alpha) = \begin{cases} 1 & , \ 0 < \alpha \leq \frac{1}{N} \\ 2 & , \ \frac{1}{N} < \alpha \leq \frac{2}{N} \\ \vdots & \\ d-1 & , \ \frac{d-2}{N} < \alpha \leq \frac{d-1}{N} \\ d & , \ \frac{d-1}{N} < \alpha \leq \frac{j}{N} \\ j+1 & , \ \frac{j}{N} < \alpha \leq \frac{j+1}{N} \\ \vdots & \\ N & , \ \frac{N-1}{N} < \alpha < 1 \end{cases}.$$

We need to calculate the state price density ξ to use the modification cost representation of Dybvig. In our discrete setting ξ is defined by:

$$\xi(\omega) = \frac{\kappa(\omega)}{p_i} = \frac{\kappa_i}{\frac{1}{N}} = N\kappa_i, \text{ for } \omega = \omega_i.$$

Thus, we receive for the distribution function and the left inverse of ξ

$$F_{\xi}(x) = \begin{cases} 0 & , x < N\kappa_1 \\ \frac{1}{N} & , N\kappa_1 \leq x < N\kappa_2 \\ \vdots & \\ \frac{N-1}{N} & , N\kappa_{N-1} \leq x < N\kappa_N \\ 1 & , x \geq N\kappa_N \end{cases}, \quad F_{\xi}^{\leftarrow}(\alpha) = \begin{cases} N\kappa_1 & , \ 0 < \alpha \leq \frac{1}{N} \\ N\kappa_2 & , \ \frac{1}{N} < \alpha \leq \frac{2}{N} \\ \vdots & \\ N\kappa_{N-1} & , \ \frac{N-2}{N} < \alpha \leq \frac{N-1}{N} \\ N\kappa_N & , \ \frac{N-1}{N} < \alpha < 1 \end{cases}.$$

Calculating the modification costs, we receive

$$\begin{aligned} P_D(F_X, F_\xi) - P_D(F_{\tilde{X}}, F_\xi) &= \int_0^1 [F_X^{\leftarrow}(\alpha) - F_{\tilde{X}}^{\leftarrow}(\alpha)] F_\xi^{\leftarrow}(\alpha) \mathrm{d}\alpha \\ &= \int_{\frac{d}{N}}^{\frac{d+1}{N}} 1 \cdot N\kappa_{d+1} \mathrm{d}\alpha + \int_{\frac{d+1}{N}}^{\frac{d+2}{N}} 2 \cdot N\kappa_{d+2} \mathrm{d}\alpha + \cdots + \int_{\frac{j-1}{N}}^{\frac{j}{N}} (j-d) \cdot N\kappa_j \mathrm{d}\alpha \\ &= \sum_{i=1}^{j-d} \kappa_{d+i} i. \end{aligned}$$

These modification costs coincide with our direct calculation in equation (2.3). An illustration of the toy example with $N = 100$ $j = 20$ and $d = 5$ is given in Figure 2.2.

We find two problems concerning a VaR shortfall risk constraint in the context of payoff modifications:

(i) The bad and thus for the investor risky states are not secured. Because of the

Distribution functions toy example (VaR-constraint)

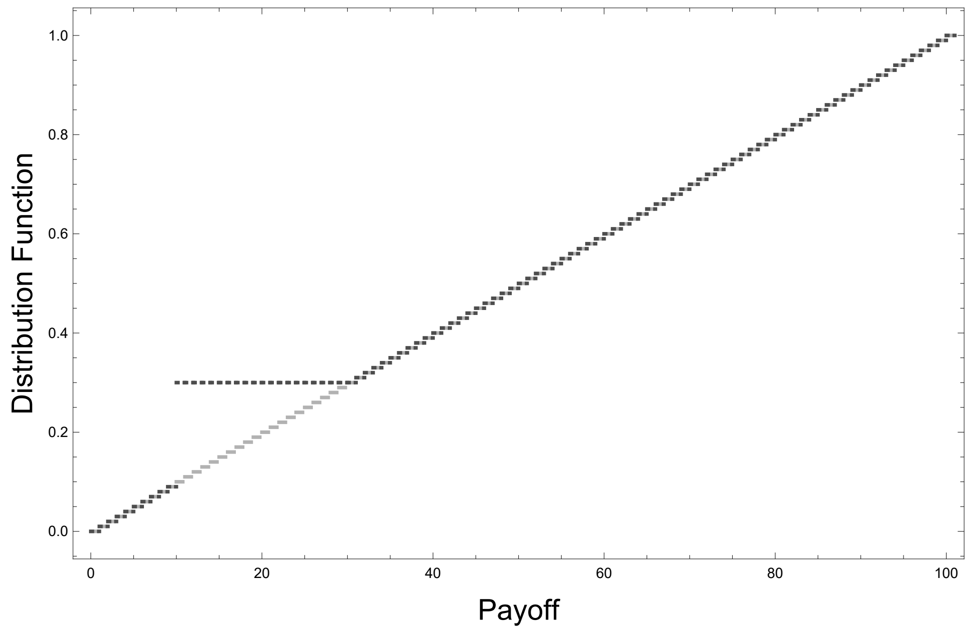


Figure 2.2: Toy example with VaR constraint: modified payoff distribution function $F_{\tilde{X}}$ vs original distribution function F_X

The toy example plotted for $N = 100, d = 10, j = 30$. The gray graph describes the distribution function $F_X(x)$, the black dotted the modified function $F_{\tilde{X}}(x)$.

VaR constraint there exists a protection, but rather on the interval $[d, j]$ and not on the more problematic states on the interval $[0, d]$. This observation can also be seen in the next section where we analyze an optimal design for an insurance contract. (ii) Increasing the VaR level α reduces the interval of the unsecured investment but still offers no protection on the worst states whenever the payoff has to be modified. This impedes the idea of protection.

This can also be transferred to a continuous-time approach e.g. in the BS-market. The famous result of Merton (1971) not only presents the optimal investment fraction in the risky asset but also presents the optimal terminal wealth that maximizes the expected utility of an investor. It is stated in terms of the state price density ξ_T. In the BS-model it holds

$$\xi_T := e^{-rT - \frac{1}{2}\frac{(\mu-r)^2}{\sigma^2}T - \frac{\mu-r}{\sigma}W_T}.$$

Thus, the optimization problem of the terminal wealth X_T under fair pricing condition is stated in the following terms:

$$\max_{X_T} \mathbb{E}_{\mathbb{P}}[u(X_T)]$$
$$\text{s.t. } \mathbb{E}_{\mathbb{P}}[\xi_T X_T] = X_0.$$

The optimal solution is given by

$$X_T^* := I(\lambda\xi_T),$$

where λ is a constant that is determined in a way s.t. the fair pricing condition

$$\mathbb{E}_{\mathbb{P}^*}[\xi_T X_T^*(\lambda)] = X_0$$

is fulfilled and $I := (u')^{-1}$ denotes the inverse function of the first derivative of the corresponding utility function. Adding a SFP constraint in terms of $\mathbb{P}(X_T < d) \leq \alpha$ to the above optimization problem gives us the problem stated in Basak and Shapiro (2001). They have rigorously proven by applying the convex-duality approach that the optimal solution is of the form:

$$X_T^{VaR,*} = \begin{cases} I(y\xi_T) & , \ \xi_T < d_1 \\ d & , \ d_1 \leq \xi_T \leq VaR_\alpha(\xi_T) \\ I(y\xi_T) & , \ \xi_T > VaR_\alpha(\xi_T) \end{cases}, \text{ where } d_1 = \frac{u'(d)}{y}.$$

The constant y is determined similar to the problem without VaR constraint s.t. the fair pricing condition is fulfilled, i.e.

$$\mathbb{E}_{\mathbb{P}^*}[\xi_T X_T^{VaR,*}(y)] = X_0, \text{ where } y \geq \lambda.$$

The solution of the VaR constrained problem is also given by a cost-efficient payoff, modified from the solution without SFP constraint similar to Proposition 2.5. For more details on this, we refer to Bernard et al. (2015).
The connection between the optimal solution under a SFP constraint and the optimal solution without a SFP constraint is given as follows: Following Basak and Shapiro (2001) the optimal solution with the VaR restriction can alternatively be interpreted resp. stated as the optimal unconstrained solution, where the initial value X_0 is reduced by the price of a put option (denoted with $1-\nu$)

$$\mathbb{E}_{\mathbb{P}^*}[\xi_T X_T^*(\lambda)] = \nu X_0$$

and an additional put option like construct that ensures the SFP constraint. The VaR optimal solution can then be stated as

$$X_T^{VaR,*} = \nu X_T^* + (d - \nu X_T^*)\mathbf{1}_{\{d_1 \leq \xi_T \leq VaR_\alpha(\xi_T)\}}.$$

Thus we can state the optimal terminal wealth with VaR constraint in terms of the Merton solution, reduced by a factor ν with an additional put option component. For more details we refer to Basak and Shapiro (2001) p. 377 and to Proposition 2.11 in the next section.

The finding of not securing the bad states of the world is not just a problem of the VaR. In the next subsection, we briefly discuss that even the ES is not able to solve this problem.

2.2.2 Cost-minimal payoff modification in the context of other (coherent) risk measures

In the discussion of risk measures the coherence axioms are of great importance. As discussed in the previous subsection, one of the most important coherent risk measures is the ES. Wei (2018) shows that measuring a risk with the ES can lead to a situation where the bad states of the insured resp. the investor is not secured just as we have seen in the last subsection with the VaR.
The first observation which can be seen easily from the definition of the expected shortfall is, that $ES_\alpha(X) \geq VaR_\alpha(X)$, because

$$ES_\alpha(X) = VaR_\alpha(X) + \frac{1}{1-\alpha}\int_{VaR_\alpha(X)}^{\infty} 1 - F_X(x)\mathrm{d}x. \tag{2.4}$$

Moreover, we see that the ES purely depends on the distribution of the random variable X. Calculating the modified payoff will preserve the cost-efficient payoff again, s.t. the bad states are not secured. This will be pointed out in the following: Again, let the regulator prescribe some constraint d regarding the expected shortfall, i.e.

$$ES_\alpha(X) \overset{!}{\leq} d.$$

Similar to the VaR case we can modify the payoff if the constraint is not fulfilled. Rewriting equation (2.4) we find an intuitive way to modify the ES s.t. the boundary can be matched:

$$ES_\alpha(X) = d_1 + \frac{1}{1-\alpha}\left\{\int_{d_1}^{d_2} 1 - F_X(x)\mathrm{d}x + \int_{d_2}^{\infty} 1 - F_X(x)\mathrm{d}x,\right\} \tag{2.5}$$

where d_1 is the VaR of the random variable X with confidence level α. For the cost-efficient modification we can adapt d_1 (i.e. the VaR of the modified payoff is smaller than the original payoff) and d_2, s.t. the ES constraint is fulfilled. Expressing the findings on the level of random variables, the following hold.

Proposition 2.6 (Cost-efficient Payoff modification - ES)
Let X be a random variable and d a prescribed constraint on the Expected Shortfall with level α, where $ES_\alpha(X) > d$. The cost-efficient modification $\tilde{X}^{ES_\alpha}$, s.t. the ES constraint is fulfilled is determined with

$$\begin{aligned}\tilde{X}^{ES}(\omega) &:= X(\omega) + [d_1 - X(\omega)]\,\mathit{1}_{\{d_1 \le X(\omega) \le d_2\}} \\ &= \begin{cases} X(\omega) & , \; X(\omega) < d_1 \\ d_1 & , \; d_1 \le X(\omega) \le d_2 \\ X(\omega) & , \; X(\omega) > d_2, \end{cases}\end{aligned}$$

where $VaR_\alpha(X) \in [d_1, d_2]$ and $VaR_\alpha(\tilde{X}) = d_1$. The modification costs are calculated as in Proposition 2.5.

This proposition shows, similar to the VaR, that the bad states of the world are not secured by a terminal ES constraint. This seems to be surprising in view of the good properties of the ES as a coherent risk measure. The paper of Wei (2018) gives a detailed proof on that. The proceeding in Proposition 2.6 to calculate the cost-efficient payoff modification is discussed in detail in Wei (2018). He even finds that the terminal wealth of the ES agent might be smaller on the bad states of the distribution than the agent who does not care about protection at all and thus not account for any shortfall constraints. Finally, Wei shows that many coherent risk measures are not appropriate in the context of portfolio selection. For more details we refer to Proposition 5.1 and Proposition 5.5 of Wei (2018). To overcome this problem there are two possibilities when the overall optimization problem (maximize over all possible payoffs X) should remain intact: The first possibility is to include another constraint into the ES resp. VaR optimization problem which is based on the state prices s.t. the company is forced to secure the bad states of the distribution. The other possibility is to change the risk measure in the shortfall constraint to one that accounts for securing the bad states of the distribution. Following the second idea, Wei finds that the Wang distortion risk measure, originally used in actuarial science, leads to a portfolio insurance strategy on the bad states when it is used as a constraint in the optimization problem. For more information we refer to Wei (2018).
Our findings give us a hint, that valuating a risk measure just by its axioms can lead to some serious problems in specific contexts, here in the context of expected utility maximization under shortfall constraints. It might be a better procedure in practice to define some properties a risk measure should fulfill and then choose a risk measure for that purpose, not the other way around. This is discussed in details in the works of Bauer and Zanjani (2016) and Kou et al. (2013).

2.3 Minimum Return Rate Guarantees under Default Risk - Optimal Design of Quantile Guarantees

Now we present the main findings and contributions of this chapter.[22] We maximize the insureds expected utility of a contract where the insured receives at least a so-called minimum return rate guarantee (MRRG) including the possibility of a default in terms of investment returns. We focus on a savings plan which is motivated by participating life insurance contracts. In reality, these contracts are more complex than our assumptions. They also include a term life insurance component and possess several premium payment options to policyholders. It is often criticized that the underlying of this kind of life insurance product is in reality typically based on book values and not market values as suggested in most research papers. However, the main effect is that the underlying possesses a lower volatility (via "smoothing") and - ceteris paribus - the value of the embedded options is lower. In any case, one can in principle account for this effect via choosing the "appropriate" volatility in the GBM - whenever the model is adjusted to empirical data via time series data. For a detailed description of participating life insurance contracts we refer e.g. to Grosen and Jørgensen (2000) and Grosen and Jørgensen (2002). Furthermore, we also define the default event exclusively in terms of the investment returns and do not consider that the insurance company may itself default.
Considering the possibility that the liabilities (guarantees) can not be honored impedes the basic idea of a guarantee. However, in reality there is no guarantee prevailing with probability one. Any guarantee may fail in times of extremely negative market conditions, i.e. guarantees are only valid under sufficiently good scenarios. Thus, one may soften the term guarantee and imagine it as honored with a high probability (quantile guarantee). In the context of participating life insurance contracts the guarantee is secured by regulatory requirements on the maximal shortfall probability. For example as discussed in the previous section Solvency II contains the condition that the shortfall probability w.r.t. a time horizon of one year is limited to 0.5%. Intuitively, it is clear that the value of a guarantee is decreasing in the shortfall probability. Default risk mitigates the guarantee component (it is less often binding and thus the guarantee is cheaper than without default risk). In contrast, control of the shortfall probability makes the guarantee more binding. Our main focus is on the optimal contract design in the presence of an upper probability bound on the shortfall probability posed by the regulator, i.e. the optimal design of quantile MRRGs.

The contributions of this section can be summarized as follows. Based on the distinction between a high and a low equity to debt ratio (compared to the combination of guarantee and participation fraction), we state the return payoff to the insured

[22] It is based on the work of Mahayni et al. (2021a).

by means of piecewise linear functions of the return of the insurers asset returns. On the one hand, this simplifies the pricing problem under default risk to the pricing of standard call (put) options. On the other hand, this already gives model independent insights, i.e. insights which are true w.r.t. any arbitrage free financial market model setup. For example, a low (high) equity to debt ratio implies a concave (piecewise concave and convex) payoff.[23] Thus, for a low equity to debt ratio, the value of the liabilities is decreasing in the riskiness of the insurer's assets. Consequently, the default risk dominates the guarantee option which contradicts the guarantee concept, i.e. if the admissible asset distributions are not restricted by an upper bound on the shortfall probability (on the guarantee). A further contribution is then given by deriving the optimal return payoff distribution to the insured. Because of the market completeness, the optimal (return) payoff to the insured can be implemented for any equity to debt ratio. Finally it is important to point out, that there are utility losses to the insured (and there is too much equity involved) if the insurer implements a suboptimal investment strategy.

This section is related to several strands of the literature including the ones on (i) pricing and hedging embedded guarantees/options, (ii) the impact of default risk (emphasizing on participating life insurance contracts), (iii) utility losses caused by guarantees and/or suboptimal investment decisions (conducted by insurance companies or pension funds), (iv) portfolio planning, (v) quantile hedging, and (vi) the analysis of piecewise convex and concave contingent payoffs.
Pricing embedded options by no-arbitrage already dates back to Brennan and Schwartz (1976). A more recent paper is Nielsen et al. (2011). Risk management and hedging aspects are discussed in Coleman et al. (2006), Coleman et al. (2007), and Mahayni and Schlögl (2008).
An early paper which already provides tools to determine closed-form solutions for the solvency restriction based on a shortfall concept under certain distribution assumptions (normal and log normal case) is given by Winkler et al. (1972) using partial moments. Non-linear optimization problems under shortfall constrains have already been solved in the past, c.f. McCabe and Witt (1980) who calculated the optimal chance-constrained expected profit of a non-life insurer.
Considering default in the context of participating life insurance contracts is firstly analyzed in Briys and De Varenne (1997) and Grosen and Jørgensen (2002). More recent papers are Schmeiser and Wagner (2015) and Hieber et al. (2019). Other papers on participating life insurance contracts excluding default risk are e.g. Bacinello (2001) who discusses amongst other results how a minimum interest rate guarantee ("technical rate") has to be set, such that the contracts are fairly priced and Gatzert et al. (2012) where the customer value of the policyholder is maximized.

Papers on utility losses caused by (suboptimal) investment strategies include Jensen

[23] In our setup, a low equity to debt ratio is always implied if there is a return guarantee which gives a return accumulation higher (or equal) one.

and Sørensen (2001), Jensen and Nielsen (2016) and Chen et al. (2019).[24] Chen et al. (2019) consider a general utility maximization under fair-pricing and budget constraints in a complete, arbitrage-free Black and Scholes model setup for an CRRA Investor. The payoff function is chosen s.t. it also includes default risk. They apply their results to equity-liked life insurances using a constant mix strategy and examine the effect of taxation.

Literature on portfolio planning with a main focus on insurance contracts with guarantees includes Huang et al. (2008), Milevsky and Kyrychenko (2008), Boyle and Tian (2008) and Mahayni and Schneider (2016). The general idea of maximizing the expected utility of the insured by choosing optimal parameter settings which fulfill fair pricing conditions has been provided in the literature before. The paper of Branger et al. (2010) analyzes different forms of point-to-point guarantees. Cliquet-style options are analyzed in Gatzert et al. (2012) and Schmeiser and Wagner (2015). In contrast to these articles we add the portfolio composition as a decision variable in the optimization problem to determine the overall expected utility-maximizing payoff of the insured in quasi-closed form.

Portfolio planning itself dates back to Merton (1969) and Merton (1971) who, amongst other results, solves the portfolio planning problem for a CRRA investor. The solution for investors who must also manage market-risk exposure using the Value-at-Risk (VaR) is firstly mentioned in Basak and Shapiro (2001). Yiu (2004) solves the problem where the VaR constraint is posed for the entire investment horizon. More recently, Gao et al. (2016) derive the solution for an investor with a dynamic mean-variance-CVaR and a dynamic mean-variance-safety-first constraint. A joint (terminal) VaR and portfolio insurance constraint is considered in Chen et al. (2018a). Multiple VaR constraints are analyzed in Chen et al. (2018b). With respect to European and American guarantees, we also refer to El Karoui et al. (2005). Quantile hedging already dates back to Föllmer and Leukert (1999). Literature on the insurance demand dates back to Leland (1980) and Benninga and Blume (1985) who show that in a complete financial market setup with risky and risk-free asset investments as also a utility function with constant risk aversion the investor will never buy portfolio insurance, instead buys the asset itself directly. Ebert et al. (2012) confirm the result for guarantee contracts, i.e. for CRRA Investors with reasonable risk aversion parameter Cumulative Prospect Theory (CPT) can not explain the demand for complex guarantee contracts. Ruß and Schelling (2018) introduce the concept of Multi Cumulative Prospect Theory (MCPT) which does not only consider the terminal value of the investment but also the annual value change. Under the MCPT the demand for complex guarantee products can be explained.

[24] In particular, Jensen and Sørensen (2001) analyze wealth losses for pension funds and emphasize that the individual investor can substantially suffer from the investment strategy conducted by the sponsor.

2.3.1 Contract design, payoffs, and fair pricing

We examine stylized versions of minimum return rate guarantees (MRRGs) which are e.g. observed in participating life insurance contracts. The insured pays a single premium at inception of the contract. The case of periodic payments is analyzed in the next chapter. The focus is on contracts which grant the insured a participation on positive investment results and include a return guarantee unless there is default risk. Since we abstract from mortality or surrender risk, there is no loss of generality due to a single premium compared to more flexible premium payments. The initial contribution of the insured is denoted by P_0. The product terminates and pays out to the insured at $T > 0$.

Stylized version of MRRG

Throughout the following, A_T denotes the terminal value of the insurance result (asset result) which is the outcome of an admissible investment strategy with initial investment A_0. In particular, the initial investment A_0 consists of the existing equity amount $E_0 \geq 0$ and the contributions of the insureds P_0, i.e. $A_0 = E_0 + P_0$. In particular, we normalize $P_0 = 1$ and set $E_0 = \alpha^{(E)}$ where $\alpha^{(E)} \in [0, 1]$ denotes the equity fraction (equity to debt ratio, respectively).

Along the lines of Schmeiser and Wagner (2015), we assume that the policyholder's account evolves from $t-1$ to t ($t \in \{1, \ldots T\}$) according to

$$P_t = P_{t-1}\left(1 + \max\left\{g, \alpha\left(\frac{A_t}{A_{t-1}} - 1\right)\right\}\right),$$

where α ($\alpha \in]0, 1[$) denotes the participation fraction and $1+g$ ($g \geq -1$) is the guaranteed accumulation factor granted for one year.[25] The special case $g = -1$ includes a contract without guarantee. For an analysis of MRRG based on guarantee rates that are linked to the interest rate evolution (so-called floating strike guarantees) we refer to the working paper of Mahayni et al. (2021b).
To simplify the expositions, we restrict ourselves to $T = 1$, i.e. we refer to the *intended* MRRG payoff P_1 to the insured, i.e. the payoff which is valid without default risk given by

$$P_1 = P_0\left(1 + \max\left\{g, \alpha\left(\frac{A_1}{A_0} - 1\right)\right\}\right). \tag{2.6}$$

[25] For different contract specifications within participation life insurance contracts cf. Nielsen et al. (2011). Further details, in particular w.r.t. participating life insurance contracts with annual return rate guarantees which are common in German-speaking countries, are given in Schmeiser and Wagner (2015).

Using $1+\max\left\{g,\alpha\left(\frac{A_1}{A_0}-1\right)\right\}=1+g+\alpha\left(\frac{A_1}{A_0}-\left(1+\frac{g}{\alpha}\right)\right)^+$ implies the following Lemma.

Lemma 2.1 (Intended payoff representation)
For $P_0=1$, the intended payoff to the insured P_1 can be represented by

$$P_1 = 1+g+\alpha\left(\frac{A_1}{A_0}-K\right)^+, \text{ where } K = 1+\frac{g}{\alpha}. \tag{2.7}$$

Thus, P_1 can be stated in terms of the payoff of (i) a long position in $e^{-r}P_0(1+g)$ zero bonds maturing in one year (r denotes the c.c. interest rate) and (ii) $\frac{\alpha P_0}{A_0}$ long calls on the synthetic asset A with maturity $T=1$ and strike $\tilde{K}=A_0(1+\frac{g}{\alpha})$. Without default risk, the MRRG payoff is illustrated in Figure 2.3. In particular, by pure dominance arguments, the (arbitrage free) value of a payoff which is always equal or sometimes even above another payoff must be higher than the value of the other payoff. Thus, two equally valuable payoffs P_1 and $\tilde{P}_1$ with $\alpha>\tilde{\alpha}$ imply that $g<\tilde{g}$.[26]

The assumption of a maturity $T=1$ implies some simplifications to our model: Because of the one period setting, the insured has no other premium payment option than an upfront premium. In the next chapter periodic premium payments are analyzed in-depth. Furthermore, the insurer cannot suffer from death or surrender of the policyholder, such that the surrender and mortality risk is excluded from our analysis. Expected utility maximizing portfolio allocation under mortality risk can be found e.g. in Milevsky and Young (2007). The impact of mortality risk on the shortfall probability is presented in Gatzert and Wesker (2014).
Thus our optimization problem in the later subsection is a purely state dependent portfolio optimization problem without time dependency. In this simplified setting, we find in the next subsection model independent insights for any arbitrage free financial market model and in Subsection 2.3.3, we can derive the utility maximizing return payoff of the insured.

MRRG under default risk

Considering default risk (DR), the insured only receives the payoff P_1 if the asset value A_1 is sufficiently high. The *actual* payoff to the insured under default risk is denoted by $L_1=P_1^{\text{With DR}}$ and is given by

$$L_1 = P_1-(P_1-A_1)^+ \text{ where} \tag{2.8}$$

$$(P_1-A_1)^+ = \max\{P_1-A_1,0\} = \max\left\{\left(1+\max\left\{g,\alpha\left(\frac{A_1}{A_0}-1\right)\right\}\right)-A_1,0\right\}$$

[26] The properties of such contracts are analyzed in detail in Nielsen et al. (2011).

MRRG (return) payoff (without default risk)

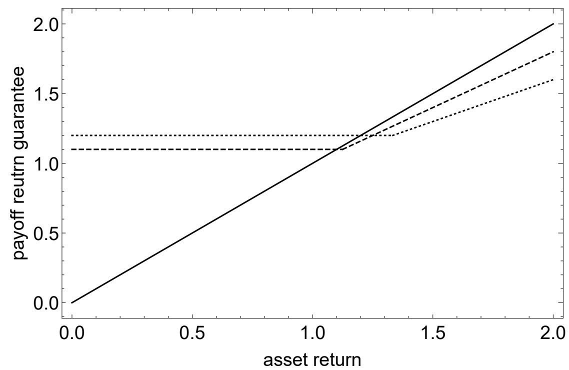


Figure 2.3: For varying asset return $\frac{A_1}{A_0}$, the figures illustrate guarantee return payoffs without default risk. The black solid line refers to $(\alpha, g) = (1, -1)$ (no guarantee), the black dashed line is given by $(\alpha, g) = (0.8, 0.1)$, and the black dotted line is based on $(\alpha, g) = (0.6, 0.2)$.

can be interpreted as the default put option of the contract provider. Although the default put option is given in terms of a nested version of the max operator (a compound option feature), it is possible to disentangle the payoff in terms of the payoffs of plain vanilla options, only. Notice that the initial value of the asset side is given by $A_0 = P_0 + E_0$. Normalizing $P_0 = 1$ and setting $E_0 = \alpha^{(E)}$ gives $A_1 = (1 + \alpha^{(E)})\frac{A_1}{A_0}$ such that

$$(P_1 - A_1)^+ = \max\left\{\left(1 + \max\left\{g, \alpha\left(\frac{A_1}{A_0} - 1\right)\right\}\right) - (1 + \alpha^{(E)})\frac{A_1}{A_0}, 0\right\}.$$

In particular, there is only one random variable $\frac{A_1}{A_0}$ involved. An intuitive interpretation of the payoff L_1 is possible if one considers the payoff of the default put option as a function of the random outcome of the investment decisions of the insurer, i.e. as a function of the asset increment $\frac{A_1}{A_0}$, and to distinguish between the strikes K_1, K_2, and K_3 defined by

$$K_1 := 1 + \frac{g}{\alpha}, \quad K_2 := \frac{1+g}{1+\alpha^{(E)}} \text{ and } K_3 := \frac{1-\alpha}{1-\alpha+\alpha^{(E)}}. \tag{2.9}$$

K_1 defines the level of $\frac{A_1}{A_0}$ such that the inner option (the call option of the insured due to the participation in the excess returns) is in the money, i.e. where the intended payoff P_1 pays out $1+\alpha\left(\frac{A_1}{A_0} - 1\right)$ instead of $1+g$. Now, the put option (of the equity holders) can be in the money in both cases, i.e. we can observe (i) the intended payoff P_1 is equal to $1+g$, but the asset side A_1 is lower, i.e. $A_1 < 1+g \Leftrightarrow \frac{A_1}{A_0} < \frac{1+g}{1+\alpha^{(E)}} = K_2$, and (ii) the intended payoff P_1 is equal to $1+\alpha\left(\frac{A_1}{A_0}-1\right)$, but the asset side A_1 is lower, i.e. $A_1 < 1+\alpha\left(\frac{A_1}{A_0}-1\right) \Leftrightarrow \frac{A_1}{A_0} < \frac{1-\alpha}{1-\alpha+\alpha^{(E)}} = K_3$.

In consequence, we can express the payoff of the default put option by means of piecewise linear functions as follows:

$$\begin{aligned}(P_1 - A_1)^+ &= \left((1+g) - (1+\alpha^{(E)})\frac{A_1}{A_0}\right) 1_{\left\{\frac{A_1}{A_0}\le\min\{K_1,K_2\}\right\}} \\ &\quad + \left(1+\alpha\left(\frac{A_1}{A_0}-1\right) - (1+\alpha^{(E)})\frac{A_1}{A_0}\right) 1_{\left\{K_1<\frac{A_1}{A_0}<\max\{K_1,K_3\}\right\}}, \\ \text{i.e. } (P_1 - A_1)^+ &= \left(1+\alpha^{(E)}\right)\left(K_2 - \frac{A_1}{A_0}\right) 1_{\left\{\frac{A_1}{A_0}\le\min\{K_1,K_2\}\right\}} \\ &\quad + \left(1+\alpha^{(E)}-\alpha\right)\left(K_3 - \frac{A_1}{A_0}\right) 1_{\left\{K_1<\frac{A_1}{A_0}<\max\{K_1,K_3\}\right\}}.\end{aligned}$$

A crucial distinction is given by a different ranking order of the strikes K_1, K_2 and K_3. However, the relation between the strikes is given by comparing the equity to debt ratio $\alpha^{(E)}$ to the guarantee g (and participation rate α). The result is summarized in the following Lemma.

Lemma 2.2 (Strikes)
Let K_1, K_2, and K_3 be defined as in Equation (2.9), then the following relations hold

(i) $K_1 = K_2 = K_3 \Leftrightarrow \alpha^{(E)} = \frac{-g(1-\alpha)}{\alpha+g}$

(ii) $K_1 > K_2 > K_3 \Leftrightarrow \alpha^{(E)} > \frac{-g(1-\alpha)}{\alpha+g}$,

(iii) $K_3 > K_2 > K_1 \Leftrightarrow \alpha^{(E)} < \frac{-g(1-\alpha)}{\alpha+g}$.

$$\text{In particular, the relation } \alpha^{(E)} > \frac{-g(1-\alpha)}{\alpha+g} \Leftrightarrow g > \frac{-\alpha\alpha^{(E)}}{1-\alpha+\alpha^{(E)}} \tag{2.10}$$

and $g \ge 0$ implies $\alpha^{(E)} \ge \frac{-g(1-\alpha)}{\alpha+g}$. In addition, notice that case (ii) ((iii), respectively) in fact means a rather high (low) equity to debt ratio compared to the guarantee g.

In summary, the payoff (return) of the default put can be represented as follows.

Proposition 2.7 (Payoff representation of the defaultable put)
The payoff of the defaultable put can be stated in terms of a piecewise linear function

in the asset increment $\frac{A_1}{A_0}$*, i.e.*

$$(P_1 - A_1)^+ = \begin{cases} \left(1+\alpha^{(E)}\right)\left(K_2 - \frac{A_1}{A_0}\right) 1_{\left\{\frac{A_1}{A_0} \le K_1\right\}} & \\ + \left(1+\alpha^{(E)} - \alpha\right)\left(K_3 - \frac{A_1}{A_0}\right) 1_{\left\{K_1 < \frac{A_1}{A_0} < K_3\right\}} & \text{for } \alpha^{(E)} \le \frac{-g(1-\alpha)}{\alpha+g} \\ \left(1+\alpha^{(E)}\right)\left(K_2 - \frac{A_1}{A_0}\right)^+ & \text{for } \alpha^{(E)} > \frac{-g(1-\alpha)}{\alpha+g} \end{cases} . \tag{2.11}$$

An intuitive way to understand the liability side under default risk is analogously given by stating the payoff L_1 depending on the asset increment $\frac{A_1}{A_0}$. First notice that, without default risk, the call option of the insured (cf. Lemma 2.1) is in the money if $\frac{A_1}{A_0} > K_1 = 1 + \frac{g}{\alpha}$. Otherwise the intended return is $1+g$. Under default risk, the insured only receives $1+g$ if this is possible, i.e. if $A_1 > P_0(1+g)$ ($P_0 = 1$, $A_0 = 1+\alpha^{(E)}$), or equivalently if $\frac{A_1}{A_0} > K_2 = \frac{1+g}{1+\alpha^{(E)}}$. For $\frac{A_1}{A_0} \le K_1 = 1 + \frac{g}{\alpha}$, the insured only receives the minimum of $1+g$ and $A_1 = (1+\alpha^{(E)})\frac{A_1}{A_0}$.

Now, consider the case that $\frac{A_1}{A_0} > K_1 = 1+\frac{g}{\alpha}$, i.e. $P_1 = 1+\alpha\left(\frac{A_1}{A_0} - 1\right)$. Again, under default risk, the insured nevertheless only receives the lower of $1+\alpha\left(\frac{A_1}{A_0} - 1\right)$ and $A_1 = (1+\alpha^{(E)})\frac{A_1}{A_0}$, which is defined by the benchmark $K_3 = \frac{1-\alpha}{1-\alpha+\alpha^{(E)}}$. In summary, we obtain

$$L_1 = \begin{cases} (1+\alpha^{(E)})\frac{A_1}{A_0} & \text{for } \frac{A_1}{A_0} < \min\{K_1, K_2\} \\ 1+g & \text{for } \min\{K_1, K_2\} \le \frac{A_1}{A_0} < K_1 \\ (1+\alpha^{(E)})\frac{A_1}{A_0} & \text{for } K_1 \le \frac{A_1}{A_0} < \max\{K_1, K_3\} \\ 1+\alpha\left(\frac{A_1}{A_0} - 1\right) & \text{for } \frac{A_1}{A_0} \ge \max\{K_1, K_2, K_3\}. \end{cases}$$

Using Lemma 2.2 immediately gives the following Proposition.[27]

Proposition 2.8 (Payoff representation of liabilities)
Let K_1*,* K_2 *and* K_3 *be defined as in Equation (2.9). For* $P_0 = 1$ *and* $\alpha^{(E)} = E_0$ *(*$A_0 = 1+\alpha^{(E)}$*) it holds*

---

[27]

$$\min\{K_1, K_2\} = \begin{cases} K_1 & \text{for } \alpha^{(E)} \le \frac{-g(1-\alpha)}{\alpha+g} \\ K_2 & \text{for } \alpha^{(E)} > \frac{-g(1-\alpha)}{\alpha+g} \end{cases}, \max\{K_1, K_3\} = \begin{cases} K_3 & \text{for } \alpha^{(E)} \le \frac{-g(1-\alpha)}{\alpha+g} \\ K_1 & \text{for } \alpha^{(E)} > \frac{-g(1-\alpha)}{\alpha+g} \end{cases},$$

$$\max\{K_1, K_2, K_3\} = \begin{cases} K_3 & \text{for } \alpha^{(E)} \le \frac{-g(1-\alpha)}{\alpha+g} \\ K_1 & \text{for } \alpha^{(E)} > \frac{-g(1-\alpha)}{\alpha+g} \end{cases}.$$

(i) ***Low equity to debt ratio:*** *For* $\alpha^{(E)} \leq \frac{-g(1-\alpha)}{\alpha+g}$*, the payoff (return) to the insured is given by*

$$L_1 = \begin{cases} (1+\alpha^{(E)})\frac{A_1}{A_0} & \textit{for } \frac{A_1}{A_0} < K_3 \\ 1+\alpha\left(\frac{A_1}{A_0}-1\right) & \textit{for } \frac{A_1}{A_0} \geq K_3, \end{cases}$$

$$\textit{i.e. } L_1 = (1+\alpha^{(E)})\frac{A_1}{A_0} - (1-\alpha+\alpha^{(E)})\left(\frac{A_1}{A_0}-K_3\right)^+. \tag{2.12}$$

(ii) ***High equity to debt ratio:*** *For* $\alpha^{(E)} > \frac{-g(1-\alpha)}{\alpha+g}$ *it holds*

$$L_1 = \begin{cases} (1+\alpha^{(E)})\frac{A_1}{A_0} & \textit{for } \frac{A_1}{A_0} < K_2 \\ 1+g & \textit{for } K_2 \leq \frac{A_1}{A_0} < K_1 \\ 1+\alpha\left(\frac{A_1}{A_0}-1\right) & \textit{for } \frac{A_1}{A_0} \geq K_1, \end{cases}$$

$$\textit{i.e. } L_1 = (1+\alpha^{(E)})\frac{A_1}{A_0} - (1+\alpha^{(E)})\left(\frac{A_1}{A_0}-K_2\right)^+ + \alpha\left(\frac{A_1}{A_0}-K_1\right)^+. \tag{2.13}$$

For a low equity to debt ratio (*Case (i)*), the above Proposition states that the liabilities of the insured are given by the payoff of

(i) $\frac{1+\alpha^{(E)}}{A_0}$ long positions in the insurer's assets A and

(ii) $\frac{1-\alpha+\alpha^{(E)}}{A_0}$ short calls with strike $K = K_3 A_0 = \frac{1-\alpha}{1-\alpha+\alpha^{(E)}}A_0$.

For a high equity to debt ratio (*Case (ii)*), the above Proposition states that the liabilities of the insured are given by the payoff of

(i) $\frac{1+\alpha^{(E)}}{A_0}$ long positions in the insurer's assets A,

(ii) $\frac{1+\alpha^{(E)}}{A_0}$ short positions in a call on A with strike $K = K_2 A_0 = \frac{1+g}{1+\alpha^{(E)}}A_0$ and

(iii) $\frac{\alpha}{A_0}$ long calls with strike $K = K_1 A_0 = \left(1+\frac{g}{\alpha}\right)A_0$.

In addition, the above Proposition immediately implies the following important properties of the liability payoffs.

Corollary 2.1 (Properties of the liability payoff)
Let L_1 *be the liability payoff stated in Proposition 2.8, then it holds*

(i) L_1 *is increasing in* g *and* $\alpha^{(E)}$*. For* $g > 0$*,* L_1 *is increasing in* α*.*

Illustration of the contract payoff under default (L_1)

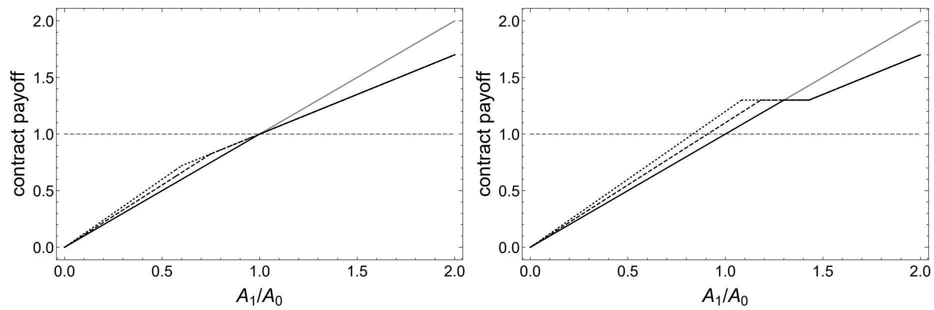


Figure 2.4: For varying the asset return $\frac{A_1}{A_0}$, the figures illustrate the payoff $L_1 = P_1 - (P_1 - A_1)^+$. It holds $0 = \alpha_1^{(E)} < \alpha_2^{(E)} < \alpha_3^{(E)}$. The black solid line refers to $\alpha^{(E)} = 0$, the black dashed line to $\alpha^{(E)} = \alpha_2^{(E)}$, and the dotted line to $\alpha^{(E)} = \alpha_3^{(E)}$. The left hand figure is based on $\alpha^{(E)} \leq \frac{-g(1-\alpha)}{\alpha+g}$ (*low equity fraction*) while the right hand figure is based on $\alpha^{(E)} > \frac{-g(1-\alpha)}{\alpha+g}$ (*high equity fraction*). In particular, the payoffs on the left hand side are piecewise concave and convex while the payoffs on the right hand side are concave.

(ii) For $\alpha^{(E)} \leq \frac{-g(1-\alpha)}{\alpha+g}$*,* L_1 *is concave in* $\frac{A_1}{A_0}$.

(iii) For $\alpha^{(E)} > \frac{-g(1-\alpha)}{\alpha+g}$*,* L_1 *is piecewise concave and piecewise convex in* $\frac{A_1}{A_0}$.

An illustration of L_1 is given in Figure 2.4. The left hand figure is based on $\alpha^{(E)} \leq \frac{-g(1-\alpha)}{\alpha+g}$ (*low equity fraction*) while the right hand figure is based on the case $\alpha^{(E)} > \frac{-g(1-\alpha)}{\alpha+g}$ (*high equity fraction*). In particular, the payoffs on the left hand side are concave while the payoffs on the right hand side are piecewise concave and convex. Intuitively, it is clear that a higher amount of equity means that the real degree of guarantee is, ceteris paribus, higher than for a lower amount of equity. This is resembled in the payoff profiles, i.e. a higher amount of equity gives more convexity to the payoff profile (implying a more valuable guarantee).

2.3.2 Fair pricing and regulatory requirements

Throughout the following analysis, we make some assumptions on the contract design (and the model setup for the financial market). We assume that the financial market model is arbitrage free. Furthermore, we assume that, because of competition, the contracts are fairly priced such that no-arbitrage is introduced (among the insurers and between the insurance products and the financial market products):

Assumption 2.1 (No arbitrage)
We assume that the financial market model is arbitrage free. Thus, the fundamental

theorem of asset pricing implies the existence of an equivalent pricing measure $\mathbb{P}^*$ such that the price of any traded asset X with payoff X_T at $T > 0$ is given by the expected discounted payoff under $\mathbb{P}^*$,i.e.

$$X_0 = \mathbb{E}_{\mathbb{P}^*}\left[e^{-\int_0^T \tilde{r}_u\,du} X_T\right], \tag{2.14}$$

where $\tilde{r}_u$ denotes the forward rate, such that $\int_0^T \tilde{r}_u\,du$ is the continuously compounded interest rate prevailing at time T.

Assumption 2.2 (Fair pricing)
We assume competition between the insurance companies (and with the opportunity to invest in the financial market). In particular, we thus assume that the insurance contracts are fairly priced, i.e. depending on the investment decisions which are carried out by the insurer on the financial market, the contract prices are given by the arbitrage free (financial market) prices.[28]

Assumption 2.3 (Stakeholders)
The policyholders are not able to participate at the arbitrage free financial market, such that they cannot replicate future cash-flows. They just have the possibility to invest in the asset side of the insurance company. The insurer itself, resp. its shareholders, of course have this access to the market.[29]

In addition, we assume later that an admissible contract design must honor regulatory requirements as e.g. posed by an upper bound on the shortfall probability. In our case these requirements are stated in the Solvency II regulatory framework where the shortfall probability of 0.5% in one year is not allowed to be excessed. For more insights and information we refer to Article 101 in EC (2009).

First, we consider the assumption on the contract pricing and the implications of postulating an arbitrage free financial model setup. Subsequently, we introduce the regulatory requirement and represent the shortfall probability in terms of the strikes introduced above.

Along the lines of Proposition 2.8, the arbitrage free value of the liabilities (and the default put, respectively) is given by the (arbitrage free) value of the corresponding portfolio of plain vanilla options. To simplify the exposition, we refer to a one year horizon, i.e. the call (or put) options have a maturity of $T = 1$. The (arbitrage

[28] It should be mentioned that in practice it would not be possible to e.g. make sure that all these contracts are initially fair: Rather, in practice, cross-subsidizing effects are unavoidable (cf. e.g. Hieber et al. (2015)).

[29] This assumption is reasonable and has often been used in other literature dealing with this topic, e.g. Schmeiser and Wagner (2015) or Briys and De Varenne (1997).

free) value of a call (put) option (with maturity $T = 1$) and strike K is denoted by $Call(K)$ ($Put(K)$). Without loss of generality, we refer to the options written on the increments $\frac{A_1}{A_0}$ (which is implied by the investment strategy of the insurer), i.e. we use the relation

$$\left(\frac{A_1}{A_0} - K\right)^+ = \frac{1}{A_0}\left(A_1 - KA_0\right)^+.$$

To be more precise, $Call(K)$ ($Put(K)$) denotes the $t = 0$ value of the $T = 1$ payoff $\left(\frac{A_1}{A_0} - K\right)^+$ ($\left(K - \frac{A_1}{A_0}\right)^+$, respectively).

Proposition 2.9 (Fair pricing conditions)
Assume that the asset A can be synthesized by a financial market strategy, i.e. the $t = 0$ price of the payoff A_1 is A_0 (A is an asset paying no dividends). In addition, assume that the financial market is arbitrage free. Then, the fair pricing condition (posed by the normalization $P_0 = 1$) is given by the condition that the market consistent price of the payoff L_1 is equal to $P_0 = 1$. In particular, depending on the equity fraction $\alpha^{(E)}$, the guarantee g, and the participation rate α, the following pricing conditions hold:

(i) ***Low equity to debt ratio:*** *For $\alpha^{(E)} \leq \frac{-g(1-\alpha)}{\alpha+g}$, it holds*

$$1 = 1 + \alpha^{(E)} - (1 - \alpha + \alpha^{(E)})Call(K_3). \tag{2.15}$$

(ii) ***High equity to debt ratio:*** *For $\alpha^{(E)} > \frac{-g(1-\alpha)}{\alpha+g}$, it holds*

$$1 = 1 + \alpha^{(E)} - (1 + \alpha^{(E)})Call(K_2) + \alpha Call(K_1). \tag{2.16}$$

where the strikes K_1, K_2 and K_3 are defined as in Equation (2.9).

Corollary 2.2 (Properties of fair contracts under default risk)
The fair pricing conditions imply the following properties

(i) For $\alpha^{(E)} = 0$, a fair contract implies $\alpha^{fair} = 1$.

(ii) In the special case that $g = -1$ (no guarantee) it also holds $\alpha^{fair} = 1$.

The proof is straightforward and the results are intuitive: Part (i) states that without equity, the insured face the whole risk of the asset investments, i.e. the (fair) liabilities are given by $L_1 = \frac{A_1}{A_0}$. In particular, without further restrictions on the distribution of $\frac{A_1}{A_0}$, i.e. restrictions on the riskiness of the investment strategy, there is no guarantee without equity. The interpretation of part (ii) is analogous. Since there is no guarantee if $g = -1$, a fair contract must imply $L_1 = \frac{A_1}{A_0}$.

Now consider the condition that there is a regulatory requirement on the shortfall probability. Assume that the regulator requires an upper bound ϵ for the probability that the intended *guaranteed* accumulation P_1 is not honored because the asset value A_1 is lower, i.e.

$$\mathbb{P}(A_1 < P_1) \leq \epsilon. \tag{2.17}$$

This is similar to our toy example in Subsection 2.2.1 where we prescribe an upper bound on the VaR.

Again, normalizing $P_0 = 1$ and using $A_1 = (1 + \alpha^{(E)})\frac{A_1}{A_0}$ implies that the event $\{A_1 < P_1\}$ can be represented in terms of the strikes $K_1 = 1 + \frac{g}{\alpha}$, $K_2 = \frac{1+g}{1+\alpha^{(E)}}$ and $K_3 = \frac{1-\alpha}{1-\alpha+\alpha^{(E)}}$:
K_1 defines the level of $\frac{A_1}{A_0}$ such that the inner option is in the money, i.e. where the intended payoff P_1 pays out $1 + \alpha\left(\frac{A_1}{A_0} - 1\right)$ instead of $1 + g$. The strike K_2 defines the level of $\frac{A_1}{A_0}$ such that the put option is in the money, i.e. the intended Payoff P_1 is equal to $1 + g$, but the asset side A_1 is lower. $K_3 = \frac{1-\alpha}{1-\alpha+\alpha^{(E)}}$ defines the level of $\frac{A_1}{A_0}$ where the liabilities can not be satisfied if the inner option is in the money, i.e.

$$\{A_1 < P_1\} = \left\{\frac{A_1}{A_0} \leq K_1; \frac{A_1}{A_0} < K_2\right\} \cup \left\{\frac{A_1}{A_0} > K_1; \frac{A_1}{A_0} < K_3\right\}. \tag{2.18}$$

With Lemma 2.2 and the representation of the shortfall event in Equation (2.18), we immediately obtain the following Proposition.

Proposition 2.10 (Shortfall probability)
The shortfall probability $SFP := \mathbb{P}(A_1 < P_1)$ is given by

$$\begin{aligned} SFP &= \mathbb{P}\left(\frac{A_1}{A_0} < \min\{K_1, K_2\}\right) + \mathbb{P}\left(K_1 \leq \frac{A_1}{A_0} \leq \max\{K_1, K_3\}\right) \\ &= \mathbb{P}\left(\frac{A_1}{A_0} < K_3\right) \mathit{1}_{\left\{\alpha^{(E)} \leq \frac{-g(1-\alpha)}{\alpha+g}\right\}} + \mathbb{P}\left(\frac{A_1}{A_0} \leq K_2\right) \mathit{1}_{\left\{\alpha^{(E)} > \frac{-g(1-\alpha)}{\alpha+g}\right\}}. \end{aligned} \tag{2.19}$$

It is worth to emphasize that, e.g. in the context of Solvency II, the upper bound on the shortfall probability determines the amount of equity which is needed to assure the solvency to a high degree, i.e. to honor the liabilities to the insured. The paper of Boonen (2017) analyzes the capital requirements under Solvency II if they were based on the ES.
Recall that $K_2 = \frac{1+g}{1+\alpha^{(E)}}$ and $K_3 = \frac{1-\alpha}{1-\alpha+\alpha^{(E)}}$. Obviously, the lower the strike is, the

Benchmark parameter

model parameter			contract parameter				upper bound on SFP
r	μ	σ	P_0	A_0	α	g	ϵ
0.03	0.07	0.2	1	$1+\alpha^{(E)}$	0.9	0.0175	0.005

Table 2.2: Benchmark parameter setting

lower is the probability that the value of a given investment strategy drops below the strike. Since the above strikes are decreasing in the equity fraction $\alpha^{(E)}$, a higher equity fraction is able to reduce the shortfall probability.[30]

2.3.3 Black and Scholes model setup and illustration

Along the lines of the previous subsections, the contracts can be fairly priced in closed form in any arbitrage free model setup which allows closed form solutions of plain vanilla options. For the sake of simplicity, we place ourselves in a Black and Scholes model setup to give some illustrations. The financial market model over the filtrated probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ is given by the Black and Scholes model, i.e. there are two investment possibilities: a risky asset S and a risk-free asset B which accumulates according to a constant interest rate r. The filtration $(\mathcal{F}_t)_{t\in[0,T]}$ is generated by the standard Brownian motion $(W_t)_{t\in[0,T]}$. Because of the completeness of the Black and Scholes model, there exists a uniquely determined equivalent martingale measure $\mathbb{P}^*$ under which the process $(W_t^*)_{t\in[0,T]}$ defines a standard Brownian motion. In particular, the risky asset $(S_t)_{t\in[0,T]}$ and risk free bond dynamics $(B_t)_{t\in[0,T]}$ are given by

$$dS_t = S_t\,(\mu\,dt + \sigma\,dW_t) = S_t\,(r\,dt + \sigma\,dW_t^*)\,,\quad S_0 = s$$
$$dB_t = B_t r\,dt,\quad B_0 = b.$$

Under the real world probability measure $\mathbb{P}$, the asset price follows a geometric Brownian motion with constant drift μ ($\mu > r$) and constant volatility σ ($\sigma > 0$). Under the uniquely defined equivalent martingale measure (pricing measure) $\mathbb{P}^*$, the asset price follows a geometric Brownian motion with constant drift r and constant volatility σ ($\sigma > 0$). The risk free bond B grows at a constant interest rate r.

Constant mix strategies

[30] However, if one assumes a complete financial market model, any reduction in the shortfall probability can also be implemented by a change in the asset distribution by means of a suitable investment strategy.

Assuming that the insurer decides to implement an investment strategy which is described by a constant fraction of wealth π invested in the risky asset (and the remaining fraction $1-\pi$ is invested in the risk free bond) implies that the asset process is also given by a lognormal process, i.e.

$$dA_t = A_t\left(\pi\frac{dS_t}{S_t} + (1-\pi)r\,dt\right).$$

Thus, w.r.t. an investment horizon of $T=1$, it holds

$$A_1 = A_0 e^{\mu_A^{(RW)}-\frac{1}{2}\sigma_A^2+\sigma_A W_1} = A_0 e^{r-\frac{1}{2}\sigma_A^2+\sigma_A W_1^*}$$
$$\text{where } \mu_A^{(RW)} = \pi\mu + (1-\pi)r \text{ and } \sigma_A = \pi\sigma.$$

$\mu^{(RW)}$ denotes the drift of the asset dynamics under the real word measure $\mathbb{P}$. Under the pricing measure $\mathbb{P}^*$, the drift is equal to r. In particular, let $N(\mu,\sigma^2)$ denote the normal distribution with mean μ and variance σ^2 and $\Phi(\cdot)$ the cumulative distribution function of the standard normal distribution. Then it holds

$$\ln\frac{A_1}{A_0} \sim N\left(\mu_A - \frac{1}{2}\sigma_A^2, \sigma_A^2\right) \text{ under } \mathbb{P},\ \ \ln\frac{A_1}{A_0} \sim N\left(r - \frac{1}{2}\sigma_A^2, \sigma_A^2\right) \text{ under } \mathbb{P}^*.$$

In consequence, the arbitrage free (competitive) price of the liabilities L_1 (the default put, respectively) can be derived by means of Proposition 2.9 where the call price formula $Call(K) = Call^{(BS)}(K,\sigma_A)$ is given by the Black and Scholes pricing formula (w.r.t. the returns), i.e.

$$Call^{(BS)}(K,\sigma_A) = \Phi(d_1(K,\sigma_A)) - e^{-r}K\Phi(d_2(K,\sigma_A)), \tag{2.20}$$
$$\text{where } d_1(K,\sigma_A) = \frac{-\ln K + r + \frac{1}{2}\sigma_A^2}{\sigma_A} \text{ and } d_2(K,\sigma_A) = d_1(K,\sigma_A) - \sigma_A.$$

Figure 2.5 gives an illustration of fair contract designs. The left figure illustrates fair tupels of the contract parameter (α, g). Along the lines of the model-free results, the (return) payoff of the MRRG under default risk is increasing in α and g. Thus, in order to stay on a fair contract design, an increasing guarantee g must be compensated by decreasing the participation rate α. In addition, the fair (α, g) combinations are lower for higher equity fractions, i.e. the black line refers to $\alpha_1^{(E)} = 0.01$, the black dashed line to $\alpha_2^{(E)} = 0.02$, and the dotted line to $\alpha_3^{(E)} = 0.05$. This result is straightforward and can, for example, be found in Grosen and Jørgensen (2002). An interesting effect arises in view of the piecewise concave and piecewise convex payoff structures (implied by $g > 0$ and $\alpha^{(E)} > 0$, cf. Corollary 2.1). Although the contract value is increasing in the equity fraction $\alpha^{(E)}$, this is not necessarily true with respect to the riskiness of the investments, that means w.r.t. π (the volatility $\sigma_A = \pi\sigma$,

Illustration of fair contracts (constant mix strategies)

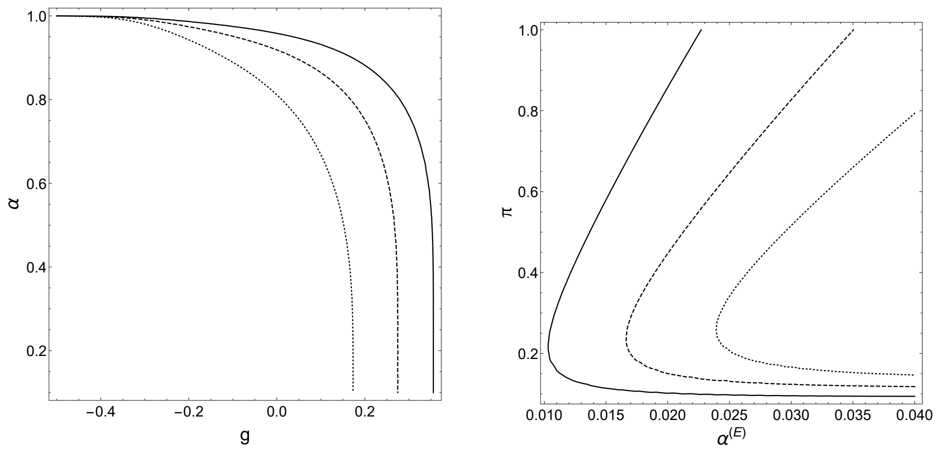


Figure 2.5: The contract and model parameters are given as in Table 2.2. The left figures illustrate fair tupels of the contract parameter (α, g). The black line refers to $\alpha_1^{(E)} = 0.01$, the black dashed line to $\alpha_2^{(E)} = 0.02$, and the dotted line to $\alpha_3^{(E)} = 0.05$. The figure on the right hand side (the black line, respectively) depicts fair contracts for the benchmark case in terms of fair combinations of the equity fraction $\alpha^{(E)}$ and the investment fraction π (defining the volatility of the assets, i.e. $\sigma_A = \pi\sigma$). The solid line refers to $\alpha = 0.9$, the dashed line refers to a lower participation fraction $\alpha = 0.85$ and the dotted line refers to $\alpha = 0.8$.

respectively). Thus, for a fixed equity fraction $\alpha^{(E)}$, there may be two investment fractions π^1 and π^2 such that the contract is fairly priced. This is illustrated in the right hand plot of Figure 2.5 which depicts fair contracts for the benchmark case in terms of fair combinations of the equity fraction $\alpha^{(E)}$ and the investment fraction π (defining the volatility of the assets, i.e. $\sigma_A = \pi\sigma$). The solid line refers to $\alpha = 0.9$, the dashed line refers to a lower participation fraction $\alpha = 0.85$ and the dotted line refers to $\alpha = 0.8$. For the shortfall probability given in Proposition 2.10, the Black and Scholes model setup immediately implies

$$SFP = \Phi(\overline{d}_0(K_3))1_{\left\{\alpha^{(E)} \leq \frac{-g(1-\alpha)}{\alpha+g}\right\}} + \Phi(\overline{d}_0(K_2))1_{\left\{\alpha^{(E)} > \frac{-g(1-\alpha)}{\alpha+g}\right\}}, \tag{2.21}$$

where $\overline{d}_0(K) := \dfrac{\ln K - (\mu_A - \frac{1}{2}\sigma_A^2)}{\sigma_A}$.

Again, notice that, e.g. in the context of Solvency II, the upper bound on the shortfall probability is posed to determine the amount of equity which is needed to assure the solvency to a high degree, i.e. to honor the liabilities to the insured. This will be of great interest in the next chapter when we analyze periodic payment streams and how the periodic payments influence the insurers capital requirements.

Fair contracts honoring the upper bound of the shortfall probability

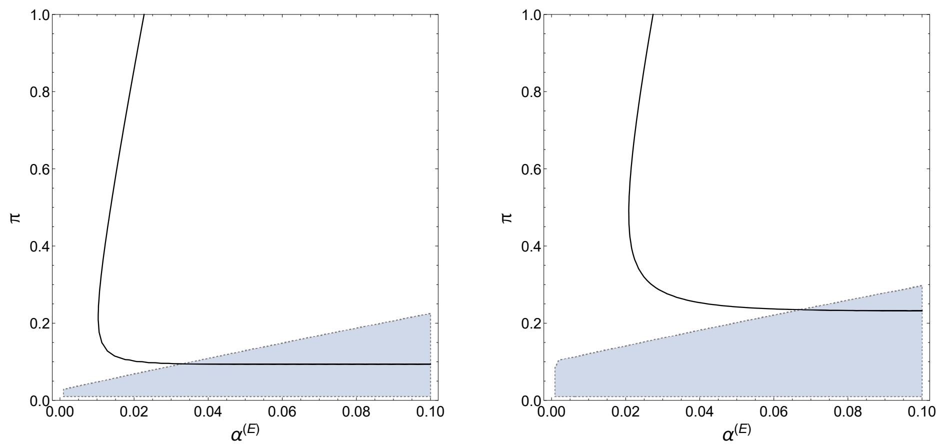


Figure 2.6: If not otherwise mentioned, the contract and model parameters are given as in Table 2.2. The black lines depict the fair contracts in terms of fair combinations of the equity fraction $\alpha^{(E)}$ and the investment fraction π (defining the volatility of the assets, i.e. $\sigma_A = \pi\sigma$). The shaded region is the region where the upper bound on the shortfall probability ($\epsilon = 0.005$) is honored. While the figure on the left hand side refers to the benchmark guarantee $g = 0.0175$, the right hand side is implied by $g = -0.0175$.

Recall that $K_2 = \frac{1+g}{1+\alpha^{(E)}}$ and $K_3 = \frac{1-\alpha}{1-\alpha+\alpha^{(E)}}$: Obviously, the lower the strike is, the lower is the probability of a constant mix strategy that its terminal value drops below the strike. Since the above strikes are decreasing in the equity fraction $\alpha^{(E)}$, a higher equity fraction is able to reduce the shortfall probability, cf. Figure 2.6 for an illustration. It is worth noticing that any reduction of the shortfall probability can also be obtained by suitably adjusting the investment strategy, i.e. the distribution of $\frac{A_1}{A_0}$.

2.3.4 Optimal design of quantile guarantees

The following section discusses, from the perspective of the insured, the optimal design of a MRRG under default risk and an upper bound on the shortfall probability. A fair contract design which provides a higher (expected) utility to the insured is also beneficial to the insurance company. The contract provider competes with other insurers and the financial market. Choosing among different contracts, the insured selects the contract which provides herself the highest (expected) utility. Throughout the following, we assume that the preferences of the insured are described by a utility function $u = u^{(CRRA)}$ implying a constant relative risk aversion (CRRA) denoted by γ, i.e. $u^{(CRRA)}(x) = \frac{x^{1-\gamma}}{1-\gamma}$ $(\gamma > 1)$ and $u^{(CRRA)}(x) = \ln x$ $(\gamma = 1)$. Assuming CRRA preferences has its merits. There are empirical investigations which justify

CRRA preference, cf. e.g. Chiappori and Paiella (2011). In addition, CRRA utility allows that the analysis is based on returns.[31] The relevant optimization problem is posed by maximizing the expected utility of the insured under constraints posed by a competitive market (fair pricing) and the restrictions posed by the regulator.[32] In the first instance, we formulate the optimization problem without stating the optimization arguments, i.e.

$$\max \mathbb{E}_{\mathbb{P}}\left[u(L_1)\right] \text{ s.t. } \mathbb{P}\left(A_1 < P_1\right) \leq \epsilon,\ \mathbb{E}_{\mathbb{P}^*}\left[e^{-r}A_1\right] = 1+\alpha^{(E)} \text{ and } \mathbb{E}_{\mathbb{P}^*}\left[e^{-r}L_1\right] = 1. \tag{2.22}$$

The first condition states the regulatory requirement on the upper bound on the shortfall of the intended payoff (guarantee) P_1. The second condition ensures that the asset value A_1 is obtainable by a self-financing investment strategy with initial investment $A_0 = 1+\alpha^{(E)}$, and the third part captures the fair pricing of the liabilities. To shed further light on the (overall) optimal design of quantile guarantees, we discuss and compare (in the Black and Scholes model setup) different approaches concerning the arguments which are optimally chosen in the maximization problem (2.22) in order to maximize the utility which is provided to the insured. As a benchmark, we consider the optimal unconstrained strategy (no upper bound on the shortfall probability). For $\alpha^{(E)} = 0$, this is the classic Merton problem (cf. Merton (1971)). The solution implies the highest possible utility and thus provides an upper bound of the expected utility of all contract designs.

We also comment on an approach suggested in Schmeiser and Wagner (2015) who assume that the insurer implements a constant mix strategy, but can decide on the fraction of asset wealth which is invested riskily. The insurer simultaneously determines the equity fraction $\alpha^{(E)}$ and the investment fraction π such that the pricing and shortfall constraints are satisfied for a given guarantee g. The utility to the insured is then maximized by selecting the guarantee g which gives the highest expected utility.

Finally, we consider the optimal solution under the pricing and shortfall constraints (without restricting the insurer's investment strategy to constant mix strategies).

[31] It is worth mentioning that CRRA preferences can not explain the existence of (quantile) guarantees using cf. Leland (1980). However, one can understand that policy makers provide tax advantages for products with downside protection for old-age provision to reduce the risk of poverty among the elderly and possible implications for tax payers - even if downside protection reduces utility on the individual level for CRRA-type policyholders. For the effect of taxation on equity-linked life insurance we refer to Chen et al. (2019)

[32] The optimization procedure with a value at risk restriction can be referred to as a chance-constrained approach. It is transferable in a non-linear (deterministic) optimization program of normal or log normal returns are assumed (cf. McCabe and Witt (1980)). Basically, we also consider log normal payoffs for $t = 1, 2, \ldots$ under a Geometric Brownian Motion (GBM) assumption. However, we have added the assumption that the insured is described by a constant relative risk aversion (CRRA) which gives further insights on the utility effects from the perspective of the insured.

The solution can be traced back to the famous result of Basak and Shapiro (2001) resp. our results from the last section. We will see that the optimal quantile hedge does not protect the insured on the bad states of the world.

The Merton solution as a benchmark

Assume that the insured is not committed to select among MRRG contracts, only. Instead, assume that she can, without transaction costs, dynamically trade on the financial market. In terms of the MRRG contracts, this is the special case that $\alpha^{(E)} = 0$ (the insured owns the asset side herself) and a vanishing shortfall probability bound $\epsilon = 1$ (she is not restricted by the regulator). The optimization problem (2.22) then boils down to

$$\max_{A_1} \mathbb{E}_{\mathbb{P}} \left[u \left(\frac{A_1}{A_0} \right) \right] \text{ s.t. } \mathbb{E}_{\mathbb{P}^*} \left[e^{-r} \frac{A_1}{A_0} \right] = 1,$$

i.e. the investor chooses the optimal payoff $L_1 = A_1$ (return, respectively, $A_0 = P_0 = 1$).[33] Assuming a Black and Scholes model setup to describe the financial market model, gives the classic Merton problem. The solution is firstly stated in Merton (1971). Under the real world measure $\mathbb{P}$, the optimal payoff $L_1^* = \frac{A_1^*}{A_0}$ is given by

$$\frac{A_1^*}{A_0} = e^{\mu_A^{(RW)} - \frac{1}{2}\sigma_A^2 + \sigma_A W_1}, \tag{2.23}$$

$$\text{where } \mu_A^{(RW)} = \pi\mu + (1-\pi)r, \ \sigma_A = \pi\sigma \text{ and } \pi = \frac{\mu - r}{\gamma\sigma^2} =: \pi^{(Mer)}.$$

In the optimum, the investor uses a constant mix strategy where the fraction π of portfolio wealth which is invested riskily is given by the quotient of the (local) excess return $(\mu - r)$ and the squared asset volatility scaled by the parameter of relative risk aversion $\gamma\sigma^2$. The certainty equivalent wealth/return CE which makes the investor indifferent to the Merton payoff is defined by the condition $u(CE) = \mathbb{E}_{\mathbb{P}}[u(A_1)]$, i.e. $CE = u^{-1}(\mathbb{E}_{\mathbb{P}}[u(A_1)])$. Straightforward calculations imply

$$CE^* = e^{r + \frac{(\mu - r)^2}{2\gamma\sigma^2}} =: CE^{(Mer)} \text{ and } y^{CE*} = \ln CE^* = r + \frac{(\mu - r)^2}{2\gamma\sigma^2}, \tag{2.24}$$

where y^{CE^*} denotes the (optimal Merton) savings rate. Notice that the above CE^* defines an upper bound to all certainty equivalents which are implied by (admissible) MRRG contracts and refer to the upper bound by $CE^{(Mer)}$. Analogously, we refer to the optimal Merton payoff (fraction) by $A_1^{(Mer)}$ ($\pi^{(Mer)}$).

[33] Recall that $\alpha^{(E)} = 0$ implies $\alpha = 1$, cf. Corollary 2.2 . With $A_0 = 1$ it follows $L_1 = A_1$.

Upper bound on SFP and restriction to constant mix strategies

Schmeiser and Wagner (2015) consider the optimization problem under a SFP condition but assume that the insurer implements a constant mix strategy. In consequence, the insurer does not consider a quantile hedge to honor the guarantee. To ensure the SFP condition for a given guarantee, the insurer is restricted to suitable combinations of investment fractions and equity capital. Amongst other results, Schmeiser and Wagner (2015) consider the optimization problem

$$\max_{g \in \mathcal{G}} \mathbb{E}_{\mathbb{P}}[u(L_1)],$$

where $\mathcal{G}$ denotes the set of admissible guarantee rates and where the equity fraction $\alpha^{(E)}$ and the investment fraction of the asset side π are determined simultaneously by the conditions[34]

$$\mathbb{P}(A_1 < P_1) \leq \epsilon \text{ and } \mathbb{E}_{\mathbb{P}^*}\left[e^{-r} L_1\right] = 1.$$

Notice that $\mathbb{P}(A_1 < P_1) = SFP$ is analytically given by Equation (2.21). The liability value $\mathbb{E}_{\mathbb{P}^*}[e^{-r} L_1]$ is stated in Proposition 2.9 in combination with Equation (2.20).[35] A few comments are worth being mentioned here: Schmeiser and Wagner (2015) consider the exact fulfillment of the shortfall probability corresponding to the minimum safety requirement where the ruin probability SFP is equal to the upper bound ϵ. Intuitively, this is meaningful if the shortfall constraint is binding in the case without equity capital, i.e. if the upper bound on the shortfall probability ϵ is sufficiently low compared to the lowest guarantee contained in the set $\mathcal{G}$. In addition, the authors consider an exogenously given participation fraction α (e.g. $\alpha = 0.9$ as implied by German legislation). However, α ($1 - \alpha$, respectively) implicitly defines a guarantee fee, i.e. the insured gives up some upside participation for downside protection. In particular, if α is already sufficiently low (compared to g), there does not exist an equity fraction $\alpha^{(E)} \geq 0$ such that the (fair) pricing condition can be satisfied, cf. Figure 2.5 and the results in Schmeiser and Wagner (2015).

As a numerical example, we refer to the benchmark parameter setting summarized in Table 2.2 and consider the above mentioned optimization problem for the guarantees g, taking the values $g \in \mathcal{G} = \{-0.1, -0.095, \ldots, 0.02, 0.025\}$ and a shortfall probability bound given by $\epsilon = 0.005$. For each $g \in \mathcal{G}$, Table 2.3 summarizes the combination of equity fraction $\alpha^{(E)}$ and investment fraction π (implying that the

[34] Notice that the condition $\mathbb{E}_{\mathbb{P}^*}[e^{-r} A_1] = 1 + \alpha^{(E)}$ is ensured since the insurer implements a constant mix strategy with initial investment $1 + \alpha^{(E)}$.

[35] Once the equity fraction $\alpha^{(E)}$ and the investment fraction of the asset side π are determined, the expected utility (and CE) can be stated in quasi closed form. Schmeiser and Wagner (2015) determine the solution by Monte Carlo simulations.

SFP is exactly met and the contract is fairly priced) as well as the certainty equivalent contract wealths CEs of insureds which are described by three different levels of relative risk aversion ($\gamma = 2$, 3.56, and 5.94). In addition, the Merton solution is summarized in the upper line. For each level of relative risk aversion, the highest certainty equivalent (CE) is marked which implies the optimal guarantee rate. Observe that the CEs obtained by the (optimal) contracts are close to (but below) the Merton solution. In addition, the corresponding investment fractions π are close to (but above) the Merton fractions. Intuitively, this is explained by the participation fraction α which is (along the lines of the benchmark parametrization) equal to $\alpha = 0.9$, i.e. the investor gives up 10% of the upside returns.

Optimal quantile payoff

As mentioned above, the Black and Scholes model is complete such that any state dependent payoff is attainable, i.e. it can be synthesized by a self-financing strategy in the asset S and the risk free investment opportunity B. In addition with the assumption that the contracts are fairly priced, we can obtain the utility maximizing quantile guarantee payoff L_1 with an initial investment of $P_0 = 1$, i.e. the optimal payoff is independent of the equity fraction $\alpha^{(E)}$. Thus, w.l.o.g. we can set $\alpha^{(E)} = 0$. Recall from Corollary 2.2 that for $\alpha^{(E)} = 0$, a fair contract implies $\alpha = 1$, i.e. $L_1 = A_1 = \frac{A_1}{A_0}$ (since $P_0 = 1$ and $A_0 = 1 + \alpha^{(E)} = 1$), such that the optimization problem (2.22) simplifies to

$$\max_{A_1} \mathbb{E}_{\mathbb{P}}\left[u(A_1)\right] \text{ s.t. } \mathbb{P}\left(A_1 < 1+g\right) \leq \epsilon \text{ and } \mathbb{E}_{\mathbb{P}^*}\left[e^{-r}A_1\right] = 1. \tag{2.25}$$

The solution to this problem can already fully be traced back to Basak and Shapiro (2001) who state the optimal payoff (in dependence of the state prices) under a terminal VaR constraint.[36]

[36] Basak and Shapiro (2001) state the optimal solution in dependence of the state prices for a general class of utility functions in a dynamic complete market setup where the investor can choose between one risk-less bond and several risky stocks.

Certainty equivalents of quantile MRRGs under the additional restriction to constant mix strategies ($\epsilon = 0.005$)

g	$\alpha^{(E)}$	π	L_0	SFP	$CE^{\gamma=2}$	$CE^{\gamma=3.56}$	$CE^{\gamma=5.94}$
					$\pi^{(Mer)} = 0.5$	$\pi^{(Mer)} = 0.28$	$\pi^{(Mer)} = \mathbf{0.169}$
					$CE^{(Mer)} = 1.0408$	$CE^{(Mer)} = 1.0363$	$CE^{(Mer)} = 1.0339$
-0.100	0.1285	0.5277	1	0.005	**1.0405**	1.0341	1.0247
					(1.0406)		
-0.090	0.1211	0.4921	1	0.005	1.0404	1.0348	1.0266
-0.080	0.1140	0.4571	1	0.005	1.0401	1.0353	1.0283
-0.075	0.1105	0.4397	1	0.005	1.0400	1.0355	1.0290
-0.070	0.1073	0.4229	1	0.005	1.0398	1.0357	1.0297
-0.065	0.1034	0.4047	1	0.005	1.0396	1.0359	1.0304
-0.060	0.1000	0.3876	1	0.005	1.0394	1.0360	1.0310
-0.050	0.0925	0.3521	1	0.005	1.0389	1.0361	1.0320
-0.045	0.0890	0.3347	1	0.005	1.0386	**1.0361**	1.0324
						(1.0362)	
-0.040	0.0850	0.3165	1	0.005	1.0383	1.0361	1.0328
-0.035	0.0812	0.2987	1	0.005	1.0380	1.0360	1.0331
-0.030	0.0775	**0.2811**	1	0.005	1.0377	1.0359	1.0334
-0.025	0.0738	0.2634	1	0.005	1.0373	1.0358	1.0336
-0.020	0.0694	0.2443	1	0.005	1.0369	1.0356	1.0337
-0.015	0.0653	0.2259	1	0.005	1.0365	1.0354	1.0338
-0.010	0.0611	0.2074	1	0.005	1.0360	1.0351	**1.0338**
							(1.0338)
-0.005	0.0569	0.1887	1	0.005	1.0356	1.0349	1.0338
0.000	0.0519	**0.1684**	1	0.005	1.0350	1.0345	1.0336
0.005	0.0471	0.1485	1	0.005	1.0345	1.0341	1.0334
0.010	0.0419	0.1278	1	0.005	1.0339	1.0336	1.0331
0.015	0.0362	0.1063	1	0.005	1.0333	1.0331	1.0328
0.020	0.0299	0.0833	1	0.005	1.0326	1.0325	1.0323
0.025	0.0219	0.0567	1	0.005	1.0318	1.0317	1.0317

Table 2.3: The table states, for the benchmark parameter setting summarized in Table 2.2, the results of the optimization problem constrained to constant mix strategies for the set of guarantees $g \in \mathcal{G}$ and a SFP bound given by $\epsilon = 0.005$. In particular, for each g, the combination of equity fraction $\alpha^{(E)}$ and investment fraction π (implying that the SFP is exactly met and the contract is fairly priced) are given in columns two and three. The last three columns summarize the associated certainty equivalent contract wealths CEs of insureds described by levels of relative risk aversion ($\gamma = 2$, 3.56, and 5.94). In addition, the Merton solution is given in the upper line. For each level of relative risk aversion, the highest certainty equivalent (CE), which can be obtained by optimally choosing the guarantee, is marked. For these cases, the CE which can be obtained without a restriction to constant mix strategies is included in italics.

Proposition 2.11 (Optimal quantile return payoff)
If the shortfall probability is not binding, i.e. if $\mathbb{P}\left(\frac{A_1^{(Mer)}}{A_0} \leq 1+g\right) \leq \epsilon$, *the optimal solution coincides with the Merton solution. If the shortfall probability is binding, i.e. if* $\mathbb{P}\left(\frac{A_1^{(Mer)}}{A_0} \leq 1+g\right) > \epsilon$, *the optimal return payoff w.r.t. the optimization problem (2.26) is given as follows*

$$\frac{A_1^*}{A_0} = \nu \frac{A_1^{(Mer)}}{A_0} + \left(1+g-\nu \frac{A_1^{(Mer)}}{A_0}\right) 1_{\left\{\underline{K}<\nu \frac{A_1^{(Mer)}}{A_0} \leq \overline{K}\right\}},$$

where $0 \leq \underline{K} \leq \overline{K} := 1+g$. $\underline{K}$ *is determined by the SFP bound* ϵ *and* ν *by the pricing condition, i.e.*

$$\mathbb{P}\left(\frac{A_1^{(Mer)}}{A_0} \leq \frac{\underline{K}}{\nu}\right) = \epsilon \text{ and } 1-\nu = e^{-r}\mathbb{E}_{\mathbb{P}^*}\left[\left(1+g-\nu \frac{A_1^{(Mer)}}{A_0}\right) 1_{\left\{\underline{K}<\nu \frac{A_1^{(Mer)}}{A_0} \leq \overline{K}\right\}}\right].$$

In the limiting cases $\epsilon \to 1$ *(no constraint on the shortfall probability) and* $\epsilon \to 0$ *(full guarantee) it holds*

(i) For $\epsilon \to 1$ *(and/or* $\mathbb{P}\left(\frac{A_1^{(Mer)}}{A_0} \leq 1+g\right) \leq \epsilon$*), it holds* $\nu = 1$*, and* $\underline{K} = \overline{K}$*, i.e. the optimal (return) payoff is given by the Merton solution* $\left(\frac{A_1^*}{A_0} = \frac{A_1^{(Mer)}}{A_0}\right)$.

(ii) For $\epsilon \to 0$*, it holds* $\underline{K} = 0$ *(and* $\overline{K} = 1+g$*) such that*

$$\frac{A_1^*}{A_0} = (1+g) + \left(\nu \frac{A_1^{(Mer)}}{A_0} - (1+g)\right)^+,$$

where ν *solves*

$$1 = e^{-r}(1+g) + \nu Call^{(BS)}\left(\frac{1+g}{\nu}, \sigma_A^{(Mer)}\right)$$

and $Call^{(BS)}$ *is given by Equation (2.20).*[37]

Instead of explicitly stating the adoption to our setup, it is worth to comment on the intuition behind the result. Obviously, if the quantile constraint is not binding, the

[37] Notice that the pricing condition is, by means of the put call parity, now given in terms of the call price.

optimal solution is given by the Merton solution. W.r.t. the other limiting case where the return payoff is constrained by a shortfall probability of zero ($\epsilon \to 0$), we also refer to El Karoui et al. (2005). The optimal unconstrained payoff is a modification of the Merton solution (unconstrained solution).[38] Intuitively, it is clear that a full hedge of the guarantee features a put option. Notice that

$$(1+g) + \left(\nu \frac{A_1^{(Mer)}}{A_0} - (1+g)\right)^+ = \nu \frac{A_1^{(Mer)}}{A_0} + \left((1+g) - \nu \frac{A_1^{(Mer)}}{A_0}\right)^+ ,$$

i.e. the return of the Merton solution is backed up by a put option with strike $K = 1+g$. The put payoff gives the tightest (and thus cheapest) possibility to obtain a full hedge of the guarantee. Thus, it enables the investor to obtain the tightest modification of the unconstrained optimal payoff. Here we see the connection to the previous section and Proposition 2.5 where we analyzed the cost-efficient payoff modification by fulfilling a VaR constraint. To honor the pricing condition, i.e. the value of the payoff must be equal to one, the investor can no longer obtain the full Merton return but only a fraction ν of it. In particular, while the value of $\frac{A_1^{(Mer)}}{A_0}$ is equal to one, the investor now receives only a fraction of the return, i.e. in the presence of a (non vanishing) guarantee, her investment amount which is not needed to finance the put is only a fraction ν ($0 < \nu < 1$).
In summary, the fraction ν is determined by a fix point problem which is due to the condition that the value of the put on the return $\nu \frac{A_1^{(Mer)}}{A_0}$ must be equal to the reduction of the initial investment $1 - \nu$ (i.e. both sides depend on ν). Intuitively it is now clear that any deviation from a perfect guarantee ($\epsilon \to 0$), an admissible shortfall probability which is higher than zero gives rise to lower hedging costs than the solution characterized above. While in the case of a zero shortfall probability the optimal payoff is given by

$$\nu \frac{A_1^{(Mer)}}{A_0} + \left((1+g) - \nu \frac{A_1^{(Mer)}}{A_0}\right) 1_{\left\{\underline{K} < \nu \frac{A_1^{(Mer)}}{A_0} \leq \overline{K}\right\}},$$

where $\underline{K} = 0$ and $\overline{K} = 1+g$, the investor is now allowed to implement a smaller *guarantee interval* $[\underline{K}, \overline{K}]$ where $0 \leq \underline{K} < \overline{K} \leq 1+g$. Notice that the upper bound on the shortfall probability implies that fixing either $\underline{K}$ or $\overline{K}$ implies the other strike such that ν is determined by the resulting fix point problem. However, the cheapest way to do so is by setting $\overline{K} = 1+g$, i.e. starting with the high asset prices (Merton returns, respectively) which are linked to the cheapest states (to be hedged).
Using the results from the last section, we can overcome this problem by e.g. using the Wang transformation risk measure as a shortfall constraint or by adding another

[38] In fact, the result does not depend on the Black and Scholes model which implies the Merton solution.

constraint to the optimization problem that guarantees the insured a protection on the bad states of the world. The paper of Chen et al. (2018a) presents an additional portfolio insurance constraint in a general setting s.t. the insured is protected on the bad states of the world.
In summary, the optimal quantile hedge is a scaled version of the Merton solution overlaid by the (cheapest) quantile hedge which honors the SFP bound.[39] In order to illustrate the improvement obtained by the optimal quantile hedge, we add in Table 2.3 the CEs associated with the optimal quantile guarantees, cf. italic numbers in brackets below the bold faced numbers referring to the optimal values under the restriction to constant mix strategies (and choosing the guarantee). Again, it is worth to emphasize that the optimal quantile payoff can be implemented for any equity fraction $\alpha^{(E)}$ of the insurer.

2.3.5 Conclusion

In this section we have analyzed the optimal design of participating life insurance contracts with minimum return rate guarantees under default risk. The benefits to the insured depend on the performance of an investment strategy which is conducted by the insurer. This strategy is initialized by an amount given by the sum of equity and the contributions of the insured. Unless there is a default event, the insured receives the maximum of a guaranteed rate and a participation in the returns. Considering default risk modifies the payoff of the insured by means of a default put implying a compound option feature (nested maximum). Based on yearly returns, we show that, in spite of the compound option feature, the (yearly return) payoff of the default put (and the liabilities to the insured) can be represented by piecewise linear functions of the investment return, i.e. the payoff of a portfolio of plain vanilla options. Thus, the liabilities are easily priced in any model setup which gives closed form solutions for standard options. In a complete market setup we then derive the optimal (expected utility maximizing) quantile guarantee payoff of an investor/insured with constant relative risk aversion. Because of the completeness assumption, the return payoff can be implemented by the insurance company for any equity to debt ratio. We illustrate the utility loss which arises if the insurer implements a suboptimal investment strategy.

[39] W.r.t. quantile hedges, the interested reader is referred to Föllmer and Leukert (1999) who show how to obtain the highest success probability when hedging a claim with a lower initial investment than the one needed for a full hedge (or the other way round). An introduction to quantile hedging and the Neyman-Pearson Lemma is also presented in Appendix A.

2.3.6 Extension and literature review on portfolio planning under risk measure constraints

It is also possible to generalize Proposition 2.11 to a maturity $T > 1$. The optimization problem can then be stated as

$$A_T^* = \underset{A_T}{\operatorname{argmax}}\ \mathbb{E}\left[u(A_T)\right] \text{ s.t. } \mathbb{P}\left(A_T < G_T\right) \leq \varepsilon \text{ and } \mathbb{E}_{\mathbb{P}^*}\left[e^{-rT}\frac{A_T}{A_0}\right] = 1, \tag{2.26}$$

where G_T is the guaranteed value of the insured at time T.

Proposition 2.12 (Optimal quantile return payoff with maturity T)
If the shortfall probability is not binding, i.e. if $\mathbb{P}\left(A_T^{(Mer)} \leq G_T\right) \leq \varepsilon$, the optimal solution coincides with the Merton solution. If the shortfall probability is binding, i.e. if $\mathbb{P}\left(A_T^{(Mer)} \leq G_T\right) > \varepsilon$, the optimal return payoff is given by

$$A_T^* = \nu A_T^{(Mer)} + \left(G_T - \nu A_T^{(Mer)}\right) 1_{\left\{\underline{K} < \nu A_T^{(Mer)} \leq \overline{K}\right\}},$$

where $0 \leq \underline{K} \leq \overline{K} := G_T$. $\underline{K}$ is determined by the SFP bound ε and ν by the pricing condition, i.e.

$$\mathbb{P}\left(\nu A_T^{(Mer)} \leq \underline{K}\right) = \varepsilon \text{ and } 1 - \nu \ = e^{-rT}\mathbb{E}_{\mathbb{P}^*}\left[\left(G_T - \nu A_T^{(Mer)}\right) 1_{\left\{\underline{K} < \nu A_T^{(Mer)} \leq \overline{K}\right\}}\right].$$

Calculating the SFP and the pricing condition in Proposition 2.12, we receive the following closed-form formulas.

Lemma 2.3 (Shortfall probability and pricing formula)
(i) The shortfall probability is given by

$$\mathbb{P}\left(\nu A_T^{(Mer)} \leq \underline{K}\right) = \mathcal{N}\left(-\frac{\ln\left(\frac{\nu A_0^{(Mer)}}{\underline{K}}\right) + (\mu - \frac{1}{2}\sigma_A^2)T}{\sigma_A\sqrt{T}}\right).$$

(ii) The t–price of the T–payoff $\left(G_T - \nu A_T^{(Mer)}\right) 1_{\left\{\underline{K} < \nu A_T^{(Mer)} \leq \overline{K}\right\}}$ is given by

$$\begin{aligned}&\mathbb{E}_{\mathbb{P}^*}\left[e^{-r(T-t)}\left(G_T - \nu A_T^{(Mer)}\right) 1_{\left\{\underline{K} < \nu A_T^{(Mer)} \leq \overline{K}\right\}}\middle|\mathcal{F}_t\right]\\ =&e^{-r(T-t)}G_T\left[\mathcal{N}\left(-d_2(\overline{K})\right) - \mathcal{N}\left(-d_2(\underline{K})\right)\right] - \nu A_t^{(Mer)}\left[\mathcal{N}\left(-d_1(\overline{K})\right) - \mathcal{N}\left(-d_1(\underline{K})\right)\right],\end{aligned} \tag{2.27}$$

where

$$d_1(K) = \frac{\ln\left(\frac{\beta A_t^{(Mer)}}{K}\right) + (r + \frac{1}{2}\sigma_A^2)(T-t)}{\sigma_A\sqrt{T-t}} \text{ and } d_2(K) = d_1(K) - \sigma_A\sqrt{T-t}.$$

In particular, for $\underline{K} = 0$ *and* $\overline{K} = G_T$ *this is the price of a plain vanilla put on* νA *with strike* $K = G_T$*, i.e.*

$$\mathbb{E}_{\mathbb{P}^*}\left[e^{-r(T-t)}\left(G_T - \nu A_T^{(Mer)}\right)^+ \middle| \mathcal{F}_t\right] = e^{-r(T-t)} G_T \mathcal{N}(-d_2(\overline{K})) - \nu A_t^{(Mer)} \mathcal{N}(-d_1(\overline{K})).$$

Setting $T = 1$ and exchanging $(1+g)$ with G_T we receive the formulas for our benchmark case with maturity $T = 1$. The proof of Lemma 2.3 is given in Appendix B.1.
Given the optimal splitting factor ν we can even calculate the optimal expected utility maximizing strategy in closed-form.

Lemma 2.4 (Optimal Expected utility)
For a CRRA utility function with risk aversion parameter γ *the optimal expected utility is given by*

$$\begin{aligned}\mathbb{E}\left[u(A_T^*)\right] &= \mathbb{E}\left[u\left(\nu A_T^{(Mer)} + \left[G_T - \nu A_T^{(Mer)}\right] 1_{\{\underline{K} \leq \nu A_T^{(Mer)} \leq \overline{K}\}}\right)\right] \\ &= \frac{(\nu A_0)^{(1-\gamma)}}{1-\gamma} e^{(1-\gamma)\left(\mu_A T - \frac{1}{2}\gamma\sigma_A^2 T\right)}\left[1 - \left(\mathcal{N}\left(\tilde{d}(\overline{K},\gamma)\right) - \mathcal{N}\left(\tilde{d}(\underline{K},\gamma)\right)\right)\right] \\ &\quad + \frac{G_T^{(1-\gamma)}}{1-\gamma}\left(\mathcal{N}(\tilde{d}(\overline{K},1)) - \mathcal{N}(\tilde{d}(\underline{K},1))\right),\end{aligned}$$

where

$$\tilde{d}(K,\gamma) := \frac{\ln\left(\frac{K}{\nu A_0}\right) - \left(\mu_A - (\gamma - \frac{1}{2})\sigma_A^2\right) T}{\sigma_A\sqrt{T}}.$$

The first term is the expected utility of the scaled Merton payoff. The second term resembles the utility loss caused by minus the probability weighted scaled Merton loss, and the third term is the probability weighted utility of the guarantee.
The proof of Lemma 2.4 is given in the Appendix B.2.

A detailed analysis of this topic including multiple intermediate VaR constraints can be found in Chen et al. (2018b). The intuition behind this is briefly stated in the following:

Consider an optimization problem stemming from a regulator who also poses an intermediate VaR constraint. In addition to the terminal SFP condition $\mathbb{P}(A_T < G_T) \leq$

ε_T, we thus introduce an additional SFP condition at t $(0 < t < T)$, i.e. the requirement $\mathbb{P}(A_t < G_t) \leq \varepsilon_t$.

In summary, the optimization problem under consideration is given by

$$\begin{aligned} A_T^* = \underset{A_T}{\operatorname{argmax}}\ \mathbb{E}\left[u(A_T)\right], \text{ s.t. } \mathbb{E}_{\mathbb{P}^*}\left[e^{-rT}\frac{A_T}{A_0}\right] = 1 \\ \text{with } \mathbb{P}(A_t < G_t) \leq \varepsilon_t \text{ and } \mathbb{P}(A_T < G_T) \leq \varepsilon_T. \end{aligned} \tag{2.28}$$

Using that the utility function is a CRRA function, it holds

$$A_T^* = \underset{A_T}{\operatorname{argmax}}\ \mathbb{E}\left[u(A_T)\right] = \underset{A_T}{\operatorname{argmax}}\ \mathbb{E}\left[u\left(\frac{A_T}{A_0}\right)\right].$$

Therefore, we can also consider

$$\begin{aligned} A_T^* = \underset{A_T}{\operatorname{argmax}}\ \mathbb{E}\left[u\left(\frac{A_t}{A_0}\frac{A_T}{A_t}\right)\right], \text{ s.t. } \mathbb{E}_{\mathbb{P}^*}\left[e^{-rT}\frac{A_T}{A_0}\right] = 1 \\ \text{with } \mathbb{P}(A_t < G_t) \leq \varepsilon_t \text{ and } \mathbb{P}(A_T < G_T) \leq \varepsilon_T. \end{aligned}$$

Let $R_{t,T} := \frac{A_T}{A_t}$. With the Bellmann principle together with the results of the previous section, it immediately follows that

$$\begin{aligned} A_T^* &= \underset{A_t}{\operatorname{argmax}}\ \mathbb{E}\left[u\left(A_t\ R_{t,T}^*(A_t)\right)\right] \\ \text{where } R_{t,T}^* &= \nu h^*\left(\frac{S_T}{S_t}\right) + \left(G_T - \nu h^*\left(\frac{S_T}{S_t}\right)\right) 1_{\left\{\underline{K} < \nu A_t h^*\left(\frac{S_T}{S_t}\right) \leq \overline{K}\right\}} \\ \text{and } h^*\left(\frac{S_T}{S_t}\right) &= e^{\left(1-\pi^{(Mer)}\right)\left(r+\frac{1}{2}\pi^{(Mer)}\sigma^2\right)T}\left(\frac{S_T}{S_t}\right)^{\pi^{(Mer)}}. \end{aligned}$$

For the exact solutions and proceedings, we refer to Chen et al. (2018b). This multiple VaR constraint setting is important for short-time VaR-type regulations in combination with long-term liability commitments and leads to a more prudent investment behavior than the optimal investment strategy under a terminal VaR constraint.

As discussed in the previous section for a CRRA investor, the expected utility maximizing terminal wealth that fulfills a terminal VaR constraint impedes the securitization aspect of the investor. The results of Chen et al. (2018b) as discussed above give reason to suspect, that this impeding behavior can be solved if we implement another risk measure as shortfall constraint resp. if we add some further constraints or restrictions to the problem. We want to respond to this question by analyzing the literature on optimal portfolio planning under risk measure constraints. Table 2.4 summarizes the literature review where the tackled optimization problem, the corresponding risk measure constraint and the model assumptions are stated. Merton (1969) and Merton (1971) are the first papers that discuss the expected utility of a CRRA resp. CARA investor. The maximum is achieved by using the constant investment fraction $\pi^{Mer} = \frac{\mu - r}{\gamma \sigma^2}$. This investment fraction is the so-called Merton fraction. As stated at the beginning of this chapter, the VaR was introduced to the public in 1994 and implemented in the Basel accord 1996. The risk measure conception of Artzner et al. (1999) has been introduced 3 years after the Basel I accord. Thus, there exists a time gap between the famous results of Merton and the first portfolio planning problem that introduced a risk measure as a constraint in literature. Basak and Shapiro (2001) have been the first ones that maximized the expected utility of an investor by choosing the optimal terminal payoff strategy s.t. a VaR constraint is fulfilled in a continuous-time setting. This optimal solution impedes the idea of protection on the bad states of the world as discussed in the previous section. Leippold et al. (2006) generalize the setting of Basak and Shapiro by analyzing an incomplete market setup. They show that the effectiveness of regulation strongly depends on market factors and the chosen model. Or to state it in other words: in another model than the Black-Scholes setting the optimal solution under the terminal VaR constraint behaves differently and may not cause the impeding behavior. In the Black-Scholes Model setting this problem can also be solved if we account for more than one VaR constraint: Yiu (2004) presents a dynamic VaR constraint that guarantees the fulfillment of a SFP for the whole investment horizon within a n-dim. Black-Scholes setting. The optimal CRRA expected utility maximizing solution protects the investor in the bad states of the world. Pirvu (2007) also analyzes a dynamic VaR constraint in an incomplete market model and finds the same results as Yiu. Cuoco et al. (2008) analyze a dynamic VaR resp. ES constraint, i.e. the wealth process has to fulfill the SFP at every point in time t. They determine the optimal investment fraction s.t. the expected utility of the terminal wealth is maximized and explore, that both risk measures lead to a risk reduction. They even formulate conditions under which the optimal solutions coincide. Thus, they can confirm, that under a dynamic risk measure constraint the impeding effect is mitigated. Shi and Werker (2012) extend the work of Cuoco et al. (2008) as they account for stochastic interest rates. They show that the results of Cuoco et al. (2008) are robust for stochastic interest rates.

Some papers include other risk measure constraints than the VaR and the ES: A terminal expected loss constraint is investigated by Gabih et al. (2005), whereas Gundel and Weber (2008) are the first ones to analyze the optimal portfolio planning problem with a terminal convex risk measure and Rogers (2009) introduces terminal coherent risk measures as a constraint. The authors find that only a terminal SFP constraint does not fulfill the protection on the bad states of the world. Thus, the question arises if any risk measure that serves as a terminal SFP constraint can protect the investor on the bad states without adding an additional feature. The answer to this question is given by Wei (2018). He introduces the weighted VaR as a terminal risk constraint in the expected utility maximization problem of the terminal wealth. As discussed in the first section of this chapter, the WVaR represents a whole class of risk measures s.t. this paper summarizes results of many other papers and gives some new insights into other risk measures. It is shown that the problem of not securing the bad states of the world in the optimal solution can be solved by using e.g. the Wang distortion risk measure. Wei states precise conditions under which a risk measure, as a terminal SFP, secures the bad states of the world. Wei (2021) introduces another risk measure that solves this problem, the so-called weighted shortfall. This risk measure is a development of the ES because the ES as terminal SFP constraint does not fulfill the assumptions in Wei (2018).

In the special context of insurance contracts in a Black-Scholes setting, some recent papers extend the results of Basak and Shapiro (2001) and find interesting insights: Chen et al. (2018a) determine in a 1-dim. Black-Scholes setup the optimal payoff for the terminal wealth and the optimal investment fraction s.t. the expected utility under a joint VaR and portfolio insurance (PI) constraint on the terminal wealth is fulfilled. They discover that adding the PI constraint to the optimization problem helps to secure the investor on the bad states of the world s.t. the optimal strategy is prudent. Chen et al. (2018b) contribute to the literature by adding multiple, intermediate VaR constraints as described above. Nguyen and Stadje (2020) analyze a terminal VaR constraint in the context of non-concave expected utility maximization including mortality risk. They figure out that the impeding behavior of the optimal solution by Basak and Shapiro (2001) does not transfer to the situation under mortality risk: the VaR constraint even leads to a more prudent risk structure. For the sake of completeness, some papers analyze a mean-variance setting and include a risk measure constraint. We refer to Alexander and Baptista (2004) and Alexander et al. (2007) for terminal VaR resp. terminal ES formulations and to Gao et al. (2016) for a dynamic risk measure assumption of the problem. Bi and Cai (2019) derive in a mean-variance setting the optimal reinsurance strategy with a terminal VaR constraint.

From this literature review we can conclude, that only a terminal SFP in the Black-Scholes world in most cases cannot protect the investor against the impeding behavior of the optimal solution. If we want to stick to a terminal SFP constraint without further restrictions or assumptions and want protection, the risk measure needs to fulfill the conditions in Theorem 4.7 resp. Proposition 5.1 in Wei (2018). If we allow for a more general model, more precise modeling of the reality (e.g. including mortality risk in the analysis), or more realistic risk management that includes more than a terminal SFP (e.g. dynamic SFP constraints), then the optimal expected utility maximizing solution protects the investor even on the bad states of the world. [40]

[40] For portfolio allocation problems under uncertainty, we refer to the literature review in Chapter 4. A literature review on portfolio allocation problems including guarantee features is given in

Authors	Risk Measure Constraint	Optimization Problem	Assumptions
Merton (1971)	-	EU, $(\pi_t)_{t\in[0,T]}$	n-dim. BS-model; CRRA and CARA utility
Basak and Shapiro (2001)	terminal VaR	EU, W_T	complete, n-dim. market model (GBM); strictly increasing, concave utility function, twice differentiable with Inada condition[41]
Alexander and Baptista (2004)	terminal VaR terminal ES	MV	n-dim. MV model; rates of return multivariate normal distributed
Yiu (2004)	dynamic VaR	EU, $(\pi_t)_{t\in[0,T]}$	complete, n-dim. market model (GBM); power utility function; include consumption
Gabih et al. (2005)	terminal EL	EU, W_T	1-dim. BS model; strictly increasing, concave utility function, twice differentiable with Inada condition
Leippold et al. (2006)	terminal VaR	EU, $(\pi_t)_{t\in[0,T]}$	incomplete 1-dim. market model; CRRA utility
Alexander et al. (2007)	terminal VaR terminal ES	MV	n-dim. MV model; returns have discrete distribution with finitely many jump points
Pirvu (2007)	dynamic VaR	EU, $(\pi_t)_{t\in[0,T]}$	incomplete, n-dim. market model (GBM) with random drift and volatility; CRRA utility function

(To be continued)

Chapter 3.

[41] The Inada conditions are given by $\lim_{x\to\infty} u'(x) = 0$ and $\lim_{x\searrow\bar{x}_u} u'(x) = \infty$, where $\bar{x}_u := \inf\{x \in \mathbb{R} : u(x) > -\infty\}$.

Authors	Risk Measure Constraint	Optimization Problem	Assumptions
Boyle and Tian (2007)	terminal VaR	EU, $(\pi_t)_{t\in[0,T]}$	complete model with n risky assets and 1 risk-free asset, upfront contributions; utility function is continuously differentiable, strictly increasing and concave and fulfills Inada
Cuoco et al. (2008)	dynamic VaR dynamic ES	EU, $(\pi_t)_{t\in[0,T]}$	complete, n-dim. market model (GBM); strictly increasing, concave utility function, continuously differentiable with Inada condition
Daníelsson et al. (2008)	terminal VaR	EU, W_T	complete, discrete n-dim. Arrow-Debreu setting
Gundel and Weber (2008)	terminal convex risk measure	EU, W_T	n-dim. semimartingale model; strictly increasing, strictly concave, continuously differentiable utility function with Inada condition
Rogers (2009)	terminal coherent risk measure	EU, W_T	complete, 1-dim. market model (GBM); concave, and strictly increasing utility function with $\lim_{x\to\infty} u'(x) = 0$
Shi and Werker (2012)	two-period VaR	EU, W_T	complete market model (GBM), stochastic interest rates (Vasicek model), CRRA utility
Kraft and Steffensen (2013)	terminal VaR terminal ES	EU, $(\pi_t)_{t\in[0,T]}$	1-dim. black-Scholes Model; CRRA utility
Gao et al. (2016)	dynamic VaR dynamic ES	MV	complete, n-dim. market model (GBM)
Zhang and Gao (2016)	terminal ES	EU, $(\pi_t)_{t\in[0,T]}$	complete, n-dim. market model (GBM); log utility
Zhao and Xiao (2016)	terminal VaR	EU, $(\pi_t)_{t\in[0,T]}$	n-dim. market model; asset price is modeled by the non-extensive statistical mechanics[42]; log-utility function
Chen et al. (2018a)	combined terminal VaR and PI	EU, $(\pi_t)_{t\in[0,T]}$ EU, W_T	1-dim. BS model; twice differentiable utility function, fulfills Inada and integrability condition

(To be continued)

Authors	Risk Measure Constraint	Optimization Problem	Assumptions
Chen et al. (2018b)	multiple VaR	EU, $(\pi_t)_{t\in[0,T]}$	complete, 1-dim. market model (GBM); strictly increasing, concave utility function, twice differentiable with Inada condition
Wei (2018)	terminal WVaR[43]	EU, W_T	complete, n-dim. market model (GBM); twice continuously differentiable, strictly increasing and strictly concave utility function that fulfills Inada condition
Bi and Cai (2019)	dynamic VaR	MV	1-dim. correlated market model with state dependent risk aversion; optimal reinsurance strategy derived
Nguyen and Stadje (2020)	terminal VaR	EU, W_T	complete, n-dim. market model (GBM); including mortality risk and non-concave utility
Wei (2021)	terminal ES terminal weighted shortfall[44]	EU, W_T	complete, n-dim. market model (GBM); strictly increasing and concave utility function, continuously differentiable that fulfills Inada condition

EU=expected utility; MV=mean-variance; PI=portfolio insurance

Table 2.4: Selected papers on optimal portfolio planning under risk measure constraints

[42] This model setup is able to account for fat tails.

[43] With the WVaR many risk measures can be represented as discussed in the previous section. We can construct spectral risk measures, distortion risk measures and coherent risk measures in this setting. Thus the paper of Wei (2018) represents many applications in the context of portfolio planning with risk measure constraints.

[44] This is a special case of a spectral risk measure.

Chapter 3

The impact of periodic premium payments and management rules on the pricing, risk management and expected utility

This chapter analyzes periodic premium payments in participating insurance contracts including a terminal guarantee. Premium payments of the insured are one of the main aspects of (life) insurance contracts. In participating life insurance contracts the insured participates with her premium payments on the asset side of the insurance company: the payments are invested in risky or risk-free assets. If the insured pays her premiums as a single payment at inception (at the end) of the contract, this is called an upfront (postponed) premium. If she splits her contributions over the contract horizon, e.g. monthly or yearly recurring payments, this is called periodic payments. In Germany, periodic payments are well accepted by policyholders: 57% of all life insurance contracts are concluded with periodic payments as stated by GDV (2021a). Thus, it is interesting to analyze periodic premium payments. Especially, since many papers in academia only assume upfront premium payments.

As stated before, the insured's payments are invested in risky and risk-free assets. Thus, the terminal contract value of the insured depends on the asset evolution. To avoid high losses and to protect the insured, typically a guarantee component is included in the contract: the insured receives at the end of the contract period the maximum of the guarantee component and the terminal wealth of the asset evolution. There are many different guarantee features, of which the most common one in literature is the terminal guarantee. Here, the value of the asset result is compared with the promised guarantee at end of the contract. But this guarantee feature is not for free: The insured participates only with a rate of $\alpha \in (0, 1)$ at her terminal account value. The remaining part $(1 - \alpha)$ corresponds to the guarantee costs. These costs depend on both, the guaranteed value and the type of premium payments: a higher guarantee of course results in higher guarantee costs. Also, the premium payments should have a huge impact on the guarantee costs: an upfront payment implies that the insured participates the whole investment horizon fully in the asset side. A periodic payment structure implies that the participation in the asset side increases with every period. The influence of periodic payments on the guarantee costs is important for the insurance company and thus an interesting aspect to analyze.

The guarantee promise of the insurance company can get problematic if the asset return evolution is not promising and the asset results are smaller than the promised guarantee: a shortfall risk occurs. For these cases the insurance companies have to build reserves and capital requirements as stated in Solvency II, s.t. this default event can only occur with a probability of 0.5%. It is interesting to analyze how the premium payments (periodic vs. one-time payment) influence the required capital. This analysis is important for insurance companies because the smaller the required capital, the more they can use to invest.

We should also include the perspective of the insured in this analysis: she has many different contract offerings on the insurance market and will pick the contract which maximizes her willingness to pay for. We interpret the willingness to pay in terms of her expected utility, i.e. the higher her expected utility the higher is her willingness to enter the contract. Thus, maximizing the expected utility of the terminal wealth of the insured is also of interest fort the insurance company. The impact of the premium payments (periodic vs. one-time payment) is also in this case of great importance.

Let us take a look at the insurance company again. As stated before, the insured decides in our setting about the payment scheme, s.t. the insurance company cannot influence the decision. But there exists one aspect that allows the insurance company to control all of the former mentioned aspects: the investment fraction in the risky asset. A change in the investment fraction directly influences the guarantee costs, the required capital to fulfill the shortfall probability constraint and the expected utility of the insured. Thus, it is interesting to analyze the effects of a tool (we call it management rule) that controls the investment fraction, depending on the former results of the asset returns and its interplay with the periodic premium payments. Therefore, we analyze the impact of periodic premium payments on participating life insurance contracts under management rules in the main part of this chapter and find implications on the pricing, the risk management and the expected utility of the insured.

Finally, one wonders what happens if we use a more complex guarantee feature instead of the terminal guarantee. We motivate two guarantee schemes (the cliquet-style and ratchet guarantee) and give a literature overview on the impact of these different guarantee features on optimal portfolio planning.

3.1 Participating Life Insurance Contracts with Periodic Premium Payments under Regime Switching

In this section, we consider (life-) insurance contracts where the insured receives the maximum of a guaranteed amount and (a fraction of) a stochastic payoff given by the outcome of a risky investment strategy.[45] Under fair pricing, the contribution of the insured defines the possible combinations of the guaranteed amount and the participation fraction in the stochastic payoff. The higher the guarantee is, the lower is the fair participation in the stochastic payoff. While there is a large strand of literature that considers the risk management, the pricing, and the benefits to the insured in the context of a single upfront contribution, fewer papers include the impact of periodic contributions, i.e. where some contributions are postponed to the future. We discuss in general the contract design of the payoff of a participating contract with terminal guarantee feature in dependence of a flexible payoff scheme to motivate our chosen stylized model setup. In this setup we analyze the effects of premium payment postponements on the risk profile of the provider as well as the benefits to the insured.
From the perspective of the product provider's risk management, there is an additional risk in case of periodic premium (cf. Bernard et al. (2017)). In case of a single upfront premium, the (fair) guarantee costs are paid by the insured at inception of the contract and can be used to mitigate (hedge) the guarantee. However, this is not possible if some of the contribution (and thus guarantee premium) is postponed to the future. Without pre-financing (and implementing a dynamic hedging strategy at the beginning), the part of the guarantee costs which is paid in the future can be lower than the amount which is then needed to hedge the guarantee (depending on the moneyness of the insurance put option). In case the contract provider is allowed to adjust the investment strategy, he can use the adjustment such that the additional risk from the periodic premium payments is mitigated. This possibility is captured in our model in terms of a *management rule* under which the insurance company can adjust the investment strategy.
Some of our results can be applied to products that belong to the class of equity-linked (or unit-linked) products with an interest rate guarantee where the insured's benefit is linked to the performance of a specific reference portfolio. Typical guaranteed equity-linked products are Equity-Index Annuities (EIA) which are very popular in the North American market or their German counterparts Select Products.[46] In addition, these results concern Variable Annuities which are insurance contracts with guarantees designed as pension products where the trusted fund is invested in a reference portfolio. However, a special focus is on traditional (German) participating

[45] This section is based on the work of Mahayni et al. (2021c).

[46] One of these products is for example "IndexSelect" by Allianz. For a detailed analysis of this product and an overview of further equity-linked products we refer to Alexandrova et al. (2017).

life insurance contracts with an annual surplus distribution which is also linked to the asset side of the insurance company. To isolate the impact of the management rule, we abstain from additional rights which may be included in those insurance products such as surrender options, paid-up and resumption options or guaranteed annuity options.[47]

The present value of the insured's contributions are normalized to one and the contribution schemes in our setting (upfront, postponed, periodic) determine when and how the policyholder participates in the asset side of the insurance company s.t. the different schemes and even the different management rules imply different guarantee costs. These costs have to be fairly priced s.t. the contracts do not account for arbitrage. For this we set ourselves in a Black-Scholes model and analyze the impact of the periodic payments (splitting factor) and the management rules on the fair pricing of the guarantee. But these two components (splitting factor and management rule) not just have an impact on the fair pricing but even more on the risk management of the contract provider as stated before. This crucial observation is analyzed in detail and we give solutions to overcome this problem. We furthermore analyze the perspective of the policyholder. Assuming that the risk preferences of the policyholder are described by a constant relative risk aversion (CRRA), we consider her portfolio planning problem in terms of maximizing her expected utility w.r.t. how to split her contributions over time. We examine the impact of periodic contributions of the policyholder on her expected utility (from which her willingness to pay can be derived) under different management rules (investment strategies, respectively) of the insurance company.

Our main contributions of this section can be summarized as follows: First, we analyze in general the account value of a participating contract with terminal guarantee in dependence of a flexible payoff scheme (n possible premium payments) and find a representation that depends on just one splitting factor. Thus we can qualitatively analyze in the following the impact of periodic premium contributions in a stylized setup with two premium payment dates. In a Black-Scholes Model setting we discuss the different management rules with which the insurance company can react to bad or good market movements and adapt the investment strategy. A main focus is on the impact of the splitting factor and the management rules on the fair pricing of the contract. We derive quasi closed-form solutions for the guarantee costs with periodic premium payments and constant management rule (investment fractions coincide) and closed-form solutions for the special cases of an upfront resp. postponed premium payment. Furthermore, we show that the guarantee costs in this case are convex and strictly monotonically increasing in the splitting factor s.t. an upper bound for the guarantee costs depending on the splitting factor is derived. For the variable management rule under a simple assumption on the adapted investment fraction we explore that also the guarantee costs of the variable management rule

[47] For an overview of embedded options in life insurance contracts we refer to Gatzert (2009).

are strictly increasing and convex in the splitting factor. Moreover, they are greater than the ones under the constant management rule for a fixed splitting factor. From the risk management perspective of the insurance company we show that periodic premium payments lead to an increase of the riskiness (compared to an upfront payment) in terms of the shortfall probability (SFP) of the insurance company resp. to an increase of the required solvency capital s.t. a SFP is matched. By implementing a variable management rule this riskiness can be reduced greatly.
Assuming that the risk preferences of the insured are described by a CRRA utility function, we also study her portfolio planning problem in terms of maximizing her expected utility concerning how to split her contributions optimal over time. We find that splitting the contributions has a huge impact on the expected utility of the insured. Under a constant management rule and no guarantee the Merton solution (i.e. the investment fraction is given by the famous Merton fraction and the contribution of the insured is given by an upfront premium payment) archives the highest expected utility. Deviations from the Merton fraction imply that splitting the premium payments becomes optimal for the insured. In case of a variable management rule we show that the overall solution is not given by the Merton solution anymore s.t. periodic payments can be used to maximize the expected utility of the insured.

This section is related to several strands of the literature including (i) periodic premium contribution schemes, (ii) pricing of embedded guarantees (options, respectively), (iii) risk management, (iv) utility losses caused by guarantees and/or suboptimal investment decisions conducted by insurance companies, (v) portfolio planning and (vi) regime-switching. Without postulating completeness, we only refer to a subset of related literature and hint at the additional literature given within the mentioned papers. Pricing of long-term guarantees by no-arbitrage dates back to Brennan and Schwartz (1976). The fair valuation of participating life insurance contracts is, to the best of our knowledge, first analyzed by Briys and De Varenne (1994) as also Briys and De Varenne (1997). Recent contributions are e.g. Kling et al. (2011), Chong (2019), Orozco-Garcia and Schmeiser (2019), Hieber et al. (2019) and Bacinello et al. (2021). The first consideration of periodic contributions dates back to Brennan and Schwartz (1976). Further discussions about single and periodic contributions can be found in e.g. Bacinello and Ortu (1993a), Bacinello and Ortu (1993b), Bacinello and Ortu (1994), Nielsen and Sandmann (1995) as also Nielsen and Sandmann (1996). They consider constant periodic contributions, while we allow for flexible periodic contributions of the policyholder. More recent works include Gatzert (2013), Bernard et al. (2017) and Eckert et al. (2021).
Periodic contributions generally do not lead to closed-form solutions because the combination of periodic contributions and asset returns of the terminal wealth results in a dependency structure. One possibility to overcome this problem is to use upper and lower bounds for the pricing. Results on this topic are e.g. given in Hürlimann (2010), Chi and Lin (2012) and Bernard et al. (2017), while the latter

two also include flexible contributions. We differentiate from their approach by deriving quasi closed-form solutions.

The consideration of default risk embedded in life insurance contracts with terminal guarantee schemes dates back to Briys and De Varenne (1994). Further contributions are given by Grosen and Jørgensen (2002) and Bernard et al. (2005). More recent works are given by Schmeiser and Wagner (2015), Hieber et al. (2019) and Mahayni et al. (2021a). Moreover, quantifying the risk resulting from long-term guarantees with appropriate risk measures is done by e.g. Barbarin and Devolder (2005), Gatzert and Kling (2007) and Devolder (2018), who analyzes the capital requirements under different risk measures. We implement a VaR-based SFP and calculate the required capital s.t. this bound is matched, depending on the management rules and the splitting factor.

To assess life insurance contracts with guarantees and participation on the surplus of the insurer, it is necessary to evaluate the asset strategy itself. Here exists a strong connection to the literature of portfolio optimization which already dates back to Merton (1971). He solves the problem of maximizing the expected utility of an investor with constant relevant risk aversion (CRRA) in a Black-Scholes model setup. The continuous utility maximizing strategy is given in terms of a constant investment fraction in the risky asset. In contrast, periodic premium contributions (e.g. yearly payments) only imply the possibility of discrete time adjustments of the insured's portfolio. This causes a utility loss compared to the optimal continuous time version. However, Rogers (2001) shows that the discretization error is negligibly small for a time lag smaller than two years, s.t. we can compare our results with the Merton solution. Further literature on portfolio optimization in a Black-Scholes model setup is given by Huang et al. (2008), Branger et al. (2010), Gatzert (2013), Schmeiser and Wagner (2015), Chen and Hieber (2016) and Mahayni et al. (2021a). We further consider a management rule, s.t. the insurer adjusts the investment strategy if positive or negative shocks on the stock market occur. Such management rules can be interpreted as a regime switch. For a detailed literature overview on regime-switching we refer to Chapter 4.

The rest of the section is organized as follows. First, we present the general contract design and the model assumptions in Subsec. 3.1.1. In Subsec. 3.1.2 we discuss the impact of the splitting factor on the fair pricing for a constant resp. a variable management rule. The impact of the splitting factor and the management rules on the risk management of the insurance company is subject to Subsec. 3.1.3 and the effects on the expected utility of the insured are presented in Subsec. 3.1.4. Finally, we conclude the section.

3.1.1 Contract design and model assumption

Since we are interested in the impact of how insurance premiums are split over time, we first describe the payoff of a participating contract with a terminal guarantee in dependence of a flexible payoff scheme.

Contract payoff

Since, in practice, insurance premiums are paid in discrete time, we consider a contract design which is based on discrete dates $\mathcal{T} = \{t_0 = 0, \ldots, t_{n-1}, t_n = T\}$. T denotes the maturity of the contract and the contributions of the insured are given by a periodic premium scheme where a_{t_i} denotes the premium paid at $t_i \in \mathcal{T} \setminus \{t_n\}$. The payoff L_T which the insured receives at $T = t_n$ is defined by the maximum of two ingredients: the portfolio (or account) value V_T of the insured and a terminal guarantee G_T, i.e.

$$L_T = \max\{V_T, G_T\} = V_T + [G_T - V_T]^+, \tag{3.1}$$

where $[G_T - V_T]^+ = \max\{G_T - V_T, 0\}$. In summary, the contract payoff can be interpreted by means of two components: the portfolio (or account) value V_T and the payoff of a European put option with maturity T, strike $K = G_T$ and underlying V (with payoff V_T). We refer to the second component as the insurance put.

Underlying of insurance put

The portfolio value V_T is linked to the investment results (account value) of the insurance company which, at t_i ($i = 0, \ldots, n$), are denoted by A_{t_i}. Furthermore, the participation in the investment results depends on the guaranteed amount G_T as well as how much premium is paid at each point in time.

Since the contract payoff defined by Eqn. (3.1) is increasing in the guaranteed amount G_T, we assume in addition that only a part $\tilde{a}_{t_i} = \alpha a_{t_i}$ is invested in the asset side of the insurance company (i.e. the remaining premium part finances the guarantee). In particular, the special case $G_T = 0$ implies $\alpha = 1$ ($\tilde{a}_{t_i} = a_{t_i}$) and α is decreasing in G_T. This observation already dates back to Nielsen and Sandmann (1996). It holds

$$V_T = V_{t_n} = \sum_{i=0}^{n-1} \tilde{a}_{t_i} \frac{A_{t_n}}{A_{t_i}}.$$

We assume that the investment decisions of the insurance company are invested in a complete and arbitrage-free financial market and the existence of a risk-free asset growing with a constant interest rate r. This implies the existence of a uniquely defined pricing measure $\mathbb{P}^*$ where the discounted account increments are a martingale, i.e.

$$e^{-r(t_n-t_i)}\mathbb{E}_{\mathbb{P}^*}\left[\left.\frac{A_{t_n}}{A_{t_i}}\right| A_{t_i}\right] = 1.$$

Abstracting from mortality risk and assuming that the periodic premiums are paid with certainty, it is convenient to define V_{t_i} as the arbitrage free t_i-price of the payoff (portfolio value) V_T, i.e.

$$V_{t_i} := e^{-r(t_n-t_i)}\mathbb{E}_{\mathbb{P}^*}\left[V_{t_n} \middle| \{A_{t_0}, \dots, A_{t_i}\}\right].$$

A nice to interpret representation of the account value is given in the following lemma.

Lemma 3.1 (Account value V)
Let A_{t_i} denote the investment result of the insurance company at t_i and a_{t_i} the premium of the insured paid at t_i. Then the account value V_{t_i} is given by

$$V_{t_i} = \alpha\left(\sum_{j=0}^{i-1} a_{t_j}\frac{A_{t_i}}{A_{t_j}} + a_{t_i} + \sum_{j=i+1}^{n-1} e^{-r(t_j-t_i)}\, a_{t_j}\right). \tag{3.2}$$

The proof of Lemma 3.1 is given in Appendix C.1. Thus, the so-called account value V_{t_i} is defined by means of the sum of three components: the already realized participation in the account value, the current premium, and the present value of future premiums. Furthermore, the account value at t_{i+1} depends on the value at t_i, the investment results of the insurance company in the period $[t_i, t_{i+1}]$ and the present value of the contributions at and after t_{i+1}. Thus the account value can be written as follows.

Lemma 3.2 (Dynamics of V)
Let $PV_{t_i} := \sum_{j=i}^{n-1} e^{-r(t_j-t_i)}a_{t_j}$ $(i = 0, \dots, n-1)$ denote the present value of the contributions at and after t_i. Then, it holds $V_{t_0} = \alpha PV_{t_0}$ and $V_{t_{i+1}}$ $(i = 0, \dots, n-1)$ is given by

$$V_{t_{i+1}} = V_{t_i}\frac{A_{t_{i+1}}}{A_{t_i}} + \alpha\left(PV_{t_{i+1}}\left(1 - e^{-r(t_{i+1}-t_i)}\frac{A_{t_{i+1}}}{A_{t_i}}\right)\right). \tag{3.3}$$

The proof of Lemma 3.2 is given in Appendix C.2.

Throughout the following, we consider the case that the present value of the periodic premiums is equal to one, i.e.

$$PV_{t_0} = \sum_{i=0}^{n-1} e^{-r(t_i-t_0)}\, a_{t_i} = 1, \tag{3.4}$$

such that $V_{t_0} = \alpha$. We now consider a so-called premium splitting factor which describes how the contributions are split over time.

Definition 3.1 (Premium splitting factor)
Normalizing the present value of the periodic premium payments to one, cf. Eqn. (3.4) implies that the periodic premium payments can be stated by means of a premium splitting factor $\{\beta_{t_0}, \dots, \beta_{t_{n-1}}\}$ *where*

$$\beta_i \geq 0 \text{ for all } i = 0, \dots, n-1, \quad \sum_{i=0}^{n-1} \beta_i = 1,$$
$$\text{such that } a_{t_i} := \beta_i e^{r(t_i - t_0)}.$$

The extreme cases are thus implied by (i) $\beta_0 = 1$ (and $\beta_i = 0$ for $i = 1, \dots n-1$) and (ii) $\beta_{n-1} = 1$ (and $\beta_i = 0$ for $i = 0, \dots n-2$) such that

(i) $(a_{t_0}, a_{t_1}, \dots, a_{t_n}) = (1, 0, \dots, 0)$ (upfront premium) and

(ii) $(a_{t_0}, a_{t_1}, \dots, a_{t_n}) = (0, 0, \dots, e^{r(t_{n-1} - t_0)})$ (postponed premium).

Now, consider the general case that $\beta_i \in]0, 1[$. Notice that with the definition of β, it follows

$$PV_{t_i} = \sum_{j=i}^{n-1} e^{-r(t_j - t_i)} a_{t_j} = \sum_{j=i}^{n-1} e^{-r(t_j - t_i)} \beta_j e^{r(t_j - t_0)} = e^{r(t_i - t_0)} \sum_{j=i}^{n-1} \beta_j. \qquad (3.5)$$

Thus it holds:

Proposition 3.1 (Dynamics of V - premium splitting factor)
It holds $V_{t_0} = \alpha$ *and* $V_{t_{i+1}}$ $(i = 0, \dots, n-1)$ *is given by*

$$V_{t_{i+1}} = \overline{\beta_{i+1}} \left(\alpha e^{r(t_{i+1} - t_0)} + \frac{A_{t_{i+1}}}{A_{t_i}} \left(V_{t_i} - \alpha e^{r(t_i - t_0)} \right) \right) + \left(1 - \overline{\beta_{i+1}} \right) V_{t_i} \frac{A_{t_{i+1}}}{A_{t_i}}, \qquad (3.6)$$
$$\text{where } \overline{\beta_{i+1}} := \sum_{j=i+1}^{n-1} \beta_j \text{ and } 1 - \overline{\beta_{i+1}} = \sum_{j=0}^{i} \beta_j.$$

The proof of Proposition 3.1 is given in Appendix C.3.

The above proposition states that $V_{t_{i+1}}$ depends on a premium splitting factor $\overline{\beta_{i+1}}$. Thus, qualitatively, the impact of periodic premium contributions (compared to a single upfront contribution) can be derived in a stylized setup with two premium payment dates.

Stylized contract design

To simplify the expositions, we consider a stylized contract design which refers to $n = 2$, where $\mathcal{T} = \{t_0 = 0, t_1 = 1, t_2 = T = 2\}$. We set $\beta_0 = \beta$ such that $\beta_1 = 1 - \beta$, i.e. the analysis reduces to one splitting factor $\beta \in [0, 1]$. Recall that the premium payment at time $t = 0$ is then given by $a_0 = \beta$ and $a_1 = (1 - \beta)e^r$ such that the present value of the contributions is normalized to one. The corner cases are now given by the upfront premium ($\beta = 1$) where all premiums are invested at $t_0 = 0$ and the postponed premium case $\beta = 0$ (all premiums are postponed to $t = 1$). With formula (3.2) we immediately receive for the portfolio value V_i ($i = 0, 1, 2$)

$$\begin{aligned} V_0 &= \alpha, \\ V_1 &= \alpha \left(\beta \frac{A_1}{A_0} + (1 - \beta)e^r \right), \\ V_2 &= \alpha \left(\beta \frac{A_2}{A_0} + (1 - \beta)e^r \frac{A_2}{A_1} \right) = V_1 \frac{A_2}{A_1}. \end{aligned}$$

Recall that our research question includes the impact of an upfront and postponed premium payment and the cases between these extremes on the pricing, the risk management and the utility of the insured. In order to ensure the comparability of different premium schemes, we assume that all contracts are fairly priced. Notice that fair pricing means that the t_0- value of the contributions is equal to the t_0-value of the contract payoff L_T, i.e. in our stylized contract design it holds

$$1 = e^{-2r} \mathbb{E}_{\mathbb{P}^*} \left[V_2 + (G_2 - V_2)^+ \right].$$

With $e^{-2r} \mathbb{E}_{\mathbb{P}^*} [V_2] = V_0 = \alpha$, it follows

$$1 - \alpha = e^{-2r} \mathbb{E}_{\mathbb{P}^*} \left[(G_2 - V_2)^+ \right], \tag{3.7}$$

i.e. $1 - \alpha$ coincides with the guarantee costs at $t = 0$, denoted with GC_0 where

$$GC_0 := e^{-2r} \mathbb{E}_{\mathbb{P}^*} \left[(G_2 - V_2)^+ \right]$$

is the price of an Asian put option.[48]

Notice that both sides of Eqn. (3.7) depend on α such that the fair α is the solution of a fix-point problem. In addition, the fair α depends on the splitting factor β, i.e. on the periodic premium payments a_0 and a_1.

[48] Asian put option means that V_2 is a weighted average of stochastic increments. More information on Asian put options can e.g. be found in Vorst (1992) and Nielsen and Sandmann (1995)

The guarantee costs in the periodic premium setting, depending on the splitting factor β, can thus be stated in terms of

$$\begin{aligned} GC_0 &= e^{-2r}\, \mathbb{E}_{\mathbb{P}^*}\left[\left(G_2 - V_1\frac{A_2}{A_1}\right)^+\right] \\ &= e^{-r}\, \mathbb{E}_{\mathbb{P}^*}\left[e^{-r}\, \mathbb{E}_{\mathbb{P}^*}\left[\left(G_2 - V_1\frac{A_2}{A_1}\right)^+ \middle| V_1\right]\right]. \end{aligned}$$

Notice that the inner expectation denotes the guarantee costs at $t = 1$, i.e.

$$GC_1 = e^{-r}\, \mathbb{E}_{\mathbb{P}^*}\left[\left(G_2 - V_1\frac{A_2}{A_1}\right)^+ \middle| V_1\right].$$

These are given by the price of a European put option with time to maturity $T = 1$, strike $K = G_2$, and underlying V, currently priced at V_1. Thus,

$$GC_0 = e^{-r}\, \mathbb{E}_{\mathbb{P}^*}\left[GC_1\right].$$

In the special case of (i) $\beta = 1$ (upfront premium), it follows $V_2 = \tilde{a}_0\frac{A_2}{A_0} = \alpha\frac{A_2}{A_0}$ and thus

$$GC_0^{(\beta=1)} = e^{-2r}\mathbb{E}_{\mathbb{P}^*}\left[\left(G_2 - \alpha\frac{A_2}{A_0}\right)^+\right].$$

Here, the Asian feature vanishes and the guarantee costs are given by a European put option with maturity $T = 2$. In the special case of (ii) (postponed premium), it follows $V_2 = \tilde{a}_1\frac{A_2}{A_1} = \alpha e^r\frac{A_2}{A_1}$, i.e.

$$\begin{aligned} GC_0^{(\beta=0)} &= e^{-2r}\mathbb{E}_{\mathbb{P}^*}\left[\left(G_2 - \alpha e^r\frac{A_2}{A_1}\right)^+\right] \\ &= \mathbb{E}_{\mathbb{P}^*}\left[e^{-r}\mathbb{E}_{\mathbb{P}^*}\left[\left(e^{-r}G_2 - \alpha\frac{A_2}{A_1}\right)^+ \middle| A_1\right]\right]. \end{aligned}$$

The Asian feature vanishes again and the guarantee costs are given by a forward starting option. Let $G_2 = e^{2g}$ where $g < r$. The condition $g < r$ ensures the existence of a fair contract (for upfront premium), cf. Bacinello (2001), it follows

$$\begin{aligned} GC_0^{(\beta=0)} &= \mathbb{E}_{\mathbb{P}^*}\left[e^{-r}\mathbb{E}_{\mathbb{P}^*}\left[\left(e^{-(r-g)}e^g - \alpha\frac{A_2}{A_1}\right)^+ \middle| A_1\right]\right] \\ &\leq \mathbb{E}_{\mathbb{P}^*}\left[e^{-r}\mathbb{E}_{\mathbb{P}^*}\left[\left(e^g - \alpha\frac{A_2}{A_1}\right)^+ \middle| A_1\right]\right]. \end{aligned}$$

In summary, the comparison of the guarantee costs linked to the corner cases (i) and (ii) reduces to comparing a guaranteed rate over different investment horizons, i.e. if the investment results $\frac{A_{i+1}}{A_i}$ are independent and identically distributed. However, in reality the investment decisions of the insurance company may depend on the market movements. This implies a dependence structure even in the case of a Black and Scholes model setup for the investment opportunity set. Further details on the financial market assumptions and the investment decisions of the insurance company are postponed to the subsequent subsection.

Model assumptions and management rule

Our financial market model over the filtrated probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ is given by the Black-Scholes model. The filtration $(\mathcal{F}_t)_{t\in[0,T]}$ is generated by the standard Brownian motion $(W_t)_{t\in[0,T]}$. Because of the completeness of the Black-Scholes model, there exists a uniquely defined equivalent martingale measure $\mathbb{P}^*$ under which the process $(W_t^*)_{t\in[0,T]}$ is a standard Brownian motion. In particular, the risk-free bond $(B_t)_{t\in[0,T]}$ grows at constant interest rate r, i.e. $B_t = e^{rt}$ and the risky asset $(S_t)_{t\in[0,T]}$ is given by

$$dS_t = S_t \left(\mu \, dt + \sigma \, dW_t\right) = S_t \left(r \, dt + \sigma \, dW_t^*\right), \quad S_0 = 1.$$

Under the real world probability measure $\mathbb{P}$, the asset price follows a geometric Brownian motion with constant drift parameter μ ($\mu > r$) and constant volatility σ ($\sigma > 0$). Under the martingale measure (pricing measure) $\mathbb{P}^*$, the asset price follows a geometric Brownian motion with constant drift r and constant volatility σ. Recall that our stylized contract design implies a maturity of $T = 2$ and a discrete payment structure. Thus we are interested in the returns $\frac{S_1}{S_0}$ and $\frac{S_2}{S_1}$ of the risky asset. The solution of the corresponding SDE is given by a log-normal distribution, s.t. it holds

$$\frac{S_{i+1}}{S_i} \overset{\text{under } \mathbb{P}}{=} e^{\mu - \frac{1}{2}\sigma^2 + \sigma(W_{i+1} - W_i)} \overset{\text{under } \mathbb{P}^*}{=} e^{r - \frac{1}{2}\sigma^2 + \sigma(W_{i+1}^* - W_i^*)}, \; i = 0, 1.$$

The insurance company can decide which fraction of wealth π_t at time t is invested in the risky asset S. According to common practice, π_t is restricted to values between zero and one ($\pi_t \in [0, 1]$), i.e. no short selling and borrowing is allowed.[49] The remaining part $(1 - \pi_t)$ is invested in the risk-free bond B. Thus, the evolution of the portfolio wealth (investment result), denoted by the stochastic process $(A_t)_{t\in[0,T]}$,

[49] Specifying the investment fraction can lead to different investment strategies, e.g. Option Based Portfolio Insurance (OBPI) and Constant Proportion Portfolio Insurance (CPPI). For a detailed discussion of these strategies we refer to Bertrand and Prigent (2001). An application of further strategies such as Buy and Hold (B & H) can e.g. be found in Branger et al. (2010).

is defined over the risky and risk-free asset:

$$dA_t = A_t \left(\pi_t \frac{dS_t}{S_t} + (1 - \pi_t) \frac{dB_t}{B_t} \right), A_0 = 1. \tag{3.8}$$

Within our stylized contract design we are interested in the returns $\frac{A_1}{A_0}$ and $\frac{A_2}{A_1}$ of the investment results. The solution of equation (3.8) is again given by a log-normal distribution s.t. it holds

$$\frac{A_{i+1}}{A_i} \overset{\text{under } \mathbb{P}}{=} e^{\mu_{A,i} - \frac{1}{2}\sigma_{A,i}^2 + \sigma_{A,i}(W_{i+1} - W_i)} \overset{\text{under } \mathbb{P}^*}{=} e^{r - \frac{1}{2}\sigma_{A,i}^2 + \sigma_{A,i}(W_{i+1}^* - W_i^*)}, \; i = 0, 1,$$

where the cumulated drift and volatility $\mu_{A,i}$ resp. $\sigma_{A,i}$ is given by

$$\begin{aligned} \mu_{A,i} &:= \pi_i \mu + (1 - \pi_i) r \\ \sigma_{A,i} &:= \pi_i \sigma. \end{aligned} \tag{3.9}$$

With the possibility of different investment fractions in $t = 0$ and $t = 1$ the insurance company can react to (good or bad) market movements and adapt the investment strategy. We account for this by introducing a so-called *management rule* depending on the asset return $\frac{A_1}{A_0}$. Notice that under the pricing measure $\mathbb{P}^*$ only the portfolio wealth' volatility $\sigma_{A,i}$ depends on the investment fraction π_i. Thus for our analysis on the contract pricing and the risk management we define the management rule in terms of the volatility $\sigma_{A,i}$. When the investment fractions are constant over time, i.e. $\pi_0 = \pi_1$ (and thus $\sigma_{A,0} = \sigma_{A,1}$), we speak of a constant management rule (CMR). For a variable management rule (VMR) we consider for the sake of simplicity a stylized version where

$$\sigma_{A,1} = \sigma_{A,0} \sum_{i=1}^{n} e_i 1_{E_i} = \sigma \left(\pi_0 \sum_{i=1}^{n} e_i 1_{E_i} \right),$$

i.e. the management rule controls the investment fraction π_1 at $t = 1$ by increasing or decreasing π_0 by the value e_i. The event E_i is a function of $\frac{A_1}{A_0}$, e.g. $n = 3$ and

$$\begin{aligned} E_1 &= \left\{ \frac{A_1}{A_0} < c_1 \right\} = \left\{ c_0 \leq \frac{A_1}{A_0} \leq c_1 \right\} \\ E_2 &= \left\{ c_1 \leq \frac{A_1}{A_0} \leq c_2 \right\} \\ E_3 &= \left\{ \frac{A_1}{A_0} > c_2 \right\} = \left\{ c_2 \leq \frac{A_1}{A_0} \leq c_3 \right\}, \end{aligned}$$

where $0 = c_0 < c_1 \leq c_2 < c_3 = +\infty$. Notice that the events E_1, E_2, E_3 are disjoint. The second equation for E_1 and E_3 holds because $\frac{A_1}{A_0}$ is absolute continuously distributed and thus atomless s.t. there is no probability mass in one point. Thus E_i

is defined by $E_i := \left\{\frac{A_1}{A_0} \in [c_{i-1}, c_i]\right\}$, s.t. the probability for each event under the pricing measure is given by

$$\mathbb{P}^*\left(c_{i-1} \leq \frac{A_1}{A_0} \leq c_i\right) = \Phi\left(\frac{\ln(c_i) - (r - \frac{1}{2}\sigma_{A,0}^2)}{\sigma_{A,0}}\right) - \Phi\left(\frac{\ln(c_{i-1}) - (r - \frac{1}{2}\sigma_{A,0}^2)}{\sigma_{A,0}}\right).$$

With this management rule setting we are able to capture the most common investment strategies, e.g. if we set $e_1 < 1$, $e_2 = 1$, $e_3 > 1$ and $c_1 < 1$, $c_2 > 1$, we are in a setting of a so-called portfolio insurance strategy. Here the insurance company increases the investment fraction at $t = 1$ if the asset return $\frac{A_1}{A_0}$ is greater than one, i.e. in case a gain is observed. The company decreases the investment fraction, if $\frac{A_1}{A_0} < 1$ and on a small interval $[c_1, c_2]$ the insurance company does not react to the market behavior and leaves the investment fraction unchanged. If all values of e_i are equal to one, we are back in the setting of a constant management rule. A discussion on how to choose the parameters e_1, e_2 and e_3 is given in the next subsection.

3.1.2 Impact of splitting factor on (fair) pricing under different management rules

We analyze the impact of the splitting factor and the management rule on the pricing of the contract. For that we first discuss guarantee costs in the Black-Scholes setting in general. After that we analyze the different management rules and splitting factors in more detail.

Fair pricing and the impact of the splitting factor

As discussed in the last subsection, the guarantee costs for a splitting factor $\beta \in [0, 1]$ are given by GC_0, i.e. by the $t = 0$ price of the payoff $(G_2 - V_2)^+$. Notice that V is the result of an admissible investment strategy with price process $V_0 = \alpha$, $V_1 = \alpha\left(\beta\frac{A_1}{A_0} + (1 - \beta)e^r\right)$, and $V_2 = V_1\frac{A_2}{A_1}$. Since the dynamics of A are given in terms of a GBM with drift μ_A (for pricing μ_A is irrelevant), volatility σ_A ($\sigma_{A,0}$ and $\sigma_{A,1}$, respectively) and $V_2 = V_1\frac{A_2}{A_1}$, we can also interpret V_2 as the result of a GBM (starting at $t = 1$ with V_1) with volatility

$$\sigma_{V,1} = \sigma_{A,1} = \pi_1\sigma.$$

Using our stylized model setup, we can state the $t = 0$ costs in terms of the Black-Scholes Put formula of a European put option $P^{\mathrm{BS}}(x, T, K, \sigma)$ with time to maturity

T and strike K with underlying X (current price x and volatility σ).[50] Applying this to our model-free results from subsection 3.1.1, we find the following formula for the guarantee costs.

Proposition 3.2 (Guarantee Costs - Black-Scholes Setup)
For $V_1 = \alpha\left(\beta\frac{A_1}{A_0} + (1-\beta)e^r\right)$ *the guarantee costs in the periodic premium setting depending on the splitting factor* β *can be stated in terms of*

$$GC_0 = e^{-r}\,\mathbb{E}_{\mathbb{P}^*}\left[GC_1\right] = e^{-r}\,\mathbb{E}_{\mathbb{P}^*}\left[P^{BS}(V_1, 1, G_2, \sigma_{A,1})\right].$$

Proposition 3.2 shows the dependence of the guarantee costs on the splitting factor β and also on the management rule, because the volatility in the put price is given by σ_{A_1}. In the special case of $\beta = 0$ it holds $V_1 = \alpha e^r$ such that V_1 is deterministic. In the special case of $\beta = 1$ it holds $V_1 = \alpha\frac{A_1}{A_0}$.
In the following we differentiate between the guarantee costs in case of a constant management rule, GC_0^{CMR}, and the guarantee costs in case of a variable management rule, GC_0^{VMR}, to separate the impact of the splitting factor and the management rule on the guarantee costs.

Splitting factor and fair pricing under constant management rules

Following a constant management rule implies $\sigma_{A,1} = \sigma_{A,0}$ such that

$$GC_1^{CMR} = P^{\mathrm{BS}}(V_1, 1, G_2, \sigma_{A,0}) = V_1 P^{\mathrm{BS}}\left(1, 1, \frac{G_2}{V_1}, \sigma_{A,0}\right)$$

and

$$GC_0^{CMR} = e^{-r}\,\mathbb{E}_{\mathbb{P}^*}\left[P^{\mathrm{BS}}(V_1, 1, G_2, \sigma_{A,0})\right].$$

Notice that GC_0^{CMR} is stated as the $t = 0$ price of a forward starting option, i.e. a put option starting at $t = 1$ with V_1. For $\beta \in]0, 1[$, the distribution of $V_1 = \alpha\left(\beta\frac{A_1}{A_0} + (1-\beta)e^r\right)$ can not be stated in closed-form.
But the special cases of (i) upfront premium ($\beta = 1$) implies $V_1 = \alpha\frac{A_1}{A_0}$ and (ii)

[50] Notice that the Black-Scholes put price is given by

$$P^{\mathrm{BS}}(x, T, K, \sigma) = e^{-rT} K\Phi(-d_2(x, T, K, \sigma)) - x\Phi(-d_1(x, T, K, \sigma))$$
$$\text{where } d_1(x, T, K, \sigma) := \frac{\ln(\frac{x}{K}) + (\frac{1}{2}\sigma^2 + r)T}{\sigma\sqrt{T}} \text{ and } d_2(x, T, K, \sigma) := d_1(x, T, K, \sigma) - \sigma\sqrt{T}.$$

postponed premium ($\beta = 0$) implies $V_1 = \alpha e^r$, thus we can calculate the guarantee costs in closed-form. It holds for $\beta = 0$

$$\begin{aligned} GC_0^{(\beta=0),CMR} &= e^{-r}\,\mathbb{E}_{\mathbb{P}^*}\left[V_1 P^{\text{BS}}\left(1,1,\frac{G_2}{V_1},\sigma_{A,1}\right)\right] \\ &= e^{-r}\,\mathbb{E}_{\mathbb{P}^*}\left[\alpha e^r P^{\text{BS}}\left(1,1,\frac{G_2}{\alpha e^r},\sigma_{A,0}\right)\right] \\ &= P^{\text{BS}}(\alpha,1,e^{-r}G_2,\sigma_{A,0}). \end{aligned}$$

In contrast, for $\beta = 1$, it follows

$$\begin{aligned} GC_0^{(\beta=1),CMR} &= e^{-r}\,\mathbb{E}_{\mathbb{P}^*}\left[P^{\text{BS}}(V_1,1,G_2,\sigma_{A,1})\right] \\ &= e^{-r}\,\mathbb{E}_{\mathbb{P}^*}\left[P^{\text{BS}}\left(\alpha\frac{A_1}{A_0},1,G_2,\sigma_{A,0}\right)\right] \\ &= P^{\text{BS}}\left(\alpha,2,G_2,\sigma_{A,0}\right) = \alpha P^{\text{BS}}\left(1,2,\frac{G_2}{\alpha},\sigma_{A,0}\right). \end{aligned}$$

The interesting question is how the guarantee costs in the CMR case are related to each other, depending on the splitting factor β. For this the following remark is crucial.

Remark 3.1 (Properties of guarantee costs in β)
Notice that the put price is a decreasing and convex function of the underlying, independent of the model assumption. Thus, for $\beta \in [0,1]$ it especially holds

$$P^{BS}(\beta x_0 + (1-\beta)x_1, T, K, \sigma) \leq \beta P^{BS}(x_0, T, K, \sigma) + (1-\beta)P^{BS}(x_1, T, K, \sigma).$$

Using the convexity property for Jensen's inequality we receive[51]

$$\begin{aligned} GC_0^{CMR} &= e^{-r}\,\mathbb{E}_{\mathbb{P}^*}\left[P^{\text{BS}}(V_1,1,G_2,\sigma_{A,0})\right] \\ &\geq e^{-r}\,P^{\text{BS}}\left(\mathbb{E}_{\mathbb{P}^*}[V_1],1,G_2,\sigma_{A,0}\right) \\ &= e^{-r}\,P^{\text{BS}}\left(\alpha e^r,1,G_2,\sigma_{A,0}\right) = GC_0^{(\beta=0),CMR}. \end{aligned}$$

Moreover, we can use the convexity property and again Jensen to show

$$\begin{aligned} P^{\text{BS}}\left(V_1,T,K,\sigma\right) &\leq \beta P^{\text{BS}}\left(\alpha\frac{A_1}{A_0},T,K,\sigma_{A,0}\right) + (1-\beta)P^{\text{BS}}(\alpha e^r,T,K,\sigma_{A,0}) \\ &= \beta P^{\text{BS}}\left(\alpha\frac{A_1}{A_0},T,K,\sigma_{A,0}\right) + (1-\beta)P^{\text{BS}}\left(\mathbb{E}_{\mathbb{P}^*}\left[\alpha\frac{A_1}{A_0}\right],T,K,\sigma_{A,0}\right) \\ &\leq \beta P^{\text{BS}}\left(\alpha\frac{A_1}{A_0},T,K,\sigma_{A,0}\right) + (1-\beta)\mathbb{E}_{\mathbb{P}^*}\left[P^{\text{BS}}\left(\alpha\frac{A_1}{A_0},T,K,\sigma_{A,0}\right)\right], \end{aligned}$$

[51] Jensen's inequality states that for a convex function u it follows $E[u(X)] \geq u(E[X])$.

Benchmark parameter

r	μ	σ	g	π
0.01	0.15	0.037	0.0025	0.3

Table 3.1: Benchmark parameter setting.

where the last inequality is again implied by Jensen. Taking expectations on both sides of the calculated inequality, we receive

$$\mathbb{E}_{\mathbb{P}^*}\left[P^{\text{BS}}\left(V_1, T, K, \sigma\right)\right] \leq \mathbb{E}_{\mathbb{P}^*}\left[P^{\text{BS}}\left(\alpha\frac{A_1}{A_0}, T, K, \sigma_{A,0}\right)\right]$$

and thus

$$GC_0^{CMR} \leq GC_0^{(\beta=1),CMR}.$$

Combining the findings in this subsection, we can state the following proposition.

Proposition 3.3 (Properties Guarantee Costs - Constant management rule)

For a constant management rule with a premium fraction $\beta \in [0,1]$, the guarantee costs $GC_0(\beta)$ are given in quasi closed-form by

$$GC_0^{CMR}(\beta) = e^{-r}\,\mathbb{E}_{\mathbb{P}^*}\left[P^{BS}(V_1, 1, G_2, \sigma_{A,0})\right].$$

For the special cases of upfront and postponed premium payments we receive

(i) $GC_0^{(\beta=0),CMR} = P^{BS}(\alpha, 1, e^{-r}G_2, \sigma_{A,0})$.

(ii) $GC_0^{(\beta=1),CMR} = P^{BS}(\alpha, 2, G_2, \sigma_{A,0})$.

Furthermore, $GC_0^{CMR}(\beta)$ is monotonically increasing and convex in β. In particular, for all $\beta \in [0,1]$ a trivial upper price bound is given by $GC_0^{(\beta=1),CMR}$ while a tighter upper price bound is implied by

$$GC_0^{CMR}(\beta) \leq \beta GC_0^{(\beta=1),CMR} + (1-\beta)GC_0^{(\beta=0),CMR}.$$

An illustration of the convexity result and pricing bounds in Proposition 3.3 is given in Figure 3.1. For the results we use the benchmark parameter setting in Table 3.1.[52] Recall that for $\beta \in]0,1[$ the contract has an Asian put feature and thus we receive here an upper pricing bound for the Asian put. The discussion on general upper and lower bounds for Asian put options can be found in Nielsen and Sandmann (2003). A

Convexity put price depending on splitting factor β - CMR

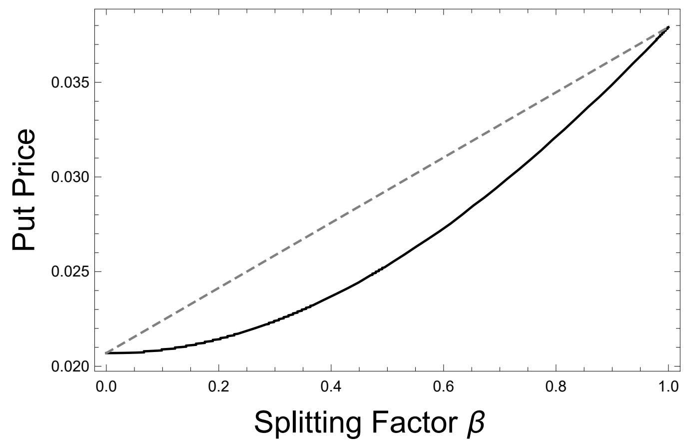


Figure 3.1: The parameters we used are $r = 0.01$, $\sigma = 0.15$, $\mu = 0.037$, $g = 0.0025$ and $\pi = 0.3$. The black line refers to the guarantee costs depending on the splitting factor β: $GC_0(\beta)$, the gray dashed line pictures the linear function between $(0, GC_0^{(\beta=0),CMR})$ and $(1, GC_0^{(\beta=1),CMR})$.

survey with pricing results on Asian Options is given in Boyle and Potapchik (2008).

Splitting factor and fair pricing under variable management rule

Following a variable management rule has the advantage for the insurance company to adjust the investment fraction at $t = 1$ and react to the last period's asset return $\frac{A_1}{A_0}$. These adjustments of course have an impact on the guarantee costs. First of all, using the same argumentation as in the last subsection, that the put-price is convex in the underlying for fixed volatility, s.t. we find

$$GC_0^{(\beta=0),VMR} \leq GC_0^{VMR}(\beta) \leq GC_0^{(\beta=1),VMR}.$$

To specify the guarantee costs of the variable management rule and to compare them with the constant management rule we need another well known result from pricing theory.

Remark 3.2 (Convexity of put price in the volatility)
Notice that the put price is convex in σ independent of the model assumption, i.e. it especially holds for all $\nu \in (0,1)$

$$P^{BS}(x, T, K, \nu\sigma_0 + (1-\nu)\sigma_1) \leq \nu P^{BS}(x, T, K, \sigma_0) + (1-\nu)P^{BS}(x, T, K, \sigma_1).$$

[52] Notice that we set the guarantee rate to $g = 0.0025$. This is the new guarantee rate standard since 2022 for German life insurers (cf. GDV (2021b)).

Convexity put price depending on volatility $\sigma_{A,1}$ - VMR case

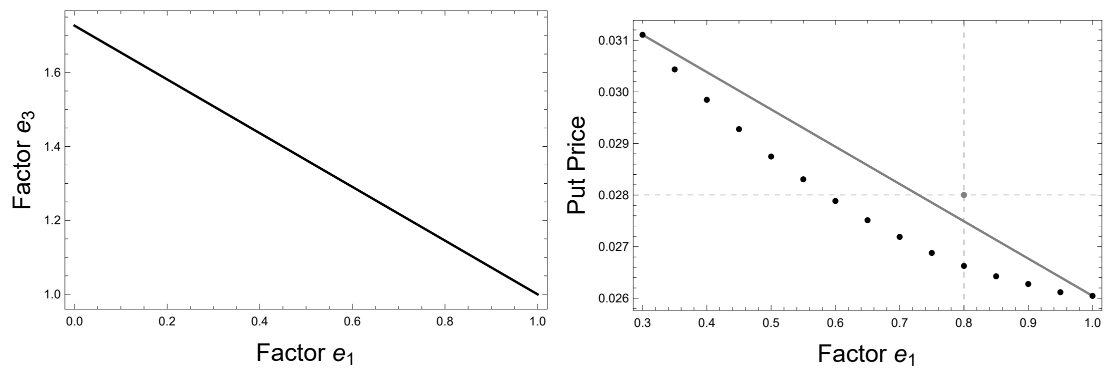


Figure 3.2: For both pictures we choose $\beta = 0.5$ and $c_1 = c_2 = 1$. The left picture shows the (e_1, e_3) combinations s.t. $\sum_{i=1}^{3} e_i \mathbb{P}^*(E_i) = 1$ for the special case that $\mathbb{P}^*(E_2) = 0$. The right picture shows the put prices depending on the choice of e_1 (black dotted line). The corresponding values of e_3 are chosen s.t. $\sum_{i=1}^{3} e_i \mathbb{P}^*(E_i) = 1$. Choosing a value for e_3 that does not solve the equation (in our example we choose $(e_1; e_3) = (0,8; 1,2)$ s.t. $\sum_{i=1}^{3} e_i \mathbb{P}^*(E_i) = 1,032$) violates the convexity property of the Black-Scholes put price in the volatility (gray dot).

Thus we first have to discuss the choice of the volatility $\sigma_{A,1}$ s.t. the convexity property is not violated and the no-arbitrage assumption (fair pricing) holds true. As already presented in the last section we model the volatility for the second period $[1, 2]$ in terms of

$$\sigma_{A,1} = \sigma_{A,0} \sum_{i=1}^{n} e_i 1_{\{E_i\}}.$$

For the choices of e_i we need to take the following consideration into account: As stated in the previous subsection we assume that $e_1 < 1, e_2 = 1$ and $e_3 > 1$ to model the different reactions to the asset returns $\frac{A_1}{A_0}$. Without further restrictions it is possible that for choices (e_1^1, e_2^1, e_3^1) and (e_1^2, e_2^2, e_3^2) the corresponding put prices coincide even if the risk structures differ.[53] This would lead to the possibility of arbitrage. Thus the put price is not concave resp. monotonically increasing in the volatility if we do not choose a suitable restriction. By looking at the expected volatility $\mathbb{E}_{\mathbb{P}^*}[\sigma_{A,1}]$, we can formulate a first restriction to the choice of e_i s.t. the riskiness of the strategies becomes comparable.

[53] For example if we choose within our benchmark parameter setting and $\beta = 0.5$ for the parameters e_1, e_2, e_3 $(0.3; 1; 1.5)$ resp. $(0.9; 1; 1.25)$ then the put prices coincides with a value of 0.03 but the strategy of the first setting is much more risky than the second one.

Assumption 3.1
To avoid violations of the convexity of the put price in the volatility we assume that

$$\mathbb{E}_{\mathbb{P}^*}\left[\sum_{i=1}^{n} e_i 1_{\{E_i\}}\right] = \sum_{i=1}^{n} e_i \mathbb{P}^*(E_i) = const.$$

The next question that arises is how to choose the constant for the expected volatility $\mathbb{E}_{\mathbb{P}^*}[\sigma_{A,1}]$.
Case 1: *const.* < 1:
If we choose for $e_1 < 1$ a value that is close to 1 and $e_2 = 1$ we find that the value of e_3 that solves the equation has to be smaller than 1. This is a contradiction to our assumption that $e_3 > 1$.
Case 2: *const.* > 1:
If we choose for $e_3 > 1$ a value that is close to 1 and $e_2 = 1$ we find that the value of e_1 that solves the equation has to be greater than 1. This is a contradiction to our assumption that $e_1 < 1$.
Thus the constant has to be 1. With this assumption the convexity property of the put price is fulfilled and the choices of e_1, e_2 and e_3 are given in a way, that our assumptions are not contradicted. With these results we can postulate our Assumption $(*)$:

Assumption $(*)$
To avoid violations of the convexity of the put price in the volatility and for the corresponding choices of e_1, e_2 *and* e_3*, we assume that*

$$\mathbb{E}_{\mathbb{P}^*}\left[\sum_{i=1}^{n} e_i 1_{\{E_i\}}\right] = \sum_{i=1}^{n} e_i \mathbb{P}^*(E_i) = 1.$$

Figure 3.2 shows under Assumption $(*)$ the corresponding choices of e_1 and e_3 for the case that $e_2 = 1$ and the corresponding put prices.

Remark 3.3
Notice that a high value of e_1 *implies a rather small value for* e_3 *s.t. the riskiness of the strategy is smaller than for a small* e_1 *(which implies a large* e_3*). This is reflected by the decreasing value of the put price in the factor* e_1*. Furthermore, if we fix* e_1 *and* e_2 *the remaining factor* e_3 *is monotone increasing in the investment fraction* π_0 *because the probability* $\mathbb{P}^*(E_3)$ *is decreasing in the investment fraction. To compensate for this lower probability the value of* e_3 *has to be increased s.t. Assumption* $(*)$ *is fulfilled. This is important for the later following expected utility section.*

Under Assumption (∗) we furthermore see that the following statement holds:

$$\mathbb{E}_{\mathbb{P}^*}[\sigma_{A,1}] = \sigma_{A,0} \sum_{i=1}^{n} e_i \mathbb{P}^*(E_i) = \sigma_{A,0}.$$

Thus the (stochastic) volatility fulfills the martingale property.

Using Assumption (∗) it follows

$$\begin{aligned} GC_0^{(\beta=0),\mathrm{VMR}} &= \alpha \sum_{i=1}^{n} P^{\mathrm{BS}} \left(1, 1, \frac{G_2}{\alpha e^r}, e_i \sigma_{A,0}\right) \mathbb{P}^*(E_i) \\ &\geq \alpha P^{\mathrm{BS}} \left(1, 1, \frac{G_2}{\alpha e^r}, \left(\sum_{i=1}^{n} e_i \mathbb{P}^*(E_i)\right) \sigma_{A,0}\right) \\ &= \alpha P^{\mathrm{BS}} \left(1, 1, \frac{G_2}{\alpha e^r}, \sigma_{A,0}\right) = GC_0^{(\beta=0),\mathrm{CMR}}. \end{aligned}$$

However, for $\beta = 1$, it follows

$$GC_0^{(\beta=1),VMR} = e^{-r}\, \mathbb{E}_{\mathbb{P}^*} \left[P^{\mathrm{BS}} \left(\alpha \frac{A_1}{A_0}, 1, G_2, \sigma_{A,0} \sum_{i=1}^{3} e_i 1_{\{E_i\}} \right) \right].$$

Again notice that the put price is convex in the volatility, thus with Jensen inequality we receive

$$\mathbb{E}_{\mathbb{P}^*} \left[P^{\mathrm{BS}} \left(\alpha \frac{A_1}{A_0}, 1, G_2, \sigma_{A,0} \sum_{i=1}^{3} e_i 1_{\{E_i\}} \right) \right] \geq P^{\mathrm{BS}} \left(\alpha \frac{A_1}{A_0}, 1, G_2, \mathbb{E}_{\mathbb{P}^*} \left[\sigma_{A,0} \sum_{i=1}^{3} e_i 1_{\{E_i\}} \right] \right).$$

Under Assumption (∗) we find

$$\begin{aligned} P^{\mathrm{BS}} \left(\alpha \frac{A_1}{A_0}, 1, G_2, \mathbb{E}_{\mathbb{P}^*} \left[\sigma_{A,0} \sum_{i=1}^{3} e_i 1_{\{E_i\}} \right] \right) &= P^{\mathrm{BS}} \left(\alpha \frac{A_1}{A_0}, 1, G_2, \sigma_{A,0} \right), \text{ i.e.} \\ \mathbb{E}_{\mathbb{P}^*} \left[P^{\mathrm{BS}} \left(\alpha \frac{A_1}{A_0}, 1, G_2, \sigma_{A,0} \sum_{i=1}^{3} e_i 1_{\{E_i\}} \right) \right] &\geq P^{\mathrm{BS}} \left(\alpha \frac{A_1}{A_0}, 1, G_2, \sigma_{A,0} \right). \end{aligned}$$

Taking expectations on both sides preserve the inequality and we receive

$$\begin{aligned} GC_0^{(\beta=1),VMR} &= e^{-r}\, \mathbb{E}_{\mathbb{P}^*} \left[P^{\mathrm{BS}} \left(\alpha \frac{A_1}{A_0}, 1, G_2, \sigma_{A,1} \right) \right] \\ &\geq e^{-r}\, \mathbb{E}_{\mathbb{P}^*} \left[P^{\mathrm{BS}} \left(\alpha \frac{A_1}{A_0}, 1, G_2, \sigma_{A,0} \right) \right] = GC_0^{(\beta=1),CMR}. \end{aligned}$$

Thus we find that under Assumption ($*$) the guarantee costs for the variable management rule are more expensive than the costs for the constant management rule for both, the upfront and the postponed premium case. For the general case $\beta \in]0,1[$ it holds again with Jensen and $\sigma_{A,0} = \mathbb{E}_{\mathbb{P}^*}[\sigma_{A,1}]$

$$P^{BS}(V_1, 1, G_2, \sigma_{A,0}) = P^{BS}(V_1, 1, G_2, \mathbb{E}_{\mathbb{P}^*}[\sigma_{A,1}]) \leq \mathbb{E}_{\mathbb{P}^*}[P^{BS}(V_1, 1, G_2, \sigma_{A,1})]$$

and thus

$$e^{-r}\mathbb{E}_{\mathbb{P}^*}[P^{BS}(V_1, 1, G_2, \sigma_{A,0})] \leq e^{-r}\mathbb{E}_{\mathbb{P}^*}[P^{BS}(V_1, 1, G_2, \sigma_{A,1})] \Leftrightarrow GC_0^{CMR} \leq GC_0^{VMR}.$$

The following proposition summarizes the results.

Proposition 3.4
For a variable management rule that fulfills Assumption ($*$) *and for every splitting factor* $\beta \in [0,1]$ *it holds*

$$\begin{aligned} (i)\, & GC_0^{(\beta=0),VMR} \leq GC_0^{VMR}(\beta) \leq GC_0^{(\beta=1),VMR} \\ (ii)\, & GC_0^{CMR}(\beta) \leq GC_0^{VMR}(\beta). \end{aligned}$$

A visualization of these results is given in Figure 3.3. The fact that the guarantee costs for the VMR are higher compared to the one for the CMR is intuitively clear because the insurance company has more possibilities resp. rights in the VMR case than in the CMR case. But more rights have to be compensated with higher costs if we want to avoid arbitrage s.t. $GC_0^{CMR}(\beta) \leq GC_0^{VMR}(\beta)$.

As seen above Assumption ($*$) is highly important for our results. Let us analyze the assumption in more detail. It holds

$$\sum_{i=1}^{3} e_i \mathbb{P}^*(E_i) = \Phi(d_1(c_1, 1, e^{2r}, \sigma_{A,0}))[e_1 - e_2] + \Phi\left(d_1(c_2, 1, e^{2r}, \sigma_{A,0})\right)[e_2 - e_3] + e_3.$$

Let us now assume that on the set E_2 it holds $\sigma_{A,1} = \sigma_{A,0}$, i.e. $e_2 = 1$. This is a realistic assumption s.t. we have three different possibilities to adjust the investment fraction in $t = 1$ depending on the return $\frac{A_1}{A_0}$: we can keep the investment fraction constant, i.e. $e_2 = 1$, we can reduce the investment fraction with $e_1 < 1$ and we can increase the investment fraction with $e_3 > 1$. Using $e_2 = 1$ we find

$$\sum_{i=1}^{n} e_i \mathbb{P}^*(E_i) \stackrel{!}{=} 1 \Leftrightarrow e_3 = \frac{\Phi(d_1(c_1, 1, e^{2r}, \sigma_{A,0}))(e_1 - 1)}{\Phi\left(d_1(c_2, 1, e^{2r}, \sigma_{A,0})\right) - 1} + 1. \qquad (3.10)$$

If we choose e.g. $e_2 = 1$ on the interval $[0.99, 1.01]$, i.e. $c_1 = 0.99$ and $c_2 = 1.01$, we are in the situation where the insurance company keeps its investment fraction

Convexity put price VMR and comparison with CMR

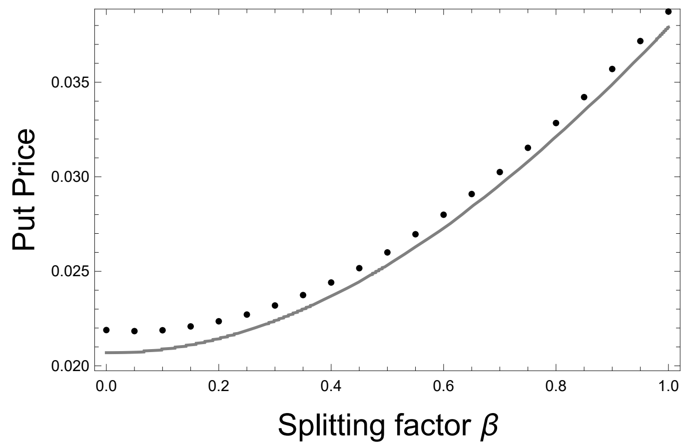


Figure 3.3: We choose within our benchmark parameter setup $\beta = 0.5$ and $c_1 = c_2 = 1, e_1 = 0.8; e_2 = 1$ and $e_3 = 1.1453$ s.t. $\sum_{i=1}^{3} e_i \mathbb{P}^*(E_i) = 1$. The gray line pictures the guarantee costs for the constant management rule $GC_0^{CMR}(\beta)$ and the black dotted line the guarantee costs for the variable management rule $GC_0^{VMR}(\beta)$, both depending on the splitting factor β.

constant over time if the asset return has a loss smaller than one percent resp. a gain smaller than one percent. If $\frac{A_1}{A_0} > 1.01$ resp. $\frac{A_1}{A_0} < 0.99$ (i.e. we are on the sets E_1 resp. E_3), the insurance company reacts and adjusts its investment fraction for the period $[1, 2]$.
Now the question is how to adjust the fractions s.t. assumption $(*)$ is fulfilled. Let us assume that the company is interested in reducing the investment fraction for the period $[1, 2]$ by 5% if the asset return has a loss of more than one percent, i.e. we set $e_1 = 0.95$. Plugging in all the information into equation (3.10), we receive that the insurance company can increase their investment fraction if the asset return $\frac{A_1}{A_0}$ has gains of more than one percent by 3.42%, i.e. $e_3 = 1.0342$ fulfills Assumption $(*)$.

3.1.3 Impact of splitting factor on risk management

In this section we discuss the risk for the insurer that stems from the periodic payment structure of the insured's contact policy. Proposition 3.3 shows that a problem of underpricing can arise if the insured is not willing to pay the whole guarantee costs at inception. Because of the structure of the periodic payments, it is natural that also the costs for the guarantee are paid proportionately to the payment structure, i.e. we have a payment stream of $\beta(1-\alpha)$ at inception and of $(1-\beta)e^r(1-\alpha)$ in $t = 1$. This is a problem for the insurance company because the risk structure at $t = 1$ differs from the one at inception. This can lead to a risk for the insurance

company of underpricing the guarantee costs, because the payment $(1-\beta)e^r(1-\alpha)$ in $t=1$ of the insured may not be sufficient for the risk structure in that period. To overcome this problem in our setting, the insurance company can adjust the investment fraction at $t=1$ to reduce the risk situation:
The insurance regulatory framework - Solvency II - contains the condition that the shortfall probability (SFP) w.r.t. a time horizon of one year is limited to 0.5%. The upper bound on the shortfall probability determines the amount of capital that is needed to assure the solvency to a high degree, i.e. to honor the liabilities to the insured. It is now of interest to analyze how much capital C at time $t=0$ is required s.t. the shortfall probability of $\varepsilon=0.005$ at $t=1$ of the periodic contract is matched. With this setting we can measure the effects of postponed premium payments for the risk structure of the insurer. Furthermore, we can analyze how a management rule affects the capital requirements, i.e. the riskiness of the insurer's assets and how we can reduce them.

Characterization shortfall probability and general problem

For the SFP the following calculations hold in general

$$\mathbb{P}\left(L_1 > Ce^r + \frac{A_1}{A_0}\right) \leq \varepsilon \Leftrightarrow \mathbb{P}\left(L_1 - \frac{A_1}{A_0} > Ce^r\right) \leq \varepsilon. \tag{3.11}$$

Thus, we need to calculate and characterize the random variable $L_1 - \frac{A_1}{A_0}$ in more detail. It holds

$$\begin{aligned} L_1 - \frac{A_1}{A_0} &= V_1 + \mathbb{E}_{\mathbb{P}^*}\left[e^{-r}\left(G_2 - V_1\frac{A_2}{A_1}\right)^+\right] - \frac{A_1}{A_0} \\ &= V_1 + GC_1 - \frac{A_1}{A_0}, \end{aligned} \tag{3.12}$$

where $GC_1 = P^{BS}(V_1, 1, G_2, \sigma_{A,1}) = V_1 P^{BS}\left(1, 1, \frac{G_2}{V_1}, \sigma_{A,1}\right)$. Moreover, it holds for the guarantee costs at $t=0$

$$1-\alpha = GC_0.$$

Notice that the fair pricing takes place under the pricing measure $\mathbb{P}^*$ but the shortfall probability is calculated under the real world measure $\mathbb{P}$, s.t. in the pricing parts of the formula we have to work with r instead of the drift μ_A.

Using the representations of GC_0 and GC_1, we can write the random variable $L_1 - \frac{A_1}{A_0}$ as follows.

Proposition 3.5 (General Representation $L_1 - \frac{A_1}{A_0}$)
The random variable $L_1 - \frac{A_1}{A_0}$ is given by

$$L_1 - \frac{A_1}{A_0} = GC_1 - \frac{V_1}{\alpha} GC_0 + (1-\beta)\left(e^r - \frac{A_1}{A_0}\right). \tag{3.13}$$

The proof of Proposition 3.5 is given in Appendix C.4. Now it is interesting to discuss the impact of the premium payments and/or the impact of the management rule on the values of the random variable $L_1 - \frac{A_1}{A_0}$. As seen in the last section, for a fixed management rule (CMR or VMR) the guarantee costs GC_0 are increasing in the splitting factor β s.t. the capital requirements are smaller if the premiums are paid upfront instead of postponed.
This is an important observation because postponed resp. periodic premium payments lead to an increasing capital requirement for the insurance company compared to upfront payments. But periodic resp. postponed premium payments can be attractive to the insured s.t. the insurance company has a target conflict: minimizing its capital requirement vs. fulfilling the insured's needs.
Here our management rule setting can be applied by the insurance company: As we have seen in the last section, for a fixed β it holds $GC_0^{VMR} > GC_0^{CMR}$, i.e. by applying a management rule we can reduce the random variable $L_1 - \frac{A_1}{A_0}$ and thus the required capital C, even though a postponed resp. a periodic payment structure is implemented. Furthermore, the participation fraction α is decreasing in GC_0, s.t. the effect is amplified.

Splitting factor and capital requirements under management rules

Recall that $GC_1 = P^{BS}(V_1, 1, G_2, \sigma_{A,1})$ and $GC_0 = e^{-r}\mathbb{E}_{\mathbb{P}^*}[GC_1]$. In the case of a variable management rule we are able to affect the investment decision at $t = 1$ by adapting the investment fraction π_1 with our management rule $\sigma_{A,1} = \sigma_{A,0}\sum_{i=1}^{3} e_i 1_{E_i}$. Again for the periodic payments ($\beta \in]0,1[$) we are not able to calculate the formulas is closed-form. In the two special cases $\beta = 0$ resp. $\beta = 1$ the following representation hold.

Proposition 3.6 (Representation $L_1 - \frac{A_1}{A_0}$ (VMR) - Special cases)
Let the insurance company follow a VMR. In the special case of a postponed premium ($\beta = 0$) it holds

$$L_1 - \frac{A_1}{A_0} = GC_1^{(\beta=0),VMR} - e^r GC_0^{(\beta=0),VMR} + e^r - \frac{A_1}{A_0}$$

and for the upfront premium case ($\beta = 1$) we receive

$$L_1 - \frac{A_1}{A_0} = GC_1^{(\beta=1),VMR} - \frac{A_1}{A_0} GC_0^{(\beta=1),VMR}.$$

The proof is straight forward. In the special case of a CMR ($\sigma_{A,0} = \sigma_{A,1}$) in combination with $\beta = 0$ resp. $\beta = 1$ we are able to simplify the representation of the random variable $L_1 - \frac{A_1}{A_0}$.

Corollary 3.1 (Representation $L_1 - \frac{A_1}{A_0}$ (CMR) - Special cases)
Let the insurance company follow a CMR. In the special case of a postponed premium ($\beta = 0$) it holds

$$L_1 - \frac{A_1}{A_0} = e^r - \frac{A_1}{A_0}$$

and for the upfront premium case ($\beta = 1$) we receive

$$\begin{aligned} L_1 - \frac{A_1}{A_0} &= GC_1^{(\beta=1),\,CMR} - \frac{A_1}{A_0} GC_0^{(\beta=1),\,CMR} \\ &= P^{BS}\left(\alpha\frac{A_1}{A_0}, 1, G_2, \sigma_{A,0}\right) - \frac{A_1}{A_0} P^{BS}\left(\alpha, 2, G_2, \sigma_{A,0}\right). \end{aligned}$$

The proof of Corollary 3.1 is stated in Appendix C.5. Notice that in the postponed premium case $L_1 - \frac{A_1}{A_0}$ is independent of the guarantee costs and given by the comparison of e^r and the random variable $\frac{A_1}{A_0}$. Thus, the required capital to fulfill the SFP of ε can be calculated in closed-form in that special case

$$\begin{aligned} &\mathbb{P}\left(L_1 - \frac{A_1}{A_0} > Ce^r\right) \leq \varepsilon \\ \Leftrightarrow &\mathbb{P}\left(\frac{A_1}{A_0} < e^r(1-C)\right) \leq \varepsilon \\ \Leftrightarrow &C^{CMR,(\beta=0)} \geq 1 - e^{-r} e^{\sigma_{A,0}\Phi^{-1}(\varepsilon) + \mu_{A,0} - \frac{1}{2}(\sigma_{A,0})^2}. \end{aligned}$$

Using the results of Proposition 3.3 resp. 3.4 with the representation of Proposition 3.5 we can conclude that $C^{CMR,(\beta=0)}$ gives us an upper bound for all required capitals C, s.t.

$$C \leq C^{CMR,(\beta=0)}.$$

A lower bound for the capital requirements is given by a VMR. But varying the weighting factors e_i of course has an impact on the pricing of the guarantee costs and thus influences the required capital C s.t. we cannot state a general lower bound for the capital requirements. But as seen in Figure 3.2 the put price increases with increasing weighting factor e_3 s.t. we can say that the lower bound gets smaller the more extreme our variable management rule is chosen. Thus we can give a recommended action to the insurance company on how to choose the management rule depending on how small the required capital should be. We can even determine the management rule s.t. the capital requirement C coincides with the one from an upfront payment without management rule s.t. the risk structures coincide. A detailed discussion on this topic is given in the next subsection.

Analysis of required capital

To determine the required capital C in regards of the shortfall probability of $\varepsilon = 0.005$ we solve

$$\mathbb{P}\left(L_1 - \frac{A_1}{A_0} > Ce^r\right) \leq \varepsilon.$$

As mentioned before for $\beta \in]0,1[$ this cannot be solved in closed-form. Thus we use Monte-Carlo simulations to visualize and discuss the effects.[54] The first results we receive by analyzing the constant management rule case, i.e. $\pi_0 = \pi_1$. The postponed premium payment leads to an increase of the capital that is required to fulfill the SFP compared to the upfront premium case (10.38 percent compared to 8.86 percent of the initial capital). Thus postponed premium payments lead to an increase in the riskiness of the insurer's portfolio. This riskiness can be reduced by introducing the splitting factor resp. the periodic premium payments: If we split the payment stream of the insured in a way that at $t = 0$ she pays 20 percent of her premiums and at $t = 1$ she pays 80 percent, then the risk in form of the required capital is reduced to 9.6 percent which is the median of the possible capital requirements in our case study. This is a high reduction compared to the fact that only 20 percent of the insured's payments are paid at inception. Thus a relatively small investment fraction at inception can reduce the riskiness of the insurers portfolio. But the best payment structure for the insurance company in terms of riskiness is the upfront payment, here the required capital is the smallest. This is not surprising because the costs for the guarantee are the highest in this case (cf. Figure 3.1). The results for the constant management rule case are visualized in Figure 3.4.

Using the variable management rule setting and suitable tuple of (e_1, e_2, e_3) s.t. assumption $(*)$ is fulfilled, we can reduce the required capital compared to the constant management rule case. This holds for all premium fractions β. In general it holds that the more extreme the management rule setting is (i.e. e_1 and e_3 differ greatly) the less capital is required s.t. the SFP is fulfilled. It is even possible to reduce the postponed capital requirement to a level of 8.86 percent in our case study (choose $e_1 = 0.63; e_2 = 1$ and $e_3 = 1.2689$), i.e. to the level of the upfront premium in the constant management rule case. This highlights the importance of the variable management rule: The insurance company has thus two possible approaches for re-

[54] We use 10^6 simulations for the results. Notice that the differentiation between pricing and real world measure is crucial fo the simulations: The random variable L_1 is calculated under the pricing measure $\mathbb{P}^*$, $\frac{A_1}{A_0}$ under the real world measure $\mathbb{P}$.

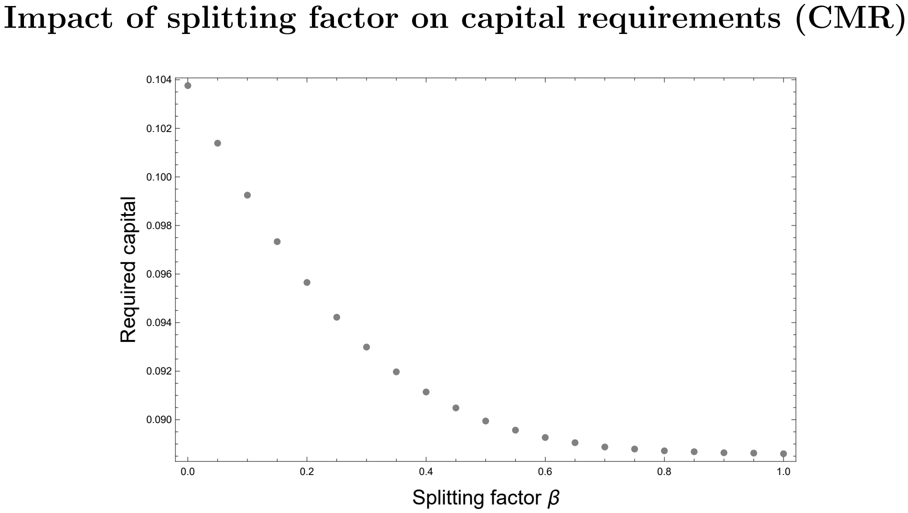


Figure 3.4: The picture shows the impact of the splitting factor on the capital requirements for a constant management rule and the benchmark parameter setting.

ducing its capital requirements if the insured intends to pay a postponed premium: Either to offer her a periodic payment contract or to implement a management rule that reduces the capital requirements to the one of an upfront premium payment as described above.

The second observation regarding the variable management rule is that also in the case of periodic premium payments or an upfront payment the required capital to fulfill the SFP is reduced. Finally, we observe that for every management rule (e_1, e_2, e_3) s.t. assumption $(*)$ is fulfilled, the splitting factor that minimizes the capital requirements differs. Thus there exists for every suitable choice of a management rule an optimal splitting factor β^* s.t. the required capital is minimized. The corresponding pictures for different management rules are stated in Figure 3.5.

3.1.4 Expected utility of the insured

We are not only interested in the impact of the management rules and splitting factor on the insurance company. Without the willingness of the insured to enter the contract, the insurance company is not able to conclude the contract. We measure the willingness to pay for a contract by the expected utility of the insured. The higher the expected utility, the higher is her willingness to enter the contract. Thus, it is interesting to analyze the impact of the splitting factor and the management rules on the expected utility of the insured to determine the optimal portfolio allocation s.t. her expected utility is maximized. For this we need to define the utility function of the insured. We consider that the insured has a constant relative risk aversion

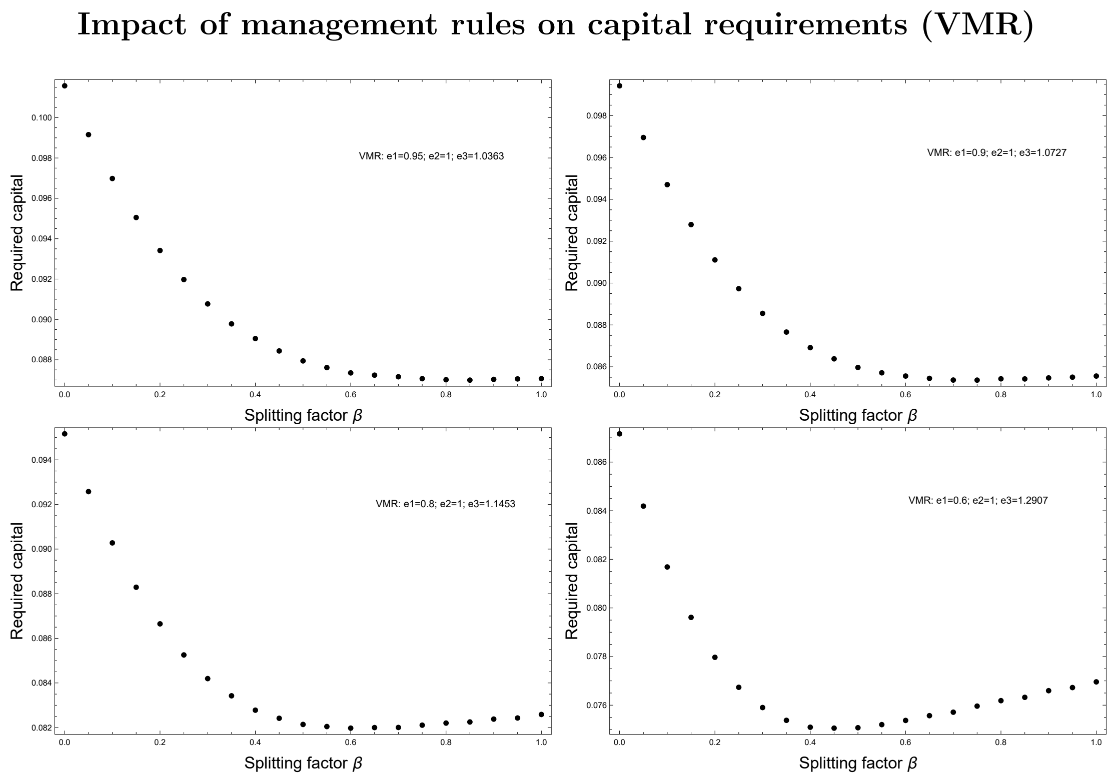


Figure 3.5: The pictures show the impact of the management rules on the capital requirements. Beside the benchmark parameter setting we use $c_1 = c_2 = 1$. The corresponding values for e_1, e_2, e_3 are stated in the pictures.

(CRRA), s.t. the utility function is given by

$$u(x) = \begin{cases} \frac{x^{1-\gamma}}{1-\gamma}, & \gamma > 1 \\ \ln(x), & \gamma = 1, \end{cases}$$

where γ denotes the relative risk aversion. Of course the fair pricing condition has to be fulfilled s.t. the optimization problem (without stating the optimization argument) is of the form

$$\begin{gathered} \max\ \mathbb{E}_{\mathbb{P}}\left[u(L_2)\right] \\ \text{s.t. } e^{-2r}\,\mathbb{E}_{\mathbb{P}^*}\left[L_2\right] = 1, \text{ where} \\ L_2 = V_2 + (G_2 - V_2)^+ \text{ and } V_2 = \alpha\left(\beta\frac{A_2}{A_0} + (1-\beta)e^r\frac{A_2}{A_1}\right). \end{gathered} \tag{3.14}$$

In the case of no guarantee ($g \to -\infty$) the put option vanishes and it holds $L_2 = V_2$ s.t. the fair pricing is always fulfilled. For the special case of an upfront premium

($\beta = 1$) it holds $V_2 = \frac{A_2}{A_0}$. This portfolio planning problem already dates back to Merton (1971). He determines the optimal investment fraction π that has to be invested in the risky asset s.t. the expected utility is maximized. Thus, he solves problem (3.14) with π as optimization argument. The solution is given by the famous Merton solution

$$\pi^{Mer} = \frac{\mu - r}{\sigma^2 \gamma}.$$

If we allow for a guarantee but stick to the case of an upfront premium payment, the optimal solution in terms of the investment fraction is also given by the Merton solution. This result can be traced back to Basak and Shapiro (2001) and found e.g. in Mahayni et al. (2021a) (cf. Chapter 2, Proposition 2.11 for the case $\varepsilon \to 1$). The only difference compared to the situation without guarantee is that the insured does not participate fully in the asset return because of the incurring guarantee costs (denoted with ν in Proposition 2.11).

It is now of interest to analyze how the splitting factor and the management rules have an impact on the optimal expected utility of the insured and how the optimal investment fraction is influenced by this. In the following analysis we speak of the certainty equivalent (CE). This is a monotone transformation of the expected utility given by

$$CE = u^{-1} \left(\mathbb{E}_{\mathbb{P}} \left[u(L_2) \right] \right).$$

Because of the monotonicity, the optimal parameter of the CE solution coincides with the optimal parameter for the expected utility solution. The certainty equivalent savings rate is furthermore given for $T = 2$ by

$$y_{CE} = \frac{1}{2} \ln(CE).$$

Impact splitting factor and constant management rule on the expected utility of the insured

We start the analysis with the constant management rule, i.e. $\sigma_{A,0} = \sigma_{A,1} = \sigma_A$ resp. $\mu_{A,0} = \mu_{A,1} = \mu_A$. Here we can identify the influence of the splitting factor on the expected utility. Let us first concentrate on the two special cases where we can calculate the expected utility in closed-form: the upfront ($\beta = 1$) and the postponed premium case ($\beta = 0$).

Proposition 3.7 (Closed-Form Solutions - CMR)
Let L_2 be the terminal wealth of the insured, $u(\cdot)$ a CRRA utility function, $G_2 = e^{2g}$ the terminal guarantee feature and $\Phi(\cdot)$ the distribution function of the standard normal distribution.

(a) For the upfront premium case ($\beta = 1$) it holds:

$$\textit{(i)}\ \mathbb{E}[u(L_2^{(upfront)})] = \tfrac{\alpha^{(1-\gamma)}}{1-\gamma} e^{(1-\gamma)(2\mu_A - \gamma\sigma_A^2)} \left\{ 1 - \Phi\left(\frac{\ln\left(\frac{G_2}{\alpha}\right) - 2[\mu_A - \sigma_A^2(\gamma - \frac{1}{2})]}{\sqrt{2}\sigma_A} \right) \right\}$$
$$+ \tfrac{1}{1-\gamma} G_2^{(1-\gamma)} \Phi\left(\frac{\ln\left(\frac{G_2}{\alpha}\right) - 2\mu_A + \sigma_A^2}{\sqrt{2}\sigma_A} \right).$$

$$\textit{(ii)}\ CE^{(upfront)} = \left\{ \alpha^{(1-\gamma)} e^{(1-\gamma)(2\mu_A - \gamma\sigma_A^2)} \left\{ 1 - \Phi\left(\frac{\ln\left(\frac{G_2}{\alpha}\right) - 2[\mu_A - \sigma_A^2(\gamma - \frac{1}{2})]}{\sqrt{2}\sigma_A} \right) \right\} \right.$$
$$\left. + G_2^{(1-\gamma)}\ \Phi\left(\frac{\ln\left(\frac{G_2}{\alpha}\right) - 2\mu_A + \sigma_A^2}{\sqrt{2}\sigma_A} \right) \right\}^{\frac{1}{1-\gamma}}.$$

(b) For the postponed case ($\beta = 0$) it holds:

$$\textit{(i)}\ \mathbb{E}[u(L_2^{(postponed)})] = \tfrac{\alpha^{(1-\gamma)}}{1-\gamma} e^{(1-\gamma)(r + \mu_A - \frac{1}{2}\gamma\sigma_A^2)} \left\{ 1 - \Phi\left(\frac{\ln\left(\frac{G_2}{\alpha}\right) - r - \mu_A - \sigma_A^2(\frac{1}{2} - \gamma)}{\sigma_A} \right) \right\}$$
$$+ \tfrac{1}{1-\gamma} G_2^{(1-\gamma)} \Phi\left(\frac{\ln\left(\frac{G_2}{\alpha}\right) - r - \mu_A + \frac{1}{2}\sigma_A^2}{\sigma_A} \right).$$

$$\textit{(ii)}\ CE^{(postponed)} = \left\{ \alpha^{(1-\gamma)} e^{(1-\gamma)(r + \mu_A - \frac{1}{2}\gamma\sigma_A^2)} \left\{ 1 - \Phi\left(\frac{\ln\left(\frac{G_2}{\alpha}\right) - r - \mu_A - \sigma_A^2(\frac{1}{2} - \gamma)}{\sigma_A} \right) \right\} \right.$$
$$\left. + G_2^{(1-\gamma)} \Phi\left(\frac{\ln\left(\frac{G_2}{\alpha}\right) - r - \mu_A + \frac{1}{2}\sigma_A^2}{\sigma_A} \right) \right\}^{\frac{1}{1-\gamma}}.$$

The proof of Proposition 3.7 is given in Appendix C.6. Notice that the participation fractions α in part a) and b) differ because of the different premium payment cases.

For the no guarantee case ($g \to -\infty$) Corollary 3.2 follows immediately:

Corollary 3.2 (Closed-Form Solutions CMR - no Guarantee)
We assume the same assumptions as in Proposition 3.7. For the special case of $g \to -\infty$ (i.e. $\alpha = 1$) it holds that $L_2 = V_2$, s.t. the following results hold.

(a) Upfront premium ($\beta = 1$):

$$\textit{(i)}\ \mathbb{E}[u(V_2^{(upfront)})] = \tfrac{1}{1-\gamma} e^{2(1-\gamma)(\mu_A - \frac{1}{2}\gamma\sigma_A^2)}.$$

$$\textit{(ii)}\ CE^{(upfront)} = e^{2(\mu_A - \frac{1}{2}\gamma\sigma_A^2)}.$$

$$\textit{(iii)}\ y_{CE}^{(upfront)}(\pi) = \mu_A - \tfrac{1}{2}\gamma\sigma_A^2.$$

(b) Postponed premium ($\beta = 0$):

$$\textit{(i)}\ \mathbb{E}[u(V_2^{(postponed)})] = \tfrac{1}{1-\gamma} e^{(1-\gamma)(r + \mu_A - \frac{1}{2}\gamma\sigma_A^2)}.$$

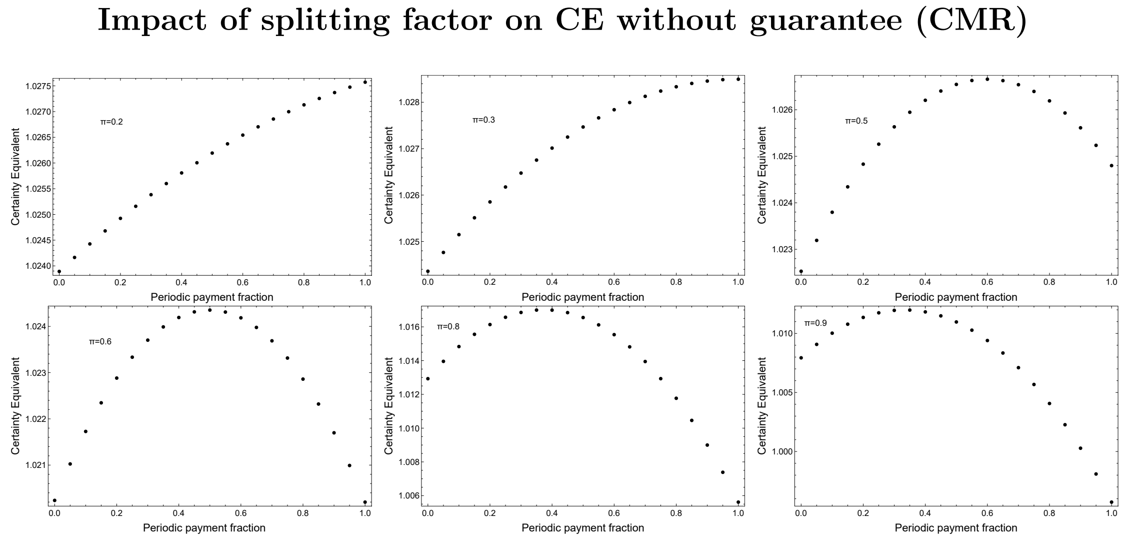


Figure 3.6: The pictures show the impact of the splitting factor β on the certainty equivalent for the case of no guarantee for different investment fractions. Beside the benchmark parameter setting we use $\gamma = 4$. The corresponding values of π are stated in the pictures.

(ii) $CE^{(postponed)} = e^{(r+\mu_A-\frac{1}{2}\gamma\sigma_A^2)}$.

(iii) $y_{CE}^{(postponed)}(\pi) = \frac{1}{2}(r + \mu_A - \frac{1}{2}\gamma\sigma_A^2)$.

The savings rates from Corollary 3.2 can be used to analyze for which investment fractions π a postponed and for which investment fractions an upfront premium payment is more beneficial. Again notice that a higher savings rate corresponds to a higher expected utility.

Proposition 3.8 (Savings Rates - CMR)
Let $y_{CE}^{(upfront)}(\pi)$ and $y_{CE}^{(postponed)}(\pi)$ be the savings rates as defined above for the case $g \to -\infty$, then it holds

$$y^{(postponed)}(\pi) > y^{(upfront)}(\pi) \Leftrightarrow \pi > 2\pi^{Mer},$$
$$y^{(postponed)}(\pi) = y^{(upfront)}(\pi) \Leftrightarrow \pi = 2\pi^{Mer},$$
$$y^{(postponed)}(\pi) < y^{(upfront)}(\pi) \Leftrightarrow \pi < 2\pi^{Mer}.$$

The proof of Proposition 3.8 is straight forward. We find that postponing the premium payments in the case of no guarantee is just favorable compared to the upfront premium case if the investment fraction of the insurance company is more than twice

as large as the Merton fraction.
Another interesting case is the one where the investment fraction is exactly twice as large as the Merton fraction. Here the savings rates and thus the expected utility of the insured coincides. We will see below that this result can be generalized in the case of periodic premium payments with no guarantee. Notice again that for a splitting factor $\beta \in]0,1[$ there exist no closed-form solutions for the expected utility of the insured. Thus we have to work with simulations. The results are presented in Figure 3.6.
First of all notice that the scale of the y-axes in the pictures differ because of the quite different outcomes for the certainty equivalents. The highest CE is achieved as discussed before for the Merton solution (here: $\pi^{Mer} = 0.3$) with an upfront premium payment. Any deviation from an upfront premium fraction leads to a utility loss and it gets higher the more we deviate from the upfront case. Thus for the Merton fraction the highest utility loss occurs if the insured invests with a postponed premium payment. This observation also holds for investment fractions $\pi < \pi^{Mer}$, because here the investment fractions are smaller than the optimal one and the insured has the willingness to stick as close as possible to the Merton solution. This is achieved by investing everything of her premium fraction at inception. For the other investment fraction cases ($\pi > \pi^{Mer}$) the situation is different: here the investment fraction is higher than the optimal Merton solution and thus it is optimal for the insured to invest a premium fraction of $\beta < 1$ to balance the overinvestment into the risky asset. As seen in Proposition 3.8 for an investment fraction of $\pi = 0.6 = 2\pi^{Mer}$ the CE for the upfront and postponed premium case coincide. Even more interesting is the observation that it holds $CE(\beta) = CE(1-\beta)$, s.t. the certainty equivalent is symmetric around the value of $\beta = 0.5$. This determines also the optimal premium fraction that maximizes the CE for that investment fraction. If the investment fraction is increased even further, the postponed premium case is preferred over the upfront premium case of the insured and the optimal splitting factor tends more and more towards the postponed premium case.
Notice that for a given splitting factor the portion β is invested in the asset side of the insurance company (and the remaining part into the risk-free asset), i.e. $\beta\pi$ is invested in the risky asset S. Thus we can determine the optimal splitting factor that maximizes the expected utility for a given investment fraction in the case of no guarantee in the way that $\beta\pi$ equals the Merton fraction or is 'as close as possible' to the Merton solution.

Lemma 3.3 (Optimal Splitting Factor CMR - no Guarantee)
For the case of no guarantee ($g \to -\infty$) and a prescribed investment fraction π the optimal splitting factor β^* that solves the optimization problem

$$\max_{\beta \in [0,1]} \mathbb{E}_{\mathbb{P}}[u(V_2)]$$

Impact of splitting factor on CE with guarantee (CMR)

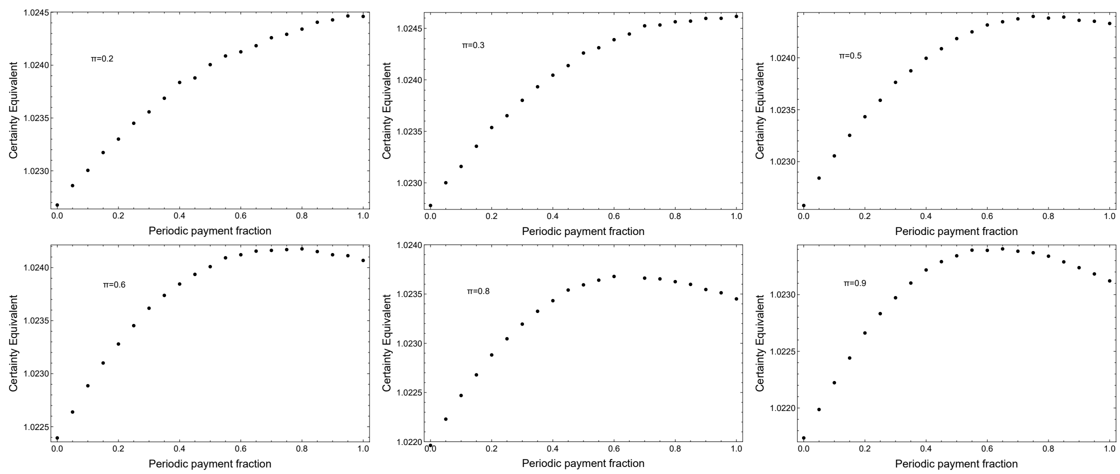


Figure 3.7: The pictures show the impact of the splitting factor β on the certainty equivalent for the case of a guarantee of $g = 0.0025$ for different investment fractions π. Beside the benchmark parameter setting we use $\gamma = 4$. The corresponding values of π are stated in the pictures.

is given by

$$\beta^* = \begin{cases} 1, & \pi \leq \pi^{Mer} \\ \frac{\pi^{Mer}}{\pi}, & \pi > \pi^{Mer}. \end{cases}$$

Including a guarantee, the optimization problem for a fixed investment fraction differs from the one above without guarantee. It is given by

$$\max_{\beta \in [0,1]} \mathbb{E}_{\mathbb{P}} \left[u(L_2) \right]$$
$$\text{s.t. } e^{-2r} \, \mathbb{E}_{\mathbb{P}^*} \left[L_2 \right] = 1.$$

We need to take care of the fair pricing condition and the put option as also the participation fraction α influence the value of the certainty equivalent, s.t. it is not possible to receive a similar result as in Lemma 3.3. But the interpretation of the results compared to the CMR case without guarantee is similar, pictured in Figure 3.7.

Again notice that the scales of the y-axes differ. Overall we find that the certainty equivalents are smaller than the ones in the no guarantee case. This is because the optimal solution of a CRRA investor is given by a pure constant mix strategy (the Merton solution). It follows that any guarantee scheme is undesired by a CRRA investor s.t. any deviation from the optimal Merton solution results in a utility loss of the insured. The overall optimal solution as mentioned above is also in this case given by the Merton fraction and an upfront premium fraction. For an investment fraction $\pi < \pi^{Mer}$ we receive the highest CE for upfront premium fractions and for $\pi > \pi^{Mer}$ the insured needs to split her contributions to receive the highest certainty equivalent. The most important difference compared to the situation without guarantee is, that the postponed premium is in none of the cases optimal. Thus we can conclude that in case of a CMR and a guarantee splitting the premium payments can be optimal for the insured in terms of the certainty equivalent, but postponing the payments is not an option for her.

Impact splitting factor and variable management rule on the expected utility of the insured

For implementing a VMR we have to take care that Assumption $(*)$ is fulfilled. We are maximizing over all splitting factors $\beta \in [0, 1]$ the expected utility for different (prescribed) investment fractions. Thus we have to adapt the choices of e_1, e_2 and e_3 as discussed in Remark 3.3. Let us again start with the analysis of the no guarantee case.
The first observation - independent of the investment fraction - is, that the certainty equivalents under a VMR are smaller than in the CMR case. This is because the VMR influences the investment fraction s.t. there exists a deviation from the Merton solution which causes certainty equivalent losses.
A second major observation and difference compared to the constant management rule is that the overall optimal solution is again achieved by the Merton fraction but with a splitting factor of $\beta < 1$. This is meaningful because in the case of a variable management rule the insurance company adapts the investment fraction over time to the market movements and thus following the 'static' Merton solution (i.e. $\beta = 1$) cannot lead to the optimal certainty equivalent. This observation is plotted in Figure 3.8. Notice that the right side of the figure also refers to an investment fraction of $\pi = 0.3$. It shows, in an enlarged way, the most interesting interval for the premium fraction. The effect is rather small for the usage of a management rule that accounts for a portfolio insurance strategy (i.e. $e_1 < 1, e_2 = 1$ and $e_3 > 1$). If we use a more risky strategy that increases the investments in bad market states and decreases the investments in a good state (i.e. $e_1 > 1, e_2 = 1$ and $e_3 < 1$ s.t. Assumption $(*)$ is fulfilled; we call it a *gambling strategy*), the effect also occurs in that magnitude. Thus we can conclude that a management rule affects the optimal solution in the way that it deviates from the Merton solution but the effects are rather small.
Besides these two observations, the certainty equivalents for different investment fractions behave similarly to the ones in the CMR case. For completeness we show other investment fraction cases in Appendix C.7.

Including a guarantee feature to the VMR setting, the differentiation between the

Impact of variable management rule on Merton solution without guarantee

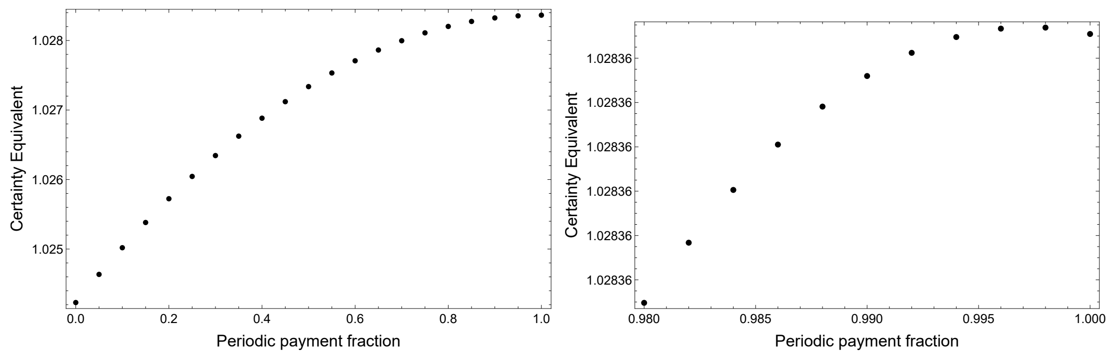


Figure 3.8: The pictures show the impact of the splitting factor β and the variable management rule on the certainty equivalent for the case of no guarantee. Besides of the benchmark parameter setting we use $\gamma = 4$, $e_1 = 0.8, e_2 = 1$ and $e_3 = 1.1453$. The corresponding value of π is the Merton fraction, i.e. $\pi = 0.3$. The left picture shows the certainty equivalent for all premium fractions $\beta \in [0, 1]$ and the right picture focuses on the fractions in the interval $[0.98, 1]$ to display the maximum CE.

portfolio insurance strategy and the more risky gambling strategy is even more important. Analyzing the portfolio insurance strategy we show that introducing the management rule leads to the fact that the certainty equivalent of the insured is maximized for an upfront premium fraction, independent of the prescribed investment fraction. This can be interpreted in the way that the risk averse insured is protected by the portfolio insurance strategy (if the asset evolution for period $[0, 1]$ has not been promising, the investment fraction into the risky asset is reduced) and thus invests all of her contributions at inception.
The gambling strategy in contrast leads to a completely different behavior: Here the optimal splitting factor is smaller than one, s.t. postponing some of the insured's payments increases her certainty equivalent. Even in the Merton solution case the deviation from an upfront premium is more present.[55] Furthermore, we find that the certainty equivalents in general are greater than the ones in the portfolio insurance setting. Thus we can conclude that the more risky gambling strategy together with splitting the contributions (investing everything at inception would be too risky for the risk averse insured) leads to a higher certainty equivalent than the more protective portfolio insurance strategy. The observations are plotted in Figure 3.9. The first row presents the portfolio insurance strategy and the second row the gambling strategy.

[55] If we use for example the strategy $e_1 = 1.2, e_2 = 1$ and $e_3 = 0.8547$ and invest the Merton fraction we receive the highest certainty equivalent for a splitting factor of $\beta = 0.813$.

Impact of variable management rule on Merton solution with guarantee

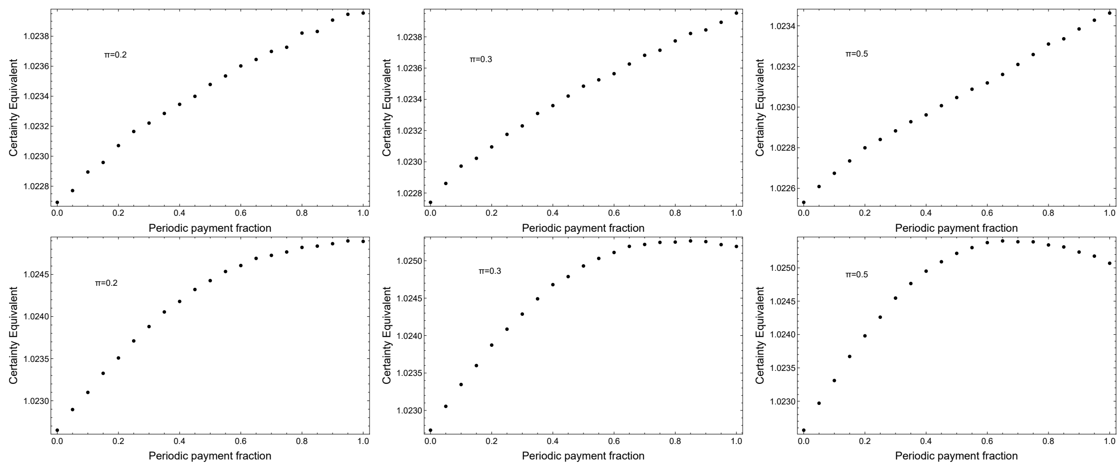


Figure 3.9: The pictures show the impact of the splitting factor β and the variable management rule on the certainty equivalent for the case of no guarantee. Beside the benchmark parameter setting we use $\gamma = 4$ and $e_1 = 0.8, e_2 = 1$. The value of e_3 is chosen depending on the investment fraction s.t. Assumption $(*)$ is fulfilled. The corresponding values of π are stated in the pictures.

Finally notice, that our results are obtained for a fixed guarantee amount g and a fixed constant relative risk aversion γ. First of all recall that the choices of e_1, e_2, e_3 for the VMR do not depend on the level of the risk aversion and the guarantee level. Thus changes in the two parameters do not influence the choices of e_1, e_2, e_3 s.t. Assumption $(*)$ is fulfilled. Varying the level of relative risk aversion leads to similar results: In the CMR case the received results qualitatively coincide for every level of risk aversion γ with our benchmark setting of $\gamma = 4$: For $\pi \leq \pi^{Mer}$ the upfront premium case is optimal and for $\pi > \pi^{Mer}$ it is optimal to split the contribution to be as close as possible on the Merton solution. Of course for $g \to -\infty$ it can be the case that because of a small level of risk aversion γ the Merton fraction is greater than 0.5 s.t. for an investment fraction of $\pi \in [0, 1]$ it is never the case that $\pi > 2\pi^{Mer}$ and thus the postponed premium case is never preferred over the upfront investment. If γ is large enough (i.e. π^{Mer} is small) and applying Lemma 3.3 we find that a postponed contribution can be preferred over the upfront premium case for even small investment fractions. Thus a postponed premium payment can even lead to an optimal solution for a given investment fraction. For the VMR we also derive similar results as in our benchmark case: The Merton solution is not optimal anymore for the no guarantee case because the insured splits her contributions to receive the highest certainty equivalent. If we allow a guarantee and implement a gambling strategy, the optimal splitting factor β is smaller than one and even the

optimal investment fraction does not coincide with the Merton fraction.
If we vary the prescribed guarantee g we find that the certainty equivalents increase with a decreasing g because the insured as a CRRA investor does not seek a guarantee. For the CMR the guarantee does not affect the results as described above: The representation of Mahayni et al. (2021a) holds for every guarantee amount g. The only difference here is the participation fraction because with changing guarantee rates the guarantee costs differ and thus the participation fraction. In case of a VMR the effects strongly depend on the value of the guarantee and the management rule choice: For a guarantee $g > 0$ the portfolio insurance strategy leads to optimal splitting factors of $\beta = 1$, independent of the investment fraction. But the higher the guarantee rate g, the closer the certainty equivalents for different splitting factors (e.g. if we set $g = 0.009$ then the difference between upfront and postponed premium in terms of the certainty equivalent is just 0.0028). For a negative guarantee rate the behavior changes and it is optimal to split the premium payments. The more negative the guarantee rate, the more extreme is this splitting reaction. Moreover, the differences in the certainty equivalents between the best and worst decision increase with decreasing guarantee. If we implement a gambling strategy there are opposing effects compared to the portfolio insurance strategy: For a negative guarantee rate it is optimal in the Merton fraction to invest everything at inception and even for an investment fraction greater than the Merton solution we receive less splitting and more investing in the first interval compared to the situation where the guarantee rate is positive. In contrast, a more positive guarantee rate (compared to our benchmark case of $g = 0.0025$) leads to more splitting and thus to the reaction that the insured invests more of her contribution at the second interval. The results for the gambling strategy are pictured in Figure 3.10. The corresponding results for the portfolio insurance strategy are stated in Appendix C.8.

3.1.5 Conclusion

We analyze the impact of periodic premium payments of the insured on the pricing of contracts, the risk management of the insurance company and the expected utility of the insured under management rules. The contract offers the insured the maximum of a guaranteed rate and a participation in the asset returns. Within a stylized two period Merton model the management rules influence the investment fraction and thus affect the risk structure of the second period. We find that the splitting factor, which determines the periodic premium payments of the insured has a huge impact on the pricing of the contract. For a constant management rule we explore that the guarantee costs, which can be stated in quasi closed-form, are monotonically increasing and convex in the splitting factor β. Including a variable management rule, which has to fulfill some assumptions to avoid violations on the fair pricing, we can compare the guarantee costs of the CMR with the ones of the VMR. We show

Impact of guarantee rate on variable management rules

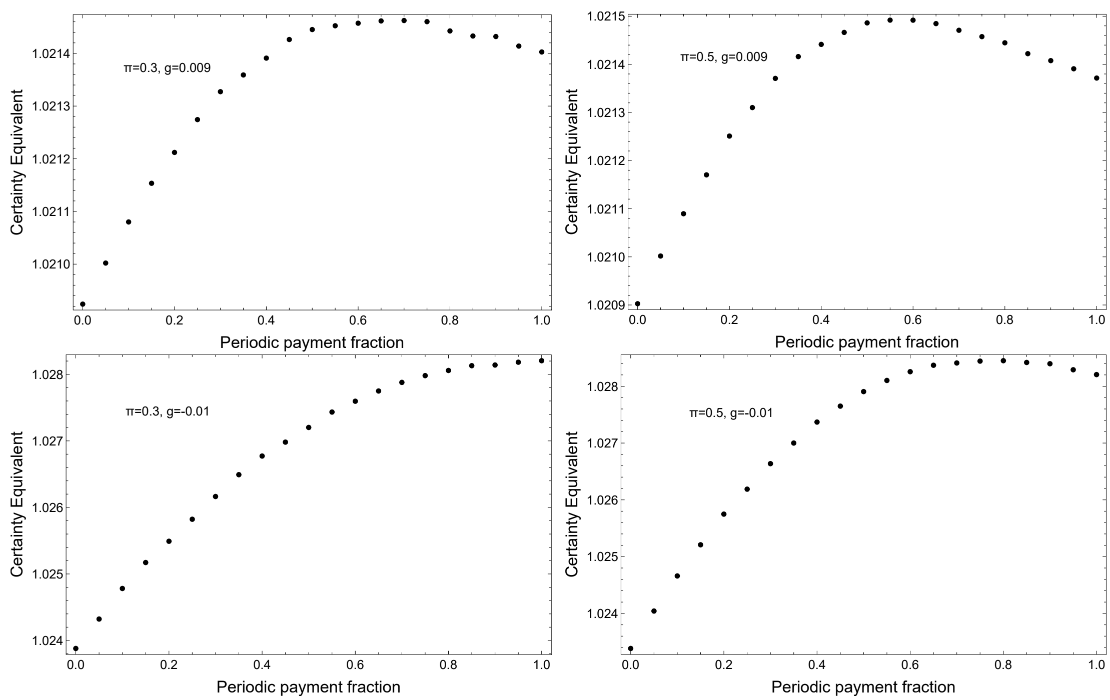


Figure 3.10: The pictures show the impact of the guarantee rate on the optimal splitting factors β for variable gambling management rule. Beside the benchmark parameter setting we use $\gamma = 4$ and $e_1 = 1.2, e_2 = 1$. The value of e_3 is chosen depending on the investment fraction s.t. Assumption $(*)$ is fulfilled. The corresponding values of π and g are stated in the pictures.

that the costs for a VMR are always greater than the ones under a CMR and that the guarantee costs are also convex and monotonically increasing in the splitting factor. For the risk management of the insurance company we find that splitting the contributions of the insured leads to an increase of the required capital for the insurance company s.t. a shortfall probability constraint is fulfilled. The required capital can be reduced by implementing a variable management rule to adapt to the riskiness of the portfolio. Finally, we discuss the impact of the splitting factor on the expected utility of the insured. For a constant management rule the upfront premium case in combination with the Merton fraction leads to the highest expected utility of the insured. Deviations from the Merton fraction imply that the optimal splitting factor has to be adapted to a value smaller than one. For variable management rules we find that an upfront contribution is not optimal for the insured and even the Merton fraction itself as investment fraction is not optimal anymore. The effects depend on the choice of the guarantee level g.

3.2 Literature review on portfolio planning including guarantee features

In the previous section, we analyzed a contract with periodic premium payments including a terminal guarantee feature. Besides the terminal guarantee there are other interesting and for the praxis relevant guarantee features: The ratchet guarantee is a frequently used feature in life insurance contracts and in the context of variable annuities (c.f. the paper of Bauer et al. (2008) and Ledlie et al. (2008)). Also, the cliquet feature is used and analyzed by many authors, see e.g. in Hieber et al. (2016) for the perspective of a life insurance analysis and Bacinello et al. (2011) for the variable annuity context. A detailed description of these guarantee features in the context of upfront premium payments is presented by Ebert et al. (2012) and Ruß and Schelling (2018). Thus an analysis of the impact of these guarantee features on the optimal portfolio planning of the insurance company resp. for the insured is another interesting research aspect: As discussed in the previous section, the CRRA investor does not want a terminal guarantee (including the guarantee feature leads to a decrease in the certainty equivalent). This is due to the fact that the guarantee can be seen as an additionally imposed constraint on the optimal investment strategy and thus reduces the value of the optimal solution. We investigate in terms of a literature overview if the other guarantee features evoke a different behavior of the insured resp. under which assumptions or models the insured wants a guarantee feature. To answer this question we analyze the portfolio planning literature that includes guarantee features. Before, let us briefly discuss the two mentioned guarantee concepts in our periodic premium payment setting. This is an extension of the existing results on cliquet and ratchet guarantees where, to the best of our knowledge, only upfront contributions are included.
Following the notation as in the previous section we can define the *Ratchet Guarantee* feature. The insured's terminal wealth including that guarantee feature, denoted with $L_2^{(Ratch)}$, is given by

$$
\begin{aligned}
L_2^{(Ratch)} &= \max\{V_1, V_2, G_2\} \\
&= V_2 + (\max\{G_2, V_1\} - V_2)^+ \\
&= \alpha\left(\beta\frac{A_2}{A_0} + (1-\beta)e^r\frac{A_2}{A_1}\right) + \\
&+ \left(\max\left\{G_2, \alpha\left(\beta\frac{A_1}{A_0} + (1-\beta)e^r\right)\right\} - \left[\alpha\left(\beta\frac{A_2}{A_0} + (1-\beta)e^r\frac{A_2}{A_1}\right)\right]\right)^+ . \quad (3.15)
\end{aligned}
$$

Here the comparison between the guarantee and the asset returns take place at every discrete point in time (in our setting at $t = 1$ and $t = 2$). The insured receives the overall maximum of the asset return evolution over time resp. the guarantee amount G_2. The costs for the ratchet guarantee scheme can be written as a put option as seen in equation (3.15). Then the ratchet guarantee costs for the insured, denoted with $GC_0^{(Ratch)}$, are given by

$$GC_0^{(Ratch)}(\beta) = e^{-2r}\mathbb{E}_{\mathbb{P}^*}\left[(\max\{G_2, V_1\} - V_2)^+\right]. \tag{3.16}$$

The other guarantee feature besides the terminal guarantee that is often applied in theory and practice is the so-called *Cliquet-Style Guarantee* feature. Combining this guarantee promise with the terminal wealth of the investment strategy, the insured's terminal payoff $L_2^{(Cliq)}$ is given by

$$L_2^{(Cliq)} = \max\{V_1, e^g\}\max\left\{\frac{V_2}{V_1}, e^g\right\} \tag{3.17}$$

$$= \max\left\{\alpha\left(\beta\frac{A_1}{A_0} + (1-\beta)e^r\right), e^g\right\}\max\left\{\frac{A_2}{A_1}, e^g\right\} \tag{3.18}$$

This guarantee feature compares at every discrete point in time the return from the investment strategy with the one period guarantee rate and takes the maximum of both components. This is done for every discrete step in time. A representation for the upfront premium case can for example be found in the paper of Ruß and Schelling (2018). For fair pricing of the guarantee costs without periodic payments we can also refer to Ruß and Schelling (2018) or Kijima and Wong (2007).

To state the guarantee costs for the cliquet-style guarantee in the periodic setting we use the fair pricing condition of the insured's liabilities. It holds:

$$e^{-2r}\mathbb{E}_{\mathbb{P}^*}\left[L_2^{(Cliq)}\right] = \alpha + GC_0^{(Cliq)}(\beta).$$

Rearranging this equation we receive

$$GC_0^{(Cliq)}(\beta) = e^{-2r}\mathbb{E}_{\mathbb{P}^*}\left[L_2^{(Cliq)}\right] - \alpha.$$

If we take the possibility into account, that the guaranteed interest rate g is vanishing, i.e. $g \to -\infty$, the terminal guarantee and cliquet style guarantee feature coincide and just gives the terminal wealth of the investment strategy. In the ratchet guarantee case there still exists a guarantee feature, even if $g \to -\infty$.[56]
After presenting and analyzing the guarantee features, let us investigate if cliquet resp. ratchet guarantees are more wanted by CRRA investors than the terminal guarantee, i.e. if the expected utility is maximized for a guarantee rate $g > -\infty$ for these guarantee features. Moreover, we discuss the assumptions and models under which a need for a guarantee component is identified. This is done in terms

[56] It could be interesting to analyze the guarantee schemes under periodic premium payments as in the previous section to analyze the impact of the splitting factor and management rules on the pricing, the risk management and the expected utility of the insured. A first draft working paper on this topic is given by Offermann and Stein (2021).

of a literature review. One of the first papers that analyzes the impact of different guarantee features on the expected utility of the insured is Tepla (2001). He analyzes the optimal investment fraction and optimal terminal wealth under a terminal guarantee with a HARA utility function. Jensen and Sørensen (2001) analyze the optimal CRRA portfolio choice under a terminal guarantee feature.[57] For a 1-dim. Black-Scholes model with deterministic interest rate and for a Vasicek term structure model, they present the impact of the terminal guarantee component on the expected utility of the insured. Allowing a guarantee feature results in a utility loss. The paper of Deelstra et al. (2003) investigates the optimal portfolio allocation of a pension fund in which terminal wealth is protected with a terminal guarantee rate. In the context of stochastic interest rates and the special case of a CRRA utility function of the investor, they determine the optimal investment strategy where short selling and borrowing are allowed. In a follow-up paper (Deelstra et al. (2004)) they furthermore investigate the optimal terminal guarantee component that maximizes the investor's expected utility. But even the optimal determined guarantee rate cannot create a higher expected utility than the case without a guarantee.

El Karoui et al. (2005) analyze under CRRA utility the optimal terminal wealth including a continuous-time guarantee component.[58] They do not compare the optimal solution under the continuous guarantee feature with the Merton solution in terms of expected utility but we can argue that if a CRRA investor does not want a terminal guarantee then she (resp. her expected utility) suffers even more under a continuous feature. Boyle and Tian (2007) analyze the optimal terminal wealth within a n-dim. market model where the guarantee component is modeled via a hedgeable random variable. They state the optimal terminal wealth s.t. this terminal guarantee is fulfilled with a prescribed probability. This presents a generalization of the results from Basak and Shapiro (2001). Branger et al. (2010) also analyzes different investment strategies for different annual guarantee schemes under CRRA utility. The utility losses for a suboptimal investment strategy resp. guarantee scheme is analyzed and illustrated. Moreover, these utility losses are also visible in the results of Schmeiser and Wagner (2015), who analyze amongst other results how to set the terminal guarantee from a regulator's point of view s.t. the CRRA expected utility of the insured is maximized. This is done with Monte-Carlo simulations. Furthermore, Mahayni et al. (2021a) also solve this problem but in quasi closed-form and discuss the impact of a terminal guarantee feature together with the impact of default risk on the optimal expected utility maximizing terminal wealth of the insured.

None of the so far analyzed papers can explain the need for a guarantee component: The CRRA investor does neither want a terminal nor a dynamic guarantee component. Also a more complex setting with stochastic interest rates or other products as e.g. pension funds cannot explain the demand for guarantees.

[57] They call it a *constant interest rate guarantee*.

[58] In the paper it is called an American guarantee feature.

A first hint is given by Chen et al. (2015). They include mortality risk in their analysis and show that under this more realistic contract modeling, the expected utility investor prefers products with a terminal guarantee. As mentioned in the previous chapter, mortality risk seems an important and natural modeling component in a life insurance contract to represent a more realistic economy.[59] Another interesting aspect in that research area that has influenced its development from that point on is given by Døskeland and Nordahl (2008). They are the first that include cumulative prospect theory (CPT)[60] into the analysis: in a 1-dim. Black-Scholes Model setup they analyze the effects of terminal and yearly guarantees on the terminal wealth of a CRRA investor. They also find that the guarantee feature implies utility losses. Applying CPT into the model, the demand for guarantees can be explained. This is the first hint under which assumptions the integration of behavioral aspects into the analysis of guarantee components seems to be expedient. For an analysis of different investment strategies under CPT we refer to Dierkes et al. (2010) and Dichtl and Drobetz (2011). The general portfolio choice problem under CPT is investigated by He and Zhou (2011).

Many other papers discuss CPT in optimal portfolio allocation to explain the demand for guarantees: Ebert et al. (2012) analyze terminal guarantees as well as ratchet and cliquet guarantees on the terminal wealth of the insured with upfront contributions in a 1-dim. Black-Scholes Model setup. They determine the expected utility of a CRRA investor and include an S-shaped utility function from Cumulative Prospect Theory (CPT) into the analysis. Their findings suggest that Prospect Theory investors favor the terminal guarantee feature compared to the more sophisticated ratchet and cliquet style guarantees.[61]

Ruß and Schelling (2018) include even Multi-Cumulative Prospect Theory (MCPT)[62] into the analysis and find under the same assumption as Ebert et al. (2012), that under MCPT the demand for cliquet style guarantee products can be explained. Thus, we can conclude that under CRRA utility assumption a terminal guarantee is in most of the cases not wanted by the investor. The only exception is a more realistic modeling including mortality risk: the combination of long time horizons, mortality risk and a high risk aversion as driving factors can explain the demand for products with guarantees. Furthermore, behavioral aspects from Prospect Theory resp. CPT and MCPT can explain the demand for terminal as also for cliquet style guarantee features. If we do not incorporate these assumptions resp. models, the investor will always suffer from guarantees in form of a loss in her expected utility.

[59] The optimal portfolio allocation problem including mortality risk for a CRRA investor is solved by Milevsky and Young (2007).

[60] Cumulative prospect theory dates back to Tversky and Kahneman (1992).

[61] Boyle and Tian (2008) are the first ones that analyze besides the terminal guarantee a ratchet guarantee feature w.r.t. the optimal expected utility maximizing investment fraction.

[62] Multi-Cumulative prospect Theory is based on CPT and considers annual changes in the contract values.

Authors	Guarantee Feature	Optimization Problem	Assumptions
Jensen and Sørensen (2001)	*terminal*	EU, $(\pi_t)_{t\in[0,T]}$	n-dim. market model with CRRA utility and upfront contributions; include constant interest rate as well as Vasicek term structure model
Tepla (2001)	*terminal*	EU, $(\pi_t)_{t\in[0,T]}$	complete, n-dim. model with strictly increasing, strictly concave and continuously differentiable utility function
Deelstra et al. (2003)	*terminal*	EU, $(\pi_t)_{t\in[0,T]}$	complete market with 1 risk-free, 1 risky asset and a zero-coupon bond; short rate process is a generalization of Vasicek resp. the Cox-Ingersoll-Ross model
Deelstra et al. (2004)	*terminal*	EU, G_T	complete market model with 1 risk-free asset, n risky assets; CRRA utility
Iwaki and Yumae (2004)	*terminal*	MV	n-dim. market model and upfront contributions
El Karoui et al. (2005)	*continuous*	EU, W_T	1-dim. Black-Scholes model with CRRA utility function
Boyle and Tian (2007)	*terminal modeled with random variable*	EU, $(\pi_t)_{t\in[0,T]}$	complete model with n risky assets and 1 risk-free asset; utility function is continuously differentiable, strictly increasing and concave and fulfills Inada; guarantee should be fulfilled within a SFP constraint
Boyle and Tian (2008)	*terminal ratchet*	EU, $(\pi_t)_{t\in[0,T]}$	strictly increasing, strictly concave and twice differentiable utility function
Døskeland and Nordahl (2008)	*terminal annual*	EU, $(\pi_t)_{t\in[0,T]}$	1-dim. black-Scholes Model, CRRA utility resp. S-shaped CPT utility function

(To be continued)

Authors	Guarantee Feature	Optimization Problem	Assumptions
Branger et al. (2010)	*annual*	EU, $(\pi_t)_{t\in[0,T]}$	1-dim. Black-Scholes model; CRRA utility
Dichtl and Drobetz (2011)	*terminal stopp loss*	EU, $(\pi_t)_{t\in[0,T]}$	1-dim. Black-Scholes model; CRRA utility and S-shaped CPT utility function
Di Giacinto et al. (2011)	*terminal*	EU, $(\pi_t)_{t\in[0,T]}$	1-dim. black-Scholes Model; utility function is twice continuously differentiable and strictly increasing and concave
Ebert et al. (2012)	*terminal ratchet cliquet*	EU, $(\pi_t)_{t\in[0,T]}$	1-dim. Black-Scholes Model, CRRA utility and S-shaped CPT utility; upfront contributions
Chen et al. (2015)	*terminal annual*	EU, $(\pi_t)_{t\in[0,T]}$	1-dim. Black-Scholes Model including mortality risk; CRRA utility and S-shaped CPT utility function
Schmeiser and Wagner (2015)	*terminal*	EU, G_T	1-dim. Black-Scholes Model, CRRA utility
Ruß and Schelling (2018)	*terminal ratchet cliquet*	EU, $(\pi_t)_{t\in[0,T]}$	1-dim. Black-Scholes Model, MCPT utility function
He et al. (2020)	*terminal*	EU, $(\pi_t)_{t\in[0,T]}$	1-dim. Black-Scholes Model; S-shaped CPT utility function
Mahayni et al. (2021a)	*terminal*	EU, G_T EU, W_T	complete, 1-dim. market model; CRRA utility

EU=expected utility, MV=Mean-Variance, CTP=cumulative prospect theory

Table 3.2: Selected papers on portfolio planning including guarantee features

Chapter 4

Parameter uncertainty, ambiguity and optimal asset allocation under time-inconsistency

In this chapter we analyze the impact of uncertainty on an optimal asset allocation problem. In general, uncertainty is connected to the situation where we do not know the exact parameters in a model resp. only know the probability of occurrence of these parameters. Not knowing about the exact model framework might also cause uncertainty. In this chapter we want to focus on a setup where the model is known but parameters are uncertain. A current example for this uncertainty aspect is given by the COVID 19 disease or the climate change: both situations imply uncertainty and make it impossible to know the exact parameters of a financial model: For example we know that the risky assets in a financial market are following a 1-dim. Black-Scholes model, but the drift and volatility parameters μ and σ are uncertain resp. only a probability distribution about the true parameters is known. An even more uncertain situation can occur if we even do not know the exact distribution of the parameters. This is referred to a situation under ambiguity.

Under these uncertainty situations it is complicated to decide on investments. There is the unrealistic possibility that an expert knows the true parameters and can invest the overall optimal solution: in our Black-Scholes model setup under CRRA preferences of the investor this is the Merton solution. Another possibility is that the investor observes the market and receives information about the parameters. This corresponds to the topic of learning: Here the investor can implement a strategy based on the learned parameters and update the strategy in recurring points in time. As a third possibility the investor might determine the optimal solution based on the current knowledge about the parameters and the probabilities of them. She invests the - at that point in time - optimal solution for the whole investment horizon and does not change the investment fraction anymore. This refers to a so-called pre-commitment strategy. An example for a pre-commitment strategy is given in the decision of building or not building a nuclear power plant. There exists uncertainty

about the effects of it and one has to decide today if it should be build or not. The decision is not revisable s.t. one has to stick to the result.
In this chapter we focus on these pre-commitment strategies. Specifically, we are working in a Black-Scholes Model setup where two possibilities for the drift and volatility are given: (μ_1, σ_1) refers to a good evolution of the market, i.e. a high drift rate and a small volatility is assumed and (μ_2, σ_2) refers to a bad evolution of the market with low drift rate and a high volatility. We do not know which possibility is corresponding to reality but we own probabilities of occurrence in terms of p for the good market evolution and $(1 - p)$ for the bad evolution. The probabilities p and $(1 - p)$ refer to a so-called a priori lottery. Under CRRA preferences it is now interesting to analyze how the optimal pre-commitment solution, depending on the a priori lottery and the remaining investment horizon, looks like. We discuss the optimal solution in detail and analyze limiting cases where the remaining investment horizon is zero or infinite.
Another aspect that is worth being analyzed is the willingness of the uniformed investor to receive the information about the true regime. We refer to it as the value of information. Furthermore, it is interesting to see, how the optimal pre-commitment solution and the value of information changes if we analyze the given setting under ambiguity. Finally, we can extend the analysis and allow that the uncertainty is not only given in $t = 0$ but rather can occur during the whole investment horizon. This refers to a setup under regime-switching, modeled with a Markov-process. Again, we can calculate the optimal pre-commitment strategy and compare it to the situation in the a priori lottery case.
Another interesting question is, what happens to the optimal solution if we change some assumptions resp. the model. This question is answered in a literature overview. We account for learning and ambiguity in portfolio allocation itself and analyze how regime switching in portfolio allocation affects the optimal solution.
Following this line of arguments we proceed as follows: First, we introduce the reader to ambiguity and learning as also discuss the impact of these topics on portfolio allocation. In the main part of this chapter we present the optimal pre-commitment strategy under an a priori lottery and under ambiguity. Furthermore, we analyze the value of information in these two cases and give an extension of the model by investigating situations in which regime-switching can occur. Finally, we discuss how the results under regime switching are affected, if we differ the assumptions or the model framework in terms of a literature overview.

4.1 Ambiguity and learning in portfolio planning - Discussion and literature review

The topic of parameter uncertainty is widely spread in literature. Especially in the context of optimal portfolio allocation, e.g. Kan and Zhou (2007) and Xia (2001), who analyze the effects of parameter uncertainty on dynamic portfolio allocation including the possibility of learning about the parameters. In this chapter parameter uncertainty is characterized in the way that the distributional parameters of the corresponding risks are uncertain, e.g. we know that the risk X is normally distributed but the corresponding drift and/or volatility parameter μ and σ is unknown or can change from one period to another. The importance of parameter uncertainty in portfolio optimization is e.g. discussed by Rogers (2001).
Uncertainty may has an effect on different aspects: It can exist w.r.t. the distributional parameters as described above. But there could also be uncertainty about the confidence interval or even about the risk measure itself because of new regulatory requirements. Uncertainty can be overcome by learning about the market over time and adapting the corresponding investment strategy. Ambiguity, on the other side, is also important when it comes to the subject of uncertainty. In the main part of this section we model the parameter uncertainty with a lottery, i.e. the investor does not know the true regime but has some probabilities for the occurrence of the true regime. If this probabilities are also uncertain, we speak of a situation under ambiguity. This will be explained in the next subsection in more detail.

4.1.1 Ambiguity in portfolio planning

The term ambiguity derives from behavioral economics research and dates back to Arrow (1951). The work of Ellsberg (1961) describes the paradox that people behave differently in situations under risk where the probabilities of the outcomes are known compared to situations where they are ambiguous and do not know the corresponding probabilities to their actions. There are many papers with behavioral focus that deal with ambiguity. We want to discuss the most important results and refer to the literature given within the stated papers:
Fox and Tversky (1995) investigate that decision-makers are ambiguity averse. This is due to a comparison of the ambiguous setting to a less ambiguous resp. more familiar situation. They call it the comparative ignorance hypothesis. Fox and Weber (2002) extend these results by observing new ways in which the decision context can affect the willingness to act under uncertainty that do not rely on the comparative evaluation scheme. A theoretical discussion and review of the ambiguity literature is given in Epstein and Schneider (2010). Trautmann et al. (2011) find that preference reversals occur in measurements of ambiguity aversion and Gollier (2011) connects ambiguity with portfolio choices. He finds that in general ambiguity aversion does

not reduce the demand for the uncertain asset. One needs to define sufficient conditions to guarantee that an increase in ambiguity aversion reduces the demand for the ambiguous asset. The combination of portfolio choice and ambiguity is an interesting research question that is also investigated in the next section.
Portfolio choices are often determined by looking at the expected utility of the corresponding investment. Decisions under ambiguity can also be modeled in a similar way, using the *smooth ambiguity model* by Klibanoff et al. (2005). They extend the expected utility approach by overlaying the situation under risk with another uncertainty situation: The expected utility approach assigns every outcome x_i with a utility function $u(x_i)$ and then weights the outcome with the corresponding probability p_i for the state of the world x_i. Thus the expected utility is given by

$$\mathbb{E}_{\mathbb{P}}[u(X)] = \sum_{i=1}^{N} p_i u(x_i) \text{ resp. } \mathbb{E}_{\mathbb{P}}[u(X)] = \int_{\Omega} u(X) \mathrm{d}\mathbb{P},$$

depending on whether it is a discrete or a continuous setup. As mentioned before the ambiguity situation occurs if there is uncertainty about the probabilities that are assigned to the different states of the world. This is taken into account by another probability measure, denoted here with $\mathbb{P}^*$, that assigns probabilities to all possible probability measures $\mathbb{P} \in \mathcal{P}$. For example, if we set ourselves in a discrete setting and have two possible probability measures, then q denotes the probability to be under the first probability measure with the probabilities $p_1^1, p_2^1, \cdots, p_N^1$ and $(1-q)$ denotes the probability to be under the second measure with probabilities $p_1^2, p_2^2, \cdots, p_N^2$. Moreover, the model of Klibanoff et al. (2005) also addresses the ambiguity aversion which has been investigated by Fox and Tversky (1995) as stated above. For this we overlay the expected utility of the decision without ambiguity with another function v. If v is concave, then it accounts for ambiguity aversion. It is thus a similar concept as in the utility context where u accounts for risk aversion if the function is concave. Thus, if we combine the two ingredients, we receive the smooth ambiguity model of Klibanoff et al. (2005). In our simple example the smooth ambiguity model is given by

$$qv\left(\sum_{i=1}^{N} p_i^1 u(x_i)\right) + (1-q)v\left(\sum_{i=1}^{N} p_i^2 u(x_i)\right).$$

To state it in general terms, the decision-maker evaluates the double expected utility given in terms of

$$\mathbb{E}_{\mathbb{P}^*}\left[v\left(\mathbb{E}_{\mathbb{P}}[u(X)]\right)\right] = \int_{\tilde{\Omega}} v\left(\int_{\Omega} u(X)\mathrm{d}\mathbb{P}\right)\mathrm{d}\mathbb{P}^*.$$

In general there are many possibilities to choose the utility and ambiguity functions u and v. In our main section of this chapter we model both functions with a CRRA function. For the risk situation we choose $u(x) = \frac{x^{1-\gamma}}{1-\gamma}$ and for the the situation under ambiguity we choose $v(x) = \frac{x^{1-\eta}}{1-\eta}$, s.t. γ denotes the risk aversion parameter and η the ambiguity aversion parameter. The smooth ambiguity model is widely spread to model ambiguity in economics. Ju and Miao (2012) as also Chen et al. (2014) analyze the asset allocation problem in case of ambiguous returns. Ahn et al. (2014) compare different ambiguity models in the context of portfolio selection and receive that the smooth ambiguity model can explain most of the decisions. In contrast Halevy (2007) shows that not all of the ambiguity situations can be explained with this model and Epstein (2010) finds some paradoxes regarding this model. The smooth ambiguity model has even been extended by Klibanoff et al. (2009). The authors include recursive preferences in their setting.

There are also other possibilities to model ambiguity. Multiple prior preferences are one of the most common ones beside the smooth ambiguity approach. They date date back to Gilboa and Schmeidler (1989). An investor with multiple prior preferences has got several priors (probability distributions). She selects the prior after choosing a portfolio that yields the lowest expected utility given her choice. Dow and da Costa Werlang (1992), Chen and Epstein (2002), Epstein and Wang (2004) as also Garlappi et al. (2007) are some examples for research papers that use this approach to model ambiguity in asset pricing and also to determine the optimal portfolio under ambiguity.

This shows that ambiguity is widely discussed in economics. The behavioral aspects and also the connection towards the financial topics (and especially in portfolio allocation) is obvious. This is also underlined by results of Hansen and Sargent (2001), Garlappi et al. (2007) and Jeong et al. (2015) who show that investors are rather ambiguity avers than ambiguity neutral and thus it is useful to include ambiguity aspects in the analysis of an optimal portfolio.
We want to discuss how ambiguity affects the optimal investment solution in the risky resp. uncertain asset compared to the Merton fraction in a situation without ambiguity. This is done in form of a literature overview. The selected papers mentioned are stated in Table 4.1, where we present the corresponding ambiguity model and the research topic of the analyzed papers.
Dow and da Costa Werlang (1992) have been the first who introduce ambiguity in the context of portfolio optimization. They analyze a two-period model in a multiple prior setting by considering a market with one ambiguous asset and one risk-free bond. They show that the investor only invests in the ambiguous asset if its price is smaller than the expected asset return value. Maenhout (2004) shows that in an optimal portfolio allocation model ambiguity aversion decreases the optimal share in equities. Furthermore, Easley and O'Hara (2009) show under a multiple prior setting that regulation of unlikely events can moderate the effects of ambiguity and increase the participation in financial market. Moreover, they show that in some markets there is no demand for a risky asset if the drift parameter μ belongs to a certain interval. This can be attributed to the ambiguity aversion of the investor

and underlines the findings of Garlappi et al. (2007): They explore that optimal portfolios overweight risk-free assets if ambiguity is considered. The results can also be traced back to Epstein and Schneider (2008) who find in a multiple-prior setting that ambiguity averse decision makers react more strongly to bad news than to good news and that they avoid volatile assets. Guidolin and Rinaldi (2010) generalize the findings of Easley and O'Hara (2009) by showing that there exists an idiosyncratic as also a systemic uncertainty in the market. Furthermore, they can confirm the results of Easley and O'Hara (2009), that there exists a price interval where it is optimal not to invest in the risky asset.

Taboga (2005) uses the smooth ambiguity approach to select a portfolio under ambiguity with a two-stage preferences approach. He finds out that the optimal solution weights all possibilities but gives more weight to the more pessimistic ones. Gollier (2011) explores the determinants of the demand for uncertain assets and of asset prices if investors are ambiguity averse with the smooth ambiguity model. As discussed above, he finds that only under sufficient conditions the investment in the more ambiguous asset is reduced if we increase the ambiguity aversion. Moreover, Maccheroni et al. (2013) include the smooth ambiguity approach into the mean-variance optimization by analyzing a risky, a risk-free and an ambiguous asset. They confirm the results of Gollier (2011), that the investment in the more ambiguous asset is reduced only under conditions. Chen et al. (2014) analyze the optimal portfolio weight and consumption in a generalized smooth ambiguity model for i.i.d. assumptions as also for a vector autoregressive model (VAR). They show that the optimal investment in case of ambiguity is more conservative than without ambiguity. Also the paper of Zhang et al. (2017) in a n-dim. discrete model setup with transaction costs and parameter uncertainty shows that the optimal dynamic trading rule gives less weight to risky asset under ambiguity.

Mostly all papers in this literature review give reason for assuming that the investor gives more weight to the pessimistic possibilities under ambiguity and thus reduces the investment in the the ambiguous asset compared to the situation without ambiguity. This result holds true independently of the concrete ambiguity model that is assumed. This tendency can be regulated as seen in Easley and O'Hara (2009), s.t. under the right regulatory, the participation in the financial market can even be enlarged. The accounting for ambiguity in optimal portfolio planning is thus an important factor. But Branger and Larsen (2013) show that it might be not sufficient just to include ambiguity in the modeling. They show that without taking the possibility of learning in the context of uncertainty under consideration, the utility loss can be high, even if we account for ambiguity. Thus an analysis of learning is of importance. This will be the topic of the next subsection.

Authors	Ambiguity Modeling	Research Topic
Dow and da Costa Werlang (1992)	Multiple Prior	Generalize EU model by analyzing the asset choice problem under one uncertain asset in a two-period model; use a non-additive probability measure, that distinguishes between quantifiable risks and unknown uncertainties
Maenhout (2004)	Multiple Prior	1-dim. model setting with an ambiguous asset and an investment horizon of T with a CRRA utility
Taboga (2005)	Smooth Ambiguity	n-dim. portfolio selection under ambiguity with a two-stage preferences approach that disentangles ambiguity and ambiguity aversion; CARA utility
Pflug and Wozabal (2007)	Multiple Prior	MV portfolio selection problem; shows the trade-off between return and risk in view of the ambiguity situation
Garlappi et al. (2007)	Multiple Prior	MV model; optimal portfolio under ambiguity overweight risk-free assets compared to models without ambiguity
Easley and O'Hara (2009)	Multiple Prior	3-dim. market model with risk-free and two risky assets where drift and volatility is ambiguous; with CARA utility the investors do not invest risky to avoid ambiguity
Guidolin and Rinaldi (2010)	Multiple Prior	Find conditions under which trading no risky asset is optimal under ambiguity; even show that there exists idiosyncratic and systemic ambiguity in the market
Gollier (2011)	Smooth Ambiguity	1 risky and 1 risk-free asset; stating sufficient conditions under which ambiguity aversion decreases the investment in uncertain asset

(To be continued)

Authors	Ambiguity Modeling	Research Topic
Branger and Larsen (2013)	Robust control Approach	1-dim. market model under an ambiguity averse investor, stock price follows jump-diffusion process; big differences between ambiguity aversion w.r.t. diffusion risk compared to jump risk
Maccheroni et al. (2013)	Smooth Ambiguity	Mean-variance setting with a risky, a risk-free and an ambiguous asset
Chen et al. (2014)	Smooth Ambiguity	Analyze i.i.d. and VAR model under generalized smooth ambiguity assumption; optimal strategy is more conservative than without ambiguity
Pınar (2014)	Multiple Prior	n-dim. MV market setting; finding closed-form solution for investor who is ambiguity averse about mean returns
Biagini and Pınar (2017)	Multiple Prior	complete, n-dim. BS-model, maximize EU under ambiguous mean and volatility; use CRRA utility and find closed-form solution
Zhang et al. (2017)	Multiple Prior	n-dim. discrete model setup, portfolio selection problem with transaction costs and parameter uncertainty; optimal dynamic trading rule gives less weight to risky resp. high volatile factors

EU=expected utility; MV= mean-variance; VAR= vector autoregressive model

Table 4.1: Selected papers on optimal portfolio planning under ambiguity

4.1.2 Learning in portfolio planning

The topic of parameter uncertainty resp. ambiguity is strongly connected to the topic of learning. Parameter uncertainty resp. ambiguity implies that investors do not know about the true parameters in a model or even cannot estimate the probabilities for the parameter distribution. But as time passes by, the investor is able to gather information about the market and thus she can learn about the parameters and adapt her strategy until they are not ambiguous or uncertain anymore. This corresponds to the topic of learning in economics and finance. We want to introduce the reader to the main aspects in this research topic and contribute with a literature review on selected papers that are connected to the topic of learning in the context

of portfolio planning.
The cornerstone of learning dates back to the well-known Bayes (updating) rule from statistics, thus this part of mathematics is also called the Bayesian statistics and was established in the 18th century. Bayes rule shows how prior beliefs can be updated into posterior knowledge after receiving new information. The basic form of the Bayes rule is given by a connection between conditional probabilities and unconditional ones. It holds

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}.$$

The left side of the equation can be interpreted as posteriori distribution. The term $\mathbb{P}(B|A)$ is used for the likelihood of the a priori distribution $\mathbb{P}(A)$ and $\mathbb{P}(B)$ is the so-called probability of the evidence. This means that in a first step we choose an a priori distribution with subjective information about the uncertain parameter Θ. In a second step we can determine from a given sample the a posteriori distribution where two sources of information are joint together: The information from the a priori distribution and the information from the sample. In a third step the resulting a posteriori distribution of the parameter can be analyzed and the process starts all over again. This can also be transferred to random variables resp. to distributions where a parameter is unknown: Let Θ be the uncertain parameter, $\Pi(\Theta)$ the a priori distribution of Θ and $f(x_1, \cdots, x_n|\Theta)$ the joint density of the sample $X_1, \cdots, X_n$. Then the a posteriori distribution is given by

$$\Pi(\Theta|x_1, \cdots, x_n) = \frac{f(x_1, \cdots, x_n|\Theta)\Pi(\Theta)}{\int f(x_1, \cdots, x_n|\Theta)\Pi(\Theta)\mathrm{d}\Theta}.$$

A typical example in this setting is that the uncertain parameter Θ is normally distributed as a priori distribution. This can also be retraced in the literature overview. For more information about the Bayesian statistics we refer e.g. to Lee (2012).
Using Bayes rule we can update the information about an unknown drift or volatility of the corresponding parameters and thus learn about the true distributional parameters. This is an important feature in many portfolio planning problems where uncertainty about parameters matters:

Let us assume that an investor can invest into a risky asset S and risk-free asset B in the 1-dim. Black-Scholes model setup, i.e.

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t, \quad \frac{dB_t}{B_t} = r dt.$$

Furthermore, the random variable $\ln(W_T) \sim \mathcal{N}(\mu, \sigma^2)$ denotes the logarithmized terminal wealth of the investor. Moreover, we assume that the investor's utility function is given by a CRRA utility with relative risk aversion parameter γ. In case that the drift and volatility parameter (μ, σ) are known (i.e. there exists no

uncertainty), the Merton fraction $\pi^{Mer} := \frac{\mu - r}{\sigma^2 \gamma}$ as investment fraction in the risky asset maximizes the expected utility of the investor.
Now we include parameter uncertainty in the analysis: There are many possibilities to model this uncertainty: the investor might know the underlying distributional process of the asset evolution but cannot observe the true drift parameter[63] or she might even be uncertain about the distributional process itself. In our small example we assume that the drift parameter μ is uncertain. Thus, the Merton solution is not applicable anymore. This is a typical assumption in the leaning literature. Including learning about the drift μ in a Bayesian setting, the optimal solution of the resulting allocation problem has been first described in terms of an SDE by Brennan (1998) and has been explicitly solved by Rogers (2001). We want to present the basic line of thoughts in the modeling of Brennan (1998) and present the explicit representation of the optimal solution afterward.
In $t = 0$, the uncertain drift μ is normally distributed with mean m_0 and volatility v_0. The conditional expectation of the drift on observable returns up to time t is given by

$$m_t := \mathbb{E}[\mu | \mathcal{F}_t^S]$$

and the variance of the drift (filtering error) by

$$v_t := \mathbb{E}[(\mu - m_t)^2 | \mathcal{F}_t^S],$$

where $\mathcal{F}_u^S := \sigma(S_t, t \leq u)$ is the Sigma-Algebra created by the risky asset S. This allows us to base our estimation for the drift on the results of S_t. The changes for the estimate m_t can be described by the following SDE as presented in Liptser and Shiriaev (1977):

$$dm_t = \frac{v_t}{\sigma^2} \left[\frac{dS_t}{S_t} - m_t dt \right] = \frac{v_t}{\sigma^2} \left[\mu dt + \sigma dW_t - m_t dt \right],$$

with the solution

$$m_t := \frac{1}{v_0 t + \sigma^2} \left(m_0 \sigma^2 + \mu v_0 t \right) + \frac{v_0}{v_0 t + \sigma^2} \sigma W_t.$$

The SDE of the variance is given by

$$dv_t = -\frac{v_t^2}{\sigma^2} dt$$

and is solved by

$$v_t := \frac{v_0 \sigma^2}{v_o t + \sigma^2}.$$

[63] Volatilities are much easier to estimate than the drift s.t. in most of the literature it is assumed that the volatility is known. See e.g. Merton (1980) and Bollerslev et al. (1992) for evidence that the variance is predictable.

Using this results we can solve the expected utility maximizing terminal wealth problem of a CRRA investor. The optimal investment fraction into the risky asset under learning at time t is then given by

$$\pi_t^{*,Learning} = \frac{m_t - r}{\gamma\sigma^2} + (1-\gamma)\frac{v_0(T-t)}{v_0 t + \sigma^2 - (1-\gamma)(v_0 T + \sigma^2)}\left(\frac{m_t - r}{\gamma\sigma^2}\right)$$
$$= \frac{m_t - r}{\sigma^2}\frac{1}{1-(1-\gamma)\frac{v_0 T+\sigma^2}{v_0 t+\sigma^2}}.$$

The optimal investment strategy under learning is given by the Merton solution, using the updated form of the drift m_t, and by an additional component. This component can be interpreted as need for hedging against unanticipated future shifts in the estimated mean. Assuming that $m_t > r$ and $\gamma > 1$, the optimal investment fraction under learning is always smaller than the optimal Merton solution without uncertainty. A detailed discussion of the solution and its implications is given in Lundtofte (2006).

This example presents the solution for a 1-dim. Black-Scholes market model with Bayes updating. A typical assumption in literature is $\mu > r$, s.t. the case of a reduced optimal investment fraction compared to the Merton solution without ambiguity is typical. Naturally, the question arises how the results differ if we change assumptions or even change the model. This is analyzed in the following literature review.

The observation that an investor cannot observe the true state of the economy has been first introduced in a continuous modeling by Williams (1977). Gennotte (1986) is the first author who calculates the optimal estimator for the unobservable expected drift and shows in a general setting that the portfolio choice can be solved in two separate steps. He finds that the uncertainty about μ leads to a reduction in the stock allocation. Detemple (1991) further generalizes the approach by looking at non-normal distributed Bayesian learning. Brennan (1998) builds upon the work of Gennotte (1986) and analyzes the effects of uncertainty as also learning about the mean return on the risky asset portfolio decision. He determines the expected utility maximizing investment fraction under learning depending on a HJB equation and discusses the effects depending on the risk aversion of the investor: Risk averse investors reduce their investment in the risky asset under learning compared to the Merton solution, risk-seeking investors increase their investments. Rogers (2001) works with the model assumptions of Brennan (1998) and calculates closed-form solutions for the optimal investment fraction into the risky asset if learning is applied by Bayesian updating. Cvitanić et al. (2006) assume that the prior distribution of the unknown drift is given by a normal distribution and for other points in time follows a linear model depending on μ_0. They determine the expected utility maximizing terminal wealth and analyze how optimal allocations depend on the correlation between the assets' expected returns. This correlation reduces uncertainty by allowing learning across assets. The change to a linear model for the unknown drift has some meaningful changes that are pointed out by Xia (2001): He extends the model of Brennan by allowing the unknown drift μ to be predicted with a linear model with n variables. He finds closed-form solutions of the optimal investment fraction for an isoelastic utility function and concludes that the optimal allocation depends on the current value of the predictive variable. The allocation can increase, decrease or vary non-monotonically compared to the optimal solution without learning. Thus, including a linear model of the unknown drift can lead to the fact that the optimal investment fraction under learning is greater than the one in the Merton case. The linear modeling of the unknown drift is also used by Branger et al. (2013). They analyze the optimal investment fraction for the expected utility maximizing terminal wealth and include ambiguity as also learning in their analysis. Learning about the drift is given by a linear model that includes observable and unobservable predictors. They derive closed-form solutions and find that both, learning and ambiguity aversion, impact the optimal investment fraction in the risky asset. Suboptimal strategies that do not include both components in the analysis lead to significant utility losses. Peijnenburg (2018) analyzes a life-cycle model that includes ambiguity aversion and the possibility of learning about the equity premium. Ambiguity aversion reduces the optimal participation fraction which leads to an underdiversification. Learning about the equity premium over the years can lead to an increase in the stock allocation.

From this observations we can conclude that in case of a risk averse investor, the uncertainty (i.e. the need for learning) implies that the optimal investment fraction that maximizes the expected utility of the investor is smaller than the Merton fraction without uncertainty. This is a plausible result: The uncertainty in combination with the risk aversion leads to a more cautious investment strategy. The opposite behavior is true if the investor is risk seeking. In case the investor can even predict a linear model for the unknown drift, then it is also not surprising that this can lead to an increase of the investment fraction in the risky asset: If the investor beliefs in the linear model and the variables give the impression of a positive market trend, then the investor increases the investment fraction. Thus, we can finally conclude from this literature overview that the more precise our information and modeling about the uncertain drift is, the better the investor assesses the situation and therefore adapts her investment fraction in the risky asset.

There are also many papers on discrete portfolio optimization under learning. For a nice literature overview regarding this topic we refer to Pastor and Veronesi (2009) p. 372. Finally notice, that the literature on optimal portfolio planning under learning is closely connected to the literature on portfolio planning under regime switches. An overview of the latter topic is given at the end of this chapter.

Authors	Learning Parameter	Optimization Problem	Assumptions
Gennotte (1986)	*Bayes* $\mu_0 \sim \mathcal{N}(m_0, v_0)$	EU	complete, $(n+s)$-dim. market model (GBM) with s technologies to invest; increasing, concave, twice differentiable von Neumann Morgenstern utility function
Detemple (1991)	*Bayes* non-normal μ	EU	n-dim. market model; von Neumann-Morgenstern utility
Brennan (1998)	*Bayes* $\mu_0 \sim \mathcal{N}(m_0, v_0)$	EU, π	1-dim. BS model with uncertainty about drift ;concave, twice differentiable utility function
Xia (2001)	*Linear model* $\mu_t = \alpha + \beta S_t$	EU, $(\pi_t)_{t\in[0,T]}$	1 risky and 1 risk-free asset with uncertainty about μ; closed-form solution under isoelastic utility
Rogers (2001)	*Bayes* $\mu_0 \sim \mathcal{N}(m_0, v_0)$	EU, $(\pi_t)_{t\in[0,T]}$	1 risky and 1 risk-free asset model including consumption with uncertainty about drift; CRRA utility
Cvitanić et al. (2006)	*Bayes* $\mu_0 \sim \mathcal{N}(m_0, v_0)$	EU, W_T	(n+2)-dim. market model; CRRA utility
Miao (2009)	*Bayes* μ is modeled by a SDE	EU	n-dim. market model; recursive multiple-prior utility process[64]
Branger et al. (2013)	*Linear model* observable and unobservable predictor for μ_t	EU, π	1-dim. BS model under drift uncertainty and ambiguity; CRRA utility
Chen et al. (2014)	*Linear model* i.i.d. and VAR	EU, $(\pi_t)_{t\in[0,T]}$	1 risky and 1 risk-free asset; uncertainty modeled under generalized smooth ambiguity assumption with CRRA functions
Peijnenburg (2018)	*Bayes* equity premium $\sim \mathcal{N}(m_0, v_0)$	EU	life-cycle model with CRRA utility
Balter et al. (2021)	*Bayes* $\mu_0 \sim \mathcal{N}(m_0, v_0)$	EU, $(\pi_t)_{t\in[0,T]}$	1-dim. BS model under drift uncertainty and ambiguity; CRRA utility

EU=expected utility; VAR=vector autoregressive model

Table 4.2: Selected papers on continuous optimal portfolio planning under learning

4.2 Optimal Asset Allocation, Time-Inconsistency and the Value of Information

4.2.1 Introduction

In this section we consider a stylized model which is tightly connected to the classical Merton problem where an investor with constant relative risk aversion maximizes her expected utility by splitting her wealth between a risky and a risk-free asset.[65] The risk-free asset grows at a constant rate and the price dynamics of the risky asset are given by a geometric Brownian motion with drift μ and volatility σ. We introduce an a priori lottery $(p, 1-p)$ in which the outcomes give rise to two regimes given in terms of (μ, σ)-tuples. Thus, we introduce a second dimension of risk which implies an outer expectation about the outcome of the lottery and an inner expectation about the expected utility within the regimes.[66] We also account for ambiguity about the "success" probability of this lottery which gives, in addition to the two dimensions of risk, a third dimension. Preferences towards risk and ambiguity are modeled by using the smooth ambiguity approach of Klibanoff et al. (2005) under a double power utility assumption. In fact, this ambiguity model implies a further outer expectation (which is analogously modeled to the risk dimensions). Optimal investment in this setting (with and without ambiguity) is thus time-inconsistent. We analyze the optimal time-inconsistent pre-commitment strategy (with and without ambiguity about the probability of the a priori lottery) and compare it to the optimal strategy if the investor can condition on the regimes (defined by the model parameters).

In the first instance, we explain time-inconsistency by the observation that the objective of maximizing the expected utility and maximizing the expected savings rate implies different investment decisions, i.e. unless the investor is described by log utility or the investment horizon converges to zero (myopia). A further limiting case is obtained by an infinite investment horizon. Our results give rise to intuitive economic explanations. If the investor can condition on the regime, she favors the regime-dependent Merton solution. If she is not able to obtain the information about the regime, her optimal (pre-commitment) strategy is between the regime-dependent Merton solutions, i.e. can be stated as a weighted average of the Merton solutions. While, in the myopic sense, the weighting factor is dominated by the regime probabilities, the impact of the regime probabilities is decreasing in the investment horizon s.t. her optimal decision is based on the worst-case, i.e. she maximizes the worst-case savings rate over the regimes. The time-inconsistency of the optimal pre-commitment strategy for $\gamma > 1$ can be traced back to the non-myopic

[64] More information about multiple priors can be found in Chen and Epstein (2002).
[65] It is based on the work of Branger et al. (2021).
[66] The outer expectation implies an aggregation of utilities.

behavior of the investor. She chooses the optimal investment over a *first period* with continuation utility over some *second period* in view. In our setup, the continuation utility is lower in Regime 2 than in Regime 1 (Regime 2 is the bad state w.r.t. to future investment opportunities). For an investor with $\gamma > 1$, the hedging motive dominates and she wants to have more wealth in the bad state. The optimal strategy thus moves towards the worst-case strategy, in which the investor foregoes some wealth in order to lower the variance and in particular to increase the level of wealth in the bad state. Our setup allows measuring the effects of time-inconsistency by means of a coefficient of time-inconsistency (normalized to values between zero and one).
We analyze the value of information, i.e. the willingness to pay for the information about the regimes. In addition, we also account for ambiguity about the regime probabilities and analyze the value of information about the probability. To simplify the expositions, we also use a stylized version in which the regime probability is a two-point random variable.

In summary, we shed light on the implications of time-inconsistency which are common in various decision problems in finance. In particular, time-inconsistency is immanent in all investment problems where maximizing expected utility implies another optimal strategy than maximizing the expected savings rate.
Thus, we provide an intuitive explanation that there is no time-consistency for a log-investor ($\gamma = 1$). Alternatively, time-consistency can technically be obtained by switching the decision objective from maximizing the expected utility of terminal wealth to maximizing EU of log-wealth (or maximizing the expected savings rate, respectively).
For $\gamma > 1$, myopia ($T \to 0$) implies that the investor acts risk neutral w.r.t. the regime dependent savings rates, i.e. the decision can be formulated by means of the expected savings rate, while an infinite investment horizon implies a maximin decision rule. Thus, the investor maximizes the worst-case savings rate. Therefore, we also add an intuitive approach why time-inconsistency implies that a risk averse investor acts like an even more risk averse investor the longer the investment horizon is. Intuitively, it is also clear that the higher the risk aversion is the faster the optimal decision converges to the result of the maximin decision rule. In addition, we show that for all pre-commitment strategies within the interval of the regime dependent Merton solutions, the savings rate is for $T \to 0$ ($T \to \infty$) given by the expected savings rate over the regimes (worst-case savings rate of the regimes). We also state intuitive results w.r.t. the value of the information about the regimes: In both limiting cases ($T \to 0$, $T \to \infty$) the maximal willingness to pay for the information about the regime is zero. In particular, the value of information is maximal for some finite investment horizon.
Using the smooth model of ambiguity, we can emphasize similar effects concerning the levels of risk and ambiguity aversion. Stating the optimal strategy by means

of a weighted average of the regime dependent Merton solutions (which depend on the level of risk aversion), the impact of the level of risk aversion and ambiguity aversion on the weighting factor are equivalent. However, we provide an intuitive explanation concerning opposing effects w.r.t. the probability distribution over the regime probabilities and the regime probabilities themselves. To simplify the exposition, we consider ambiguity about the a priori lottery $(p, 1-p)$ in terms of two possible values for p, i.e. p_a or p_b $(p_a > p_b)$. Intuitively, it is clear that the ambiguity situation is more severe the higher the spread between the probabilities p_a and p_b is. The highest spread is obtained for $p_a = 1$ and $p_b = 0$. Moreover, a higher ambiguity situation implies a lower risk situation, i.e. for $p_a = 1$ $(p_b = 0)$ the regime is known and implies Regime 1 $(p_a = 1)$ resp. Regime 2 $(p_b = 0)$.

Our first dimension of risk already dates back to Merton (1971). He solves the problem of maximizing the expected utility of an investor with constant relevant risk aversion (CRRA) in a Black-Scholes model setup. As mentioned above, our setup is similar to the problem of a social planner who aggregates the utilities of investors with different beliefs or different levels of risk aversion. Chen et al. (2021a) consider a collective of investors in a pension fund with heterogeneous risk preferences. The investment decision is delegated to a fund manager who promises a minimum guarantee for the investment. In a further study, Chen et al. (2021b) solve an optimal collective investment problem under portfolio insurance constraints assuming that investors have different levels of risk aversion and differ in their willingness to pay management fees. Further studies that take heterogeneous risk aversions of investors of a collective investment problem into account are Alserda et al. (2019). For the special case of two investors with different utility functions (log and isoelastic utility), Dumas (1989) analyzes the allocation of wealth and aggregation of capital. Garlappi et al. (2017) analyze a dynamic collective investment problem with a group of agents having heterogeneous beliefs. They find that group decisions are dynamically inconsistent and lead to inefficient underinvestments. Jackson and Yariv (2014) show that with any heterogeneity in time preferences there exists a present bias in aggregating utilities. Furthermore, heterogeneity leads to time-inconsistency even though the individual preferences are time-consistent. This finding is confirmed for a household consumption problem by Adams et al. (2014).
Due to the assumption of a double risk situation our problem is similar to a setup in which two investors with heterogeneous beliefs are restricted to follow the same strategy. Thus, time-inconsistency arises naturally. The strand of literature referring to time-inconsistency in optimal asset allocation problems can be traced back to Strotz (1955). He proposes two strategies of dealing with time-inconsistency – a strategy of pre-commitment and a strategy of consistent planning. Balter et al. (2021) compare a pre-commitment strategy with a dynamically consistent one in the context of ambiguity and learning and determine a point of regret for a pre-commitment investor. Björk and Murgoci (2014) as also Björk et al. (2017) account

for time-inconsistency in stochastic control problems. They derive a game theoretical solution within a discrete-time and continuous-time framework. These two papers build on the work of Basak and Chabakauri (2010) who study dynamic portfolio choice under mean-variance preferences. They show that the optimal investment strategy is time-inconsistent and find a distinction between pre-commitment, dynamically consistent and myopic strategies. Further literature in this context is given by Cong and Oosterlee (2016), Pedersen and Peskir (2017), Dai et al. (2021) as also the recent papers of Vigna (2020) and van Staden et al. (2021) who compare dynamically consistent and pre-commitment strategies in a mean-variance setup.
Accounting for ambiguity (which dates back to Arrow (1951)) adds a third dimension of risk in our analysis. There is strong empirical evidence for the existence of ambiguity in decision making: Antoniou et al. (2015), Brenner and Izhakian (2018) and Dimmock et al. (2016) find that ambiguity is priced in the equity market. An increase of ambiguity leads to underinvestments. Ambiguity in portfolio choice is widely spread in literature. Guidolin and Rinaldi (2013) provide an overview of the portfolio choice facing ambiguity literature. Biagini and Pınar (2017) derive a robust solution of the Merton problem of an ambiguity averse investor. Borgonovo and Marinacci (2015) give results for certainty equivalents in a multi-event problem in the presence of risk and ambiguity aversion. Jin and Zhou (2015) analyze a portfolio choice problem in an expected utility and mean-variance framework by maximizing the worst sharpe ratio. Further literature in the context of ambiguity in a mean-variance framework is given by Maccheroni et al. (2013), Pflug and Wozabal (2007) and Pınar (2014).

The outline of this section is as follows. In Subsec. 4.2.2 we give a brief review over the basic Merton results, i.e. about the optimal expected utility maximizing strategy, the utility and savings rate which is obtained by it. Subsequently, we introduce a stylized modification of the Merton problem, i.e. we introduce an a priori lottery which defines the drift and volatility tuple. In Subsec. 4.2.3 we derive the expected utility maximizing pre-commitment strategy. Comparing the case with observed regimes to the case with unobserved regimes gives us the value of information about the regime which we analyze in more detail in Subsec. 4.2.4. Moreover, in Subsec. 4.2.5 we also account for ambiguity about the regime probabilities. We show the analogies and differences stemming from risk and ambiguity aversion. Finally, we conclude the section.

4.2.2 A priori lottery

We modify the classic Merton problem by introducing an a priori lottery where the outcome is one of two regimes. Once the regime is known, the investment problem boils down to the Merton problem. To simplify the exposition, we give a review of the Merton problem as well as some basic results.

Throughout the following, we consider an investor with constant relative risk aversion (CRRA), i.e. her utility function is

$$u(x) = \begin{cases} \frac{x^{1-\gamma}}{1-\gamma} & \gamma > 1 \\ \ln x & \gamma = 1 \end{cases},$$

where γ denotes her relative risk aversion. The investor is equipped with an initial amount of V_0 which we can normalize to $V_0 = 1$ because of the CRRA framework. We only consider $\gamma \geq 1$ as the usual choice in asset allocation. Her investment decisions are given in terms of the fraction π of her portfolio wealth which she invests in a risky asset, the stock S. The remaining fraction is invested in a risk-free asset growing with constant interest rate r. In the benchmark model of Black-Scholes, the dynamics of the stock price are

$$dS_t = \mu S_t dt + \sigma S_t dW_t \text{ where } S_0 = s_0.$$

If the investor chooses the weight π_t for the risky investment in the stock, the dynamics of her wealth are given by

$$dV_t = [r + \pi_t(\mu - r)]V_t dt + \pi_t \sigma V_t dW_t.$$

The optimal investment strategy (which maximizes the expected utility) is given by the well-known constant Merton fraction[67]

$$\pi^{\text{Mer}} = \frac{\mu - r}{\gamma\sigma^2}. \tag{4.1}$$

This strategy also maximizes the certainty equivalent CE_T and the savings rate y_T which are in general defined by

$$u(CE_T) = E[u(V_T)] \quad \text{and} \quad y_T = \frac{1}{T}\ln CE_T.$$

The savings rate y (we drop the index T since the savings rate in our setting is independent of T) for a constant portfolio weight π is given by [68]

$$y(\pi) = r + \pi(\mu - r) - \frac{1}{2}\gamma\pi^2\sigma^2, \tag{4.2}$$

[67] In particular, notice that the optimal strategy implies a constant investment fraction. With no uncertainty about the future dynamics there are no state variables to condition on, and with CRRA, there is also no need to condition on current wealth. Moreover, a time-dependent strategy would increase the variance without increasing the mean, and is thus dominated by a time-independent strategy.

[68] Since we have assumed normally distributed log returns, the maximization of utility in the base case is equivalent to a mean-variance portfolio selection problem.

and the maximal savings rate is

$$y(\pi^{\text{Mer}}) = r + \frac{1}{2\gamma} \cdot \frac{(\mu - r)^2}{\sigma^2} = r + \frac{1}{2\gamma} \cdot \lambda^2,$$

where $\lambda = \frac{\mu - r}{\sigma}$ denotes the constant market price of risk. In view of the introduction of two regimes, we state the loss rate implied by a strategy π instead of the optimal strategy π^{Mer}. The loss rate is given by

$$l(\pi) = y(\pi^{\text{Mer}}) - y(\pi) = \frac{1}{2}\gamma\sigma^2 \left(\pi - \pi^{\text{Mer}}\right)^2. \tag{4.3}$$

In general, the difference between the savings rates for two different strategies π_a and π_b is

$$\begin{aligned} y(\pi_a) - y(\pi_b) &= (\pi_a - \pi_b)\mu - \frac{1}{2}\gamma(\pi_a^2 - \pi_b^2)\sigma^2 \\ &= \gamma\sigma^2(\pi_a - \pi_b)\left(\pi^{\text{Mer}} - \frac{\pi_a + \pi_b}{2}\right). \end{aligned}$$

In the following, we consider two regimes that reflect different dynamics of the stock price or different beliefs of the investor about these dynamics. In regime i ($i \in \{1, 2\}$) the expected return and the volatility of the stock are denoted by μ_i and σ_i; the corresponding probability measure is denoted by P_i. The risk-free rate is constant and equal to r in all regimes. In particular, we now assume

$$dS_{t,i} = \mu_i S_{t,i} dt + \sigma_i S_{t,i} dW_{t,i} \text{ for } i = 1, 2$$

where $S_{0,1} = S_{0,2} = s_0$ and $W_{\cdot,i}$ is a Brownian motion under the probability measure P_i. We interpret the two regimes as a good (Regime 1) and a bad (Regime 2) one, i.e. it holds that $\lambda_1 \geq \lambda_2$. This is true for $\mu_1 \geq \mu_2$ and $\sigma_1 \leq \sigma_2$, but there are also combinations with $\mu_1 > \mu_2$ and $\sigma_1 > \sigma_2$ for which this holds true.[69]

Throughout the following, we use the convention that $y(\pi, i)$ ($i \in \{1, 2\}$) denotes the savings rate within regime i. With Eqn. (4.2), it holds (for $i \in \{1, 2\}$)

$$y(\pi, i) = r + \pi(\mu_i - r) - \frac{1}{2}\gamma\pi^2\sigma_i^2.$$

The investment fraction $\pi \neq 0$ which implies $y(\pi, 1) = y(\pi, 2)$ is denoted by π^{equal}. Straightforward calculations give

$$y(\pi, 1) - y(\pi, 2) = \begin{cases} \frac{1}{2}\gamma(\sigma_1^2 - \sigma_2^2)\pi(\pi^{\text{equal}} - \pi) & \sigma_1 \neq \sigma_2 \\ \pi\,(\mu_1 - \mu_2) & \sigma_1 = \sigma_2 \end{cases}, \tag{4.4}$$

[69] In a graph with the ratio σ_1/σ_2 of volatilities on the x-axis and the ratio $(\mu_1 - r)/(\mu_2 - r)$ of expected excess returns on the y-axis, the condition $\lambda_1 > \lambda_2$ corresponds to the points above the 45-degree line.

where

$$\pi^{\text{equal}} = \frac{\mu_1 - \mu_2}{\frac{1}{2}\gamma(\sigma_1^2 - \sigma_2^2)}. \tag{4.5}$$

The two regimes thus result in the same savings rate for the trivial choice $\pi = 0$ (then the savings rate coincides with the risk-free rate) and for $\pi = \pi^{\text{equal}}$. If Regime 1 comes with the higher expected return and the lower volatility, then π^{equal} is negative. For positive portfolio weights it then always holds true that the savings rate is higher in Regime 1 than in Regime 2. If Regime 1 has a larger expected return and a higher volatility but still the higher market price of risk, the two regimes give the same savings rate for $\pi = \pi^{\text{equal}} > 0$, and the difference of the savings rates switches sign for this choice of π. The relevant cases later on will be the ones in which π^{equal} is between π_1^{Mer} and π_2^{Mer}. This holds true for

$$\frac{\sigma_1\sigma_2}{0.5(\sigma_1^2 + \sigma_2^2)} < \frac{(\mu_1 - r)/\sigma_1}{(\mu_2 - r)/\sigma_2} < \frac{0.5(\sigma_1^2 + \sigma_2^2)}{\sigma_1\sigma_2}.$$

If $\sigma_1 = \sigma_2$, the two regimes are either identical and the savings rates coincide for every π, or Regime 1 always gives a higher savings rate than Regime 2.

In our stylized modification of the Merton problem ("initial lottery") the regime is determined by the lottery $L = (p, 1-p)$ at time 0 and then stays constant over time. We thus add a second dimension to the risk situation. The expected utility of the investor (immediately before the lottery L takes place) is[70]

$$EU_{T,p} = p\,\mathbb{E}_{P_1}[u(V_T)] + (1-p)\mathbb{E}_{P_2}[u(V_T)]. \tag{4.6}$$

For the portfolio choice of the investor we distinguish two cases. In the first case (information about the regime), the regime is observable and the investor can condition her strategy on the regime, i.e., $\tilde{\pi} = (\pi_1, \pi_2)$. In the second case (no information about the regime), the regime can not be observed and the investor has to rely on some regime-independent strategy $\tilde{\pi} = (\pi, \pi)$. Notice that conditioning on the regime gives, in the (EU maximizing) optimum, on each regime a constant investment strategy. Thus, there is no restriction of generality by considering strategies $\tilde{\pi} = (\pi_1, \pi_2)$. Without the information about the regime, there is some restriction imposed by considering strategies $\tilde{\pi} = (\pi, \pi)$. Furthermore, the optimal strategies do not depend on the state of the economy and they are constant. Thus we can w.l.o.g. restrict our analysis to constant strategies which depend on the regime at most.

Comparing the case with observed regimes to the case with unobserved regimes gives us the value of information about the regime which we analyze in more detail

[70] While the utilities refer to the ones obtained in different regimes, our setup is similar to the problem of a social planner who aggregates the utilities of investors with different beliefs or different levels of risk aversion (cf. literature given in the introduction of the section).

in Section 4.2.4. We restrict the analysis to the case in which the investor receives this information immediately after the initial lottery in which the regime is drawn. We abstract from the case of learning in which the investor would gradually update her subjective probabilities of the two regimes and learns the true regime in the long run. This assumption is not a restriction to our model: The work of Bäuerle and Grether (2017) analyze a Bayesian investor who can learn about an uncertain drift μ and wants to maximize the CRRA expected utility of her terminal wealth. They find that the optimal fraction invested in the risky asset of a risk averse investor converges for $T \to \infty$ to the smallest possible Merton ratio, i.e. to the worst-case scenario. In particular, the paper reveals that the effect of learning does not play a role for a long-term investor.

$EU_{T,p}$ aggregates over utilities in the two regimes. With

$$\mathbb{E}_{P_i}[u(V_T)] = \begin{cases} \frac{1}{1-\gamma} e^{y(\pi,i)(1-\gamma)T} & \gamma > 1 \\ y(\pi,i)T & \gamma = 1 \end{cases},$$

it follows

$$EU_{T,p} = \begin{cases} \frac{1}{1-\gamma}\left[p e^{y(\pi,1)(1-\gamma)T} + (1-p)e^{y(\pi,2)(1-\gamma)T}\right] & \gamma > 1 \\ \left[p y(\pi,1) + (1-p) y(\pi,2)\right] T & \gamma = 1 \end{cases}.$$

Notice that the aggregation is highly non-linear unless $\gamma = 1$.

Due to the double risk situation, i.e. a lottery over two different Merton problems, risk aversion γ comes into play twice. First, to determine the expected utility $\mathbb{E}_{P_i}$ of a strategy conditional on the regime. The investor uses a CRRA-utility function with risk aversion γ, s.t. γ captures the aversion against normally distributed return innovations. The larger γ, the lower the savings rate, and the smaller the optimal portfolio weight. Second, γ is used again when the investor aggregates the utilities over the two regimes, i.e. when he calculates $EU_{T,p}$ given the savings rates $y(\pi,i)$ in the two regimes $i = 1, 2$. The larger γ, the lower the savings rate $y_{T,p}$ resulting out of $y(\pi,1)$ and $y(\pi,2)$.

Maximizing expected utility in (4.6) is equivalent to maximizing the certainty equivalent savings rate $y_{T,p}$ where

$$y_{T,p}(\pi) := \frac{1}{T} \ln\left(u^{-1}\left(EU_{T,p}\right)\right), \tag{4.7}$$

i.e.

$$y_{T,p}(\pi) = \begin{cases} \frac{1}{(1-\gamma)T} \ln\left[p e^{y(\pi,1)(1-\gamma)T} + (1-p)e^{y(\pi,2)(1-\gamma)T}\right] & \gamma \neq 1 \\ p y(\pi,1) + (1-p) y(\pi,2) & \gamma = 1 \end{cases}. \tag{4.8}$$

In general, the savings rate of the initial lottery does not coincide with the expected savings rate given by $p\,y(\pi,1)+(1-p)y(\pi,2)$ which is independent of the investment horizon T. Jensen's inequality implies

$$y_{T,p}(\pi)\begin{cases} > p\,y(\pi,1)+(1-p)y(\pi,2) & \gamma<1 \\ = p\,y(\pi,1)+(1-p)y(\pi,2) & \gamma=1 \\ < p\,y(\pi,1)+(1-p)y(\pi,2) & \gamma>1 \end{cases}.$$

In the special case of log-utility, the savings rate for an initial lottery coincides with the expected savings rates of the Merton problems in the two regimes. This implies that the strategy that maximizes expected utility also yields the highest expected savings rate. This is not true for $\gamma \neq 1$, i.e. unless one considers the myopic case $T \to 0$. The dependence of the savings rate for the limiting cases $T \to 0$ and $T \to \infty$ is given in the following proposition:

Proposition 4.1 (Decreasing certainty equivalent savings rate in maturity) *For $\gamma > 1$, the certainty equivalent savings rate $y_{T,p}(\pi)$ for a constant strategy π and the savings rate $y_{T,p}(\pi^{*,pre}_{T,p})$ for the optimal pre-commitment strategy $\pi^{*,pre}_{T,p}$ are both a decreasing function of T. The limiting values of the certainty equivalent savings rate $y_{T,p}(\pi)$ are*

$$\lim_{T\to 0} y_{T,p}(\pi) = p\,y(\pi,1)+(1-p)y(\pi,2)$$
$$\lim_{T\to \infty} y_{T,p}(\pi) = \min\{y(\pi,1), y(\pi,2)\}.$$

The proof of Proposition 4.1 is given in Appendix D.1.

For $\gamma > 1$ (and $T > 0$) the savings rate of a strategy π equals the expected savings rate over the regimes only for the boundary cases $p = 0$ and $p = 1$, i.e. if the second dimension of the risk situation vanishes. The larger the risk aversion γ, the lower the savings rate that is resulting out of the savings rates $y(\pi,i)$ in the two regimes. Furthermore, the savings rate increases in the probability of the good state.

The difference between $y_{T,p}(\pi)$ and $p\,y(\pi,1)+(1-p)y(\pi,2)$ also depends on p and is maximized for

$$p^*_{SR}(\pi,T) = \frac{1}{1-e^{(y(\pi,1)-y(\pi,2))(1-\gamma)T}} + \frac{1}{(y(\pi,1)-y(\pi,2))(1-\gamma)T}.$$

For $T \to 0$, we have that $\lim_{T\to 0} p^*_{SR} = 0.5$, i.e. in the myopic case the highest difference is observed in the case when the uncertainty about the regime is largest. It also holds that $p^*_{SR}(\pi,T)$ is increasing in T. With the savings rate $y_{T,p}(\pi)$ approaching the lower of the two savings rates, the difference to the higher expected savings rate is maximized when the latter has more and more weight on the larger of the two savings rates.

4.2.3 Pre-commitment and the impact of time-inconsistency

The investor cannot observe the true regime, i.e. she cannot choose a regime-dependent strategy today, and we assume that she also cannot learn about the regime over time. We denote her optimal strategy as the optimal pre-commitment strategy. With no state variables and CRRA utility, this strategy is time- and state-independent but depends on the length of the investment horizon T only.[71] To facilitate the exposition, we directly specify the optimal pre-commitment strategy $\pi^{*,\text{pre}}_{T,p}$ by maximizing the expected utility over a constant (regime independent) investment fraction π, i.e.

$$\pi^{*,\text{pre}}_{T,p} := \text{argmax}_\pi EU_{T,p} = \text{argmax}_\pi y_{T,p}(\pi),$$

where $EU_{T,p}$ is given in Eqn. (4.6) and $y_{T,p}$ is given in Eqn. (4.8).

For the interpretation, we give the optimal pre-commitment strategy in terms of the regime dependent Merton solutions. In analogy to Eqn. (4.1), we define π^{Mer}_i ($i \in \{1, 2\}$) by

$$\pi^{\text{Mer}}_i = \frac{\mu_i - r}{\gamma \sigma_i^2}.$$

Since the highest possible savings rate in Regime i is obtained by π^{Mer}_i, it follows that[72]

$$\pi^{*,\text{pre}}_{T,p} \in \left[\min\{\pi^{\text{Mer}}_1, \pi^{\text{Mer}}_2\}, \max\{\pi^{\text{Mer}}_1, \pi^{\text{Mer}}_2\}\right] =: \mathcal{A}.$$

We can therefore write the optimal pre-commitment strategy $\pi^{*,\text{pre}}_{T,p}$ as a weighted average of the regime dependent Merton solutions, i.e.

$$\pi^{*,\text{pre}}_{T,p} := \alpha^*_{T,p} \pi^{Mer}_1 + (1 - \alpha^*_{T,p}) \pi^{Mer}_2, \tag{4.9}$$

where $\alpha^*_{T,p}$ gives the optimal weight of the Merton solution for Regime 1. We stress the impact of the investment horizon and the regime probabilities by the notation $\alpha^*_{T,p}$. In addition, $\alpha^*_{T,p}$ may depend on all model and preference parameters. In the following proposition, we give the implicit function for $\alpha^*_{T,p}$ which involves $\pi^{*,\text{pre}}_{T,p}$ and which follows from the first order condition for the optimal pre-commitment strategy.[73]

Proposition 4.2 (Optimal pre-commitment strategy)
Along the lines of Eqn. (4.9), the optimal pre-commitment strategy $\pi^{*,pre}_{T,p}$ is given by

[71] For a proof see Balter et al. (2021).
[72] Notice that the assumption $\lambda_1 \geq \lambda_2$ does not necessarily imply $\pi^{\text{Mer}}_1 \geq \pi^{\text{Mer}}_2$.
[73] The result can easily be generalized to more than two regimes.

the weighting factor $\alpha^*_{T,p}$ *where*

$$\alpha^*_{T,p} = \frac{p\sigma_1^2 f_1(\pi^{*,pre}_{T,p}, T)}{p\sigma_1^2 f_1(\pi^{*,pre}_{T,p}, T) + (1-p)\sigma_2^2 f_2(\pi^{*,pre}_{T,p}, T)}$$
$$\textit{and } f_i(\pi, T) = e^{y(\pi,i)(1-\gamma)T}, \; i = 1, 2.$$

The proof of Proposition 4.2 is given in Appendix D.2.

Different from the optimal Merton strategy, the optimal pre-commitment strategy depends on T for $\gamma \neq 1$ and is thus time-inconsistent: If the investor conducts the strategy herself, she will regret her decision as time moves by and the remaining investment horizon becomes shorter.

The time-inconsistency of the optimal pre-commitment strategy can be traced back to the non-myopic behavior of the investor. For $\gamma > 1$, she chooses the optimal investment over a *first period* with continuation utility over some *second period* in view. In line with that intuition, the time dependence of the pre-commitment strategy is driven by the savings rates in the two regimes. The optimal weight thus also depends on the regime-specific savings rates captured in the functions f_i.

We first look at the limiting values for the optimal pre-commitment strategy:

Proposition 4.3 (Limiting values of optimal pre-commitment strategy) *The limiting values for the optimal pre-commitment strategy are given by*

$$\lim_{T\to 0} \alpha^*_{T,p} = \frac{p\sigma_1^2}{p\sigma_1^2 + (1-p)\sigma_2^2}$$

and

$$\lim_{T\to\infty} \pi^{*,pre}_{T,p} = \begin{cases} \pi_1^{Mer} & y(\pi,1) < y(\pi,2) \; \forall \pi \in \mathcal{A} \\ \pi_2^{Mer} & y(\pi,2) < y(\pi,1) \; \forall \pi \in \mathcal{A} \\ \pi^{equal} & \textit{otherwise} \end{cases}, \qquad (4.10)$$

where π^{equal} *is defined in Equation* (4.5)*.*

The limiting case $T \to 0$ gives the myopic investment decision. The optimal weights of the Merton solutions are the same for all levels of risk aversion γ. With myopia, there is no more dependence on continuation utilities, but the weights are driven by the probabilities $(p, 1-p)$ and the volatilities σ_i respectively their difference only.

To get the intuition for the functional form of $\alpha^*_{0,p}$, note that the savings rate for $T \to 0$ coincides with the expected savings rate $p\, y(\pi, 1) + (1-p)y(\pi, 2)$. With

$$y(\pi, i) = y(\pi_i^{\mathrm{Mer}}, i) - l(\pi, i),$$

Benchmark parameter

μ_1	μ_2	σ_1	σ_2	r
0.03	0.01	0.10	0.20	0.00

Table 4.3: Benchmark parameter constellation.

the optimal pre-commitment strategy does not only maximize the expected savings rate, but also minimizes the expected loss rate. The loss rate in regime i is given by Eq. (4.3). It depends on the squared difference between the strategy π and the optimal Merton strategy and scales with the squared volatility σ_i^2. The deviation between the strategy π and the Merton-strategy π_i^{Mer} thus enters the expected loss rate with a factor that depends on the probability of the regime and the regime-dependent variance. Consequently, the weighting factor of the Merton-strategy π_i^{Mer} in the optimal pre-commitment strategy is proportional to the regime-probabilities and variances, too. This is also shown in the left graph of Figure 4.1, which plots the limiting $\alpha_{0,p}^*$ as a function of $\sigma_2 - \sigma_1$ and for different values of p. The benchmark parameter constellation in terms of the regime parameters μ_i and σ_i $(i = 1, 2)$ is given in Table 4.3. The graph confirms that the weight for regime 1 is increasing in the probability of this regime and is decreasing in the volatility σ_2 in the other regime.

The other limiting case is given by $T \to \infty$. For an infinite investment horizon, the savings rate of a strategy π is given by the lower of the regime-dependent savings rates (cf. Proposition 4.1). The optimal pre-commitment strategy is thus given by the strategy $\pi \in \mathcal{A}$ which maximizes this worst-case savings rate. The worst-case regime on the interval $\mathcal{A}$ can always be Regime 1, always be Regime 2, or switch from one regime to the other. Accordingly, the optimal limiting strategy which achieves the maximal worst-case utility is the Merton strategy for Regime 1, the Merton strategy for Regime 2, or π_{equal} (see Eq. (4.10)).

The limiting optimal pre-commitment strategy for $T \to \infty$ does not depend on the probabilities as long as both regimes have a positive probability. The reason is that it coincides with the worst-case strategy, which depends on parameters within the regimes but not on the probability of the regimes.

Next, we look at the behavior of the weighting factors for general T. A special case is given by $\gamma = 1$ for which $f_i(\pi, T) \equiv 1$. The weighting factor then simplifies to

$$\alpha_{T,p}^* = \frac{p\sigma_1^2}{p\sigma_1^2 + (1-p)\sigma_2^2}.$$

For the myopic investor, the optimal strategy is time-consistent and coincides with the optimal myopic strategy for $T \to 0$. The reason is that the savings rate is equal to the expected savings rate not just in the limit but for all T. An easy intuition why the assumption of log-utility avoids problems stemming from time-inconsistency is that log utility implies a myopic behavior. The investor thus always chooses the strategy that is optimal over the next instant and neither takes the remaining investment horizon nor the continuation utility into account.[74]

For $\gamma > 1$, the weighting factor of a regime i does not only depend on its probability and its volatility but also on a discount function f_i. This function is the smaller, the larger the savings rate in a regime is, and thus downplays the weight of the "good" regime so that the optimal strategy is shifted towards the worst-case strategy. The impact of the savings rate is the larger, the longer the investment horizon and the higher the risk aversion is. In addition, it is the larger, the more the savings rates in the two regimes differ from each other.

To aggregate the impact of the discount functions f_1 and f_2, we define the function δ as

$$\delta(\pi, T) := 1 - \frac{f_1(\pi, T)}{f_2(\pi, T)} = 1 - e^{[y(\pi,1)-y(\pi,2)](1-\gamma)T}.$$

It describes the relative difference between expected utilities in the two regimes for a given strategy π. For $y(\pi, 1) > y(\pi, 2)$ and $\gamma > 1$, we have $\delta \in [0, 1)$. The lower limit of $\delta = 0$ is attained for the limiting case $T \to 0$ while the upper limit of one is approached for $T \to \infty$. For $y(\pi, 1) = y(\pi, 2)$ and thus $\pi = \pi^{\text{equal}}$, δ is identically equal to zero.

With this definition of δ, we can rewrite the portfolio weight of the optimal pre-commitment strategy as

$$\alpha^*_{T,p}(\pi^{*,pre}_{T,p}) = \frac{p\sigma_1^2(1 - \delta(\pi^{*,pre}_{T,p}, T))}{p\sigma_1^2(1 - \delta(\pi^{*,pre}_{T,p}, T)) + (1-p)\sigma_2^2}. \tag{4.11}$$

For $T = 0$ and thus $\delta = 0$, the optimal strategy is the myopic one. For $T > 0$, the strategy becomes time-inconsistent if $\gamma \neq 1$ and $y(\pi^{*,pre}_{T,p}, 1) \neq y(\pi^{*,pre}_{T,p}, 2)$. The degree of time-inconsistency is captured by δ which ultimately goes to one. The portfolio weight α then approaches zero, and the pre-commitment strategy approaches the worst-case strategy π_2^{Mer}. The special case $y(\pi^{*,pre}_{T,p}, 1) = y(\pi^{*,pre}_{T,p}, 2)$ holds for $\pi^{*,pre}_{T,p} = \pi^{equal}$. With a zero difference between the savings rates, δ is identically equal to zero. The weight α is then time-independent, and the resulting strategy is time-consistent.

For $y(\pi, 1) > y(\pi, 2)$ and $\gamma > 1$, δ increases in the difference of the savings rates given in Eqn. (4.4). This difference in turn increases in the difference of the volatilities

[74] One could thus avoid time-inconsistency of the optimal strategy by setting the objective function equal to the expected savings rates, which is equivalent to maximizing the expected utility of log wealth.

$\alpha^*_{0,p}$ for different p and $\alpha^*_{T,p}$ for different levels of risk aversion γ

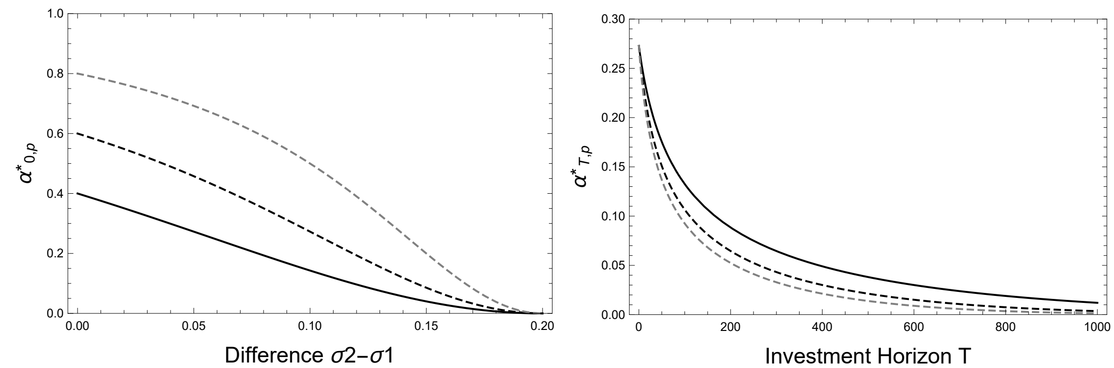


Figure 4.1: The left picture displays the optimal weight on the Merton solution in Regime 1 at $T = 0$ depending on the differences $\sigma_2 - \sigma_1$ (where $\sigma_2 = 0.2$). The black line pictures $p = 0.4$, the black dashed $p = 0.6$ and the gray dashed $p = 0.8$. The right picture shows the optimal weight on the Merton solution in Regime 1 $\alpha^*_{T,p}$ depending on the investment horizon T. The black line pictures $\gamma = 2$, the black dashed $\gamma = 4$ and the gray dashed $\gamma = 16$.

and the difference of the portfolio weights from π^{equal}. Intuitively, the *force* to the worst-case regime matters more, the higher the difference of good and bad regime is, which is illustrated in the right hand side of Figure 4.1.

Furthermore, δ increases in the probability p of the good regime and in the risk aversion γ. These relations are illustrated in Figure 4.2 and Figure 4.3. Intuitively, risk aversion causes the convergence of the optimal strategy towards the worst-case strategy, which is reached for the limiting value $\delta = 1$. The larger the risk aversion, the faster this convergence takes place, and thus the larger the corresponding $\delta(\pi^{*,pre}_{T,p}, T)$. To get the intuition for the impact of p, note that time-inconsistency shifts the importance from the good regime towards the bad regime over time. This shift is the more severe, the higher the myopic importance of the good regime is (i.e. p).

We can also use δ to link the time-inconsistence of the optimal pre-commitment strategy to the hedging needs of the investor. As pointed out above, the reason for time-inconsistency is that the investor takes the difference between the continuation utilities into account. For $y(\pi, 1) > y(\pi, 2)$ and thus $\delta \in [0, 1)$, regime 2 is the worse one. By moving the strategy towards the worst-case strategy, the investor foregoes some returns to lower the variance and in particular to increase the level of wealth in the bad state. The strength of the hedging demand increases in the risk aversion, the planning horizon and the difference between the regimes. In line with this intuition, δ increases in γ, T, and the difference between the regimes.

Impact of investment horizon T on time-inconsistency

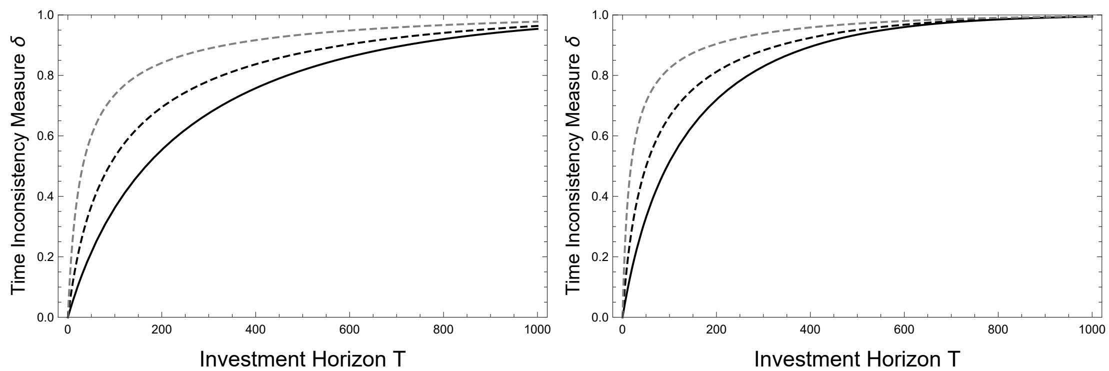


Figure 4.2: The left (right) figure refers to a level of risk aversion $\gamma = 2$ ($\gamma = 8$). The black graph refers to the optimal pre-commitment strategy $\pi_{T,p}^{*,\mathrm{pre}}$ for $p = 0.2$. The dashed black (dashed gray) graph refers to the optimal pre-commitment strategy for $p = 0.5$ ($p = 0.8$).

$\pi_{T,p}^{*,pre}$ **for different investment horizons T depending on probability** p

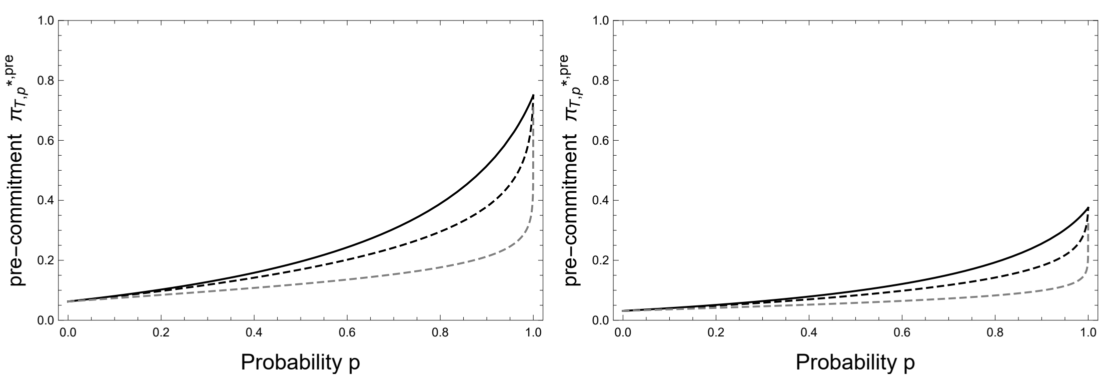


Figure 4.3: The left figure refers to $\gamma = 4$ ($\pi_1^{Mer} = 0.75, \pi_2^{Mer} = 0.0625$), the right figure to $\gamma = 8$ ($\pi_1^{Mer} = 0.375, \pi_2^{Mer} = 0.03125$). The black graphs picture $T = 2$, the black dashed $T = 20$ and the gray dashed $T = 100$.

Notice that δ is not only related to the weights of the optimal pre-commitment strategy, but can also be used to write the savings rate in terms of the time-inconsistency

$$y_{T,p}(\pi) = y(\pi, 2) + \frac{1}{(1-\gamma)T} \ln\left[1 - p\,\delta(\pi, T)\right]. \tag{4.12}$$

If Regime 2 is the bad one, the second term is non-negative and measures the additional contribution of the good state to the savings rate. This contribution depends on p as well as on δ. It increases in the probability p of the good regime. It also increases in δ, which in turn increases in the difference between the savings rates in the good and bad regime.

For the optimal pre-commitment strategy, the savings rate is

$$y_{T,p}(\pi^{*,pre}_{T,p}) = y(\pi^{*,pre}_{T,p}, 2) + \frac{1}{(1-\gamma)T} \ln\left[1 - p\delta(\pi^{*,pre}_{T,p}, T)\right]. \tag{4.13}$$

For $\pi^{*,\text{pre}}_{T,p} = \pi^{\text{equal}}$, the savings rates in the two states coincide. δ is then equal to zero and the optimal savings rate is equal to the (identical) savings rates in the two states. Otherwise, the second term is positive. The savings rate then increases in the probability of Regime 1 and in the utility gain in Regime 1 relative to Regime 2.

Along the lines of Proposition 4.1, myopia (which we see for $\gamma = 1$ and $T = 0$) implies that the investor acts risk neutral w.r.t. the regime dependent savings rates, i.e. the savings rate in case of risk over the regimes is equal to the expected savings rate. An infinite investment horizon implies a worst-case decision, i.e. the decision is based on the worst-case savings rate.

In our stylized setup, the optimal pre-commitment strategy converges toward the worst-case strategy and is time-inconsistent. The effect of time-inconsistency can be captured by the function δ which depends on the difference between the savings rates in the two regimes and which drives the optimal weighting factor α. If $\pi_1^{\text{Mer}} > \pi_2^{\text{Mer}}$ and $\sigma_1 \leq \sigma_2$, the weighting factors for the Merton strategies in the myopic case are proportional to the regime probability and the variance of the regime, so that a higher risk in a regime biases the optimal pre-commitment strategy towards the Merton strategy for this regime. Moreover, the weight of the good regime decreases in the investment horizon, i.e. the investor puts more and more weight on the worst-case strategy. The effect is more pronounced, the higher the level of risk aversion γ is.

4.2.4 Value of information

We now compare the utility of the optimal pre-commitment strategy to the one of the optimal strategy under full information if the investor can condition her strategy on the regime. Thus, the optimal strategy π^* under full information maximizes expected utility when the outcome of the a priori lottery is known, i.e.

$$\begin{aligned}\pi^* = (\pi_1^*, \pi_2^*) &:= \underset{(\pi_1,\pi_2)}{\operatorname{argmax}} \{p\mathbb{E}_{P_1}[u(V_T(\pi_1))] + (1-p)\mathbb{E}_{P_2}[u(V_T(\pi_2))]\} \\ &= \underset{\pi_1}{\operatorname{argmax}}\; p\mathbb{E}_{P_1}[u(V_T(\pi_1))] + \underset{\pi_2}{\operatorname{argmax}}\; (1-p)\mathbb{E}_{P_2}[u(V_T(\pi_2))].\end{aligned}$$

The second line follows from the fact that the two terms in the weighted sum depend on either π_1 or π_2. It immediately follows with the Merton result:

Proposition 4.4 (Optimal strategy under full information)
In case of an initial lottery L over the regimes 1 and 2, the expected utility and the expected savings rate of a CRRA investor who can condition the strategy on the true regime are maximized for

$$\pi^* = (\pi_1^{Mer}, \pi_2^{Mer}). \tag{4.14}$$

The maximal savings rate $y_{T,p}(\pi^*)$ when we can condition on the regimes is

$$y_{T,p}(\pi^*) = \begin{cases} \frac{1}{(1-\gamma)T} \ln \left[pe^{y(\pi_1^{\text{Mer}},1)(1-\gamma)T} + (1-p)e^{y(\pi_2^{\text{Mer}},2)(1-\gamma)T}\right] & \gamma \neq 1 \\ p\, y(\pi_1^{\text{Mer}},1) + (1-p)y(\pi_2^{\text{Mer}},2) & \gamma = 1 \end{cases},$$

and the maximal certainty equivalent CE_T^* is

$$CE_T^* = \begin{cases} \left[pe^{y(\pi_1^{\text{Mer}},1)(1-\gamma)T} + (1-p)e^{y(\pi_2^{\text{Mer}},2)(1-\gamma)T}\right]^{\frac{1}{1-\gamma}} & \gamma \neq 1 \\ e^{\left[p\, y(\pi_1^{\text{Mer}},1)+(1-p)y(\pi_2^{\text{Mer}},2)\right]T} & \gamma = 1 \end{cases}.$$

The value of the regime information can be measured by the quotient of the certainty equivalents associated with the optimal strategies with and without the regime information, i.e. by the ratio $CE_{T,p}^{*,\text{pre}}/CE_T^*$ of the certainty equivalents. This ratio gives the percentage of wealth the investor foregoes if she does not learn about the regime immediately after the initial lottery has taken place. Alternatively, one can measure the value of information in terms of the difference in the savings rates $y_{T,p}(\pi^*) - y_{T,p}(\pi_{T,p}^{*,\text{pre}})$, i.e. by the annual rate of return that the investor foregoes.

For $\gamma = 1$, the difference of the savings rate is simply given by the expected loss rate in the two regimes, i.e. it holds

$$y_{T,p}(\pi^*) - y_{T,p}(\pi^{*,\text{pre}}) = p\, l(\pi_{T,p}^{*,\text{pre}},1) + (1-p)\, l(\pi_{T,p}^{*,\text{pre}},2),$$

where the loss rate is given in Equation (4.3). For $\gamma > 1$, it holds that

$$y_{T,p}(\pi^*) - y_{T,p}(\pi_{T,p}^{*,\text{pre}}) = \frac{1}{(1-\gamma)T} \ln \left[\frac{pe^{y(\pi_1^{\text{Mer}},1)(1-\gamma)T} + (1-p)e^{y(\pi_2^{\text{Mer}},2)(1-\gamma)T}}{pe^{y(\pi_{T,p}^{*,pre},1)(1-\gamma)T} + (1-p)e^{y(\pi_{T,p}^{*,pre},2)(1-\gamma)T}}\right].$$

It follows immediately

Proposition 4.5 (Value of information)

(i) The difference of the savings rates is given by

$$y_{T,p}(\pi^*) - y_{T,p}(\pi^{*,pre}_{T,p}) = \begin{cases} \beta_{T,p}(1)\, l(\pi^{*,pre}_{T,p},1) + (1-\beta_{T,p}(1))\, l(\pi^{*,pre}_{T,p},2) & \gamma = 1 \\ \frac{1}{(1-\gamma)T} \ln\left[\beta_{T,p}(\gamma) e^{l(\pi^{*,pre}_{T,p},1)(1-\gamma)T} + (1-\beta_{T,p}(\gamma))\, e^{l(\pi^{*,pre}_{T,p},2)(1-\gamma)T}\right] & \gamma \neq 1 \end{cases},$$

where

$$\beta_{T,p}(\gamma) := \frac{p e^{y(\pi^{*,pre}_{T,p},1)(1-\gamma)T}}{p e^{y(\pi^{*,pre}_{T,p},1)(1-\gamma)T} + (1-p) e^{y(\pi^{*,pre}_{T,p},2)(1-\gamma)T}} = \frac{p(1-\delta(\pi^{*,pre}_{T,p},T))}{p(1-\delta(\pi^{*,pre}_{T,p},T)) + 1 - p}.$$

(ii) The ratio of the certainty equivalents is given by

$$\frac{CE^*_{T,p}}{CE_{T,p}(\pi^{*,pre}_{T,p})} = e^{\left[y_{T,p}(\pi^*) - y_{T,p}(\pi^{*,pre}_{T,p})\right]T}$$
$$= \begin{cases} e^{\left[\beta_{T,p}(1)\, l(\pi^{*,pre}_{T,p},1) + (1-\beta_{T,p}(1))\, l(\pi^{*,pre}_{T,p},2)\right]T} & \gamma = 1 \\ \left[\beta_{T,p}(\gamma) e^{l(\pi^{*,pre}_{T,p},1)(1-\gamma)T} + (1-\beta_{T,p}(\gamma))\, e^{l(\pi^{*,pre}_{T,p},2)(1-\gamma)T}\right]^{\frac{1}{1-\gamma}} & \gamma \neq 1 \end{cases}.$$

The proof of Proposition 4.5 is given in Appendix D.3. The loss in the savings rate and in the certainty equivalent is thus equal to some weighted average of the regime-specific loss rates, where the weights depend on the certainty equivalents in the two regimes.

Intuitively, we expect the value of information to increase with the planning horizon, i.e. with the length of the time horizon over which it is relevant. However, the savings rates converge towards the worst-case savings rates when the investment horizon goes to infinity, which lowers the value of information (the difference of the savings rates) about the true regime for an increasing investment horizon. By the same argument the difference of the certainty equivalents first increases and then decreases in T.

The limiting behavior of the gains from information is summarized in the following proposition:

Proposition 4.6 (Limits for value of information)

(i) For the limits of the difference of the savings rates it holds

$$\lim_{T\to 0}\left[y_{T,p}(\pi^*) - y_{T,p}(\pi^{*,pre}_{T,p})\right] = \frac{1}{2}\gamma p(1-p)\left(\pi_1^{Mer} - \pi_2^{Mer}\right)^2 \frac{\sigma_1^2\sigma_2^2}{p\sigma_1^2 + (1-p)\sigma_2^2}.$$
$$\lim_{T\to\infty}\left[y_{T,p}(\pi^*) - y_{T,p}(\pi^{*,pre}_{T,p})\right] = \begin{cases} 0 & \pi^{equal} \notin \mathcal{A} \\ y(\pi_2^{Mer},2) - y(\pi^{equal},\cdot) & \pi^{equal} \in \mathcal{A}. \end{cases}$$

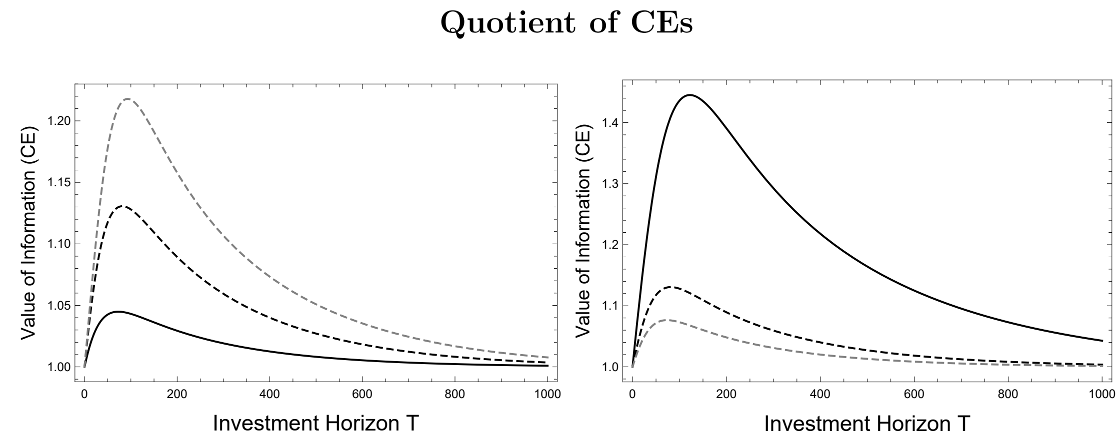


Figure 4.4: The left figure refers to $p = 0.2$ (black), $p = 0.5$ (black dashed), and $p = 0.7$ (gray dashed) where $\gamma = 4$. The right figure to $\gamma = 2$ (black), $\gamma = 4$ (black dashed), $\gamma = 6$ (gray dashed) where $p = 0.5$.

(ii) For the limits of the ratio of the certainty equivalents it holds

$$\lim_{T\to 0} \frac{CE^*_{T,p}}{CE_{T,p}(\pi^{*,pre}_{T,p})} = 1$$

$$\lim_{T\to \infty} \frac{CE^*_{T,p}}{CE_{T,p}(\pi^{*,pre}_{T,p})} = \begin{cases} 1 & \pi^{equal} \notin \mathcal{A} \\ \infty & \pi^{equal} \in \mathcal{A}. \end{cases}$$

The proof of Proposition 4.6 is given in Appendix D.4.

Figure 4.4 gives the ratio of the certainty equivalents as a function of T. For $T \to 0$, both CEs converge to the initial investment. Their ratio thus goes to one. In the case $T \to \infty$, both CEs converge to infinity. For the parameters in Table 4.3, the worst-case strategy converges towards the Merton strategy in the worse regime, and the ratio of the CEs goes to one. There is thus an investment horizon $\hat{T}$ for which the value of information obtains its maximum. As the figure shows, this $\hat{T}$ increases in p and decreases in γ.

The right graph furthermore shows that the value of information is decreasing in relative risk aversion γ. To get the intuition, note that the portfolio weights are proportional to $1/\gamma$. A higher risk aversion thus induces the investor to take smaller portfolio weights, which limits the savings rates, i.e. the gains from investing. This consequently also lowers the difference between the savings rates, i.e. the value of information.

Figure 4.5 shows the difference in the savings rates as a function of the probability p. In the limiting cases $p = 0$ and $p = 1$, there is no difference between the cases with and without full information, so that the difference of the savings rates is

$y_{T,p}(\pi^*) - y_{T,p}(\pi^{*,pre})$ **for varying** p **with different investment horizons** T

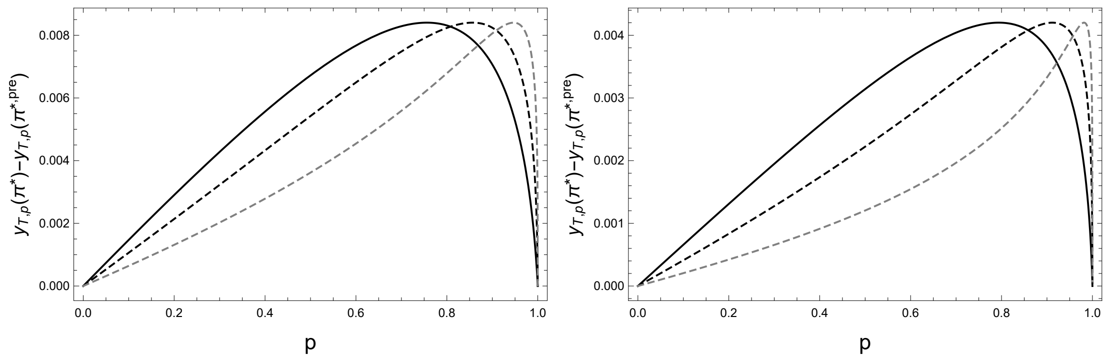


Figure 4.5: The left figure refers to a risk aversion of $\gamma = 2$, the right to $\gamma = 4$.The black lines picture $T = 20$, the black dashed $T = 50$ and the gray dashed $T = 100$.

equal to zero. For intermediate values of p, the difference is first increasing and then decreasing in p. To get the intuition, we write the difference of the savings rates as

$$y_{T,p}(\pi^*) - y_{T,p}(\pi^{*,\text{pre}}) = y_{T,p}(\pi^*) - \left[y(\pi^{*,pre}_{T,p}, 2) - \frac{1}{(1-\gamma)T} \ln \left[1 - p\delta(\pi^{*,pre}_{T,p}, T)\right] \right].$$

The first term, the savings rate in the full information case, is increasing in the probability p of the good regime. The same holds for the second term, the savings rate in the limited information case. A larger p moves the optimal pre-commitment strategy away from the worst-case strategy, which lowers its savings rate in regime 2 but increases the gain from regime 1. While the increase of the savings rate in the full information case dominates for small p, the opposite holds for large p. When T approaches zero, the value of information for $T \to 0$ is largest for $p = \sigma_2/(\sigma_1 + \sigma_2)$. For equal volatilities, this simplifies to $p = 0.5$ for which uncertainty about the true regime is largest. The dependence on p vanishes for $T \to \infty$ when the savings rates are determined by the worst-case values.

4.2.5 A priori lottery and ambiguity

We now discuss the impact of ambiguity on time-inconsistency and the value of information. We use the smooth ambiguity approach of Klibanoff et al. (2005) to model the investor's ambiguity aversion, i.e. the impact of uncertainty on the probability p of the a priori lottery on the investor's preferences.

The investor's time $t = 0$ certainty equivalent to receiving V_T is given by

$$v^{-1}\left(E_p\left[v\left(u^{-1}\left(EU_{T,p}\right)\right)\right]\right) \tag{4.15}$$

for two increasing utility and ambiguity functions u and v. The corresponding savings rate is given by

$$y_T^{\text{amb}}(\pi) = \frac{1}{T} \ln \left(v^{-1} \left(E_p \left[v \left(u^{-1} \left(EU_{T,p} \right) \right) \right] \right) \right).$$

For $v = u$, the investor is ambiguity neutral and we are back in a situation with risk.

The optimal pre-commitment strategy under ambiguity $\pi^{*,\text{pre, amb}}$ is defined by

$$\begin{aligned} \pi_{T,p}^{*,\text{pre, amb}} &:= \text{argmax}_\pi \; E_p \left[v \left(u^{-1} \left(EU_{T,p} \right) \right) \right] \\ &= \text{argmax}_\pi \; y_T^{\text{amb}}(\pi). \end{aligned}$$

Throughout the following, we assume that both u and v are CRRA functions, i.e.

$$u(x) = \begin{cases} \frac{x^{1-\gamma}}{1-\gamma} & \gamma > 1 \\ \ln x & \gamma = 1 \end{cases} \quad \text{and } v(x) = \begin{cases} \frac{x^{1-\eta}}{1-\eta} & \eta > 1 \\ \ln x & \eta = 1 \end{cases},$$

where γ and η capture the (constant) relative aversions towards risk and ambiguity.

For the sake of simplicity, we model ambiguity by a situation with two different probability distributions $(p_a, 1 - p_a)$ and $(p_b, 1 - p_b)$ over the regimes 1 and 2. The investor assigns the probabilities $\tilde{p}$ and $1 - \tilde{p}$ to these two distributions. W.l.o.g., we assume $p_b \leq p_a$. Thus, without ambiguity aversion, we are back in a decision problem under risk with a lottery $(q, 1 - q)$ where

$$q := \tilde{p} p_a + (1 - \tilde{p}) p_b. \tag{4.16}$$

For a given portfolio weight π, the expected utility of the investor is

$$E_p \left[v \left(u^{-1} \left(EU_{T,p} \right) \right) \right] = \frac{1}{1-\eta} \left[\tilde{p} e^{y_{T,p_a}(\pi)(1-\eta)T} + (1 - \tilde{p}) e^{y_{T,p_b}(\pi)(1-\eta)T} \right],$$

where $y_{T,p_a}(\pi)$ is the savings rate in a risk situation described by the distribution $(p_a, 1 - p_a)$ and $y_{T,p_b}(\pi)$ is the savings rate in a risk situation described by the distribution $(p_b, 1 - p_b)$. Note that the aggregation over the two probability distributions in case of ambiguity has the same functional form as the aggregation over the two regimes in case of risk.

The savings rate in case of ambiguity is

$$y_{T,\tilde{p}}^{\text{amb}}(\pi) := \frac{1}{(1-\eta)T} \ln \left[\tilde{p} e^{y_{T,p_a}(\pi)(1-\eta)T} + (1 - \tilde{p}) e^{y_{T,p_b}(\pi)(1-\eta)T} \right].$$

Plugging in the corresponding formulas for the savings rates in case of risk over the regimes gives

$$y^{\text{amb}}_{T,\tilde{p}}(\pi) = \begin{cases} \frac{1}{(1-\eta)T} \ln \Big[\tilde{p} \left[p_a e^{y(\pi,1)(1-\gamma)T} + (1-p_a) e^{y(\pi,2)(1-\gamma)T}\right]^{\frac{1-\eta}{1-\gamma}} \\ \quad +(1-\tilde{p}) \left[p_b e^{y(\pi,1)(1-\gamma)T} + (1-p_b) e^{y(\pi,2)(1-\gamma)T}\right]^{\frac{1-\eta}{1-\gamma}}\Big] & \gamma \neq 1 \\ \frac{1}{(1-\eta)T} \ln \left[\tilde{p} e^{\bar{y}_{p_a}(\pi)(1-\eta)T} + (1-\tilde{p}) e^{\bar{y}_{p_b}(\pi)(1-\eta)T}\right] & \gamma = 1 \end{cases}.$$

For the limiting case $T \to 0$, it holds that

$$\begin{aligned} \lim_{T\to 0} y^{\text{amb}}_{T,\tilde{p}}(\pi) &= \tilde{p} \lim_{T\to 0} y_{T,p_a}(\pi) + (1-\tilde{p}) \lim_{T\to 0} y_{T,p_b}(\pi) \\ &= \tilde{p} \left(p_a y(\pi,1) + (1-p_a) y(\pi,2)\right) + (1-\tilde{p}) \left(p_b y(\pi,1) + (1-p_b) y(\pi,2)\right) \\ &= q\, y(\pi,1) + (1-q)\, y(\pi,2), \end{aligned}$$

and for the limiting case $T \to \infty$, it holds that

$$\lim_{T\to\infty} y^{\text{amb}}_{T,\tilde{p}}(\pi) = \min\{y(\pi,1), y(\pi,2)\}.$$

In both limiting cases the savings rate no longer depends on the risk aversion over the regimes and the ambiguity aversion over the lotteries.

For $T > 0$ and $\eta > \gamma$, Jensen's inequality[75] implies that

$$\begin{aligned} y^{\text{amb}}_{T,\tilde{p}}(\pi) &< \tilde{p}\, y_{T,p_a}(\pi) + (1-\tilde{p}) y_{T,p_b}(\pi) \\ &\leq \tilde{p}\, \left(p_a y(\pi,1) + (1-p_a) y(\pi,2)\right) + (1-\tilde{p})\, \left(p_b y(\pi,1) + (1-p_b) y(\pi,2)\right) \\ &= q\, y(\pi,1) + (1-q)\, y(\pi,2). \end{aligned}$$

Similar to risk aversion, ambiguity aversion thus also reduces the savings rate relative to the expected savings rate under the lottery $(q, 1-q)$.

The optimal pre-commitment strategy of the ambiguity-averse investor is given by

$$\pi^{*,\text{pre,amb}}_{T,\tilde{p}} := \text{argmax}_\pi \frac{1}{(1-\eta)T} \ln \left[\tilde{p} e^{y_{T,p_a}(\pi)(1-\eta)T} + (1-\tilde{p}) e^{y_{T,p_b}(\pi)(1-\eta)T}\right].$$

The strategies that maximize the two savings rates in the above expression separately from each other are $\pi^{*,\text{pre}}_{T,p_a}$ and $\pi^{*,\text{pre}}_{T,p_b}$. Similar to the optimal pre-commitment strategy in case of risk (which is between the optimal strategies π_1^{Mer} and π_2^{Mer} in the two

[75] $EU = \frac{1}{1-\gamma} e^{y(\pi,i)(1-\gamma)T}$ is a concave function of the portfolio weight, thus the expectation $EU_{T,p}$ over the regimes is concave in the portfolio weights. Thus utility under ambiguity is for $\eta > \gamma$ concave in the portfolio weights.

Impact of ambiguity on pre-commitment strategy

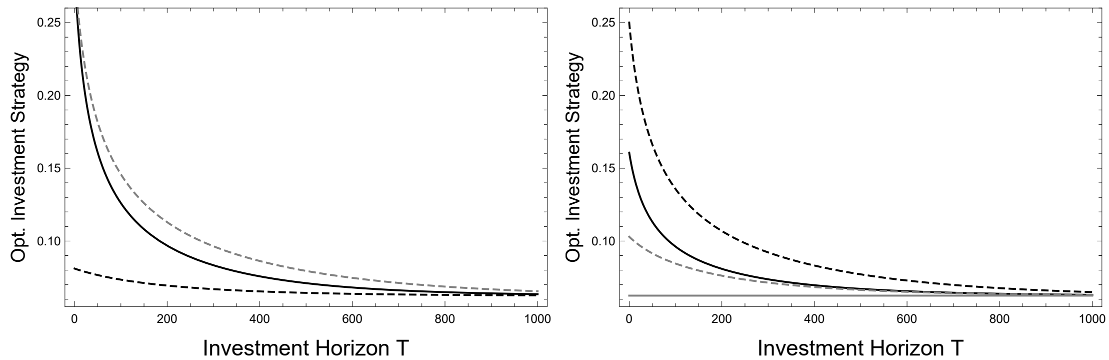


Figure 4.6: The gray line displays $\pi_2^{Mer} = 0.0625$. The black graph refers to the optimal pre-commitment strategy under ambiguity aversion $\pi_{T,\tilde{p}}^{*,\text{pre, amb}}$ with $\tilde{p} = 0.5$, $p_a = 0.6$ and $p_b = 0.2$. The dashed black (dashed gray) graph refers to the optimal pre-commitment strategy without ambiguity for $\gamma = 4$ under the given probability distribution over the regimes with $\tilde{p} = 0.6$ ($\tilde{p} = 0.2$). The left (right) figure refers to a level of ambiguity $\eta = 4.000001$ ($\eta = 16$).

regimes), the optimal pre-commitment strategy in case of ambiguity is between $\pi_{T,p_a}^{*,\text{pre}}$ and $\pi_{T,p_b}^{*,\text{pre}}$, i.e. it holds that

$$\pi_{T,\tilde{p}}^{*,\text{pre, amb}} \in \left[\min\left\{\pi_{T,p_a}^{*,\text{pre}}, \pi_{T,p_b}^{*,\text{pre}}\right\}, \max\left\{\pi_{T,p_a}^{*,\text{pre}}, \pi_{T,p_b}^{*,\text{pre}}\right\}\right] =: \mathcal{A}_T^{\text{amb}}.$$

For $T \to \infty$, the limiting value of the optimal pre-commitment strategy does no longer depend on the regime probability s.t.

$$\lim_{T\to\infty} \mathcal{A}_T^{\text{amb}} = \begin{cases} \{\pi_1^{\text{Mer}}\} & y(\pi,1) < y(\pi,2)\ \forall \pi \in \mathcal{A} \\ \{\pi_2^{\text{Mer}}\} & y(\pi,2) < y(\pi,1)\ \forall \pi \in \mathcal{A} \\ \{\pi^{\text{equal}}\} & \textit{otherwise} \end{cases}. \tag{4.17}$$

Again, we first look at the limiting values of the optimal pre-commitment strategies under ambiguity.

Proposition 4.7 (Limiting results ambiguity)
For $T \to 0$, it holds that

$$\lim_{T\to 0} \pi_T^{*,pre,\ amb} = \frac{q\sigma_1^2}{q\sigma_1^2 + (1-q)\sigma_2^2}\pi_1^{Mer} + \frac{(1-q)\sigma_2^2}{q\sigma_1^2 + (1-q)\sigma_2^2}\pi_2^{Mer} = \lim_{T\to 0} \pi_{T,q}^{*,pre}.$$

For $T \to \infty$, it holds that

$$\lim_{T\to\infty} \pi_T^{*,pre,\ amb} = \begin{cases} \pi_2^{Mer} & y(\pi,2) < y(\pi,1) \quad \forall \pi \in \mathcal{A} \\ \pi_1^{Mer} & y(\pi,1) < y(\pi,2) \quad \forall \pi \in \mathcal{A} \\ \pi^{equal} & \textit{otherwise} \end{cases}.$$

Impact of probability $\tilde{p}$ on optimal strategies $\pi^{*,\text{pre, amb}}_{T,\tilde{p}}$ and $\pi^{*,\text{pre}}_{T,q}$

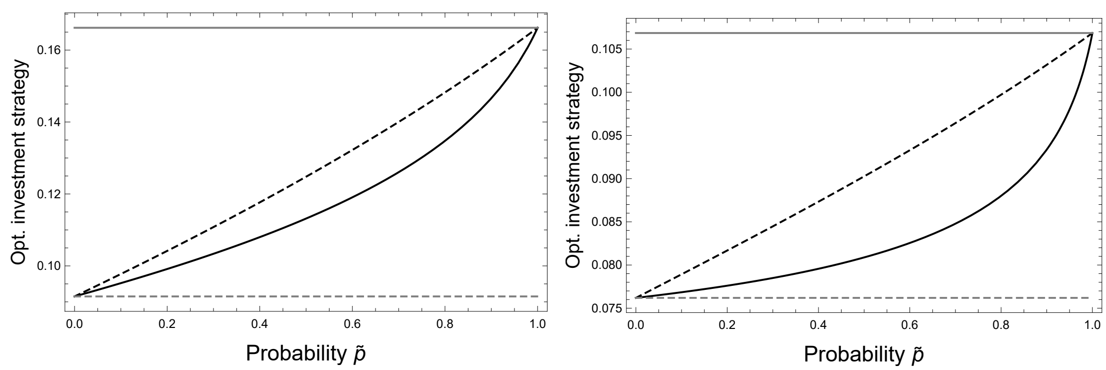


Figure 4.7: The pictures are created for $\gamma = 4$, $\eta = 16$, $p_a = 0.6, p_b = 0.2$ with an investment horizon of $T = 50$ (left figure) and $T = 200$ (right figure). The black graphs show the optimal pre-commitment strategy under risk and ambiguity $\pi^{*,\text{pre, amb}}_{T,\tilde{p}}$, the black dashed graphs show the optimal pre-commitment strategy under risk $\pi^{*,\text{pre}}_{T,q}$ with probability q, the gray lines the optimal pre-commitment strategy under risk with probability p_a, the gray dashed the optimal pre-commitment under risk with probability p_b.

For $T \to 0$, the optimal pre-commitment strategy under risk and ambiguity coincides with the optimal pre-commitment strategy under risk with probability distribution $(q, 1-q)$. For $T \to \infty$, we are again in the case that the investor maximizes the worst-case savings rate over the regimes, i.e. she uses the Merton strategy of the bad regime. In both cases the strategy depends on γ (via the Merton fraction) but not on the ambiguity aversion parameter η. An illustration is given in Figure 4.6.

In the limiting cases the optimal strategies coincide for all combinations $\tilde{p}$, p_a, p_b that imply the same q. We now compare the optimal strategy $\pi^{*,\text{pre, amb}}_{T,\tilde{p}}$ under risk and ambiguity with the optimal strategy $\pi^{*,\text{pre}}_{T,q}$ under risk. Notice that $\pi^{*,\text{pre}}_{T,q}$ is the optimal strategy of an ambiguity neutral investor. Thus, for an ambiguity-averse investor it holds $\pi^{*,\text{pre, amb}}_{T,\tilde{p}} \leq \pi^{*,\text{pre}}_{T,q}$, i.e.

$$\pi^{*,\text{pre, amb}}_{T,\tilde{p}} \in \left[\min\left\{\pi^{*,\text{pre}}_{T,p_a}, \pi^{*,\text{pre}}_{T,p_b}\right\}, \pi^{*,\text{pre}}_{T,q}\right].$$

In addition, recall that for $T < \infty$, the optimal pre-commitment strategy (weighting factor) is increasing in p. Thus, for $p_b < p_a$

$$\pi^{*,\text{pre, amb}}_{T,\tilde{p}} \in \left[\pi^{*,\text{pre}}_{T,p_b}, \pi^{*,\text{pre}}_{T,q}\right].$$

For given T and varying probability $\tilde{p}$, an illustration of $\pi^{*,\text{pre}}_{T,p_b}$ and $\pi^{*,\text{pre}}_{T,q}$ and thus the domain $\pi^{*,\text{pre, amb}}_{T,\tilde{p}}$ is presented in Figure 4.7. In analogy to the previous section, we to state the optimal pre-commitment strategy $\pi_{T,\tilde{p}}^{*,pre,amb}$ under ambiguity as a weighted average of the two regime dependent Merton fractions π_1^{Mer} and π_2^{Mer}.

Proposition 4.8 (Optimal Pre-Commitment Strategy under Ambiguity) *The optimal pre-commitment strategy $\pi_{T,\tilde{p}}^{*,pre,amb}$ under ambiguity aversion η and risk aversion γ ($\eta > \gamma$) solves the equation*

$$\begin{aligned}\pi = \tilde{\alpha}_{T,\tilde{p}}(\pi)\left(\alpha_{T,p_a}(\pi)\pi_1^{Mer} + (1-\alpha_{T,p_a}(\pi))\pi_2^{Mer}\right)\\ + (1-\tilde{\alpha}_{T,\tilde{p}}(\pi))\left(\alpha_{T,p_b}(\pi)\pi_1^{Mer} + (1-\alpha_{T,p_b}(\pi))\pi_2^{Mer}\right).\end{aligned} \tag{4.18}$$

The weight $\tilde{\alpha}_{T,\tilde{p}}(\pi)$ is given by

$$\begin{aligned}\alpha_{T,p_i}(\pi) &= \frac{p_i\sigma_1^2(1-\delta_T^{pre}(\pi))}{p_i\sigma_1^2(1-\delta_T^{pre}(\pi)) + (1-p_i)\sigma_2^2}, \text{ for } i = a, b\\ \tilde{\alpha}_{T,\tilde{p}}(\pi) &= \frac{\tilde{p}(1-\delta_T^{amb}(\pi))}{\tilde{p}(1-\delta_T^{amb}(\pi)) + (1-\tilde{p})},\\ \delta_T^{amb}(\pi) &= 1 - \frac{p_a\sigma_1^2(1-\delta_T^{pre}(\pi)) + (1-p_a)\sigma_2^2}{p_b\sigma_1^2(1-\delta_T^{pre}(\pi)) + (1-p_b)\sigma_2^2}\left[\frac{p_a(1-\delta_T^{pre}(\pi)) + (1-p_a)}{p_b(1-\delta_T^{pre}(\pi)) + (1-p_b)}\right]^{\frac{\eta-\gamma}{\gamma-1}},\end{aligned}$$

the weights $\alpha_{T,p_a}(\pi)$ and $\alpha_{T,p_b}(\pi)$ are given in Equation (4.11) applied to π (instead of the optimal pre-commitment strategy).

The proof of Proposition 4.8 is given in Appendix D.5. Notice that in the special case $p_a = p_b$ there is no (third dimension) ambiguity since the lotteries $(p_a, 1-p_a)$ and $(p_b, 1-p_b)$ coincide. It is intuitively clear that the optimality condition simplifies to the optimal pre-commitment strategy under the lottery $(p_a, 1-p_a) = (p_b, 1-p_b)$

$$\pi_{T,\tilde{p}}^{*,pre,amb} = \pi_{T,q}^{*,pre} (= \pi_{T,p_a}^{*,pre} = \pi_{T,p_b}^{*,pre}).$$

Now, consider the case where the second risk dimension vanishes, i.e. $p_a = 1, p_b = 0$. Here Proposition 4.8 simplifies to

$$\begin{aligned}\pi_{T,\tilde{p}}^{*,pre,amb} &= \hat{\alpha}_{T,\tilde{p}}^{*}\pi_1^{Mer} + (1-\hat{\alpha}_{T,\tilde{p}}^{*})\pi_2^{Mer}\\ \text{where } \hat{\alpha}_{T,\tilde{p}}^{*}(\pi) &= \frac{\tilde{p}\sigma_1^2 e^{y(\pi,1)(1-\eta)T}}{\tilde{p}\sigma_1^2 e^{y(\pi,1)(1-\eta)T} + (1-\tilde{p})\sigma_2^2 e^{y(\pi,2)(1-\eta)T}}\end{aligned}$$

In consequence, for a given q (i.e. the probability of Regime 1 if the lotteries are merged), the ambiguity aversion is the more important, the higher the difference of p_a and p_b is (equality holds for $p_a = p_b$). In contrast, this decreases the risk situation (i.e. if $p_b = 0$ and $p_a = 1$, we have no risk situation about the regime). Thus, there is an opposing effect of ambiguity situation and risk situation about the regimes.

$\pi^{*,pre,amb}_{T,\tilde{p}}$, const. q for different p_a and p_b combinations

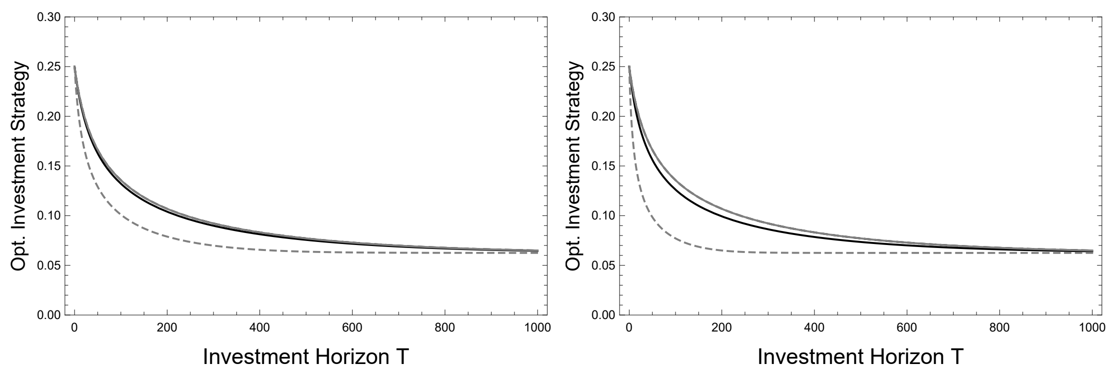


Figure 4.8: The left (right) hand side refers to $\gamma = 4, \eta = 8$ ($\gamma = 4, \eta = 16$). Both illustrations are plotted with a constant $q = \tilde{p} = 0.6$. The black graphs display $\pi^{*,pre,amb}_{T,\tilde{p}}$ for $p_a = 0.7, p_b = 0.45$. For $p_a = p_b = 0.6$ it holds $\pi^{*,pre,amb}_{T,\tilde{p}} = \pi^{*,pre}_{T,q}$ (gray and black dashed graphs). The gray dashed graphs refer to $p_a = 1, p_b = 0$.

In addition, Proposition 4.8 sheds light on the importance of the ratio of the levels of ambiguity aversion η and risk aversion γ. Intuitively, it is clear that the optimal pre-commitment strategy under all three dimensions $\pi^{*,pre,amb}_{T,\tilde{p}}$ is decreasing in the risk aversion γ. A higher γ leads to a reduction of the regime based Merton fractions (first risk dimension). In addition, a higher level of risk aversion yields a faster convergence towards the Merton fraction associated with the worst-case regime (second risk dimension). Concerning the third dimension (ambiguity), the speed of convergence towards the worst-case strategy (maximin strategy) is monotonically increasing in the difference resp. the ratio of η and γ, i.e. the higher the difference between the two parameters, the faster the convergence. For fixed investment horizon T the probability distributions $(p_a, 1-p_a)$ and $(p_b, 1-p_b)$ become less important (more important) for the investment decision if the difference between γ and η gets smaller (bigger). For given q, an illustration of the convergence behavior of the optimal pre-commitment strategy under ambiguity and risk aversion to the maximin strategy is given in Figure 4.8.

In addition to the value of the information about the regime (*complete* information), we consider now the value of the information about the lottery (second risk dimension), i.e. the willingness to pay for resolving the ambiguity (the knowledge whether $(p_a, 1-p_a)$ or $(p_b, 1-p_b)$ applies). Here, the investor knows whether to use the strategy $\pi^{*,\text{pre}}_{T,p_a}$ or $\pi^{*,\text{pre}}_{T,p_b}$. Under *complete* information, she knows whether to use

Value of information: $\pi^{*,pre}_{T,pa}$ and $\pi^{*,pre}_{T,pb}$ known, const. q

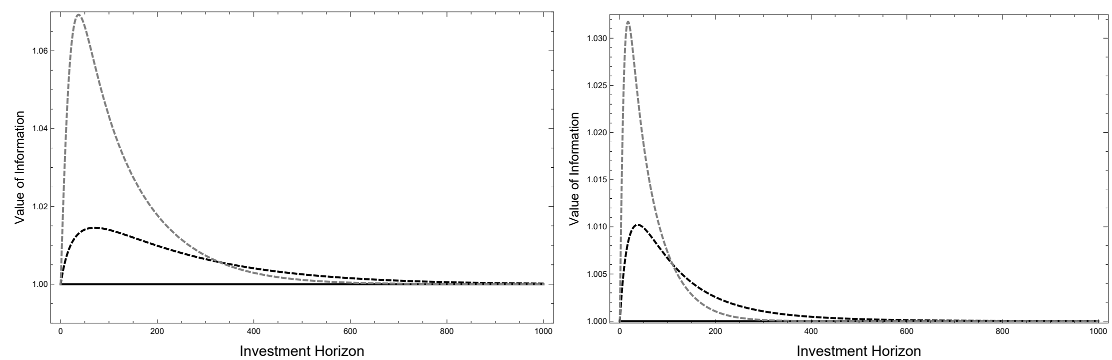


Figure 4.9: Left side: $\gamma = 4, \eta = 8$. Right side: $\gamma = 4, \eta = 16$. $q = \tilde{p} = 0.6$. Black line: $p_a = p_b = 0.6$, black dashed: $p_a = 0.8, p_b = 0.3$, gray dashed: $p_a = 1, p_b = 0$

π_1^{Mer} or π_2^{Mer}. We consider the following ratios certainty equivalents

$$\text{VoI}^{\text{Lot}}(T) = \frac{\left[\tilde{p}\, CE_{T,p_a}(\pi^{*,pre}_{T,p_a})^{1-\eta} + (1-\tilde{p})CE_{T,p_b}(\pi^{*,pre}_{T,p_b})^{1-\eta}\right]^{\frac{1}{1-\eta}}}{\left[\tilde{p}\, CE_{T,p_a}(\pi^{*,pre,amb}_{T,\tilde{p}})^{1-\eta} + (1-\tilde{p})CE_{T,p_b}(\pi^{*,pre,amb}_{T,\tilde{p}})^{1-\eta}\right]^{\frac{1}{1-\eta}}}$$

and

$$\text{VoI}^{\text{Complete}}(T) = \frac{q\, CE_T(\pi_1^{\text{Mer}}, 1) + (1-q)CE_T(\pi_2^{\text{Mer}}, 2)}{\left[\tilde{p}\, CE_{T,p_a}(\pi^{*,pre,amb}_{T,\tilde{p}})^{1-\eta} + (1-\tilde{p})CE_{T,p_b}(\pi^{*,pre,amb}_{T,\tilde{p}})^{1-\eta}\right]^{\frac{1}{1-\eta}}},$$

i.e. $\text{VoI}^{\text{Lot}}(T) - 1$ denotes the willingness to pay for the knowledge of the lottery and $\text{VoI}^{\text{Complete}}(T) - 1$ denotes the willingness to pay for the knowledge of the regime.

An illustration of the value of resolving the ambiguity situation, i.e. $VoI^{\text{Lot}}(T)$, is given in Figure 4.9. Notice that (for all p_a, p_b combinations) the limits $T \to 0$ and $T \to \infty$ of $VoI_{\text{Lot}}(T)$ are equal to 1, i.e. the willingness to pay is zero. Intuitively, the same reasonings apply to the third (ambiguity) dimension as for the second risk dimension (cf. Sec. 4.2.4). Obviously, the willingness to pay is also zero. In the special case $p_a = p_b$ (no ambiguity) it holds $VoI^{\text{Lot}}(T) = 1$ (independent of the investment horizon T). Similar to Sec. 4.2.4, there is an investment horizon $\hat{T}$ for which the value of information achieves its maximum. It is the highest for $p_a = 1, p_b = 0$ where as argued above the ambiguity influence is the highest. However, beyond $\hat{T}$, the value of information drops faster for a higher difference between η and γ since a higher difference implies a faster convergence against the worst-case as already mentioned before.

4.2.6 Conclusion

We consider a stylized setup of an investment decision to shed light on the impact and problems of time-inconsistency. In the first instance, we introduce time-inconsistency using a double risk situation. While the outer risk is given by a simple a priori lottery, the inner risk situation is a regime, coinciding with the classic Merton problem. Although our stylized setup is artificial (in the sense that we do not allow for learning about the regimes), it fits many (dynamic) decision problems (cf. Introduction).

The double risk situation allows an intuitive interpretation of the results. Technically, we can separate the outer and inner risk situation. Since the outer risk situation increases in time (the investment horizon), the optimal decision of the investor converges to the optimal decision within the worst-case regime, i.e. the investor chooses the expected utility maximizing strategy for this regime. For a finite investment horizon, the optimal investment decision is explained by a weighted average of the optimal regime dependent (Merton) solutions. While in the myopic case, the weights resemble the probabilities given by the lottery, there is a time dependent reduction of the probability of the good regime, i.e. as the investment horizon increases the worst-case regime gets more important. Thus, in the limiting case of an infinite investment horizon, the decision is only based on the worst-case regime. In particular, we provide a measure (normalized to $[0, 1]$) for the impact of time-inconsistency. The measure is increasing in the level of risk aversion, because the impact of time makes the risk situation higher, such that the impact is the higher, the higher the risk aversion. Furthermore, it is increasing in the probability of the good regime: the shift towards the worst-case regime is the more severe, the lower the probability of the worst-case regime is. Concerning the willingness to pay for the information about the regime, we have the (obvious) result that the higher the risk aversion is, the higher is the willingness to pay for the information. However, we show that the willingness to pay obtains a maximum, i.e. first increases in the investment horizon and then decreases to zero. Thus, there is an investment horizon where the willingness to pay is maximized. In addition to the two dimensions of risk, we also introduce an additional dimension stemming from ambiguity about the regime probabilities. Using the smooth ambiguity model of Klibanoff et al. (2005), implies a further outer expectation (accounting for the ambiguity aversion). Again, we can separate the effects of the two risk situations as well as the ambiguity aversion. We explain why the impact of time-inconsistency gets more ambiguous. This is explained by the observation that varying the ambiguity situation may also change the risk situation.

4.3 Extension and literature review on portfolio allocation under uncertainty aspects

4.3.1 Possibility of one regime switch

A possible extension to the model in the previous section is given by allowing one regime switch on the interval $[0, T]$.[76] Let us assume the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$, where the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ is generated by the standard Brownian motion $(W_t)_{t\in[0,T]}$ and the continuous time observable Markov process $(Y_t)_{t\in[0,T]}$, i.e. $\mathcal{F}_t = \sigma(W_s, Y_s; 0 \leq s \leq t)$. We assume that Y_t has two possible states, i.e. $Y_t(\omega) \in \{1, 2\}$. The unconditional probability at $t = 0$ of the Markov process is given by

$$\mathbb{P}\left(Y_0(\omega) = 1\right) = p, \quad \mathbb{P}\left(Y_0(\omega) = 2\right) = 1 - p.$$

Thus, the Markov process starts in state 1 with probability p and in state 2 with $1 - p$. With this setup we take into account that the first state of the Markov Chain Y_0 is uncertain. We further assume that both, the drift parameter μ and the volatility σ, depend on the Markov Process, i.e. $\mu_t = \mu(Y_t)$ and $\sigma_t = \sigma(Y_t)$. When the actual regime is given by state 1, i.e. $Y_t(\omega) = 1$ we write $\mu_t = \mu(Y_t = 1) = \mu_1$ and $\sigma_t = \sigma(Y_t = 1) = \sigma_1$ (resp. μ_2 and σ_2 for state 2).
Furthermore, the time of a regime switch is stochastic, modeled by an exponentially distributed random variable $\tau \sim Exp(\lambda)$, where λ is the scale parameter. The density and distribution function of τ are given by

$$f_\tau^\lambda(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}, \; F_\tau^\lambda(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}. \tag{4.19}$$

Our financial market model contains two assets, a risky asset S and a risk-free asset B. Both are adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$, i.e. by evaluating the asset evolution, we know all informations about the Brownian motion and of the Markov process until time t.The evolution of the risk free asset $(B_t)_{t\in[0,T]}$ is given by

$$dB_t = B_t r dt, \tag{4.20}$$

[76] In general, regime-switching is often connected to uncertain distributional parameters, e.g. the drift μ or the volatility σ. Depending on the economic circumstances there can occur a good regime (resp. a bad regime) which will be represented by a high drift and a small volatility (resp. a small drift and a high volatility). A regime 'switches' if the circumstances change. This change is mathematically modeled with a Markov chain resp. a Markov process. If the current state of the regime is known, we speak of an observable Markov chain, if it is not known, we speak of an unobservable or hidden Markov chain. A short introduction to the Markov chain regime-switching problematic can be found in the Appendix A2. Furthermore, we give a review on the optimal portfolio planning literature under regime-switching at the end of this chapter.

where r defines the risk-free interest rate. The solution of this SDE is given by the continuously compounded interest of the initial value B_0. The dynamics of the risky asset $(S_t)_{t\in[0,T]}$ are defined by

$$dS_t = S_t\mu(Y_t)dt + S_t\sigma(Y_t)dW_t, \tag{4.21}$$

where the drift μ and the volatility σ both depend on the Markov process $(Y_t)_{t\in[0,T]}$. Using these definitions the evolution of the portfolio wealth $(V_t)_{t\in[0,T]}$ of the investment strategy is given by the following dynamics:

$$\begin{aligned} dV_t &= V_t\left(\pi_t\frac{dS_t}{S_t} + (1-\pi_t)\frac{dB_t}{B_t}\right) \\ &= V_t\left(\pi_t\mu(Y_t)dt + \pi_t\sigma(Y_t)dW_t + (1-\pi_t)rdt\right) \\ &= V_t\left(\{\pi_t[\mu(Y_t)-r]+r\}dt + \pi_t\sigma(Y_t)dW_t\right) \\ &= V_t\left(\mu_{A,t}(Y_t)dt + \sigma_{A,t}(Y_t)dW_t\right), \end{aligned} \tag{4.22}$$

where π_t is the investment fraction in the risky asset S. For the special case where the investment fraction is constant over time, i.e. $\pi_t = \pi$, for all $t \in [0,T]$, we follow a so called Constant Mix strategy (CM). In this case it holds $\mu_{A,t}(Y_t) = \mu_A(Y_t) = \pi\mu(Y_t) + (1-\pi)r; \sigma_{A,t}(Y_t) = \sigma_A(Y_t) = \pi\sigma(Y_t)$, for all $t \in [0,T]$.
In consequence, for the solution of the SDE following a CM strategy it holds under the real world measure $\mathbb{P}$:

$$V_t = V_0 e^{(\mu_A(Y_t)-\frac{1}{2}\sigma_A^2(Y_t))t+\sigma_A(Y_t)W_t}.$$

If the regime switches at the random point in time $\tau \leq T$ from regime 1 to regime 2, it holds:

$$\begin{aligned} X_1 &= \frac{V_T}{V_0} = \frac{V_\tau}{V_0}\frac{V_T}{V_\tau} \\ &= e^{[\pi(\mu_1-r)+r-\frac{1}{2}\pi^2\sigma_1^2]\tau+\sigma_1\pi W_\tau} e^{[\pi(\mu_2-r)+r-\frac{1}{2}\pi^2\sigma_2^2](T-\tau)+\sigma_2\pi(W_T-W_\tau)} \\ &= e^{[\pi(\mu_2-r)+r-\frac{1}{2}\pi^2\sigma_2^2]T} e^{[\pi(\mu_1-\mu_2)-\frac{1}{2}\pi^2(\sigma_1^2-\sigma_2^2)]\tau+\pi(\sigma_1 W_\tau+\sigma_2(W_T-W_\tau))}. \end{aligned} \tag{4.23}$$

If the regime switches randomly at a point in time τ with $\tau > T$, the terminal wealth can be stated as

$$\begin{aligned} X_2 &= \frac{V_T}{V_0} \\ &= e^{[\pi(\mu_1-r)+r-\frac{1}{2}\pi^2\sigma_1^2]T+\sigma_1\pi W_T}. \end{aligned} \tag{4.24}$$

Now our aim is to calculate and maximize the expected utility resp. the certainty equivalent of the investor. As in the previous section we use a power utility with constant relative risk aversion parameter γ, the *CRRA utility* function, defined by

$$u(x) := \begin{cases} \frac{x^{1-\gamma}}{1-\gamma}, & \text{for } \gamma > 1 \\ \ln(x), & \text{for } \gamma = 1. \end{cases} \tag{4.25}$$

Using a CRRA utility function has its merits in the context of portfolio optimization. It allows us an analysis only based on the asset returns. We can state the expected utility optimization problem of the investors terminal wealth by

$$\max_{\pi_t} \mathbb{E}_{\mathbb{P}} \left[u(V_T) \right]. \tag{4.26}$$

The famous result from Merton (1971) provides the optimal investment fraction π_t to maximize the investors expected utility in the absence of regime switches. Here the optimal portfolio investment strategy is given by a CM strategy with optimal investment fraction

$$\pi_t^* = \pi^* = \pi^{\text{Mer}} = \frac{\mu - r}{\gamma \sigma^2}. \tag{4.27}$$

If a regime switch occurs, the Merton solution of course cannot be optimal anymore because of the changing drift and volatility parameters. Intuitively one could suggest that the overall optimal expected utility in a regime-switching environment is achieved by adapting the investment fraction to the Merton solution which fits the current parameters in the regime and changing it after a regime switch occurs. This intuition is true if and only if the regime switch is observable. Sotomayor and Cadenillas (2009) in a more general setting as also Ocejo (2018) show that the investment strategy which maximizes the expected utility of the investor in (4.26) is given by

$$\pi_t^* = \frac{\mu(Y_t) - r}{\gamma \sigma(Y_t)^2}. \tag{4.28}$$

Here the parameters of the asset dynamic (in contrast to the Merton solution in Equation (4.27)) are depending on the current regime at time t.
In the situation of one regime switch and knowing the start regime we can see that the optimal investment strategy of the investor is given by

$$\pi_t^* = 1_{\{t<\tau\}} \pi^{\text{Mer}_1} + 1_{\{t \geq \tau\}} \pi^{\text{Mer}_2} = 1_{\{t<\tau\}} \frac{\mu_1 - r}{\gamma \sigma_1^2} + 1_{\{t \geq \tau\}} \frac{\mu_2 - r}{\gamma \sigma_2^2}. \tag{4.29}$$

In the next step, it is interesting to calculate and analyze the utility losses which occur from a suboptimal investment strategy. For this, we need to calculate the expected utility of the investor's investment strategy. We receive results in closed-form.

4.3.2 Log Utility

For the special case $\gamma = 1$ we receive the following results.

Proposition 4.9 (Expected Utility under Regime-Switching - Log Utility) *Let u be the Log-utility function, τ the exponentially distributed random point in time with intensity parameter λ where the regime switches from state 1 to state 2 and T the maturity of the contract. Then the expected utility of the terminal wealth V_T can be stated as*

$$\begin{aligned}\mathbb{E}_{\mathbb{P}^*}\left[u\left(\frac{V_T}{V_0}\right)\right] &= \Big[\nu(T,\lambda)\sigma_1^2\pi\pi_1^{Mer} + (1-\nu(T,\lambda))\sigma_2^2\pi\pi_2^{Mer} \\ &\qquad -\frac{1}{2}\pi^2(\nu(T,\lambda)\sigma_1^2 + (1-\nu(T,\lambda))\sigma_2^2) + r\Big]T \\ &= \nu(T,\lambda)y(\pi,1)T + (1-\nu(T,\lambda))y(\pi,2)T,\end{aligned}$$

where

$$\begin{aligned}\nu(T,\lambda) :=& e^{-\lambda T} + \frac{1}{\lambda T} - \frac{1}{\lambda T}e^{-\lambda T} = \frac{1}{\lambda T} + \mathbb{P}(\tau > T)\left(1 - \frac{1}{\lambda T}\right), \\ 1-\nu(T,\lambda) :=& \mathbb{P}(\tau \le T)\left(1 - \frac{1}{\lambda T}\right).\end{aligned}$$

Let us briefly comment on the interpretation of the term $\nu(T,\lambda)y(\pi,1)T$:
With regard to our assumption, we are starting in Regime 1. Because of taking expectations the first (and only) regime switch is expected at $t = \frac{1}{\lambda}$. It could also be possible that the switch takes place after T, i.e. we stay the whole time of the investment horizon in Regime 1. This is captured by the term $\mathbb{P}(\tau > T)$. If this scenario occurs then we are (under expectation) not only in the first time interval $[0, \frac{1}{\lambda}]$ in Regime 1, but also the remaining time $(T - \frac{1}{\lambda})$. This is captured by $\nu(T,\lambda)T$ resp. its utility $\nu(T,\lambda)y(\pi,1)T$. The term $(1-\nu(T,\lambda))y(\pi,2)T$ can be interpreted similar:
We are starting in Regime 1, i.e. under expectation a switch is at most possible at $t = \frac{1}{\lambda}$. Thus it is only possible to stay the remaining time $(T - \frac{1}{\lambda})$ in Regime 2 if this scenario occurs, i.e. the probability $\mathbb{P}(\tau \le T)\left(T - \frac{1}{\lambda}\right)$. But this is $(1-\nu)T$ resp. the utility $(1-\nu(T,\lambda))y(\pi,2)T$.

Proposition 4.10 (Pre-Commitment under Regime-Switching - Log Case) *The pre-commitment strategy that maximizes the expected utility of the log investor is given by*

$$\begin{aligned}\pi^{pre,*} &= \frac{\nu(T,\lambda)\sigma_1^2\pi_1^{Mer} + (1-\nu(T,\lambda))\sigma_2^2\pi_2^{Mer}}{\nu(T,\lambda)\sigma_1^2 + (1-\nu(T,\lambda))\sigma_2^2} \\ &= \alpha(\nu(T,\lambda))\pi_1^{Mer} + (1-\alpha(\nu(T,\lambda)))\pi_2^{Mer}, \ \textit{with} \\ \alpha(\nu(T,\lambda)) &:= \frac{\nu(T,\lambda)\sigma_1^2}{\nu(T,\lambda)\sigma_1^2 + (1-\nu(T,\lambda))\sigma_2^2}.\end{aligned}$$

4.3.3 Power Utility

For the general power utility setting, we are also able to receive closed-form results. We further assume w.l.o.g. that $r = 0$ as in the previous section.

Proposition 4.11 (Expected Utility Regime-Switching - Power Utility) *Let u be the CRRA utility function with relative risk aversion parameter $\gamma > 1$, τ the exponentially distributed random point in time with intensity parameter λ where the regime switches from state 1 to state 2 and T the maturity of the contract. Then the expected utility of the terminal wealth V_T can be stated as*

$$\mathbb{E}_{\mathbb{P}}\left[u\left(\frac{V_T}{V_0}\right)\right] = \frac{1}{1-\gamma}\left[a(\pi,\lambda)e^{y(\pi,1)(1-\gamma)T}e^{-\lambda T} + (1-a(\pi,\lambda))e^{y(\pi,2)(1-\gamma)T}\right], \text{ with}$$
$$a(\pi,\lambda) = \frac{\xi_1-\xi_2}{\xi_1-\xi_2-\lambda} = \frac{y(\pi,1)(1-\gamma)-y(\pi,2)(1-\gamma)}{y(\pi,1)(1-\gamma)-y(\pi,2)(1-\gamma)-\lambda}.$$

The proof of Proposition 4.11 is given in Appendix D.7.

Proposition 4.12 (Pre-Commitment in Regime-Switching - Power Utility) *For $\gamma > 1$ (and $r = 0$) the optimal pre-commitment strategy is given by the implicit function*

$$\begin{aligned}
\pi^{pre,*} &= \alpha\pi_1^{Mer} + (1-\alpha)\pi_2^{Mer}, \text{ where}\\
\alpha &= \frac{\sigma_1^2 g_1}{\sigma_1^2 g_1 + \sigma_2^2 g_2}, 1-\alpha = \frac{\sigma_2^2 g_2}{\sigma_1^2 g_1 + \sigma_2^2 g_2} \text{ with}\\
g_1 &= e^{-\lambda T}e^{y(\pi^{*,pre},1)(1-\gamma)T}\left(\frac{T(\xi_1-\xi_2)(\xi_1-\xi_2-\lambda)}{\lambda}-1\right) + e^{y(\pi^{*,pre},2)(1-\gamma)T}\\
g_2 &= e^{-\lambda T}e^{y(\pi^{*,pre},1)(1-\gamma)T} + e^{y(\pi^{*,pre},2)(1-\gamma)T}(T(\lambda-\xi_1+\xi_2)-1)
\end{aligned}$$

The proof of Proposition 4.12 is given in Appendix D.8.

Remark 4.1
(i) Using the relation $\mathbb{E}_{\mathbb{P}}\left[u\left(\frac{V_T}{V_0}\right)\right] = u(CE)$ we can easily calculate the certainty equivalent (CE) of the terminal wealth where

$$CE = \left[\frac{\xi_1-\xi_2}{\xi_1-\xi_2-\lambda}e^{(\xi_1-\lambda)T} + \frac{-\lambda}{\xi_1-\xi_2-\lambda}e^{\xi_2 T}\right]^{\frac{1}{1-\gamma}}.$$

(ii) Setting $\gamma = 0$ we get the closed-form formula for the expected value of the terminal wealth $\mathbb{E}_{\mathbb{P}}[V_T]$.

Defining $y_{T,\lambda}(\pi) := \frac{1}{T}\ln\left(u^{-1}\left(EU_{T,\lambda}\right)\right)$ we receive with $r = 0$:

$$y_{T,\lambda}(\pi) = \begin{cases} \nu(T,\lambda)y(\pi,1) + (1-\nu(T,\lambda))y(\pi,2), & \gamma = 1 \\ \frac{1}{(1-\gamma)T}\ln\left[a(\pi,\lambda)e^{y(\pi,1)(1-\gamma)T}e^{-\lambda T} + (1-a(\pi,\lambda))e^{y(\pi,2)(1-\gamma)T}\right], & \gamma > 1 \end{cases}. \tag{4.30}$$

Using these results, we are able to analyze the impact of pre-commitment strategies in a regime-switching environment similar to the previous section within the a-priori setup. This will be an interesting and promising analysis for further research.

4.3.4 Regime-switching and portfolio planning - A literature review

Let us end the chapter by giving a literature review on aspects of portfolio planning including the regime-switching topic. As stated before in Chapter 2, Merton (1969) and Merton (1971) are the first two articles that face the challenge of determining the optimal investment fraction in the risk asset s.t. the expected utility of an investor is maximized. Over the years, there are many research papers that include the impact of regime-switching in the optimal investment decision. In this literature overview we want to answer the question how the optimal solution under regime-switching differs compared to the standard Merton solution in case of only one regime. A review of the most common literature is presented in Table 4.4. It includes the corresponding Markov Chain (hidden or observable) and the parameters that should be modeled under the regime switch as well as the optimization problem and the assumptions of the authors. As noticed before, the topic of regime-switching in portfolio planning is closely related to the ones of learning and ambiguity in this research area. Thus we only mention the papers that specifically touch the regime-switching aspects.
The first paper, to the best of our knowledge, that has made regime-switching in the economics research area popular is Hamilton (1989). He considers an observable discrete-time Markov-switching autoregressive time series model. In the context of portfolio allocation Zariphopoulou (1992) has been the first that investigates the EU maximizing terminal wealth problem including consumption where the drift μ is modeled with a MC. The paper works with a utility function that fulfills the Inada conditions s.t. there is no closed-form solution possible. Elliott and Van der Hoek (1997) analyze a hidden Mean-Variance model and determine the optimal strategy for a one-period contract. Ang and Bekaert (2002) investigate the optimal investment fraction in discrete time where the asset returns are modeled with an observable Markov chain. They test empirically the influence of regime-switching on the optimal terminal wealth. If the corresponding portfolio does not account for a risk-free asset, the effects of different regimes are small. If we introduce a risk-free asset, the observation changes and ignoring the regimes is costly. Zhou and Yin

(2003) work within a continuous MV market model where the drift, the interest rate and even the volatility is modeled with an observable MC. They find that for the special case of a constant interest rate the results show similarity to the results without regime-switching. This observation can be found in many papers:
Bäuerle and Rieder (2004) show that for CRRA utility and a 1-dim. BS model that integrates regime-switching for r, μ and σ the optimal investment fraction is given by the Merton solution, depending on the current state of the MC, i.e.

$$\pi_t^{*,obs.} = \frac{\mu(t) - r(t)}{\sigma(t)^2\gamma}. \tag{4.31}$$

The authors also tackle the problem where the drift can switch within a hidden MC model in Rieder and Bäuerle (2005). They find explicit solutions in the CRRA utility case. The optimal solution is given by two factors: a myopic part that refers to the solution under observable states and by a hedging demand part that depends on a PDE. An optimal solution under a hidden Markov Chain for the drift, depending on a PDE, is also given by Honda (2003). Sotomayor and Cadenillas (2009) work in a n-dim. observable MC market model with a general assumption on the utility function. They discuss the general problem and present closed-form solutions for specific HARA functions and also consumption is included in the model. Ocejo (2018) presents the same results as Bäuerle and Rieder (2004) but achieved the solution based on Laplace transforms. The work of Zhang et al. (2010) analyze an enlarged 1-dim. model s.t. it is complete and solve the expected utility maximizing investment fraction problem for the terminal wealth in closed-form for CRRA utility and a MC with n possible states for the interest rate r, the drift μ and the volatility σ. The results coincide with equation (4.31).

Liu (2011) even includes ambiguity aversion in a model with a risky and risk-free asset, where the drift parameter follows a hidden MC. The derived optimal solution contains three parts: the myopic Merton solution depending on the regime as stated in equation (4.31), an intertemporal hedging component and a hedging component that refers to the ambiguity situation. Capponi and Figueroa-López (2014) analyze a setting where to the standard model a risky bond is introduced. The bond can default if the counterparty cannot serve it. For the special case of CRRA utility, they find the expected utility maximizing investment fractions for the products. The investment into the risky asset is given by 4.31, the optimal investment into the risky (defaultable) bond is given by the unique solution of a non-linear system of equations. Fu et al. (2014) introduce observable regime-switching to an incomplete market where also derivatives are included. For the CRRA utility case they find closed-form solutions for the optimal investment fraction that maximize the terminal wealth. It is the sum of a modified Merton solution in combination with delta hedging strategy.
There are also Mean-Variance models that include regime-switching. Frauendorfer

et al. (2007) derive in a complex Ornstein-Uhlenbeck model with an observable MC, that models the drift and volatility parameter, an efficient portfolio that is optimal to a MV criterion. A discrete MV setting with observable MC is presented in Yin and Zhou (2004), whereas Costa and Araujo (2008) as also Xie (2009) analyze a continuous (generalized) MV model.
Referring to this literature overview, we can conclude that there is a crucial difference when it comes to portfolio planning within a regime-switching model: If the states of the Markov-Chain are observable, the optimal investment fraction corresponds to the optimal Merton solution, depending on the current state of the economy. Thus, we invest at time t the investment fraction given in equation (4.31) including the current states of μ, σ and r. If we are in the situation that the states of the Markov Chain are not observable, the optimal solution consists of more than one term, depending on the problem we are facing. It always contains the optimal observable solution, stated in equation (4.31) and some additional terms that refer to a hedging demand because of the unobservable situation.

Authors	Markov Process	Optimization Problem	Assumptions
Zariphopoul (1992)	*observable* μ	EU, $(\pi_t)_{t\in[0,T]}$	1 risky and 1 risk-free asset with consumption, ; strictly increasing and concave utility that fulfills Inada conditions
Elliott and Van der Hoek (1997)	*hidden*	MV	n-dim. discrete model for the rates of return
Ang and Bekaert (2002)	*observable*	EU, $(\pi_t)_{t\in\{0,\cdots,T\}}$	discrete, n-dim. market model, CRRA utility
Honda (2003)	*hidden* μ	EU, $(\pi_t)_{t\in[0,T]}$	1-dim. BS model; HARA utility
Zhou and Yin (2003)	*observable* r, μ, σ	MV	n-dim. continuous market model

(To be continued)

Authors	Markov Process	Optimization Problem	Assumptions
Bäuerle and Rieder (2004)	*observable* r, μ, σ	EU, $(\pi_t)_{t\in[0,T]}$	1 risky and 1 risk-free asset; CRRA utility
Sass and Haussmann (2004)	*hidden* μ	EU, $(\pi_t)_{t\in[0,T]}$	n-dim. market model, d possible states for the MC; strictly increasing and concave, twice continuously differentiable utility function with Inada, explicit solution determined numerically
Rieder and Bäuerle (2005)	*hidden* μ	EU, $(\pi_t)_{t\in[0,T]}$	1 risky and 1 risk-free asset, d possible states of the MC; CRRA utility
Frauendorfer et al. (2007)	*observable* μ, σ	MV	n-dim. Ornstein-Uhlenbeck that models pension funds
Jang et al. (2007)	*observable* r, μ, σ	EU, $(\pi_t)_{t\in[0,T]}$	1 risky and 1 risk-free asset, 2 possible states of the MC; CRRA utility; model transaction costs
Nagai and Runggaldier (2007)	*hidden* μ	EU, $(\pi_t)_{t\in[0,T]}$	n-dim. market model and d possible states for the MC;CRRA utility; give insights with a PDE approach
Taksar and Zeng (2007)	*hidden* μ, σ	EU, $(\pi_t)_{t\in\{0,\cdots,T\}}$	n-dim. market model and d possible states of the MC; discrete approximation for optimal terminal wealth strategy with applicable representation for log-utility
Costa and Araujo (2008)	*observable* μ, σ	generalized MV	n-dim. market model with d possible states of the MC
Xie (2009)	*observable* r, μ, σ	MV	1-dim. market model that includes liability payments, d possible states of the MC
Sotomayor and Cadenillas (2009)	*observable* r, μ, σ	EU, $(\pi_t)_{t\in[0,T]}$	n-dim. market model with consumption and d possible states of the MC; strictly increasing and concave utility that fulfills Inada conditions
Elliott et al. (2010)	*hidden* μ	MV	1 risky, 1 risk-free asset, d possible states of the MC

(To be continued)

Authors	Markov Process	Optimization Problem	Assumptions
Zhang et al. (2010)	*observable* r, μ, σ	EU, $(\pi_t)_{t\in[0,T]}$	1-dim. incomplete market with d possible states of the MC, enlarged to complete one; CRRA utility
Korn et al. (2011)	*hidden* r, μ, σ	EU, $(\pi_t)_{t\in[0,T]}$	1 risky, 1 risk-free asset and 1 cuponbond, also include a pension fund, d possible states of the MC; continuous, strictly increasing, strictly concave and continuously differentiable utility function that fulfills Inada
Liu (2011)	*hidden* μ	EU, $(\pi_t)_{t\in[0,T]}$	1 risky and 1 risk-free bond, include consumption and ambiguity, 2 possible states of the MC; CRRA utility
Çanakoğlu and Özekici (2012)	*observable* μ, σ	EU, $(\pi_t)_{t\in[0,T]}$	n-dim. market model, stochastic interest rate, regime-switching via MP; HARA utility
Shen and Siu (2012)	*observable* short rate, σ	EU, $(\pi_t)_{t\in[0,T]}$	1 risky, 1 risk-free asset and 1 zerobond; stochastic interest rate modeled with Vasicek; continuous, (strictly) increasing, (strictly) concave and continuously differentiable utility function that fulfills Inada
Capponi and Figueroa-López (2014)	*observable* r, μ, σ	EU, $(\pi_t)_{t\in[0,T]}$	1 risky, 1 risk-free asset and 1 risky bond, d possible states for the MC; strictly increasing and concave utility, explicit solution for CRRA
Fu et al. (2014)	*observable* r, μ, σ	EU, $(\pi_t)_{t\in[0,T]}$	model with 1 risky, 1 risk-free asset and an option, d possible states of the MC; strictly increasing and strictly concave utility function
Ocejo (2018)	*observable* r, μ, σ	EU, $(\pi_t)_{t\in[0,T]}$	cf. Sotomayor and Cadenillas (2009); presents solution based on Laplace transforms
Bo et al. (2019)	*hidden* μ, σ default coefficient	EU, $(\pi_t)_{t\in[0,T]}$	n-dim. market model with defaultable stocks and infinitely many possible states for MC

EU=expected utility; MC=Markov Chain; MV=Mean-variance

Table 4.4: Selected papers on optimal portfolio planning under regime-switching

At the end of this chapter, let us take a look at the risk measure literature that contributes to the regime-switching topic. The first paper that includes a possible regime switch in the assets log-returns and its impact on the VaR as also the ES is Hardy (2001). In a Bayesian updating process, she calculates closed-form formulas for both risk measures. Kawata and Kijima (2007) develop a simple regime-switching model to estimate portfolio VaR. This model is able to correct the underestimation problem of risk that has been reported in literature. Elliott and Miao (2009) propose regime-switching models to measure the VaR and ES for a single financial asset as well as portfolios. They capture the volatility clustering phenomenon by assuming the returns follow Student-t distributions. All three papers assume an observable Markov chain. Elliott and Siu (2010) minimize the portfolio risk by applying convex risk measures in an observable MC regime-switching and game-theoretical environment. For the entropic risk measure they discuss special cases.

In the context of portfolio allocation under risk measure constraints there exist some more papers that investigate the impact of a regime-switching environment on the optimal portfolio planning rule. Yiu et al. (2010) include consumption and a Value at Risk constraint. They calculate the optimal investment fraction under an observable Markov Chain depending on a HJB equation and present numerical results: They show that if a regime switches toward a bad regime (high volatility), the corresponding Value at Risk level decreases. This is due to the fact that the bad regime influences the optimal investment fraction as seen in equation (4.31): If the volatility increases, the investment in the risky asset decreases and thus the risk itself decreases because more is invested in the risk-free bond. Therefore, the capital requirements decrease. Liu et al. (2012) solve this problem for the special case of CRRA utility in the absence of consumption with a dynamic VaR constraint in closed-form. It is given by a generalization of the Merton solution that includes the risk constraint and the Markov Chain. The paper of Liu et al. (2014) analyzes a dynamic convex risk measure as a constraint within an observable Markov model. They solve the problem using game theory.

Zhu et al. (2016) present one of the first papers that introduces a VaR-SFP constraint in an optimization problem within a hidden Markov Chain setting. They discuss the optimal portfolio strategy numerically: The investment in the risky asset is reduced to fulfill the VaR constraint. Hu and Wang (2017) present the optimal consumption and investment fraction for a regime-switching model with a VaR constraint that includes liabilities. For a special case of exponential utility functions and two states MC, they derive explicit solutions: The optimal investment fraction is of the same structure as described above in the hidden Markov-Chain setting without risk constraints: One part of the solution is given by equation (4.31), scaled with a term that corresponds to the risk constraint and a second part that is given by a PDE. Finally, Yan et al. (2020) analyze an investment-reinsurance policy for an insurance contract and minimize the SFP modeled with a VaR under an observable

Markov Chain. Concluding this literature overview on the impact of SFP constraints on the optimal investment fractions in a regime-switching portfolio allocation problem, we find commonalities to the problems without risk constraint: in the observable Markov Chain setting, the optimal solution is based on the Merton fraction, adapted to the current state of the Markov Chain. If the Markov Chain is unobservable, the optimal solution is given by a combination of a myopic component of the observable case and an additional (PDE) hedging component. The impact of the risk constraint scales the optimal solution. The different regimes affect the riskiness of the optimal strategy and thus the required capital to fulfill the SFP constraint: a higher volatility regime leads to a less risky investment fraction and thus to a reduction in the capital requirement.

Authors	Risk Measure	Optimization Problem	Assumptions
Hardy (2001)	VaR, ES	-	log-normal distributed returns, observable MC
Kawata and Kijima (2007)	VaR	-	log-normal distributed returns, observable MC
Elliott and Miao (2009)	VaR	-	Student-t distributed returns
Elliott and Siu (2010)	convex risk measures	minimize portfolio risk	1 risky and 1 risk free asset, observable MC on r, μ, σ with d possible states
Yiu et al. (2010)	dynamic VaR	EU, $(\pi_t)_{t\in[0,T]}$ *observable MC* r, μ, σ	1 risky and 1 risk-free asset, d possible states of the MC; include consumption, two time differentiable, strictly increasing and concave utility function with Inada
Liu et al. (2012)	dynamic VaR	EU, $(\pi_t)_{t\in[0,T]}$ *observable MC* r, μ, σ	1 risky and 1 risk-free asset, d possible states of the MC; CRRA utility
Liu et al. (2014)	dynamic convex risk measure	EU, $(\pi_t)_{t\in[0,T]}$ *observable MC* r, μ, σ	1 risky and 1 risk-free asset, d possible states of the MC; exponential utility function

(To be continued)

Authors	Risk Measure	Optimization Problem	Assumptions
Zhu et al. (2016)	VaR	EU,$(\pi_t)_{t\in[0,T]}$ *hidden MC* μ	n-dim. market model with consumption, d possible states of the MC; strictly increasing and concave, twice continuously differentiable with Inada
Hu and Wang (2017)	VaR	EU, $(\pi_t)_{t\in[0,T]}$ *observable MC* r, μ, σ	1 risky, 1 risk-free asset and liabilities, include consumption and d possible states of the MC; 2 time continuously differentiable, strictly increasing and concave utility function with Inada
Setyani et al. (2018)	VaR	EU, $(\pi_t)_{t\in[0,T]}$ *observable MC* r, μ, σ	1 risky and 1 risk-free asset, include consumption and 2 possible states of the MC; strictly increasing and concave utility
Yan et al. (2020)	VaR	minimize ruin probability *observable MC*	1 risky and 1 risk-free asset, 2 possible states of the continuous MC; analyzing an investment-reinsurance policy

EU=expected utility; MC=Markov Chain; MV=Mean-variance

Table 4.5: Selected papers on portfolio allocation with risk measures (constraints) under regime-switching

General Conclusion and Further Research

The present thesis consists of three chapters that contribute to the literature of quantitative risk management on pricing, shortfall probability management, optimal asset allocation and uncertainty.

We contribute to the pricing literature in Chapter 2 by deriving model-independent insights of a minimum return rate guarantee (MRRG) product under default risk which implies a nested (compound) option feature. Despite this feature, the payoff of the default put and the liabilities to the insured can be represented by piecewise linear functions of the investment return, i.e. the payoff of a portfolio of simple put and call options. This implies that the liabilities are easily priced in any model setup which gives closed-form solutions for standard options.
Furthermore, in Chapter 3 we calculate fairly priced guarantee costs of a MRRG contract, where periodic premium payments and a management rule, that controls the investments in the risky asset depending on the former asset evolution, are included. We find in a two-period Black-Scholes model setup that the splitting factor, which determines the periodic premium payments of the insured, has a huge impact on the pricing of the contract: For a constant management rule, we show that the guarantee costs, which can be stated in quasi closed-form, are monotonically increasing and convex in the splitting factor. Including a variable management rule, which has to fulfill some assumptions to avoid violations on the fair pricing, we can compare the guarantee costs of the constant management rule with the ones of the variable management rule: The costs for a variable rule are always greater than the ones under a constant rule and the costs are also convex and monotonically increasing in the splitting factor.

Contributions to the shortfall probability aspect are also given in Chapter 3 by analyzing the effects of periodic premium payments and management rules on the required capital of the insurance company. We find that splitting the contributions of the insured leads to an increase of the required capital for the insurance company s.t. the shortfall probability constraint is fulfilled. The required capital can be reduced by implementing a variable management rule to adapt to the riskiness of the portfolio. Moreover, we derive in Chapter 2 the expected utility-maximizing payoff of a MRRG contract under default risk that fulfills a prescribed shortfall probability bound. We discuss the impeding behavior of the optimal solution that the insured is not secured on the bas states of the world. This behavior stems from cost-efficient payoff modifications.Moreover, we discuss the utility loss for the insured which arises if the insurer implements a suboptimal investment strategy. This is also strongly connected to the optimal asset allocation contributions in the thesis.

Asset allocation contributions are furthermore given in Chapter 3 by analyzing the expected utility maximizing periodic premium fraction of a MRRG contract for a given investment fraction under management rules. For a constant management rule, the upfront premium case in combination with the Merton fraction leads to the optimal expected utility of the insured. Deviations from the Merton fraction imply that the optimal splitting factor has to be adapted to a value smaller than one. For variable management rules, we find that an upfront contribution is not optimal for the insured and even the Merton fraction itself as investment fraction is not an optimal choice anymore.

Additionally, we consider in Chapter 4 a stylized setup of an investment decision to shed light on the impact and problems of time-inconsistency using a double risk situation. The outer risk is given by a simple a priori lottery, whereas the inner risk situation is a regime that coincides with the classic Merton problem. For a finite investment horizon, the optimal investment decision is explained by a weighted average of the optimal regime-dependent (Merton) solutions. While in the myopic case, the weights resemble the probabilities given by the lottery, there is a time-dependent reduction of the probability of the good regime, i.e. as the investment horizon increases the worst-case regime gets more important. Thus, in the limiting case of an infinite investment horizon, the decision is only based on the worst-case regime. In particular, we provide a measure (normalized to $[0, 1]$) for the impact of time-inconsistency. The measure is increasing in the level of risk aversion and with the probability of the good regime.

The uncertainty aspect in the thesis is also covered in Chapter 4. We model the risk situation with an a priori lottery and overlay it with an ambiguous situation where the probabilities for the different regimes are unclear. Using the smooth ambiguity model of Klibanoff et al. (2005) implies a further outer expectation (accounting for the ambiguity aversion). Again, we can separate the effects of the two risk situations as well as the ambiguity aversion.

All of the derived results are embedded into a literature overview of asset allocation problems under risk measure constraints, guarantee features, ambiguity, learning and regime-switching: We find solutions to the impeding behavior that the insured is not protected on the bad states of the world if we implement a terminal SFP con-

straint. Using a dynamic risk constraint resp. a more realistic model that includes mortality risk can solve the problem.
The inclusion of mortality risk can also explain the demand for a guarantee feature in the CRRA utility case. Furthermore, if we include behavioral aspects from Prospect Theory we can even explain the demand for more complex guarantee features.
Moreover, we find that uncertainty aspects as ambiguity reduce the investment fraction in the risky asset compared to the Merton solution. Regime-switching, depending on an observable resp. unobservable Markov chain, also influence the optimal investment fraction that maximizes the expected utility of the investor.

Concerning parameter uncertainty and optimal asset allocation under time inconsistency, it is interesting to study the effects of regime switches on the optimal, time-inconsistent investment strategy in more detail. We have already received some first results as mentioned in Section 4.3, that build the basis for a detailed discussion and analysis.
This research topic contains further interesting research questions that we want to tackle in the future: The setting allows us to study asset allocation problems with restrictions on admissible portfolios if the restrictions can be changed by regulatory decisions. The regime setting then captures different constraints as an upper limit on downside risk measures like Value at Risk and Expected Shortfall or restrictions on the insurance companies' investment decision (e.g. upper bounds on wealth fractions invested riskily or an upper bound for the position in brown assets). Solutions of the resulting asset allocation problems allow us to analyze how uncertainty about a change in regulatory rules – e.g. a discussion about new rules in banking or insurance – changes the investment decisions. Moreover, there are possible interesting applications: Asset allocation problems in insurance economics are often subject to additional restrictions on the downside risk of the optimal portfolio as described in this thesis. The specific form and level of these restrictions/guarantees are subject to regulatory changes. We could capture this uncertainty in our regime-switching model in which the regimes correspond to different constraints imposed on the optimal portfolio. Another application could be that we assume, that the regimes describe different climate futures chosen by the decisions of policymakers like business-as-usual or sticking to the limit of 2°C. The regimes differ w.r.t. the dynamics of asset prices: changes in regimes can induce price jumps as e.g. a gain of green assets at the expense of brown assets. Our implemented regime-switching model allows us to study the impact of regulatory risks on the optimal allocation to brown and green assets.

The topic of climate risks is in general a very interesting research area: Two current papers of the Basel Committee on Banking Supervision explore how climate-related risks can affect banks. Basel Committee on Banking Supervision (2021b) points out that all other financial risks are affected by climate risks and that the current Basel framework may not sufficiently address climate risks. Basel Committee on Banking Supervision (2021a) discusses, amongst other aspects, the risk measurement approaches for climate risks. Thus the discussed risk measures in this thesis could be rethought on the specific properties of climate risks or could be used to measure the impact of climate change like the climate VaR, developed by Dietz et al. (2016). Other challenging 'new' types of risks are cyber risks. They become more and more relevant because of an increasing number of cyberattacks in the last years.[77] Eling and Wirfs (2019) calculate actual costs of cyber risk events based on over 26.000 cyber events. They show that cyber risks have to be modeled via extreme value theory (EVT) to evaluate the costs of cyber risks. This could give a new initiation of a debate on the VaR because as we have seen in the thesis the VaR behaves similarly compared to the ES in the context of heavy-tailed risks.

Moreover, as stated in Section 3.2, a detailed analysis of MRRG contracts with periodic premium payments under different guarantee features could be an interesting topic for further research. Several papers analyze cliquet and ratchet guarantee contracts but all of these works, to the best of our knowledge, consider upfront premium payments.

Finally, the results and representations of the MRRG contract under default risk in Chapter 2 can also be used for future research. It could be interesting to extend our problem to more than one policyholder. If a default occurs, we have to think about a sharing rule between the policyholders for the actual payoff. This leads to the problem of bankruptcy rules. An article that builds a starting point for our analysis is given by Boonen (2019).

[77] For a literature overview of quantifying IT risks and for a definition of cyber risks we refer to Mukhopadhyay et al. (2013) and Biener et al. (2015).

Appendix

Appendix A1: Theoretical Aspects of Chapter 2 - Neyman-Pearson Lemma and Quantile Hedging

Neyman-Pearson Lemma and Quantile Hedging

As we have seen in Chapter 2, quantile guarantees build an optimal design for MR-RGs under default risk. The concept of quantile hedging is well known in literature and dates back to Föllmer and Leukert (1999). The idea behind the quantile hedge is as follows:
In a complete market model setup, every option can be perfectly hedged with a self-financing strategy $(\xi_t)_{t\in[0,T]}$ with initial value A_0. If the investor is only willing to invest the value V_0 with $V_0 < A_0$ the question arises what the best strategy or hedge under these circumstances looks like.[78] Another situation could be that the investor is not interested in a full hedge and wants to fix a shortfall probability bound $\varepsilon \in (0,1)$ on a strategy. What is the minimum initial value $\tilde{V}_0$, s.t. this shortfall bound is fulfilled?[79] The answers to this are given by the quantile hedging theory. Developing this theory, based on the paper of Föllmer and Leukert (1999), will be the main goal of this section.
For proving that the quantile hedge is the optimal solution for these problems the authors use a powerful tool from statistical test theory, the so-called Neyman-Pearson Lemma. This theorem is used by many other authors in highly relevant research topics in mathematics and economics for proving optimality of a solution in an optimization problem. Thus, before going more into details on the quantile hedging theory, we want to motivate and present the main results from the Neyman-Pearson theory that has been first developed in the work of Neyman and Pearson (1933). Later we will see, that the formulation of the Neyman-Pearson problem is of the

[78] Föllmer and Leukert (1999) call this problem maximizing the probability of success.
[79] Föllmer and Leukert (1999) speak of maximizing the expected success ratio.

same structure as the main part in proving the optimality of the quantile hedge. Knowing about the results of the Neyman-Pearson theory will help us to prove the optimality of a quantile hedge. The following section is based on the work of Witting (1985).

Neyman-Pearson Theory - General Motivation of Test Theory

A main topic in statistics is the so-called Test Theory. This setting aims to determine whether a parameter ϑ of an unknown distribution $\mathbb{P}_\vartheta$ belongs to a parameter set Θ_0 or Θ_1. The set Θ_0 corresponds to the so-called null hypothesis H_0, the set Θ_1 to the alternative H_1. That means if we choose $\vartheta \in \Theta_0$, the created distribution P_ϑ belongs to H_0. If we take $\eta \in \Theta_1$, P_η belongs to H_1. The form of the null hypothesis and the alternative will be clarified in the following. Now a statistical test is used to decide, based on a sample $\{x_1, \ldots, x_n\}$, if the null hypothesis is true or needs to be rejected. In case the null hypothesis is rejected we decide to go for the alternative H_1.

There can occur two possible errors, the so-called *type I error*, if the null hypothesis is true but has been rejected from the test, and the *type II error*, where the null hypothesis is not rejected by the test although it is not true anymore.

In general, the aim is to control the type I error and minimize under that constraint the type II error. A more detailed and mathematical explanation of this will be given in the next section.

As mentioned before, a test function resp. a test is needed to decide if H_0 is true or needs to be rejected. Two different tests are relevant for our purpose: the randomized and the non-randomized test. For defining these two tests, we use the most basic test as an example, a so-called one-sided test. This example will also be useful and attainable in the Neyman-Pearson and quantile hedging context.

The null hypothesis and alternative for a one-sided test can be formulated as follows:

$$H_0 : \text{The true parameter } \vartheta \text{ is smaller or equal than the parameter } \tilde{\vartheta}, \text{ i.e. } \vartheta \leq \tilde{\vartheta}$$
$$H_1 : \text{The true parameter } \vartheta \text{ is greater than the parameter } \tilde{\vartheta}, \text{ i.e. } \vartheta > \tilde{\vartheta}.$$

Now it is reasonable to ask how the test function for this hypothesis test is formed:[80] From the given sample $\{x_1, \ldots, x_n\}$ we denote the set of points, where we decide for the alternative, with S. The complement S^C denotes the points, where we assume the null hypothesis is true. For deciding we need to estimate the unknown parameter ϑ. This is done by a function T, a so-called *statistic*. The specific form of T differs, in the Neyman-Pearson context we will precise this function in the next section.

[80] This is stated in Witting (1985), p.35-41.

If $T(x) \leq s$ for a suitable $s \in \mathbb{R}$ and $x \in \{x_1, \ldots, x_n\}$, we believe in H_0.
If $T(x) > s$ for a suitable $s \in \mathbb{R}$ and $x \in \{x_1, \ldots, x_n\}$, we believe in H_1.

Thus, the subsets S and S^C can be written as follows:

$$S := \{x : T(x) > s\};\ S^C := \{x : T(x) \leq s\}.$$

With this results the non-randomized test can be formulate as:

Definition 4.1 (Non-randomized Test)
A ***non-randomized test*** *is a function* $\varphi : \{x_1, \ldots, x_n\} \to \{0; 1\}$*, defined as*

$$\varphi(x) := 1_{\{x:\, T(x)>s\}}(x) = \begin{cases} 1, & x \in S \\ 0, & x \in S^C \end{cases}.$$

This non-randomized test can be interpreted as the probability to decide for alternative H_1. In test theory the main aim is to control the type I error, i.e. to restrict the probability for rejecting the null hypothesis although it is still true by a level of $\alpha \in (0, 1)$. In general it is not possible to exhaust the full boundary condition with a non-randomized test.[81] Therefore, it is useful to take a look at a so-called *randomized test*. Here the test function φ is generalized s.t. all values between zero and one are possible outcomes. In case of a one-sided test, the randomized test can be defined as follows:

Definition 4.2 (Randomized Test)
A ***randomized test*** *is a function* $\varphi : \{x_1, \ldots, x_n\} \to [0, 1]$*, defined as*

$$\varphi(x) := 1_{(s,\infty]}(T(x)) + \gamma(x) 1_{\{s\}}(T(x)) = \begin{cases} 1, & T(x) > s \\ \gamma(x), & T(x) = s \\ 0, & T(x) < s \end{cases},$$

where γ *is a measurable mapping from* $\{x : T(x) = s\} \to [0, 1]$*, the so-called* ***randomization****.*

The interpretation of a randomized test is as follows:
For $T(x) > s$, we decide for the alternative H_1. In case of $T(x) < s$, we choose the null hypothesis H_0. For $T(x) = s$, we do a randomization, i.e. we are choosing

[81] For an example see Witting (1985), page 36-37.

randomly H_0 or H_1. This random choice is done by doing a Bernoulli experiment with the distribution $B(1, \gamma(x))$. If the event happens (with probability $\gamma(x)$), we decide for the alternative H_1. Otherwise we choose the null hypothesis H_0.
By using a randomized test, it is possible to exhaust the type I error probability α. In case of a one-sided test, H_0 and H_1 only contain one distribution each. As we will see in the next section, this is the case in the Neyman-Pearson context. For this we need to choose the constant s and the randomization γ in the following way:

$$s := \inf\{t : P_0(T > t) \leq \alpha\}, \text{ where } P_0 \text{ is the distribution in } H_0,$$

$$\gamma(x) = \bar{\gamma} = \begin{cases} \frac{\alpha - P_0(T>s)}{P_0(T=s)}, & \text{for } P_0(T = s) > 0 \\ 0, & \text{for } P_0(T = s) = 0. \end{cases} \tag{4.32}$$

The randomization γ is constant and the number s is just the generalized inverse function of the distribution function of T under the probability measure P_0. For the complete formulation of this theorem and its poof we refer to Theorem 1.38 in Witting (1985).
Before we formulate the Neyman-Pearson Lemma, one more definition is needed, the so-called *density quotient*. The famous Radon-Nikodým theorem gives a connection between two probability measures on a measurable space $(\Omega, \mathcal{F})$ with the so-called property of absolute continuity. This leads to the Radon-Nikodým density. If the property of absolute continuity is not fulfilled, there exists a more general result, containing the so-called density quotient. Before defining the density quotient, we want to state the famous Randon-Nikodým theorem: These results will be also needed in the quantile hedging part of this section. The following definition and theorem is taken from Föllmer and Schied (2016).[82]

Definition 4.3 (Absolute continuous measure)
Let $\mathbb{P}$ and $\mathbb{Q}$ be two probability measures on the measurable space $(\Omega, \mathcal{F})$. $\mathbb{Q}$ is ***absolute continuous*** *with respect to $\mathbb{P}$, in symbols $\mathbb{P} << \mathbb{Q}$, when the following holds:*

$$\mathbb{P}(A) = 0 \Rightarrow \mathbb{Q}(A) = 0, \text{ for all } A \in \mathcal{F}.$$

With this definition we can formulate the Radon-Nikodým theorem, which characterizes absolute continuous measures:

Theorem 4.1 ***(Radon-Nikodým)***
Let $\mathbb{P}$ and $\mathbb{Q}$ be two probability measures on the measurable space $(\Omega, \mathcal{F})$. Then the following statements are equivalent:

(i) $\mathbb{P} << \mathbb{Q}$

[82] For a more general approach we refer to the book of Elstrodt (2018), Chapter 7.

(ii) There exists a random variable $\xi : \Omega \to [0, \infty]$*, such that*

$$\mathbb{Q}(A) = \int_A \xi \, \mathrm{d}\mathbb{P} = \mathbb{E}_{\mathbb{P}}[1_A \xi], \text{ for all } A \in \mathcal{F}.$$

The random variable ξ is called the Radon-Nikodým density, symbolized with $\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}}$. For a proof of this theorem we refer e.g. to Bauer (2011), p. 101-104.
If neither $\mathbb{P} << \mathbb{Q}$ nor $\mathbb{Q} << \mathbb{P}$ holds, the representation in the above theorem does not hold anymore. But there exists a generalization, that holds for every probability measures $\mathbb{P}$ and $\mathbb{Q}$, the so-called *Lebesgue Decomposition*. In this theorem the density quotient replaces the roll of the Radon-Nikodým density.

Theorem 4.2 ***(Lebesgue Decomposition)***
Let $\mathbb{P}$ *and* $\mathbb{Q}$ *be two probability measures on the measurable space* $(\Omega, \mathcal{F})$*. Then there exists a random variable* $L \geq 0$ *and a set* $N \in \mathcal{F}$ *with* $\mathbb{Q}(N) = 0$*, s.t. the following representation holds:*

$$\mathbb{P}(A) = \mathbb{P}(A \cap N) + \int_A L \, \mathrm{d}\mathbb{Q}, \text{ for all } A \in \mathcal{F}.$$

The function L *is called* ***density quotient*** *from* $\mathbb{P}$ *w.r.t.* $\mathbb{Q}$*.*

Now we have discussed all theoretical parts which are required for understanding the Neyman-Pearson Lemma. This theory will be presented in the following section.

Neyman-Pearson Lemma

Coming back to the main problem in test theory, namely minimizing the type II error by controlling the type I error, the Neyman-Pearson Lemma gives an explicit representation of an optimal test that solves this problem.[83] The Neyman-Pearson Lemma is set in the most basic case of test theory: There are only two probability measures involved, i.e. the null hypothesis and the alternative are given by:

$$H_0 = \{\mathbb{P}\} \text{ and } H_1 = \{\mathbb{Q}\}.$$

The type II error occurs if $\mathbb{Q}$ is the true probability measure, but the null hypothesis is not rejected. The type I error occurs, if $\mathbb{P}$ is the true measure, but H_0 is rejected. Mathematically, we can state the test problem in a simplified way as follows:

$$\begin{aligned} &\mathbb{Q}(H_0 \text{ is not rejected}) \to \min \quad \Leftrightarrow \quad && \mathbb{Q}(H_0 \text{ is rejected}) \to \max \\ &\text{s.t. } \mathbb{P}(H_0 \text{ is rejected}) \leq \alpha && \text{s.t. } \mathbb{P}(H_0 \text{ is rejected}) \leq \alpha. \end{aligned} \tag{4.33}$$

[83] This theorem was first introduced by Neyman and Pearson (1933).

This structure of the Neyman-Pearson problem is also well known in the context of portfolio optimization under some constraint. For example, in the quantile hedging case the success probability of a hedge should be maximized (this corresponds to the type II error minimization), s.t. a shortfall constraint is fulfilled (this corresponds to the type I error control).
Thus knowing about the optimal solution in the Neyman-Pearson problem will provide immediately the optimal solution in the quantile hedging case. Writing the optimization problem in (4.33) in a more complex way, we can introduce the former defined test function φ to determine the optimal test φ^*, that solves the problem.

$$\begin{aligned} \int \varphi(x)q(x)\mathrm{d}\mu &\to \sup_{\varphi\in\Phi} \\ s.t. \int \varphi(x)p(x)\mathrm{d}\mu &\le \alpha, \end{aligned} \tag{4.34}$$

where p and q are Lebesgue densities of the probability measures $\mathbb{P}$ and $\mathbb{Q}$, μ is a dominating measure, Φ is a set of all test functions φ and α the control probability of the type I error.
Intuitively, φ will be set as large as possible (that means $\varphi = 1$) in case the quotient $\frac{q(x)}{p(x)}$ is large and φ will be set as small as possible (that means $\varphi = 0$) in case $\frac{q(x)}{p(x)}$ is small (cf. Witting (1985), p 192).
Now the Neyman-Pearson Lemma can be formulated, which states an optimal randomized test, a so-called **best-α-test**, that solves problem (4.34).

Proposition 4.13 (Neyman-Pearson Lemma)
Let $\mathbb{P}$ and $\mathbb{Q}$ be two probability measures on the measurable space $(\Omega, \mathcal{F})$ and L a density quotient of $\mathbb{Q}$ w.r.t. $\mathbb{P}$. Then there exists a best-α-test φ^* for testing the null hypothesis $H_0 = \{\mathbb{P}\}$ against the alternative $H_1 = \{\mathbb{Q}\}$ for a given level $\alpha \in (0,1)$. The following statements hold:

(i) The best-α-test φ^* is given by

$$\varphi^*(x) = 1_{\{L(x)>s\}}(x) + \bar{\gamma} 1_{\{L(x)=s\}}(x),$$

where $\bar{\gamma}$ and s are determined as in (4.32).

(ii) The density quotient L is given by

$$L(x) := \frac{q(x)}{p(x)} 1_{\{p(x)>0\}}(x) + \infty 1_{\{p(x)=0, q(x)>0\}}(x).$$

(iii) With the best-α-test φ^* it holds: $\mathbb{E}_{\mathbb{P}}[\varphi^*] = \alpha$.

For the proof of the proposition we refer to Witting (1985), p. 193-194.

Remark 4.2

(I) *To avoid the notation of* ∞ *in the density quotient* L*, the test* φ^* *can also be written in the following way:* $\varphi^*(x) = 1_{\{q(x)>s\cdot p(x)\}}(x) + \bar{\gamma} 1_{\{q(x)=s\cdot p(x)\}}(x)$*. This is possible because the following equation holds:*

$$\{x : L(x) > s\} = \{x : q(x) > s \cdot p(x)\}.$$

(II) *For a constant randomization of* $\bar{\gamma} = 0$*, we are in the setting of a non-randomized test. This is the most basic form of the Neyman-Pearson Lemma.* φ^* *has then the following form:*

$$\varphi^*(x) = 1_{\{q(x)>s\cdot p(x)\}}(x).$$

The following property of an optimal Neyman-Pearson test is also of interest in the context of portfolio optimization:

Corollary 4.1
Let φ^* *be the best-*α*-test from the Neyman-Pearson Lemma.*

(i) *If the best-*α*-test is given by a non-randomized test* φ^* *and for any measurable set* $A \in \mathcal{F}$ *holds* $\mathbb{P}(A) \leq \mathbb{P}(\{q(x) > s \cdot p(x)\}) = \mathbb{E}_{\mathbb{P}}[\varphi^*]$*, then also* $\mathbb{Q}(A) \leq \mathbb{Q}(\{q(x) > s \cdot p(x)\}) = \mathbb{E}_{\mathbb{Q}}[\varphi^*]$.

(ii) *If the best-*α*-test is given by a randomized test* φ^* *and for any* $\varphi \in \Phi$ *holds* $\int \varphi \, \mathrm{d}\mathbb{P} \leq \int \varphi^* \, \mathrm{d}\mathbb{P}$*, then also* $\int \varphi \, \mathrm{d}\mathbb{Q} \leq \int \varphi^* \, \mathrm{d}\mathbb{Q}$.

For the proof of this corollary, we refer to Föllmer and Schied (2016), p. 494-495.

Some examples and applications of the Neyman-Pearson Lemma in the context of statistics are given by Dudewicz and Mishra (1988) on pages 444-471. For our purpose this theorem is a useful tool for proving the optimality of the quantile hedge. This will be part of the next section, where we also show, how the Neyman-Pearson Lemma is used to prove the optimality of this hedge.

Quantile Hedging

After reviewing the most important results and techniques regarding the Neyman-Pearson Theory, we can formulate the quantile hedging part in this section. Quantile hedging goes back to Föllmer and Leukert (1999) and is widely applied to proof optimality results: Melnikov and Smirnov (2012) for example use this tool to provide a quasi closed-form solution for the dynamic CVaR hedging. Melnikov and Tong (2014) apply quantile hedging on equity-linked life insurance contracts including transaction costs, whereas Wang (2009) applies it to the so-called guaranteed minimum death benefits contracts (GMDB), which are a component in many variable annuity contracts. Quantile hedging is also applicable to a Markovian regime switching model, as shown by Lien et al. (2021).

Model Setup

Let us start with the basic assumptions and notations in the context of quantile hedging. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ be a filtered probability space and $(V_t)_{t\in[0,T]}$ a value process represented as

$$V_t = V_0 + \int_0^t \xi_s \mathrm{d}X_s,$$

where $(X_t)_{t\in[0,T]}$ is a semi-martingale and $\big(V_0, (\xi_t)_{t\in[0,T]}\big)$ a self-financing and admissible strategy. V_0 is the starting capital and $(\xi_t)_{t\in[0,T]}$ a predictable process that corresponds to the number of assets the investor owns at time t. We further assume the the model is complete, i.e. there exists a uniquely defined equivalent martingale measure $\mathbb{P}^*$.

The contingent claim H, which is $\mathcal{F}_T$ measurable w.r.t. $\mathbb{P}^*$ (we write $H \in L_1(\mathbb{P}^*)$) can be perfectly hedged because of the complete model assumption. Thus there exists a predictable process ξ^H, s.t.

$$\mathbb{E}_{\mathbb{P}^*}\left[H|\mathcal{F}_t\right] = H_0 + \int_0^t \xi_s^H \mathrm{d}X_s \, \mathbb{P} - a.s.,$$

i.e. the option can be replicated with the self-financing strategy $\left(H_0, \xi^H\right)$, where it holds $H_0 = \mathbb{E}_{\mathbb{P}^*}\left[H|\mathcal{F}_0\right] = \mathbb{E}_{\mathbb{P}^*}\left[H\right]$.

This is the standard setting for a perfect hedge. Let us now assume that the investor does not want to invest H_0 at $t = 0$ resp. cannot afford this amount of money but $\tilde{V}_0$ with $\tilde{V}_0 < H_0$. In this case a perfect hedge is not possible anymore. We are interested in maximizing the success probability of the hedge, i.e.

$$\mathbb{P}\left(\left\{V_0 + \int_0^t \xi_s \mathrm{d}X_s \geq H\right\}\right) \to \max, \; s.t. V_0 \leq \tilde{V}_0. \tag{4.35}$$

The first step to solve this problem is to reduce it to the construction of the so-called *success set* $\{V_T > H\}$.

The construction of success sets

Proposition 4.14
Let $\tilde{A} \in \mathcal{F}_T$ be a solution of

$$\mathbb{P}(A) \to \max,\ s.t.\, \mathbb{E}_{\mathbb{P}^*}[H\boldsymbol{1}_A] \leq \tilde{V}_0. \tag{4.36}$$

Furthermore, let $\tilde{\xi}$ be the perfect hedge for $\tilde{H} = H\boldsymbol{1}_{\tilde{A}} \in L_1(\mathbb{P}^*)$, i.e.

$$\mathbb{E}_{\mathbb{P}^*}[H\boldsymbol{1}_{\tilde{A}}|\mathcal{F}_t] = \underbrace{\mathbb{E}_{\mathbb{P}^*}[H\boldsymbol{1}_{\tilde{A}}]}_{\tilde{H}_0} + \int_0^t \tilde{\xi}_s \mathrm{d}X_s\ \mathbb{P} - a.s.$$

Then $(\tilde{V}_0, \tilde{\xi})$ solves problem (4.35) and it holds $\{V_T \geq H\} = \tilde{A}\,\mathbb{P} - a.s.$

A proof is given by Föllmer and Leukert (1999), p. 255. To construct the maximal success set we need to apply the Neyman-Pearson Lemma. For that we first construct the Radon-Nikodým density:

$$\frac{\mathrm{d}\mathbb{Q}^*}{\mathrm{d}\mathbb{P}^*} := \frac{H}{\mathbb{E}_{\mathbb{P}^*}[H]} = \frac{H}{H_0}.$$

In general, the Radon-Nikodým density helps to calculate probabilities in the form of

$$Q(A) = \int_A \frac{\mathrm{d}Q}{\mathrm{d}P}\mathrm{d}P = \mathbb{E}_{\mathbb{P}}\left[\frac{\mathrm{d}Q}{\mathrm{d}P}\mathbf{1}_A\right].$$

In our setting it holds

$$\mathbb{Q}^*(A) = \mathbb{E}_{\mathbb{P}^*}\left[\frac{\mathrm{d}\mathbb{Q}^*}{\mathrm{d}\mathbb{P}^*}\mathbf{1}_A\right] = \mathbb{E}_{\mathbb{P}^*}\left[\frac{H}{H_0}\mathbf{1}_A\right] = \frac{1}{H_0}\mathbb{E}_{\mathbb{P}^*}\left[H\mathbf{1}_A\right],$$

s.t. equation (4.36) can be written as $\mathbb{Q}^*(A) \leq \frac{\tilde{V}_0}{H_0} := \alpha$. Moreover, we define

$$\tilde{a} := \inf\left\{a : \mathbb{Q}^*\left(\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{P}^*} > aH\right) \leq \alpha\right\} \text{ and } \tilde{A} := \left\{\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{P}^*} > \tilde{a}H\right\}.$$

With this we receive the following proposition.

Proposition 4.15
Assume that the set $\tilde{A}$ solves $\mathbb{Q}^*(\tilde{A}) = \alpha$. Then the optimal strategy that solves (4.36) is given by $(\tilde{V}_0, \tilde{\xi})$.

Following Proposition 4.15, we find that if the second condition in (4.36) is fulfilled with equality, then it is automatically the optimal solution. In general it is not true that one can always find a set $A \in \mathcal{F}_T$ s.t. $\mathbb{Q}^*(A) = \alpha$ holds. For this we need to apply the Neyman-Pearson theory to construct the so-called expected success ratio.

Expected success ratio

Let $\varphi : \Omega \to [0,1]$ be a $\mathcal{F}_T$ measurable function and $\mathcal{R}$ the class of all this functions. Taking a look at the optimization problem

$$\max_{\varphi \in \mathcal{R}} \mathbb{E}_{\mathbb{P}}[\varphi], \ s.t. \underbrace{\int \varphi \mathrm{d}\mathbb{Q}^* \leq \alpha}_{\mathbb{E}_{\mathbb{Q}^*}[\varphi]}, \tag{4.37}$$

it is of the form of the Neyman-Pearson Lemma. Thus we can apply Proposition 4.13 which states that the optimal solution is of the following form

$$\tilde{\varphi} = \mathbf{1}_{\left\{\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{P}^*} > \tilde{a}H\right\}} + \gamma \mathbf{1}_{\left\{\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{P}^*} = \tilde{a}H\right\}}, \text{ where}$$
$$\gamma = \begin{cases} 0 & , \mathbb{Q}^*\left(\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{P}^*} = \tilde{a}H\right) = 0 \\ \frac{\alpha - \mathbb{Q}^*\left(\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{P}^*} > \tilde{a}H\right)}{\mathbb{Q}^*\left(\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{P}^*} = \tilde{a}H\right)} & , otherwise \end{cases}.$$

Let us now define the success ratio that is given for every strategy (V_0, ξ) of a value process V by

$$\varphi_V := \mathbf{1}_{\{V_T \geq H\}} + \frac{V_T}{H} \mathbf{1}_{\{V_T < H\}}.$$

Notice that $\varphi_V \in \mathcal{R}$ and $\{\varphi_V = 1\}$ coincides with the success set $\{V_T \geq H\}$ of the strategy (V_0, ξ). Now the aim is to find a strategy that maximizes the expected success ratio $\mathbb{E}_{\mathbb{P}}[\varphi]$.

Proposition 4.16
Let $\tilde{\xi}$ be the perfect hedge for $\tilde{H} := H\tilde{\varphi}$, then $(\tilde{V}_0, \tilde{\xi})$ maximizes the expected success ratio $\mathbb{E}_{\mathbb{P}}[\varphi]$ under all admissible strategies (V_0, ξ) with $V_0 \leq \tilde{V}_0$ and the success ratio of $(\tilde{V}_0, \tilde{\xi})$ is given by $\tilde{\varphi}$.

Remark 4.3
If it holds $\mathbb{Q}^*(\tilde{A}) = \alpha$ then it follows that $\tilde{\varphi} = \mathbf{1}_{\{\tilde{A}\}}$. This is the well known strategy as in Proposition 4.15. Thus, the setting including the expected success ratio is a real generalization.

This is the solution for the situation where we want to maximize the success probability for a given initial capital V_0. Another interesting aspect is to determine the amount of money at inception, s.t. a given success probability is fulfilled.

Cost minimization for given Success Probability

Let us assume we have a prescribed shortfall probability $\varepsilon \in (0,1)$ and we search for the smallest initial capital V_0 s.t. there exists a strategy (V_0, ξ) with

$$\mathbb{P}\left(V_0 + \int_0^T \xi_s \mathrm{d}X_s \geq H\right) \geq 1-\varepsilon.$$

This problem can also be formulated equivalently in the Neyman-Pearson context, i.e. search for a set $A \in \mathcal{F}_T$ s.t.

$$\mathbb{E}_{\mathbb{P}^*}[H\mathbf{1}_A] \to \min \ s.t. \ \mathbb{P}(A) \geq 1-\varepsilon. \tag{4.38}$$

With the previously derived results $\mathbb{E}_{\mathbb{P}^*}[H\mathbf{1}_A]$ can be written as $H_0\mathbb{Q}^*(A)$. Thus (4.38) can equivalently be written as

$$\mathbb{Q}^*(A^C) \to \max \ s.t. \ \mathbb{P}(A^C) \leq \varepsilon. \tag{4.39}$$

The solution can be derived with the Neyman-Pearson Lemma: For this let

$$\begin{aligned} \tilde{b} :=& \inf\left\{b : \mathbb{P}\left(\frac{\mathrm{d}\mathbb{Q}^*}{\mathrm{d}\mathbb{P}} > b\right) \leq \varepsilon\right\} \\ \text{and } \tilde{B}^C :=& \left\{\frac{\mathrm{d}\mathbb{Q}^*}{\mathrm{d}\mathbb{P}} > \tilde{b}\right\} = \left\{\frac{\mathrm{d}\mathbb{Q}^*}{\mathrm{d}\mathbb{P}^*} > \tilde{b}\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{P}^*}\right\} \\ =& \left\{\frac{H}{\mathbb{E}_{\mathbb{P}^*}[H]} > \tilde{b}\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{P}^*}\right\} = \left\{\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{P}^*} < \frac{H}{\tilde{b}\,\mathbb{E}_{\mathbb{P}^*}[H]}\right\}. \end{aligned}$$

<u>Case 1: $\mathbb{P}(\tilde{B}) = 1-\varepsilon$ (i.e. $\mathbb{P}(\tilde{B}^C) = \varepsilon$):</u>

In this case $\tilde{B}^C$ solves problem (4.39) and thus $\tilde{B}$ solves (4.38). Using the previous results imply that the optimal strategy is then given by the replication of the option $H\mathbf{1}_{\tilde{B}}$.

<u>Case 2: $\mathbb{P}(\tilde{B}) \neq 1-\varepsilon$:</u>

Here we can apply the success ratio discussed in the last subsection. Define

$$\begin{aligned} \tilde{\varphi} :=& \mathbf{1}_{\left\{\frac{\mathrm{d}\mathbb{Q}^*}{\mathrm{d}\mathbb{P}} < \tilde{B}\right\}} + \gamma \mathbf{1}_{\left\{\frac{\mathrm{d}\mathbb{Q}^*}{\mathrm{d}\mathbb{P}} = \tilde{B}\right\}}, \text{ where} \\ \gamma =& \begin{cases} \frac{(1-\varepsilon)-\mathbb{P}\left(\frac{\mathrm{d}\mathbb{Q}^*}{\mathrm{d}\mathbb{P}} < \tilde{B}\right)}{\mathbb{P}\left(\frac{\mathrm{d}\mathbb{Q}^*}{\mathrm{d}\mathbb{P}} = \tilde{B}\right)}, & \mathbb{P}\left(\frac{\mathrm{d}\mathbb{Q}^*}{\mathrm{d}\mathbb{P}} < \tilde{B}\right) < 1-\varepsilon \\ 0, & else \end{cases}. \end{aligned}$$

Applying the Neyman-Pearson Lemma, we receive the following proposition.

Proposition 4.17
Let $\tilde{\xi}$ be the perfect hedge of $\tilde{H} := H\tilde{\varphi}$ and $\tilde{V}_0 := \mathbb{E}_{\mathbb{P}^*}[\tilde{H}]$. Then $(\tilde{V}_0, \tilde{\xi})$ has minimal costs under all strategies with expected success ratio $\mathbb{E}_P[\varphi] \geq 1 - \varepsilon$ and the success ratio is given by $\tilde{\varphi}$ with $\mathbb{E}_{\mathbb{P}}[\varphi] = 1 - \varepsilon$.

An application of the Quantile hedging theory in the context of the Black-Scholes Model is given in Föllmer and Leukert (1999), p. 260-262.

Appendix A2: Theoretical Aspects of Chapter 4 - Regime Switching setup via Markov Chains

Observable Markov Chains

Uncertainty of distribution parameters can have a huge impact on the corresponding capital requirements of a risk measure. We want to discuss this impact in a Black-Scholes Model setup. Regime switching is often modeled via Markov Chains.
We consider a regime switching model with two regimes.[84] We work on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t^O)_{t\in[0,T]}, \mathbb{P})$, where the filtration $(\mathcal{F}_t^O)_{t\in[0,T]}$ is generated by the standard Brownian motion $(W_t)_{t\in[0,T]}$ and the continuous time observable Markov Chain[85] $(Y_t^O)_{t\in[0,T]}$, i.e. $\mathcal{F}_t^O = \sigma(W_s, Y_s^O; 0 \leq s \leq t)$. We assume that Y_t^O has two possible regimes $\{s_1, s_2\}$.
The unconditional probability at $t = 0$ of the Markov Chain is given by

$$\mathbb{P}\left(Y_0^O(\omega) = s_1\right) = p, \quad \mathbb{P}\left(Y_0^O(\omega) = s_2\right) = 1 - p,$$

so the Markov Chain starts in state s_1 with probability p and in s_2 with $1 - p$.
With the transition probabilities

$$q_{ij}(t, u) := \mathbb{P}(Y_u = s_j | Y_t = s_i)$$

the *generator* A of the continuous time Markov Chain is defined as

$$A_t = \begin{pmatrix} a_{11}(t) & a_{12}(t) \\ a_{21}(t) & a_{22}(t) \end{pmatrix}, \text{ where}$$

$$a_{ij}(t) := \lim_{h\to 0} \frac{q_{ij}(t, t+h) - \delta_{ij}}{h}, \text{ where } \delta_{ij} := \begin{cases} 1, i = j \\ 0, i \neq j \end{cases}.$$

Our financial market model contains two assets, a risky asset S and a risk-free asset B. Both are adapted to the filtration $\mathcal{F} = (\mathcal{F}_t^O)_{t\geq 0}$. Using this filtration by evaluating the asset evolution, we have all information about the Brownian motion and the Markov process up to time t. Thus, the current state of the Markov process is known when we evaluate S_t.
The evolution of the risk-free asset without regime-switching is given by

$$dB_t = B_t r dt, \tag{4.40}$$

[84] Empirical evidence for a model with two regimes (a so called bear and bull state) is e.g. given in Guidolin and Timmermann (2008).

[85] There also exists the concept of unobservable or as synonymous term hidden Markov Chain. A short introduction to this setup is given in the following subsection.

where r defines the risk-free interest rate. The solution of this SDE is given by the continuously compounded interest of the initial value B_0, i.e. $B_t = B_0 e^{rt}$.

The dynamics of the risky asset $(S_t)_{t\in[0,T]}$ are defined by

$$dS_t = S_t\mu(Y_t^O)dt + S_t\sigma(Y_t^O)dW_t, \tag{4.41}$$

where the drift μ and the volatility σ both depend on the Markov Chain Y^O.

As we are in a situation that the Markov Chain is observable and just has two possible states s_1 and s_2, we can calculate the solution of Equation (4.40) depending on the state of the world the system is at time t:
For the case that $Y_t^O(\omega) = s_1$, we write $\mu(Y_t^O(\omega)) = \mu(s_1) = \mu_1$, for $Y_t^O(\omega) = s_2$ we write $\mu(Y_t^O(\omega)) = \mu(s_2) = \mu_2$. Analogue, we write σ_1, when we see that the observable Markov Chain is at time t in state s_1 and σ_2 when the Chain is at time t in state s_2.

Using these definitions, the evolution of the investor's portfolio wealth $(V_t)_{t\in[0,T]}$ is given by the following dynamics:

$$\begin{aligned} dV_t &= V_t\left(\pi_t\frac{dS_t}{S_t} + (1-\pi_t)\frac{dB_t}{B_t}\right) \\ &= V_t\left(\pi_t\mu(Y_t)dt + \pi_t\sigma(Y_t)dW_t + (1-\pi_t)rdt\right) \\ &= V_t\left(\{\pi_t[\mu(Y_t) - r] + r\}dt + \pi_t\sigma(Y_t)dW_t\right) \\ &= V_t(\mu_{A,t}(Y_t)dt + \sigma_{A,t}(Y_t)dW_t), \end{aligned} \tag{4.42}$$

where π_t is the investment fraction of the risky asset S.
In this section we want to analyze the special case where the investment fraction is constant over time, i.e. $\pi_t = \pi$, for all $t \in [0,T]$. Here we follow a so-called Constant Mix strategy (CM) where the drift and volatility is also constant over time ($\mu_{A,t}(Y_t) = \mu_A(Y_t) = \pi\mu(Y_t) + (1-\pi)r; \sigma_{A,t}(Y_t) = \sigma_A(Y_t) = \pi\sigma(Y_t)$, for all $t \in [0,T]$). In consequence for the solution following a CM strategy it holds under the real world measure $\mathbb{P}$. Furthermore we simplify the setting s.t. the regime uncertainty only occurs at $t = 0$, s.t. we can write

$$V_t = V_0 e^{(\mu_A(Y_0) - \frac{1}{2}\sigma_A^2(Y_0))t + \sigma_A(Y_0)W_t}.$$

Unobservable Markov Chains

In the observable Markov Chain setup we worked on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t^O)_{t\in[0,T]}, \mathbb{P})$ with filtration $\mathcal{F}_t^O$. In the case of an unobservable Markov Chain[86] we no longer have the information about the current state of the chain

[86] In literature a synonymous term used for an unobservable Markov Chain is hidden Markov Chain, e.g. in Liu (2011).

which is why we have to adjust the filtration to $\mathcal{F}^H = (\mathcal{F}_t^H)_{t\geq 0}$, where $\mathcal{F}_t^H = \sigma(S_s; 0 \leq s \leq t)$. In contrast to the filtration $\mathcal{F}_t^O$, where the information about the current state of the Markov Chain and the Brownian Motion are included, the filtration $\mathcal{F}_t^H$ is the sigma algebra generated by the risky asset S. This means that we have all information at time t about the asset price S_t, but we neither have information about the actual regime nor information about the Brownian Motion. To overcome this, we use so called 'filtered probabilities'. Let us denote $(Y_t^H)_{t\in[0,T]}$ as unobservable Markov Chain if the current state of the chain $Y_{t_0}^H(\omega)$ is unknown. In case of a two regime world with regimes s_1 and s_2, the filtered probabilities are defined by

$$p_t = P(Y_t(\omega) = s_1|\mathcal{F}_t^H) \overset{\text{Markov Property}}{=} P(Y_t(\omega) = s_1|S_t) \tag{4.43}$$

$$1 - p_t = P(Y_t(\omega) = s_2|\mathcal{F}_t^H) \overset{\text{Markov Property}}{=} P(Y_t(\omega) = s_2|S_t), \tag{4.44}$$

where p_0 (resp. $1 - p_0$) is the probability that at time $t = 0$ the unobservable Markov Chain starts in state s_1 (resp. s_2). With this setting we can work in a so-called 'Markovian equivalent economy'(see e.g. Gennotte (1986)).
Honda (2003) models transition probabilities λ based on an exponential distribution, where the parameter $\lambda_{12} = \lambda_{21} = \lambda$. This means that the transition probability for a regime switch from s_1 to s_2 is the same as switching from s_2 to s_1. With this assumption the stochastic process of the filtered probabilities evolves as follows

$$dp_t = \lambda(1 - 2p_t)dt + p_t(1 - p_t)\frac{\mu(s_2) - \mu(s_1)}{\sigma}d\overline{W}_t = \mu_p(p_t)dt + \sigma_p(p_t)d\overline{W}_t, \tag{4.45}$$

where $\mu_p(p_t) = \lambda(1-2p_t)$ and $\sigma_p(p_t) = p_t(1-p_t)\frac{\mu(s_2)-\mu(s_1)}{\sigma}$. An interpretation of this formula is given in Honda (2003), p. 49.
$\overline{W}_t$ is a Brownian Motion with respect to $(\mathcal{F}_t^H)_{t\geq 0}$, constructed by

$$\bar{W}_t = \int_0^t \frac{dS_s - S_s\hat{\mu}(p_t)}{S_s\sigma}ds, \tag{4.46}$$

where $\hat{\mu}(p_t) = p_t\mu(s_1) + (1 - p_t)\mu(s_2)$. Combining formulas (4.45) and (4.46), the risky asset S can be represented as follows:

$$dS_t = S_t\mu_p(p_t)dt + S_t\sigma_p(p_t)d\overline{W}_t. \tag{4.47}$$

Thus we can construct the 'Markovian equivalent economy' combining the risk free bond, the filtered probabilities given in (4.45) and the risky asset given in (4.47) with the filtration $(\mathcal{F}_t^H)_{t\geq 0}$ on the probability space $(\Omega, \mathcal{F}, P)$.

Appendix B: Proofs of Chapter 2

B.1: Proof of Lemma 2.3

Part (i) follows immediately with the observation that under the measure $\mathbb{P}$ $\beta A_T^{(Mer)}$ is normally distributed with mean $(\mu - \frac{1}{2}\sigma_A^2)T$ and standard deviation $\sigma_A\sqrt{T}$.
For (ii) notice that

$$\mathbb{E}_{\mathbb{P}^*}\left[e^{-r(T-t)}\left(G_T - \beta A_T^{(Mer)}\right)1_{\left\{\underline{K}<\beta A_T^{(Mer)}\leq\overline{K}\right\}}\middle|\mathcal{F}_t\right] = E_1 - E_2,$$

where

$$E_1 := \mathbb{E}_{\mathbb{P}^*}\left[e^{-r(T-t)}G_T 1_{\left\{\underline{K}<\beta A_T^{(Mer)}\leq\overline{K}\right\}}\middle|\mathcal{F}_t\right],$$
$$E_2 := \mathbb{E}_{\mathbb{P}^*}\left[e^{-r(T-t)}\beta A_T^{(Mer)} 1_{\left\{\underline{K}<\beta A_T^{(Mer)}\leq\overline{K}\right\}}\middle|\mathcal{F}_t\right].$$

E_1 is immediately implied by the observation that under $\mathbb{P}^*$ and given the information $\mathcal{F}_t$, $\beta A_T^{(Mer)}$ is normally distributed with mean $\beta A_t^{(Mer)} + (r - \frac{1}{2}\sigma_A^2)T$ and standard deviation $\sigma_A\sqrt{T}$, i.e.

$$\begin{aligned}
E_1 = e^{-r(T-t)}G_T &\left[\mathcal{N}\left(\frac{\ln\left(\frac{\overline{K}}{\beta A_t^{(Mer)}}\right) - (r - \frac{1}{2}\sigma_A^2)(T-t)}{\sigma_A\sqrt{T-t}}\right)\right. \\
&\left. -\mathcal{N}\left(\frac{\ln\left(\frac{\underline{K}}{\beta A_t^{(Mer)}}\right) - (r - \frac{1}{2}\sigma_A^2)(T-t)}{\sigma_A\sqrt{T-t}}\right)\right] \\
= e^{-r(T-t)}G_T &\left[\mathcal{N}\left(-d_2(\overline{K})\right) - \mathcal{N}\left(-d_2(\underline{K})\right)\right].
\end{aligned}$$

In addition, notice that

$$\begin{aligned}
E_2 &= \mathbb{E}_{\mathbb{P}^*}\left[e^{-r(T-t)}\beta A_T^{(Mer)} 1_{\left\{\underline{K}<\beta A_T^{(Mer)}\leq\overline{K}\right\}}\middle|\mathcal{F}_t\right] \\
&= \mathbb{E}_{\mathbb{P}^*}\left[e^{-r(T-t)}\beta A_t^{(Mer)} e^{(r-\frac{1}{2}\sigma_A^2)(T-t)+\sigma_A(W_T^*-W_t^*)} 1_{\left\{\underline{K}<\beta A_T^{(Mer)}\leq\overline{K}\right\}}\middle|\mathcal{F}_t\right] \\
&= \beta A_t^{(Mer)}\mathbb{E}_{\mathbb{P}^*}\left[e^{-\frac{1}{2}\sigma_A^2(T-t)+\sigma_A(W_T^*-W_t^*)} 1_{\left\{\underline{K}<\beta A_T^{(Mer)}\leq\overline{K}\right\}}\middle|\mathcal{F}_t\right].
\end{aligned}$$

With Girsanov's Theorem,

$$\begin{aligned}
&\mathbb{E}_{\mathbb{P}^*}\left[e^{-\frac{1}{2}\sigma_A^2(T-t)+\sigma_A(W_T^*-W_t^*)} 1_{\left\{\underline{K}<\beta A_T^{(Mer)}\leq\overline{K}\right\}}\middle|\mathcal{F}_t\right] \\
=&\tilde{\mathbb{P}}\left(\underline{K} < \beta A_t^{(Mer)} e^{(r+\frac{1}{2}\sigma_A^2)(T-t)+\sigma_A(\tilde{W}_T-\tilde{W}_t)} \leq \overline{K}\right),
\end{aligned}$$

where $\tilde{W}_t = W_t^* - \sigma_A t$ is a Brownian motion under $\tilde{\mathbb{P}}$, i.e. the above probability is given by

$$\begin{aligned}
&\tilde{\mathbb{P}}\left(\underline{K} < \beta A_t^{(Mer)} e^{(r+\frac{1}{2}\sigma_A^2)(T-t)+\sigma_A(\tilde{W}_T-\tilde{W}_t)} \leq \overline{K}\right) \\
=&\mathcal{N}\left(\frac{\ln\left(\frac{\overline{K}}{\beta A_t^{(Mer)}}\right) - (r+\frac{1}{2}\sigma_A^2)(T-t)}{\sigma_A\sqrt{T-t}}\right) - \mathcal{N}\left(\frac{\ln\left(\frac{\underline{K}}{\beta A_t^{(Mer)}}\right) - (r+\frac{1}{2}\sigma_A^2)(T-t)}{\sigma_A\sqrt{T-t}}\right) \\
=&\mathcal{N}\left(-d_1(\overline{K})\right) - \mathcal{N}\left(-d_1(\underline{K})\right).
\end{aligned}$$

B.2: Proof of Lemma 2.4

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

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}[u(A_T^*)] &= \mathbb{E}_{\mathbb{P}}\left[u\left(\beta A_T^{(Mer)} + \left\{G_T - \beta A_T^{(Mer)}\right\} 1_{\left\{\underline{K}<\beta A_T^{(Mer)}\leq\overline{K}\right\}}\right)\right] \\
&= \mathbb{E}_{\mathbb{P}}\left[u\left(\beta A_T^{(Mer)} 1_{\left\{\beta A_T^{(Mer)}\leq\underline{K}\right\}\cup\left\{\beta A_T^{(Mer)}>\overline{K}\right\}} + G_T 1_{\left\{\underline{K}<\beta A_T^{(Mer)}\leq\overline{K}\right\}}\right)\right] \\
&= \mathbb{E}_{\mathbb{P}}\left[u\left(\beta A_T^{(Mer)}\right) 1_{\left\{\beta A_T^{(Mer)}\leq\underline{K}\right\}\cup\left\{\beta A_T^{(Mer)}>\overline{K}\right\}} + u\left(G_T\right) 1_{\left\{\underline{K}<\beta A_T^{(Mer)}\leq\overline{K}\right\}}\right] \\
&= E_3 + E_4,
\end{aligned}$$

where

$$\begin{aligned}
E_3 &:= \mathbb{E}_{\mathbb{P}}\left[\frac{\left(\beta A_T^{(Mer)}\right)^{(1-\gamma)}}{1-\gamma} 1_{\left\{\beta A_T^{(Mer)}\leq\underline{K}\right\}\cup\left\{\beta A_T^{(Mer)}>\overline{K}\right\}}\right], \\
E_4 &:= \mathbb{E}_{\mathbb{P}}\left[\frac{G_T^{(1-\gamma)}}{1-\gamma} 1_{\left\{\underline{K}<\beta A_T^{(Mer)}\leq\overline{K}\right\}}\right].
\end{aligned}$$

Let us first calculate E_3. Notice that the following property for indicator functions holds:

$$1_{A\cup B} = 1_A + 1_B + 1_{A\cap B}.$$

In our case the set A is given by $\left\{\beta A_T^{(Mer)} \leq \underline{K}\right\}$ and B by $\left\{\beta A_T^{(Mer)} > \overline{K}\right\}$. They are disjoint because $\underline{K} \leq \overline{K}$. So E_3 can be stated as follows:

$$\begin{aligned}
E_3 = &\mathbb{E}_{\mathbb{P}}\left[\frac{\left(\beta A_T^{(Mer)}\right)^{(1-\gamma)}}{1-\gamma} 1_{\left\{\beta A_T^{(Mer)}\leq\underline{K}\right\}}\right] + \mathbb{E}_{\mathbb{P}}\left[\frac{\left(\beta A_T^{(Mer)}\right)^{(1-\gamma)}}{1-\gamma} 1_{\left\{\beta A_T^{(Mer)}>\overline{K}\right\}}\right] \\
=: &E_5 + E_6.
\end{aligned}$$

E_5 and E_6 are similar to compute, thus we only calculate E_5 in detail. For E_6 we can proceed analogously.

$$\begin{aligned} E_5 &= \frac{\beta^{(1-\gamma)}}{1-\gamma}\mathbb{E}_{\mathbb{P}}\left[\left(A_T^{(Mer)}\right)^{(1-\gamma)} 1_{\left\{\beta A_T^{(Mer)} \le \underline{K}\right\}}\right] \\ &= \frac{\beta^{(1-\gamma)}}{1-\gamma} A_0^{(1-\gamma)}\mathbb{E}_{\mathbb{P}}\left[e^{(1-\gamma)(\mu_A - \frac{1}{2}\sigma_A^2)T + (1-\gamma)\sigma_A W_T} 1_{\left\{\beta A_T^{(Mer)} \le \underline{K}\right\}}\right], \end{aligned} \tag{4.48}$$

where the last equality holds because $A_T^{(Mer)} = A_0 e^{(\mu_A - \frac{1}{2}\sigma_A^2)T + \sigma_A W_T}$.
Now our aim is to use Girsanov's theorem to calculate the expected value. For this we need the Radon-Nikodym density which is in our setting given by

$$Z_T := e^{-\frac{1}{2}(1-\gamma)^2\sigma_A^2 T + \sigma_A(1-\gamma)W_T}.$$

Rewriting (4.55), we get

$$\begin{aligned} E_5 &= \frac{(\beta A_0)^{(1-\gamma)}}{1-\gamma} e^{(1-\gamma)(\mu_A - \frac{1}{2}\sigma_A^2)T} e^{\frac{1}{2}(1-\gamma)^2\sigma_A^2 T}\mathbb{E}_{\mathbb{P}}\left[e^{-\frac{1}{2}(1-\gamma)^2\sigma_A^2 T} e^{(1-\gamma)\sigma_A W_T} 1_{\left\{\beta A_T^{(Mer)} \le \underline{K}\right\}}\right] \\ &= \frac{(\beta A_0)^{(1-\gamma)}}{1-\gamma} e^{(1-\gamma)[\mu_A T - \frac{1}{2}\gamma\sigma_A^2 T]}\mathbb{E}_{\mathbb{P}}\left[Z_T 1_{\left\{\beta A_T^{(Mer)} \le \underline{K}\right\}}\right]. \end{aligned} \tag{4.49}$$

With Girsanov's Theorem,

$$\mathbb{E}_{\mathbb{P}}\left[Z_T 1_{\left\{\beta A_T^{(Mer)} \le \underline{K}\right\}}\right] = \tilde{\mathbb{P}}\left(\beta A_T^{(Mer)} \le \underline{K}\right),$$

where $\tilde{\mathbb{P}}$ is the uniquely determined equivalent martingale measure of $\mathbb{P}$ and $\tilde{W}_T$ is a BM under $\tilde{\mathbb{P}}$ given by $\tilde{W}_T = W_T - (1-\gamma)\sigma_A T$.
Again notice, that

$$\ln\left(\frac{A_T^{(Mer)}}{A_0}\right) = \left(\mu_A - \frac{1}{2}\sigma_A^2\right)T + \sigma_A W_T = \cdots = \left[\mu_A - \left(\gamma - \frac{1}{2}\right)\sigma_A^2\right]T + \sigma_A\tilde{W}_T.$$

Then we get

$$\begin{aligned} \tilde{\mathbb{P}}\left(\beta A_T^{(Mer)} \le \underline{K}\right) &= \tilde{\mathbb{P}}\left(\frac{\ln\left(\frac{A_T^{(Mer)}}{A_0}\right) - \left[\mu_A - \left(\gamma - \frac{1}{2}\right)\sigma_A^2\right]T}{\sigma_A\sqrt{T}} \le \frac{\ln\left(\frac{\underline{K}}{\beta A_0}\right) - \left[\mu_A - \left(\gamma - \frac{1}{2}\right)\sigma_A^2\right]T}{\sigma_A\sqrt{T}}\right) \\ &= \mathcal{N}\left(\frac{\ln\left(\frac{\underline{K}}{\beta A_0}\right) - \left[\mu_A - \left(\gamma - \frac{1}{2}\right)\sigma_A^2\right]T}{\sigma_A\sqrt{T}}\right). \end{aligned} \tag{4.50}$$

Combining (4.56) and (4.50), it holds

$$E_5 = \frac{(\beta A_0)^{(1-\gamma)}}{1-\gamma} e^{(1-\gamma)[\mu_A T - \frac{1}{2}\gamma\sigma_A^2 T]} \mathcal{N}\left(\frac{\ln\left(\frac{\underline{K}}{\beta A_0}\right) - \left[\mu_A - \left(\gamma - \frac{1}{2}\right)\sigma_A^2\right] T}{\sigma_A \sqrt{T}} \right). \quad (4.51)$$

Analogously, we get

$$\begin{aligned} E_6 = \dots = & \frac{(\beta A_0)^{(1-\gamma)}}{1-\gamma} e^{(1-\gamma)[\mu_A T - \frac{1}{2}\gamma\sigma_A^2 T]} \tilde{\mathbb{P}}\left(\beta A_T^{(Mer)} > \overline{K}\right) = \dots \\ = & \frac{(\beta A_0)^{(1-\gamma)}}{1-\gamma} e^{(1-\gamma)[\mu_A T - \frac{1}{2}\gamma\sigma_A^2 T]} \left\{ 1 - \mathcal{N}\left(\frac{\ln\left(\frac{\overline{K}}{\beta A_0}\right) - \left[\mu_A - \left(\gamma - \frac{1}{2}\right)\sigma_A^2\right] T}{\sigma_A \sqrt{T}} \right) \right\}. \end{aligned} \quad (4.52)$$

Using the results stated in (4.51) and (4.52), E_3 is given by

$$E_3 = \frac{(\beta A_0)^{(1-\gamma)}}{1-\gamma} e^{(1-\gamma)[\mu_A T - \frac{1}{2}\gamma\sigma_A^2 T]} \left\{ 1 - \left[\mathcal{N}\left(\tilde{d}(\overline{K},\gamma)\right) - \mathcal{N}\left(\tilde{d}(\underline{K},\gamma)\right) \right] \right\}. \quad (4.53)$$

It remains to calculate E_4:

$$\begin{aligned} E_4 &= \mathbb{E}_{\mathbb{P}}\left[\frac{G_T^{(1-\gamma)}}{1-\gamma} 1_{\left\{\underline{K} < \beta A_T^{(Mer)} \leq \overline{K}\right\}} \right] = \frac{G_T^{(1-\gamma)}}{1-\gamma} \mathbb{E}_{\mathbb{P}}\left[1_{\left\{\underline{K} < \beta A_T^{(Mer)} \leq \overline{K}\right\}} \right] \\ &= \frac{G_T^{(1-\gamma)}}{1-\gamma} \mathbb{P}\left(\underline{K} < \beta A_T^{(Mer)} \leq \overline{K} \right) \\ &= \frac{G_T^{(1-\gamma)}}{1-\gamma} \mathbb{P}\left(\frac{\ln\left(\frac{\underline{K}}{\beta A_0}\right) - (\mu_A - \frac{1}{2}\sigma_A^2)T}{\sigma_A\sqrt{T}} < \xi \leq \frac{\ln\left(\frac{\overline{K}}{\beta A_0}\right) - (\mu_A - \frac{1}{2}\sigma_A^2)T}{\sigma_A\sqrt{T}} \right), \end{aligned}$$

where $\xi \sim \mathcal{N}(0,1)$.
So it holds

$$E_4 = \frac{G_T^{(1-\gamma)}}{1-\gamma} \left\{ \mathcal{N}(\tilde{d}(\overline{K},1)) - \mathcal{N}(\tilde{d}(\underline{K},1)) \right\}. \quad (4.54)$$

The final solution is given by combining (4.53) and (4.54). □

Appendix C: Proofs of Chapter 3

C.1: Proof of Lemma 3.1

The account value V_{t_i} is given by

$$
\begin{aligned}
V_{t_i} &= e^{-r(t_n-t_i)}\mathbb{E}_{\mathbb{P}^*}\left[V_{t_n}\middle|\{A_{t_0},\ldots,A_{t_i}\}\right] \\
&= e^{-r(t_n-t_i)}\mathbb{E}_{\mathbb{P}^*}\left[\sum_{j=0}^{n-1}\tilde{a}_{t_j}\frac{A_{t_n}}{A_{t_j}}\middle|\{A_{t_0},\ldots,A_{t_i}\}\right] \\
&= e^{-r(t_n-t_i)}\mathbb{E}_{\mathbb{P}^*}\left[\sum_{j=0}^{i-1}\tilde{a}_{t_j}\frac{A_{t_i}}{A_{t_j}}\frac{A_{t_n}}{A_{t_i}}+\sum_{j=i}^{n-1}\tilde{a}_{t_j}\frac{A_{t_n}}{A_{t_j}}\middle|\{A_{t_0},\ldots,A_{t_i}\}\right] \\
&= \sum_{j=0}^{i-1}\tilde{a}_{t_j}\frac{A_{t_i}}{A_{t_j}}+\sum_{j=i}^{n-1}e^{-r(t_j-t_i)}\ \tilde{a}_{t_j} \\
&= \alpha\left(\sum_{j=0}^{i-1}a_{t_j}\frac{A_{t_i}}{A_{t_j}}+\sum_{j=i}^{n-1}e^{-r(t_j-t_i)}\ a_{t_j}\right) \\
&= \alpha\left(\sum_{j=0}^{i-1}a_{t_j}\frac{A_{t_i}}{A_{t_j}}+a_{t_i}+\sum_{j=i+1}^{n-1}e^{-r(t_j-t_i)}\ a_{t_j}\right).
\end{aligned}
$$

C.2: Proof of Lemma 3.2

Using the representation of Lemma 3.1, the account value V_{t_i} is given by

$$
V_{t_i} = \alpha\left(\sum_{j=0}^{i-1}a_{t_j}\frac{A_{t_i}}{A_{t_j}}+a_{t_i}+\sum_{j=i+1}^{n-1}e^{-r(t_j-t_i)}\ a_{t_j}\right).
$$

Furthermore, notice that at t_{i+1} $(i \leq n-2)$ it holds

$$
\begin{aligned}
V_{t_{i+1}} &= \alpha \left(\sum_{j=0}^{i} a_{t_j} \frac{A_{t_{i+1}}}{A_{t_j}} + a_{t_{i+1}} + \sum_{j=i+2}^{n-1} e^{-r(t_j - t_{i+1})} \, a_{t_j} \right) \\
&= \alpha \left(\frac{A_{t_{i+1}}}{A_{t_i}} \sum_{j=0}^{i-1} a_{t_j} \frac{A_{t_i}}{A_{t_j}} + a_{t_i} \frac{A_{t_{i+1}}}{A_{t_i}} + a_{t_{i+1}} + e^{r(t_{i+1}-t_i)} \sum_{j=i+2}^{n-1} e^{-r(t_j - t_i)} \, a_{t_j} \right) \\
&= \alpha \left(\left(\sum_{j=0}^{i-1} a_{t_j} \frac{A_{t_i}}{A_{t_j}} + a_{t_i} \right) \frac{A_{t_{i+1}}}{A_{t_i}} + a_{t_{i+1}} + e^{r(t_{i+1}-t_i)} \sum_{j=i+2}^{n-1} e^{-r(t_j - t_i)} \, a_{t_j} \right) \\
&= V_{t_i} \frac{A_{t_{i+1}}}{A_{t_i}} + \alpha \left(- \left(\sum_{j=i+1}^{n-1} e^{-r(t_j - t_i)} a_{t_j} \right) \frac{A_{t_{i+1}}}{A_{t_i}} + a_{t_{i+1}} + e^{r(t_{i+1}-t_i)} \sum_{j=i+2}^{n-1} e^{-r(t_j - t_i)} \, a_{t_j} \right) \\
&= V_{t_i} \frac{A_{t_{i+1}}}{A_{t_i}} + \alpha \left(- \left(\sum_{j=i+1}^{n-1} e^{-r(t_j - t_i)} a_{t_j} \right) \frac{A_{t_{i+1}}}{A_{t_i}} + e^{r(t_{i+1}-t_i)} \sum_{j=i+1}^{n-1} e^{-r(t_j - t_i)} \, a_{t_j} \right) \\
&= V_{t_i} \frac{A_{t_{i+1}}}{A_{t_i}} + \alpha \left(\left(\sum_{j=i+1}^{n-1} e^{-r(t_j - t_i)} a_{t_j} \right) \left(e^{r(t_{i+1}-t_i)} - \frac{A_{t_{i+1}}}{A_{t_i}} \right) \right) \\
&= V_{t_i} \frac{A_{t_{i+1}}}{A_{t_i}} + \alpha \left(\left(\sum_{j=i+1}^{n-1} e^{-r(t_j - t_{i+1})} a_{t_j} \right) \left(1 - e^{-r(t_{i+1}-t_i)} \frac{A_{t_{i+1}}}{A_{t_i}} \right) \right),
\end{aligned}
$$

such that

$$
\begin{aligned}
V_{t_{i+1}} &= V_{t_i} \frac{A_{t_{i+1}}}{A_{t_i}} + \alpha \left(\left(\sum_{j=i+1}^{n-1} e^{-r(t_j - t_{i+1})} a_{t_j} \right) \left(1 - e^{-r(t_{i+1}-t_i)} \frac{A_{t_{i+1}}}{A_{t_i}} \right) \right) \\
&= V_{t_i} \frac{A_{t_{i+1}}}{A_{t_i}} + \alpha \left(PV_{t_{i+1}} \left(1 - e^{-r(t_{i+1}-t_i)} \frac{A_{t_{i+1}}}{A_{t_i}} \right) \right).
\end{aligned}
$$

C.3: Proof of Proposition 3.1

Using the representation of the account value in Lemma 3.2 together with Eqn. (3.5) we find that

$$\begin{aligned}
V_{t_{i+1}} &= V_{t_i}\frac{A_{t_{i+1}}}{A_{t_i}} + \alpha\left(e^{r(t_{i+1}-t_0)}\sum_{j=i+1}^{n-1}\beta_j\left(1-e^{-r(t_{i+1}-t_i)}\frac{A_{t_{i+1}}}{A_{t_i}}\right)\right)\\
&= \alpha\overline{\beta_{i+1}}e^{r(t_{i+1}-t_0)} + \left(V_{t_i}-\alpha e^{r(t_i-t_0)}\overline{\beta_{i+1}}\right)\frac{A_{t_{i+1}}}{A_{t_i}}\\
&= \alpha e^{r(t_i-t_0)}\left[e^{r(t_{i+1}-t_i)}\overline{\beta_{i+1}} - \frac{A_{t_{i+1}}}{A_{t_i}}\overline{\beta_{i+1}}\right] + V_{t_i}\frac{A_{t_{i+1}}}{A_{t_i}}\\
&= e^{r(t_i-t_0)}\left[\overline{\beta_{i+1}}\left(\alpha e^{r(t_{i+1}-t_i)} - \alpha\frac{A_{t_{i+1}}}{A_{t_i}}\right)\right] + V_{t_i}\frac{A_{t_{i+1}}}{A_{t_i}}\left(\overline{\beta_{i+1}} + \left(1-\overline{\beta_{i+1}}\right)\right)\\
&= e^{r(t_i-t_0)}\left[\overline{\beta_{i+1}}\left(\alpha e^{r(t_{i+1}-t_i)} - \alpha\frac{A_{t_{i+1}}}{A_{t_i}} + V_{t_i}\frac{A_{t_{i+1}}}{A_{t_i}}\right)\right] + V_{t_i}\frac{A_{t_{i+1}}}{A_{t_i}}\left(1-\overline{\beta_{i+1}}\right)\\
&= e^{r(t_i-t_0)}\left[\overline{\beta_{i+1}}\left(\alpha e^{r(t_{i+1}-t_i)} + \frac{A_{t_{i+1}}}{A_{t_i}}(e^{-r(t_i-t_0)}V_{t_i}-\alpha)\right) + \left(1-\overline{\beta_{i+1}}\right)\left(e^{-r(t_i-t_0)}V_{t_i}\frac{A_{t_{i+1}}}{A_{t_i}}\right)\right]\\
&= \overline{\beta_{i+1}}\left(\alpha e^{r(t_{i+1}-t_0)} + \frac{A_{t_{i+1}}}{A_{t_i}}\left(V_{t_i}-\alpha e^{r(t_i-t_0)}\right)\right) + \left(1-\overline{\beta_{i+1}}\right)V_{t_i}\frac{A_{t_{i+1}}}{A_{t_i}}.
\end{aligned}$$

C.4: Proof of Proposition 3.5

Using the representation of L_1, we find that

$$\begin{aligned}
L_1 - \frac{A_1}{A_0} &= V_1 + V_1 P^{BS}\left(1,1,\frac{G_2}{V_1},\sigma_{A,1}\right) - \frac{A_1}{A_0}\\
&= V_1\left(1 + P^{BS}\left(1,1,\frac{G_2}{V_1},\sigma_{A,1}\right)\right) - \frac{A_1}{A_0}.
\end{aligned}$$

With $V_1 = \alpha(\beta\frac{A_1}{A_0} + (1-\beta)e^r)$ and some simple calculations it follows

$$\begin{aligned}
L_1 - \frac{A_1}{A_0} =& \beta\left(\alpha\frac{A_1}{A_0}P^{BS}\left(1,1,\frac{G_2}{V_1},\sigma_{A,1}\right) - (1-\alpha)\frac{A_1}{A_0}\right) +\\
&(1-\beta)\left(\alpha e^r P^{BS}\left(1,1,\frac{G_2}{V_1},\sigma_{A,1}\right) + \alpha e^r - \frac{A_1}{A_0}\right).
\end{aligned}$$

Recall that $(1-\alpha) = GC_0$ and thus $L_1 - \frac{A_1}{A_0}$ can be rewritten as

$$\begin{aligned} L_1 - \frac{A_1}{A_0} =& \alpha \left(\beta \frac{A_1}{A_0} + (1-\beta) e^r \right) P^{BS} \left(1, 1, \frac{G_2}{V_1}, \sigma_{A,1} \right) - \\ & \left(\beta \frac{A_1}{A_0} GC_0 + (1-\beta) e^r GC_0 \right) + (1-\beta) \left(e^r - \frac{A_1}{A_0} \right) \\ =& GC_1 - \frac{V_1}{\alpha} GC_0 + (1-\beta) \left(e^r - \frac{A_1}{A_0} \right). \end{aligned}$$

C.5: Proof of Corollary 3.1

The general formula for the random variable $L_1 - \frac{A_1}{A_0}$ is received by inserting the corresponding formulas for the CMR case from the guarantee section 3. For the special case of a postponed premium payment it holds $\frac{V_1}{\alpha} = e^r$ and thus

$$e^r GC_0^{(\beta=0),\,CMR} = \mathbb{E}_{\mathbb{P}^*}[GC_1^{(\beta=0),\,CMR}].$$

Furthermore, it holds

$$GC_1^{(\beta=0),\,CMR} = P^{BS}(e^r \alpha, 1, G_2, \sigma_{A,0}).$$

This is a deterministic value s.t.

$$e^r GC_0^{(\beta=0),\,CMR} = GC_1^{(\beta=0),\,CMR}.$$

Inserting the result in the general formula gives the claim. For the upfront premium case insert the definition of V_1 with $\beta = 1$ and the claimed result holds.

C.6: Proof of Proposition 3.7

ad a):

$$\begin{aligned} \mathbb{E}_{\mathbb{P}}[u(L_2)] &= \mathbb{E}_{\mathbb{P}} \left[u \left(\max \left\{ \alpha \frac{A_2}{A_0}, G_2 \right\} \right) \right] \\ &= \mathbb{E}_{\mathbb{P}} \left[u \left(\alpha \frac{A_2}{A_0} 1_{\left\{ \alpha \frac{A_2}{A_0} > G_2 \right\}} + G_2 1_{\left\{ \alpha \frac{A_2}{A_0} \leq G_2 \right\}} \right) \right] \\ &= \mathbb{E}_{\mathbb{P}} \left[u \left(\alpha \frac{A_2}{A_0} 1_{\left\{ \alpha \frac{A_2}{A_0} > G_2 \right\}} \right) \right] + \mathbb{E}_{\mathbb{P}} \left[u \left(G_2 1_{\left\{ \alpha \frac{A_2}{A_0} \leq G_2 \right\}} \right) \right] \\ &= E_1 + E_2, \end{aligned}$$

where

$$E_1 := \mathbb{E}_{\mathbb{P}}\left[u\left(\alpha\frac{A_2}{A_0}1_{\left\{\alpha\frac{A_2}{A_0}>G_2\right\}}\right)\right],$$
$$E_2 := \mathbb{E}_{\mathbb{P}}\left[u\left(G_2 1_{\left\{\alpha\frac{A_2}{A_0}\le G_2\right\}}\right)\right].$$

Let us first calculate the value of E_1 and recall, that $u(x) = \frac{x^{(1-\gamma)}}{1-\gamma}$.

$$\begin{aligned} E_1 &= \frac{1}{1-\gamma}\mathbb{E}_{\mathbb{P}}\left[\left(\alpha\frac{A_2}{A_0}\right)^{(1-\gamma)} 1_{\left\{\alpha\frac{A_2}{A_0}>G_2\right\}}\right] \\ &= \frac{\alpha^{1-\gamma}}{1-\gamma}\mathbb{E}_{\mathbb{P}}\left[e^{(2\mu_A-\sigma_A^2)(1-\gamma)+(1-\gamma)\sigma_A W_2}\ 1_{\left\{\alpha\frac{A_2}{A_0}>G_2\right\}}\right], \end{aligned} \tag{4.55}$$

where the last equality holds because $\frac{A_2}{A_0} = e^{2\mu_A-\sigma_A^2+\sigma_A W_2}$. Now our aim is to use Girsanov's theorem to calculate the expected value. For this we need the Radon-Nikodym density which is in our setting given by

$$Z_2 := e^{-\sigma_A^2(1-\gamma)^2+\sigma_A(1-\gamma)W_2}.$$

Rewriting (4.55), we get

$$\begin{aligned} E_1 &= \frac{\alpha^{1-\gamma}}{1-\gamma}e^{(2\mu_A-\sigma_A^2)(1-\gamma)}e^{(1-\gamma)^2\sigma_A^2}\ \mathbb{E}_{\mathbb{P}}\left[e^{-\sigma_A^2(1-\gamma)^2+\sigma_A(1-\gamma)W_2}1_{\left\{\alpha\frac{A_2}{A_0}>G_2\right\}}\right] \\ &= \frac{\alpha^{1-\gamma}}{1-\gamma}e^{(2\mu_A-\sigma_A^2)(1-\gamma)}e^{(1-\gamma)^2\sigma_A^2}\ \mathbb{E}_{\mathbb{P}}\left[Z_2 1_{\left\{\alpha\frac{A_2}{A_0}>G_2\right\}}\right]. \end{aligned} \tag{4.56}$$

With Girsanov's Theorem,

$$\mathbb{E}_{\mathbb{P}}\left[Z_2 1_{\left\{\alpha\frac{A_2}{A_0}>G_2\right\}}\right] = \tilde{\mathbb{P}}\left(\alpha\frac{A_2}{A_0} > G_2\right),$$

where $\tilde{\mathbb{P}}$ is the uniquely determined equivalent martingale measure of $\mathbb{P}$ and $\tilde{W}_T$ is a BM under $\tilde{\mathbb{P}}$ given by $\tilde{W}_2 = W_2 - 2(1-\gamma)\sigma_A$. Using this result, we get

$$\begin{aligned} \tilde{\mathbb{P}}\left(\alpha\frac{A_2}{A_0} > G_2\right) &= 1-\tilde{\mathbb{P}}\left(\frac{A_2}{A_0} \le \frac{G_2}{\alpha}\right) = 1-\tilde{\mathbb{P}}\left(\frac{\tilde{W}_2}{\sqrt{2}} \le \frac{\ln(\frac{G_2}{\alpha}) - 2[\mu_A - (\gamma-\frac{1}{2})\sigma_A^2]}{\sqrt{2}\ \sigma_A}\right) \\ &= 1-\Phi\left(\frac{\ln(\frac{G_2}{\alpha}) - 2[\mu_A - \sigma_A^2(\gamma-\frac{1}{2})]}{\sqrt{2}\sigma_A}\right). \end{aligned} \tag{4.57}$$

Combining (4.56) and (4.57), it holds

$$E_1 = \frac{\alpha^{1-\gamma}}{1-\gamma} e^{(1-\gamma)(2\mu_A - \gamma\sigma_A^2)} \left\{ 1 - \Phi\left(\frac{\ln(\frac{G_2}{\alpha}) - 2[\mu_A - \sigma_A^2(\gamma - \frac{1}{2})]}{\sqrt{2}\sigma_A} \right) \right\}. \tag{4.58}$$

It remains to calculate E_2:

$$\begin{aligned} E_2 &= \frac{1}{1-\gamma} G_2^{(1-\gamma)} \mathbb{E}_{\mathbb{P}} \left[1_{\left\{\alpha \frac{A_2}{A_0} \le G_2\right\}} \right] = \frac{1}{1-\gamma} G_2^{(1-\gamma)} \ \mathbb{P}\left(\alpha \frac{A_2}{A_0} \le G_2 \right) \\ &= \frac{1}{1-\gamma} G_2^{(1-\gamma)} \ \mathbb{P}\left(\frac{W_2}{\sqrt{2}} \le \frac{\ln(\frac{G_2}{\alpha}) - 2\mu_A + \sigma_A^2}{\sqrt{2}\sigma_A} \right) \\ &= \frac{1}{1-\gamma} G_2^{(1-\gamma)} \ \Phi\left(\frac{\ln(\frac{G_2}{\alpha}) - 2\mu_A + \sigma_A^2}{\sqrt{2}\sigma_A} \right). \end{aligned} \tag{4.59}$$

Combining (4.58) and (4.59) we get the final solution.
For the CE just use the relation $CE = u^{-1}(\mathbb{E}_{\mathbb{P}}[u(L_2)])$ with $u^{-1}(x) = ((1-\gamma)x)^{\frac{1}{1-\gamma}}$.

ad (b):

$$\begin{aligned} \mathbb{E}_{\mathbb{P}}[u(L_2)] &= \mathbb{E}_{\mathbb{P}}\left[u\left(\max\left\{ \alpha e^r \frac{A_2}{A_1}, G_2 \right\} \right) \right] \\ &\ \vdots \\ &= \mathbb{E}_{\mathbb{P}}[E_3] + \mathbb{E}_{\mathbb{P}}[E_4], \end{aligned}$$

where

$$\begin{aligned} E_3 &:= \frac{\alpha^{1-\gamma}}{1-\gamma} \mathbb{E}_{\mathbb{P}} \left[\left(e^r \frac{A_2}{A_1} \right)^{1-\gamma} 1_{\left\{\alpha e^r \frac{A_2}{A_1} > G_2\right\}} \right] \\ E_4 &:= \frac{1}{1-\gamma} \mathbb{E}_{\mathbb{P}} \left[(G_2)^{1-\gamma} \, 1_{\left\{\alpha e^r \frac{A_2}{A_1} \le G_2\right\}} \right]. \end{aligned}$$

Let us calculate E_3 where we use that $\frac{A_2}{A_1} = e^{\mu_A - \frac{1}{2}\sigma_A^2 + \sigma_A(W_2 - W_1)}$. So it holds

$$\begin{aligned} E_3 &= \frac{\alpha^{1-\gamma}}{1-\gamma} \mathbb{E}_{\mathbb{P}} \left[\left(e^r \frac{A_2}{A_1} \right)^{1-\gamma} 1_{\left\{\alpha e^r \frac{A_2}{A_1} > G_2\right\}} \right] \\ &= \frac{\alpha^{1-\gamma}}{1-\gamma} e^{r(1-\gamma)} e^{(1-\gamma)(\mu_A - \frac{1}{2}\sigma_A^2)} \mathbb{E}_{\mathbb{P}} \left[e^{\sigma_A(1-\gamma)(W_2 - W_1)} 1_{\left\{\alpha e^r \frac{A_2}{A_1} > G_2\right\}} \right] \end{aligned} \tag{4.60}$$

As shown in the proof of part (a), we now want to use Girsanov's Theorem. The Radon-Nikodym density here is given by

$$\hat{Z}_2 := e^{-\frac{1}{2}(1-\gamma)^2\sigma_A^2 + \sigma_A(1-\gamma)(W_2 - W_1)}.$$

Rewriting (4.60) and with the fact that

$$\sigma_A(\hat{W}_2 - \hat{W}_1) = \sigma_A(W_2 - W_1) - \sigma_A^2(1-\gamma),$$

where $\hat{W}_T$ is a BM under the uniquely determined equivalent martingale measure $\hat{\mathbb{P}}$, we get

$$\begin{aligned}
E_3 &= \frac{\alpha^{1-\gamma}}{1-\gamma} e^{r(1-\gamma)} e^{(1-\gamma)(\mu_A - \frac{1}{2}\gamma\sigma_A^2)} \mathbb{E}_{\mathbb{P}}\left[\hat{Z}_2 1_{\left\{\alpha e^r \frac{A_2}{A_1} > G_2\right\}}\right] \\
&= \frac{\alpha^{1-\gamma}}{1-\gamma} e^{r(1-\gamma)} e^{(1-\gamma)(\mu_A - \frac{1}{2}\gamma\sigma_A^2)} \left\{1 - \hat{\mathbb{P}}\left(\alpha e^r \frac{A_2}{A_1} \le G_2\right)\right\} \\
&\vdots \\
&= \frac{\alpha^{1-\gamma}}{1-\gamma} e^{(1-\gamma)(r+\mu_A - \frac{1}{2}\gamma\sigma_A^2)} \left\{1 - \Phi\left(\frac{\ln(\frac{G_2}{\alpha}) - r - \mu_A - \sigma_A^2(\frac{1}{2}-\gamma)}{\sigma_A}\right)\right\} \qquad (4.61)
\end{aligned}$$

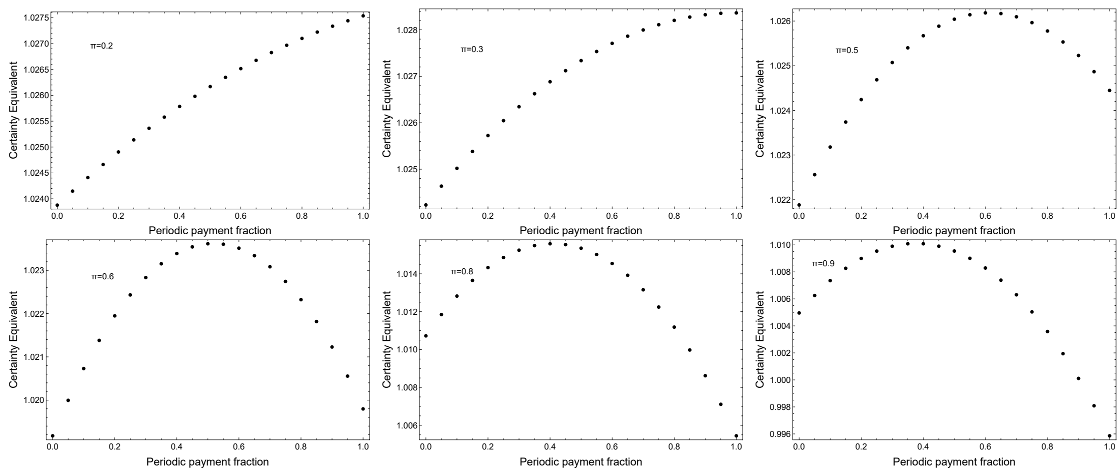
It remains to calculate E_4:

$$\begin{aligned}
E_4 &= \frac{1}{1-\gamma}\mathbb{E}_{\mathbb{P}}\left[(G_2)^{1-\gamma} 1_{\left\{\alpha e^r \frac{A_2}{A_1} \le G_2\right\}}\right] \\
&= \frac{1}{1-\gamma} G_2^{(1-\gamma)} \mathbb{P}\left(\alpha e^r \frac{A_2}{A_1} \le G_2\right) \\
&= \frac{1}{1-\gamma} G_2^{(1-\gamma)} \mathbb{P}\left(W_2 - W_1 \le \frac{\ln(\frac{G_2}{\alpha}) - r - \mu_A + \frac{1}{2}\sigma_A^2}{\sigma_A}\right) \\
&= \frac{1}{1-\gamma} G_2^{(1-\gamma)} \Phi\left(\frac{\ln(\frac{G_2}{\alpha}) - r - \mu_A + \frac{1}{2}\sigma_A^2}{\sigma_A}\right). \qquad (4.62)
\end{aligned}$$

Combining (4.61) and (4.62) we get the final result.
The certainty equivalent can be calculated analogously to part (a). □

C.7: Supplementary figure - Certainty Equivalent VMR no guarantee

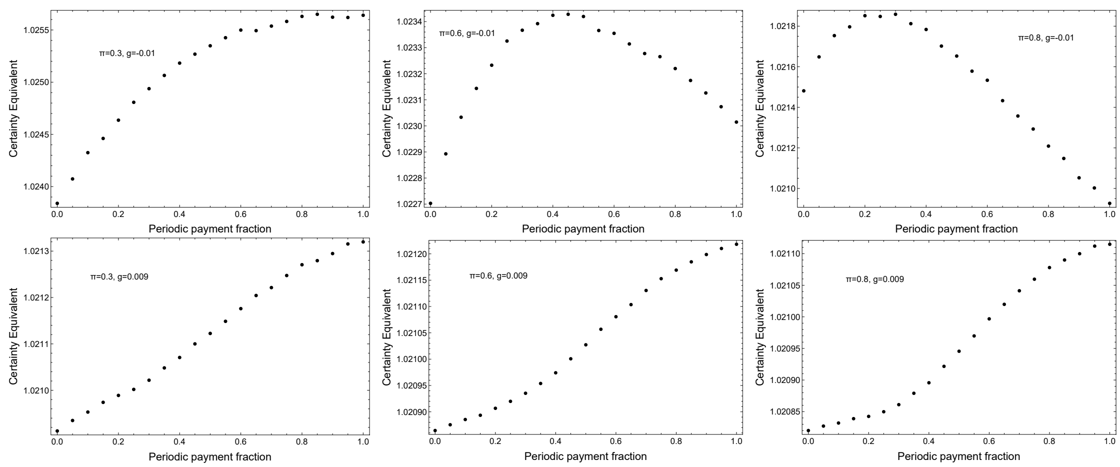
Impact of variable management rule on Merton solution without guarantee



The pictures show the impact of the splitting factor β and the variable management rule on the certainty equivalent for the case of no guarantee. In addition to the benchmark parameter setting, we use $\gamma = 4$ and $e_1 = 0.8, e_2 = 1$. The value of e_3 is chosen depending on the investment fraction s.t. Assumption $(*)$ is fulfilled. The corresponding values of π are stated in the pictures.

C.8: Supplementary figure - Certainty Equivalent VMR no guarantee

Impact of guarantee rate on variable management rules



The pictures show the impact of the guarantee rate on the optimal splitting factors β for a variable portfolio insurance management rule. In addition to the benchmark parameter setting, we use $\gamma = 4$ and $e_1 = 0.8, e_2 = 1$. The value of e_3 is chosen depending on the investment fraction s.t. Assumption $(*)$ is fulfilled. The corresponding values of π and g are stated in the pictures.

Appendix D: Proofs of Chapter 4

D.1: Proof of Proposition 4.1

For $\gamma > 1$, the certainty equivalent savings rate $y_{T,p}(\pi)$ is given by

$$y_{T,p}(\pi) = \frac{1}{(1-\gamma)T} \ln\left[pe^{y(\pi,1)(1-\gamma)T} + (1-p)e^{y(\pi,2)(1-\gamma)T}\right].$$

For the partial derivative with respect to T, it holds that

$$\begin{aligned}
\frac{\partial y_{T,p}(\pi)}{\partial T} &= \frac{1}{T}\frac{pe^{y(\pi,1)(1-\gamma)T}y(\pi,1) + (1-p)e^{y(\pi,2)(1-\gamma)T}y(\pi,2)}{pe^{y(\pi,1)(1-\gamma)T} + (1-p)e^{y(\pi,2)(1-\gamma)T}} \\
&\quad - \frac{1}{(1-\gamma)T^2}\ln\left[pe^{y(\pi,1)(1-\gamma)T} + (1-p)e^{y(\pi,2)(1-\gamma)T}\right] \\
&= \frac{1}{(1-\gamma)T^2}\frac{pe^{y(\pi,1)(1-\gamma)T}y(\pi,1)(1-\gamma)T + (1-p)e^{y(\pi,2)(1-\gamma)T}y(\pi,2)(1-\gamma)T}{pe^{y(\pi,1)(1-\gamma)T} + (1-p)e^{y(\pi,2)(1-\gamma)T}} \\
&\quad - \frac{1}{(1-\gamma)T^2}\ln\left[pe^{y(\pi,1)(1-\gamma)T} + (1-p)e^{y(\pi,2)(1-\gamma)T}\right] \\
&= \frac{1}{(1-\gamma)T^2}\frac{pe^{x_1}x_1 + (1-p)e^{x_2}x_2}{pe^{x_1} + (1-p)e^{x_2}} - \frac{1}{(1-\gamma)T^2}\ln\left[pe^{x_1} + (1-p)e^{x_2}\right],
\end{aligned}$$

where we define $x_i = y(\pi,i)(1-\gamma)T$. We can then write the derivative as a function of the random variable X with realizations x_1 (with probability p) and x_2 (with probability $(1-p)$):

$$\begin{aligned}
\frac{\partial y_{T,p}(\pi)}{\partial T} &= \frac{1}{(1-\gamma)T^2}\left[\frac{E\left[e^X X\right]}{E\left[e^X\right]} - \ln E\left[e^X\right]\right] \\
&= \frac{1}{(1-\gamma)T^2 E\left[e^X\right]}\left\{\underbrace{E\left[e^X X\right] - E\left[e^X\right]\ln E\left[e^X\right]}_{>0 \quad (z\ln z \text{ is convex function for } z>0)}\right\} \\
&< 0.
\end{aligned}$$

Next, we turn to the certainty equivalent savings rate for the optimal pre-commitment strategy where we have to take into account that the optimal pre-commitment strategy depends on T:

$$\begin{aligned}
\frac{\partial y_{T,p}(\pi^{*,pre}_{T,p})}{\partial T} &= \left.\frac{\partial y_{T,p}(\pi)}{\partial T}\right|_{\pi=\pi^{*,pre}_{T}} + \underbrace{\left.\frac{\partial y_{T,p}(\pi)}{\partial \pi}\right|_{\pi=\pi^{*,pre}_{T,p}}}_{=0 \quad \text{(FOC)}}\frac{\partial \pi^{*,pre}_{T,p}}{\partial T} \\
&< 0.
\end{aligned}$$

D.2: Proof of Proposition 4.2

We need to show that $\pi_{T,p}^{*,pre} = \alpha_{T,p}^{*}\pi_1^{Mer}+(1-\alpha_{T,p}^{*})\pi_2^{Mer}$, where $\pi_{T,p}^{*,pre} = \underset{\pi}{\operatorname{argmax}}\{y_{T,p}(\pi)\}$. For $\gamma > 1$ the certainty equivalent savings rate $y_{T,p}(\pi)$ is given by

$$y_{T,p}(\pi) = \frac{1}{(1-\gamma)T} \ln \left(p e^{y(\pi,1)(1-\gamma)T} + (1-p) e^{y(\pi,2)(1-\gamma)T} \right).$$

Calculating the FOC, we receive

$$\begin{aligned} \frac{\partial y_{T,p}}{\partial \pi} &= \frac{p\gamma\sigma_1^2(\pi_1^{Mer}-\pi)e^{y(\pi,1)(1-\gamma)T} + (1-p)\gamma\sigma_2^2(\pi_2^{Mer}-\pi)e^{y(\pi,2)(1-\gamma)T}}{p e^{y(\pi,1)(1-\gamma)T} + (1-p)e^{y(\pi,2)(1-\gamma)T}} \overset{!}{=} 0 \\ \Leftrightarrow \pi &\overset{!}{=} \frac{p\sigma_1^2\pi_1^{Mer}e^{y(\pi,1)(1-\gamma)T} + (1-p)\sigma_2^2\pi_2^{Mer}e^{y(\pi,2)(1-\gamma)T}}{p\sigma_1^2 e^{y(\pi,1)(1-\gamma)T} + (1-p)\sigma_2^2 e^{y(\pi,2)(1-\gamma)T}}, \end{aligned} \tag{4.63}$$

i.e. the optimal pre-commitment strategy $\pi_{T,p}^{*,pre}$ is implicitly defined as solution of (4.63). Separating the fraction leads to

$$\begin{aligned} \pi_{T,p}^{*,pre} =& \frac{p\sigma_1^2 e^{y(\pi_{T,p}^{*,pre},1)(1-\gamma)T}}{p\sigma_1^2 e^{y(\pi_{T,p}^{*,pre},1)(1-\gamma)T} + (1-p)\sigma_2^2 e^{y(\pi_{T,p}^{*,pre},2)(1-\gamma)T}}\pi_1^{Mer} + \\ & \frac{p\sigma_1^2 e^{y(\pi_{T,p}^{*,pre},1)(1-\gamma)T}}{p\sigma_1^2 e^{y(\pi_{T,p}^{*,pre},1)(1-\gamma)T} + (1-p)\sigma_2^2 e^{y(\pi_{T,p}^{*,pre},2)(1-\gamma)T}}\pi_2^{Mer} \\ \Leftrightarrow \pi_{T,p}^{*,pre} =& \alpha_{T,p}^{*}\pi_1^{Mer} + (1-\alpha_{T,p}^{*})\pi_2^{Mer}. \end{aligned}$$

D.3: Proof of Proposition 4.5

The value of information is given by the difference of the certainty equivalent savings rates $y_{T,p}(\pi^*) - y_{T,p}(\pi^{*,\mathrm{pre}})$. For $\gamma > 1$ it holds

$$y_{T,p}(\pi^*) - y_{T,p}(\pi^{*,\mathrm{pre}}) = \frac{1}{(1-\gamma)T} \ln \left[\frac{p e^{y_{T,p}(\pi_1^{Mer},1)(1-\gamma)T} + (1-p)e^{y_{T,p}(\pi_2^{Mer},2)(1-\gamma)T}}{p e^{y_{T,p}(\pi_{T,p}^{*,pre},1)(1-\gamma)T} + (1-p)e^{y_{T,p}(\pi_{T,p}^{*,pre},2)(1-\gamma)T}} \right].$$

The inner part of the log function can be written as

$$\begin{aligned} & \frac{p e^{y_{T,p}(\pi_{T,p}^{*,pre},1)(1-\gamma)T}}{p e^{y_{T,p}(\pi_{T,p}^{*,pre},1)(1-\gamma)T} + (1-p)e^{y_{T,p}(\pi_{T,p}^{*,pre},2)(1-\gamma)T}} \frac{p e^{y_{T,p}(\pi_1^{Mer},1)(1-\gamma)T}}{p e^{y_{T,p}(\pi_{T,p}^{*,pre},1)(1-\gamma)T}} + \\ & \frac{(1-p)e^{y_{T,p}(\pi_{T,p}^{*,pre},2)(1-\gamma)T}}{p e^{y_{T,p}(\pi_{T,p}^{*,pre},1)(1-\gamma)T} + (1-p)e^{y_{T,p}(\pi_{T,p}^{*,pre},2)(1-\gamma)T}} \frac{(1-p)e^{y_{T,p}(\pi_2^{Mer},2)(1-\gamma)T}}{(1-p)e^{y_{T,p}(\pi_{T,p}^{*,pre},2)(1-\gamma)T}} \\ =& y_{T,p}(\pi_{T,p}^{*,pre},1)e^{l\left(\pi_{T,p}^{*,pre},1\right)(1-\gamma)T} + (1-\beta_{T,p}(\gamma))e^{l\left(\pi_{T,p}^{*,pre},2\right)(1-\gamma)T}. \end{aligned}$$

For $\gamma = 1$ the value of information $y_{T,p}(\pi^*) - y_{T,p}(\pi^{*,\text{pre}})$ is given by

$$\begin{aligned} y_{T,p}(\pi^*) - y_{T,p}(\pi^{*,\text{pre}}) =& p(y_{T,p}(\pi_1^{Mer}, 1) - y_{T,p}(\pi_{T,p}^{*,pre}, 1)) + \\ & (1-p)(y_{T,p}(\pi_2^{Mer}, 2) - y_{T,p}(\pi_{T,p}^{*,pre}, 2)). \end{aligned}$$

Using the fact that $\beta_{T,p}(1) = p$ gives the claimed representation.

D.4: Proof of Proposition 4.6

For $\gamma > 1$ it holds

$$\lim_{T\to 0} \left\{ y_{T,p}(\pi^*) - y_{T,p}(\pi_{T,p}^{*,pre}) \right\} = \lim_{T\to 0} y_{T,p}(\pi^*) - \lim_{T\to 0} y_{T,p}(\pi_{T,p}^{*,pre}).$$

Notice that

$$\begin{aligned} \lim_{T\to 0} y_{T,p}(\pi_{T,p}^{*,pre}) &= \lim_{T\to 0} \left\{ py(\pi_{T,p}^{*,pre}, 1) + (1-p)y(\pi_{T,p}^{*,pre}, 2) \right\} \\ &= py\left(\lim_{T\to 0} \pi_{T,p}^{*,pre}, 1 \right) + (1-p)y\left(\lim_{T\to 0} \pi_{T,p}^{*,pre}, 2 \right) \end{aligned}$$

and

$$\begin{aligned} &\lim_{T\to 0} \pi_{T,p}^{*,pre} = \lim_{T\to 0} \alpha_{T,p}^* \pi_1^{Mer} + (1 - \lim_{T\to 0} \alpha_{T,p}^*)\pi_2^{Mer}, \text{ where} \\ &\lim_{T\to 0} \alpha_{T,p}^* = \frac{p\sigma_1^2}{p\sigma_1^2 + (1-p)\sigma_2^2}. \end{aligned}$$

Using the results of Proposition 4.1 and the fact that

$$l(\pi_{T,p}^{*,pre}, i) = y(\pi_i^{Mer}, i) - y(\pi_{T,p}^{*,pre}, i) = \frac{1}{2}\gamma\sigma_i^2(\pi_{T,p}^{*,pre} - \pi_i^{Mer})^2,$$

we get

$$\lim_{T\to 0} \left\{ y_{T,p}(\pi^*) - y_{T,p}(\pi_{T,p}^{*,pre}) \right\} = pl(\lim_{T\to 0} \pi_{T,p}^{*,pre}, 1) + (1-p)l(\lim_{T\to 0} \pi_{T,p}^{*,pre}, 2). \tag{4.64}$$

Calculating the two loss rates, we receive with the above stated results

$$l(\lim_{T\to 0} \pi_{T,p}^{*,pre}, 1) = \frac{1}{2}\gamma(1-p)^2\sigma_2^2 \left(\frac{\sigma_1\sigma_2}{p\sigma_1^2 + (1-p)\sigma_2^2} \right)^2 \left(\pi_1^{\text{Mer}} - \pi_2^{\text{Mer}} \right)^2 \tag{4.65}$$

$$l(\lim_{T\to 0} \pi_{T,p}^{*,pre}, 2) = \frac{1}{2}\gamma p^2\sigma_1^2 \left(\frac{\sigma_1\sigma_2}{p\sigma_1^2 + (1-p)\sigma_2^2} \right)^2 \left(\pi_1^{\text{Mer}} - \pi_2^{\text{Mer}} \right)^2. \tag{4.66}$$

Combining (4.64), (4.65) and (4.66) we finally get

$$\begin{aligned}\lim_{T\to 0}\left\{y_{T,p}(\pi^*)-y_{T,p}(\pi^{*,pre}_{T,p})\right\} =&\frac{1}{2}\gamma\left(\pi_1^{\text{Mer}}-\pi_2^{\text{Mer}}\right)^2\frac{p^2(1-p)\sigma_1^2+p(1-p)^2\sigma_2^2}{\left(p\sigma_1^2+(1-p)\sigma_2^2\right)^2}\;\sigma_1^2\sigma_2^2\\ =&\frac{1}{2}\gamma p(1-p)\left(\pi_1^{\text{Mer}}-\pi_2^{\text{Mer}}\right)^2\frac{\sigma_1^2\sigma_2^2}{p\sigma_1^2+(1-p)\sigma_2^2}.\end{aligned}$$

For $\gamma = 1$ we are immediately in the situation of equation (4.64), s.t. the same result holds.

For the case $T \to \infty$ we distinguish between:

$$\lim_{T\to\infty}\pi^{*,pre}_{T,p}\neq\pi^{equal}\ \text{and}\ \lim_{T\to\infty}\pi^{*,pre}_{T,p}=\pi^{equal}.$$

For $\lim_{T\to\infty}\pi^{*,pre}_{T,p}\neq\pi^{equal}$ it holds with equation (4.13)

$$\begin{aligned}&\lim_{T\to\infty}\{y_{T,p}(\pi^*)-y_{T,p}(\pi^{*,pre})\}\\ &=\lim_{T\to\infty}\left\{\frac{1}{(1-\gamma)T}\ln\left[\frac{pe^{(y(\pi_1^{Mer},1)-y(\pi^{*,pre}_{T,p},1))(1-\gamma)T}+(1-p)e^{(y(\pi_2^{Mer},2)-y(\pi^{*,pre}_{T,p},2))(1-\gamma)T}}{1-p\delta(\pi^{*,pre}_{T,p},T)}\right]\right\}\\ &=\lim_{T\to\infty}\left\{\frac{1}{(1-\gamma)T}\right\}\ln\left[\frac{1}{1-p}\right]=0.\end{aligned}$$

For $\pi^{*,pre}_{T,p}=\pi^{equal}$ it holds that $y_{T,p}(\pi^{*,pre}_{T,p},1)=y_{T,p}(\pi^{*,pre}_{T,p},2)$ (for this we write $y(\pi^{equal},\cdot)$ because the regime i does not matter here), s.t. $\delta(\pi^{*,pre}_{T,p},T)=0$, for all T and thus with equation (4.13) it holds

$$\begin{aligned}\lim_{T\to\infty}\{y_{T,p}(\pi^*)-y_{T,p}(\pi^{*,pre})\}&=\lim_{T\to\infty}\left\{y_{T,p}(\pi^*)-y(\pi^{equal},\cdot)\right\}\\ &=\min\left\{y(\pi_1^{Mer},1),y(\pi_2^{Mer},2)\right\}-y(\pi^{equal},\cdot).\end{aligned}$$

D.6: Proof of Proposition 4.8

We want to show that $\pi_T^{*,pre,amb}=\underset{\pi}{\operatorname{argmax}}\left\{y^{amb}_{T,\tilde{p}}(\pi)\right\}$ is given by

$$\alpha^*_{T,\tilde{p}}\left(\alpha_{T,p_a}\pi_1^{Mer}+(1-\alpha_{T,p_a})\pi_2^{Mer}\right)+(1-\alpha_{T,\tilde{p}})\left(\alpha_{T,p_b}\pi_1^{Mer}+(1-\alpha_{T,p_b})\pi_2^{Mer}\right).$$

The certainty equivalent savings rate for $\gamma > 1$ with ambiguity $y^{amb}_{T,p}(\pi)$ is given by

$$y^{\text{amb}}_{T,\tilde{p}}(\pi):=\frac{1}{(1-\eta)T}\ln\left[\tilde{p}e^{y_{T,p_a}(\pi)(1-\eta)T}+(1-\tilde{p})e^{y_{T,p_b}(\pi)(1-\eta)T}\right].$$

Calculating the FOC we receive

$$\frac{\partial y^{\text{amb}}_{T,\tilde{p}}(\pi)}{\partial \pi} = \frac{\tilde{p}\frac{\partial y_{T,p_a}(\pi)}{\partial \pi} e^{y_{T,p_a}(\pi)(1-\eta)T} + (1-\tilde{p})\frac{\partial y_{T,p_b}(\pi)}{\partial \pi} e^{y_{T,p_b}(\pi)(1-\eta)T}}{\tilde{p}e^{y_{T,p_a}(\pi)(1-\eta)T} + (1-\tilde{p})e^{y_{T,p_b}(\pi)(1-\eta)T}} \overset{!}{=} 0. \qquad (4.67)$$

Within the results in the proof of Proposition 4.2 it furthermore holds

$$\frac{\partial y_{T,p_i}(\pi)}{\partial \pi} = \frac{p_i \gamma \sigma_1^2(\pi_1^{Mer} - \pi)e^{y(\pi,1)(1-\gamma)T} + (1-p_i)\gamma\sigma_2^2(\pi_2^{Mer} - \pi)e^{y(\pi,2)(1-\gamma)T}}{p_i e^{y(\pi,1)(1-\gamma)T} + (1-p_i)e^{y(\pi,2)(1-\gamma)T}}, \text{ for } i = a, b.$$

Using this result, we can solve the FOC (4.67) for π and can formulate that $\pi_T^{*,pre,amb}$ has to fulfill the equation

$$\begin{aligned}\pi =& \frac{\tilde{p}}{1-\tilde{p}}\left(\frac{p_a e^{y(\pi,1)(1-\gamma)T} + (1-p_a)e^{y(\pi,2)(1-\gamma)T}}{p_b e^{y(\pi,1)(1-\gamma)T} + (1-p_b)e^{y(\pi,2)(1-\gamma)T}}\right)^{\frac{\eta-\gamma}{\gamma-1}} \times \\ & \frac{p_a\sigma_1^2 e^{y(\pi,1)(1-\gamma)T}(\pi_1^{Mer} - \pi) + (1-p_a)\sigma_2^2 .6^{y(\pi,2)(1-\gamma)T}(\pi_2^{Mer} - \pi)}{p_b\sigma_1^2 e^{y(\pi,1)(1-\gamma)T} + (1-p_b)\sigma_2^2 e^{y(\pi,1)(1\gamma)T}} + \\ & \frac{p_b\sigma_1^2\pi_1^{Mer} e^{y(\pi,1)(1-\gamma)T} + (1-p_b)\sigma_2^2\pi_2^{Mer} e^{y(\pi,2)(1-\gamma)T}}{p_b\sigma_1^2 e^{y(\pi,1)(1-\gamma)T} + (1-p_b)\sigma_2^2 e^{y(\pi,2)(1-\gamma)T}}.\end{aligned}$$

Simplifying this equation we receive

$$\begin{aligned}\pi =& \frac{\tilde{p}\xi_a}{\tilde{p}\xi_a + (1-\tilde{p})\xi_b}\left[\frac{p_a\sigma_1^2 e^{y(\pi,1)(1-\gamma)T}\pi_1^{Mer} + (1-p_a)\sigma_2^2 e^{y(\pi,2)(1-\gamma)T}\pi_2^{Mer}}{p_a\sigma_1^2 e^{y(\pi,1)(1-\gamma)T} + (1-p_a)\sigma_2^2 e^{y(\pi,2)(1-\gamma)T}}\right] + \\ & \left(1 - \frac{\tilde{p}\xi_a}{\tilde{p}\xi_a + (1-\tilde{p})\xi_b}\right)\left[\frac{p_b\sigma_1^2 e^{y(\pi,1)(1-\gamma)T}\pi_1^{Mer} + (1-p_b)\sigma_2^2 e^{y(\pi,2)(1-\gamma)T}\pi_2^{Mer}}{p_b\sigma_1^2 e^{y(\pi,1)(1-\gamma)T} + (1-p_b)\sigma_2^2 e^{y(\pi,2)(1-\gamma)T}}\right], \text{ where}\end{aligned}$$

$$\begin{aligned}\xi_a &= \left(p_a\sigma_1^2 e^{y(\pi,1)(1-\gamma)T} + (1-p_a)\sigma_2^2 e^{y(\pi,2)(1-\gamma)T}\right)\left[p_a e^{y(\pi,1)(1-\gamma)T} + (1-p_a)e^{y(\pi,2)(1-\gamma)T}\right]^{\frac{\eta-\gamma}{\gamma-1}}, \\ \xi_b &= \left(p_b\sigma_1^2 e^{y(\pi,1)(1-\gamma)T} + (1-p_b)\sigma_2^2 e^{y(\pi,2)(1-\gamma)T}\right)\left[p_b e^{y(\pi,1)(1-\gamma)T} + (1-p_b)e^{y(\pi,2)(1-\gamma)T}\right]^{\frac{\eta-\gamma}{\gamma-1}}.\end{aligned}$$

Recall that $\delta^{pre}(\pi, T) = 1 - e^{(y(\pi,1)-y(\pi,2))(1-\gamma)T}$ and define $\delta^{amb}(\pi, T) := 1 - \frac{\xi_a}{\xi_b}$, s.t. we can finally write the equation as

$$\begin{aligned}\pi =& \frac{\tilde{p}(1-\delta_T^{amb}(\pi))}{\tilde{p}(1-\delta_T^{amb}(\pi)) + (1-\tilde{p})}\left[\frac{p_a\sigma_1^2(1-\delta_T^{pre}(\pi))\pi_1^{Mer}}{p_a\sigma_1^2(1-\delta_T^{pre}(\pi)) + (1-p_a)\sigma_2^2} + \frac{(1-p_a)\sigma_2^2\pi_2^{Mer}}{p_a\sigma_1^2(1-\delta_T^{pre}(\pi)) + (1-p_a)\sigma_2^2}\right] + \\ & \frac{1-\tilde{p}}{\tilde{p}(1-\delta_T^{amb}(\pi)) + (1-\tilde{p})}\left[\frac{p_b\sigma_1^2(1-\delta_T^{pre}(\pi))\pi_1^{Mer}}{p_b\sigma_1^2(1-\delta_T^{pre}(\pi)) + (1-p_b)\sigma_2^2} + \frac{(1-p_b)\sigma_2^2\pi_2^{Mer}}{p_b\sigma_1^2(1-\delta_T^{pre}(\pi)) + (1-p_b)\sigma_2^2}\right] \\ =& \alpha^*_{T,\tilde{p}}\left(\alpha_{T,p_a}\pi_1^{Mer} + (1-\alpha_{T,p_a})\pi_2^{Mer}\right) + (1-\alpha_{T,\tilde{p}})\left(\alpha_{T,p_b}\pi_1^{Mer} + (1-\alpha_{T,p_b})\pi_2^{Mer}\right).\end{aligned}$$

D.7: Proof of Proposition 4.11

With the results of equations (4.23) and (4.24) we can state the expected utility of a CRRA investor in terms of

$$\mathbb{E}_{\mathbb{P}}\left[u\left(\frac{V_T}{V_0}\right)\right] = \mathbb{E}_{\mathbb{P}}\left[u\left(1_{\{\tau\le T\}}X_1 + 1_{\{\tau>T\}}X_2\right)\right]$$
$$= \frac{1}{1-\gamma}\left\{\mathbb{E}_{\mathbb{P}}\left[1_{\{\tau\le T\}}X_1^{(1-\gamma)}\right] + \mathbb{E}_{\mathbb{P}}\left[1_{\{\tau>T\}}X_2^{(1-\gamma)}\right]\right\}. \tag{4.68}$$

Let's start with the calculation of $E_1 := \mathbb{E}_{\mathbb{P}}\left[1_{\{\tau\le T\}}X_1^{(1-\gamma)}\right]$ in (4.68):

$$E_1 = \mathbb{E}_{\mathbb{P}}\left[1_{\{\tau\le T\}}e^{[\pi(\mu_2-r)+r-\frac{1}{2}\pi^2\sigma_2^2]T(1-\gamma)}e^{[\pi(\mu_1-\mu_2)-\frac{1}{2}\pi^2(\sigma_1^2-\sigma_2^2)]\tau(1-\gamma)+\pi(\sigma_1W_\tau+\sigma_2(W_T-W_\tau))(1-\gamma)}\right]$$
$$= e^{[\pi(\mu_2-r)+r-\frac{1}{2}\pi^2\sigma_2^2]T(1-\gamma)}\mathbb{E}_{\mathbb{P}}\left[1_{\{\tau\le T\}}e^{[\pi(\mu_1-\mu_2)-\frac{1}{2}\pi^2(\sigma_1^2-\sigma_2^2)]\tau(1-\gamma)}\mathbb{E}_{\mathbb{P}}\left[e^{\pi(\sigma_1W_\tau+\sigma_2(W_T-W_\tau))(1-\gamma)}|\tau\right]\right], \tag{4.69}$$

where the last equation holds because of the basic properties of the conditional expectation. Furthermore the Brownian motion W_τ is independent from $W_T - W_\tau$. Together with the fact that $\mathbb{E}\left[e^X\right] = e^{\mu+\frac{1}{2}\sigma^2}$ for $X \sim N(\mu,\sigma^2)$, equation (4.69) can be written as

$$e^{[\pi(\mu_2-r)+r-\frac{1}{2}\pi^2\sigma_2^2]T(1-\gamma)}\mathbb{E}_{\mathbb{P}}\left[1_{\{\tau\le T\}}e^{[\pi(\mu_1-\mu_2)-\frac{1}{2}\pi^2(\sigma_1^2-\sigma_2^2)]\tau(1-\gamma)}\mathbb{E}_{\mathbb{P}}\left[e^{\pi(\sigma_1W_\tau+\sigma_2(W_T-W_\tau))(1-\gamma)}|\tau\right]\right]$$
$$= e^{[\pi(\mu_2-r)+r-\frac{1}{2}\pi^2\sigma_2^2]T(1-\gamma)}\mathbb{E}_{\mathbb{P}}\left[1_{\{\tau\le T\}}e^{[\pi(\mu_1-\mu_2)-\frac{1}{2}\pi^2(\sigma_1^2-\sigma_2^2)]\tau(1-\gamma)}e^{\frac{1}{2}\pi^2\sigma_1^2(1-\gamma)^2\tau+\frac{1}{2}\pi^2\sigma_2^2(1-\gamma)^2(T-\tau)}\right]$$
$$= e^{\xi_2T}\mathbb{E}_{\mathbb{P}}\left[1_{\{\tau\le T\}}e^{\nu\tau}\right]. \tag{4.70}$$

Now it holds that the function $g(x) := 1_{\{x\le T\}}e^{\nu x}$ is measurable, so the transformation $g(\tau)$ is still a random variable. Together with $\tau \sim Exp(\lambda)$ and its absolute continuous density function $f_\tau^\lambda(x)$we get

$$\mathbb{E}_{\mathbb{P}}\left[g(\tau)\right] = \int_{-\infty}^{\infty} g(x)f_\tau^\lambda(x)dx. \tag{4.71}$$

Combining (4.70) and (4.71), we get

$$e^{\xi_2T}\mathbb{E}_{\mathbb{P}}\left[1_{\{\tau\le T\}}e^{\nu\tau}\right] = e^{\xi_2T}\int_{-\infty}^{\infty} 1_{\{x\le T\}}e^{\nu x}f_\tau^\lambda(x)dx$$
$$= e^{\xi_2T}\int_0^T e^{\nu x}\lambda e^{-\lambda x}dx$$
$$= \lambda e^{\xi_2T}\frac{1}{\nu-\lambda}\left[e^{(\nu-\lambda)T}-1\right]. \tag{4.72}$$

For the calculation of $E_2 := \mathbb{E}_{\mathbb{P}}\left[1_{\{\tau>T\}}X_2^{(1-\gamma)}\right]$ in (4.68) it holds:

$$\begin{aligned}
E_2 &= \mathbb{E}_{\mathbb{P}}\left[1_{\{\tau>T\}}e^{[\pi(\mu_1-r)+r-\frac{1}{2}\pi^2\sigma_1^2]T(1-\gamma)+\sigma_1\pi W_T(1-\gamma)}\right] \\
&= e^{[\pi(\mu_1-r)+r-\frac{1}{2}\pi^2\sigma_1^2]T(1-\gamma)}\mathbb{E}_{\mathbb{P}}\left[1_{\{\tau>T\}}e^{\sigma_1\pi W_T(1-\gamma)}\right] \\
&= e^{[\pi(\mu_1-r)+r-\frac{1}{2}\pi^2\sigma_1^2]T(1-\gamma)}\mathbb{E}_{\mathbb{P}}\left[1_{\{\tau>T\}}\mathbb{E}_{\mathbb{P}}\left[e^{\sigma_1\pi W_T(1-\gamma)}|\tau\right]\right] \\
&= e^{\xi_1 T}\mathbb{E}_{\mathbb{P}}\left[1_{\{\tau>T\}}\right] \\
&= e^{\xi_1 T}\left[1-F_\tau^\lambda(T)\right] \\
&= e^{(\xi_1-\lambda)T}. \qquad (4.73)
\end{aligned}$$

Combining (4.72) and (4.73) gives the final result. □

D.8: Proof of Proposition 4.12:

Proof Maximizing EU pre-commitment strategy Proposition 1:
To maximize the expected utility we have to minimize the expression

$$\frac{\xi_1-\xi_2}{\xi_1-\xi_2-\lambda}e^{(\xi_1-\lambda)T}+\frac{-\lambda}{\xi_1-\xi_2-\lambda}e^{\xi_2 T}.$$

The first order condition is given by

$$\frac{\partial}{\partial\pi}\left\{\frac{\xi_1-\xi_2}{\xi_1-\xi_2-\lambda}\right\}e^{(\xi_1-\lambda)T}+\frac{\xi_1-\xi_2}{\xi_1-\xi_2-\lambda}\frac{\partial}{\partial\pi}\left\{e^{(\xi_1-\lambda)T}\right\}+$$
$$\frac{\partial}{\partial\pi}\left\{\frac{-\lambda}{\xi_1-\xi_2-\lambda}\right\}e^{\xi_2 T}+\frac{-\lambda}{\xi_1-\xi_2-\lambda}\frac{\partial}{\partial\pi}\left\{e^{\xi_2 T}\right\}\stackrel{!}{=}0.$$

Let us calculate the corresponding derivatives first (r=0):

- $\frac{\partial\xi_1}{\partial\pi}=(1-\gamma)\gamma\sigma_1^2(\pi_1^{Mer}-\pi)$
- $\frac{\partial\xi_2}{\partial\pi}=(1-\gamma)\gamma\sigma_2^2(\pi_2^{Mer}-\pi)$
- $\frac{\partial(\xi_1-\xi_2)}{\partial\pi}=(1-\gamma)\gamma\left[\sigma_1^2\pi_1^{Mer}-\sigma_2^2\pi_2^{Mer}-\pi(\sigma_1^2-\sigma_2^2)\right]$
- $\frac{\partial}{\partial\pi}\left\{\frac{\xi_1-\xi_2}{\xi_1-\xi_2-\lambda}\right\}=\frac{\frac{\partial(\xi_1-\xi_2)}{\partial\pi}(\xi_1-\xi_2-\lambda)-(\xi_1-\xi_2)\frac{\partial(\xi_1-\xi_2-\lambda)}{\partial\pi}}{(\xi_1-\xi_2-\lambda)^2}=\frac{-\lambda\frac{\partial(\xi_1-\xi_2)}{\partial\pi}}{(\xi_1-\xi_2-\lambda)^2}$
  $=\frac{-\lambda(1-\gamma)\gamma\left[\sigma_1^2(\pi_1^{Mer}-\pi)-\sigma_2^2(\pi_2^{Mer}-\pi)\right]}{(\xi_1-\xi_2-\lambda)^2}$
- $\frac{\partial}{\partial\pi}\left\{e^{(\xi_1-\lambda)T}\right\}=\frac{\partial\xi_1 T}{\partial\pi}e^{(\xi_1-\lambda)T}=(1-\gamma)\gamma\sigma_1^2(\pi_1^{Mer}-\pi)Te^{(\xi_1-\lambda)T}$

- $\frac{\partial}{\partial\pi}\left\{\frac{-\lambda}{\xi_1-\xi_2-\lambda}\right\} = \frac{\lambda\frac{\partial(\xi_1-\xi_2-\lambda)}{\partial\pi}}{(\xi_1-\xi_2-\lambda)^2} = \frac{\lambda\frac{\partial(\xi_1-\xi_2)}{\partial\pi}}{(\xi_1-\xi_2-\lambda)^2} = -\frac{\partial}{\partial\pi}\left\{\frac{\xi_1-\xi_2}{\xi_1-\xi_2-\lambda}\right\}$
- $\frac{\partial}{\partial\pi}\left\{e^{\xi_2 T}\right\} = \frac{\partial\xi_2 T}{\partial\pi}e^{\xi_2 T} = (1-\gamma)\gamma\sigma_2^2(\pi_2^{Mer}-\pi)Te^{\xi_2 T}$

Because of the fact that $\frac{\partial}{\partial\pi}\left\{\frac{-\lambda}{\xi_1-\xi_2-\lambda}\right\} = -\frac{\partial}{\partial\pi}\left\{\frac{\xi_1-\xi_2}{\xi_1-\xi_2-\lambda}\right\}$ and $\frac{-\lambda}{\xi_1-\xi_2-\lambda} = 1 - \frac{\xi_1-\xi_2}{\xi_1-\xi_2-\lambda}$ we can write the first order condition as follows

$$\begin{aligned}
&\frac{\partial}{\partial\pi}\left\{\frac{\xi_1-\xi_2}{\xi_1-\xi_2-\lambda}\right\}(e^{(\xi_1-\lambda)T}-e^{\xi_2 T}) + \frac{\xi_1-\xi_2}{\xi_1-\xi_2-\lambda}\left(\frac{\partial}{\partial\pi}\left\{e^{(\xi_1-\lambda)T}\right\} - \frac{\partial}{\partial\pi}\left\{e^{\xi_2 T}\right\}\right) + \\
&\frac{\partial}{\partial\pi}\left\{e^{\xi_2 T}\right\} \stackrel{!}{=} 0 \\
\Leftrightarrow &\frac{-\lambda(1-\gamma)\gamma\left[\sigma_1^2(\pi_1^{Mer}-\pi)-\sigma_2^2(\pi_2^{Mer}-\pi)\right]}{(\xi_1-\xi_2-\lambda)^2}(e^{(\xi_1-\lambda)T}-e^{\xi_2 T}) + \\
&\frac{\xi_1-\xi_2}{\xi_1-\xi_2-\lambda}\left((1-\gamma)\gamma\sigma_1^2(\pi_1^{Mer}-\pi)Te^{(\xi_1-\lambda)T} - (1-\gamma)\gamma\sigma_2^2(\pi_2^{Mer}-\pi)Te^{\xi_2 T}\right) + \\
&(1-\gamma)\gamma\sigma_2^2(\pi_2^{Mer}-\pi)Te^{\xi_2 T} \stackrel{!}{=} 0 \\
\Leftrightarrow &\pi^{pre,*} = \frac{[\sigma_1^2(\pi_1^{Mer}-\pi^{pre,*})-\sigma_2^2(\pi_2^{Mer}-\pi^{pre,*})](e^{(\xi_1-\lambda)T}-e^{\xi_2 T})}{Te^{\xi_2 T}\sigma_2^2(\xi_1-\xi_2-\lambda)} \\
&-\frac{(\xi_1-\xi_2)\sigma_1^2(\pi_1^{Mer}-\pi^{pre,*})e^{(\xi_1-\lambda)T}}{\lambda\sigma_2^2 e^{\xi_2 T}} + \pi_2^{Mer} \\
\Leftrightarrow &\pi^{pre,*} = \pi_1^{Mer}\frac{[\lambda\sigma_1^2(e^{(\xi_1-\lambda)T}-e^{\xi_2 T}) - T(\xi_1-\xi_2-\lambda)(\xi_1-\xi_2)\sigma_1^2 e^{(\xi_1-\lambda)T}}{T\lambda e^{\xi_2 T}\sigma_2^2(\xi_1-\xi_2-\lambda)} + \\
&\pi_2^{Mer}\frac{\lambda\sigma_2^2(e^{\xi_2 T}-e^{(\xi_1-\lambda)T}) + T\lambda e^{\xi_2 T}\sigma_2^2(\xi_1-\xi_2-\lambda)}{T\lambda e^{\xi_2 T}\sigma_2^2(\xi_1-\xi_2-\lambda)} + \\
&\pi^{pre,*}\frac{\left(\lambda e^{\xi_2 T}-\lambda e^{(\xi_1-\lambda)T}\right)(\sigma_1^2-\sigma_2^2) + T(\xi_1-\xi_2-\lambda)(\xi_1-\xi_2)\sigma_1^2 e^{(\xi_1-\lambda)T}}{T\lambda e^{\xi_2 T}\sigma_2^2(\xi_1-\xi_2-\lambda)} \\
\Leftrightarrow &\pi^{pre,*} = \alpha_1\pi_1^{Mer} + \alpha_2\pi_2^{Mer} + \alpha_3\pi^{pre,*} \\
\Leftrightarrow &\pi^{pre,*} = \pi_1^{Mer}\frac{\alpha_1}{1-\alpha_3} + \pi_2^{Mer}\frac{\alpha_2}{1-\alpha_3}, \text{where } \alpha_1+\alpha_2+\alpha_3 = 1 \text{ with} \\
&\alpha_1 = \frac{[\lambda\sigma_1^2(e^{(\xi_1-\lambda)T}-e^{\xi_2 T}) - T(\xi_1-\xi_2-\lambda)(\xi_1-\xi_2)\sigma_1^2 e^{(\xi_1-\lambda)T}}{T\lambda e^{\xi_2 T}\sigma_2^2(\xi_1-\xi_2-\lambda)} \\
&\alpha_2 = \frac{\lambda\sigma_2^2(e^{\xi_2 T}-e^{(\xi_1-\lambda)T}) + T\lambda e^{\xi_2 T}\sigma_2^2(\xi_1-\xi_2-\lambda)}{T\lambda e^{\xi_2 T}\sigma_2^2(\xi_1-\xi_2-\lambda)} \\
&\alpha_3 = \frac{\left(\lambda e^{\xi_2 T}-\lambda e^{(\xi_1-\lambda)T}\right)(\sigma_1^2-\sigma_2^2) + T(\xi_1-\xi_2-\lambda)(\xi_1-\xi_2)\sigma_1^2 e^{(\xi_1-\lambda)T}}{T\lambda e^{\xi_2 T}\sigma_2^2(\xi_1-\xi_2-\lambda)}.
\end{aligned}$$

Bibliography

Acerbi, C. (2002), Spectral measures of risk: A coherent representation of subjective risk aversion, *Journal of Banking & Finance* **26**(7), 1505–1518. (Cited on pages 7, 17, 20, and 21.)

Acerbi, C. and Tasche, D. (2002), On the coherence of expected shortfall, *Journal of Banking & Finance* **26**(7), 1487–1503. (Cited on page 13.)

Adam, A., Houkari, M. and Laurent, J.-P. (2008), Spectral risk measures and portfolio selection, *Journal of Banking & Finance* **32**(9), 1870–1882. (Cited on page 20.)

Adams, A., Cherchye, L., De Rock, B. and Verriest, E. (2014), Consume Now or Later? Time Inconsistency, Collective Choice, and Revealed Preference, *American Economic Review* **104**(12), 4147–83. (Cited on page 125.)

Ahn, D., Choi, S., Gale, D. and Kariv, S. (2014), Estimating ambiguity aversion in a portfolio choice experiment, *Quantitative Economics* **5**(2), 195–223. (Cited on page 114.)

Alexander, G. J. and Baptista, A. M. (2004), A comparison of VaR and CVaR constraints on portfolio selection with the mean-variance model, *Management science* **50**(9), 1261–1273. (Cited on pages 61 and 62.)

Alexander, G. J., Baptista, A. M. and Yan, S. (2007), Mean–variance portfolio selection with 'at-risk'constraints and discrete distributions, *Journal of Banking & Finance* **31**(12), 3761–3781. (Cited on pages 61 and 62.)

Alexandrova, M., Bohnert, A., Gatzert, N. and Russ, J. (2017), Equity-linked life insurance based on traditional products: the case of Select Products, *European Actuarial Journal* **7**(2), 379–404. (Cited on page 67.)

Ali, M., Alam, N. and Rizvi, S. A. R. (2020), Coronavirus (COVID-19)—An epidemic or pandemic for financial markets, *Journal of Behavioral and Experimental Finance* **27**, 100341. (Cited on page 2.)

Alserda, G. A., Dellaert, B. G., Swinkels, L. and van der Lecq, F. S. (2019), Individual pension risk preference elicitation and collective asset allocation with heterogeneity, *Journal of Banking & Finance* **101**, 206–225. (Cited on page 125.)

Ang, A. and Bekaert, G. (2002), International asset allocation with regime shifts, *The review of financial studies* **15**(4), 1137–1187. (Cited on pages 156 and 158.)

Antoniou, C., Harris, R. D. and Zhang, R. (2015), Ambiguity Aversion and Stock Market Participation: An Empirical Analysis, *Journal of Banking & Finance* **58**, 57–70. (Cited on page 126.)

Arrow, K. J. (1951), Alternative Approaches to the Theory of Choice in Risk-Taking Situations, *Econometrica: Journal of the Econometric Society* **19**(4), 404–437. (Cited on pages 112 and 126.)

Artzner, P., Delbaen, F. and Koch-Medina, P. (2009), Risk Measures and Efficient use of Capital, *ASTIN Bulletin: The Journal of the IAA* **39**(1), 101–116. (Cited on page 5.)

Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1999), Coherent measures of risk, *Mathematical Finance* **9**(3), 203–228. (Cited on pages 1, 5, 7, 9, 12, and 60.)

Bacinello, A. R. (2001), Fair pricing of life insurance participating policies with a minimum interest rate guarantee, *ASTIN Bulletin: The Journal of the IAA* **31**(2), 275–297. (Cited on pages 33 and 75.)

Bacinello, A. R. and Ortu, F. (1993a), Pricing equity-linked life insurance with endogenous minimum guarantees, *Insurance: Mathematics and Economics* **12**(3), 245–257. (Cited on page 69.)

Bacinello, A. R. and Ortu, F. (1993b), Pricing guaranteed securities-linked life insurance under interest rate risk, *Actuarial Approach for Financial Risks* **1**, 35–55. (Cited on page 69.)

Bacinello, A. R. and Ortu, F. (1994), Single and Periodic premiums for Guaranteed Equity-Linked Life Insurance under Interest Rate Risk- The lognormal and Vasicek Case, *In: L. Peccati, M. Viren (Eds.). Financial Modelling: Recent Research.* Physica-Verlag. 1-25. (Cited on page 69.)

Bacinello, A. R., Chen, A., Sehner, T. and Millossovich, P. (2021), On the market-consistent valuation of participating life insurance heterogeneous contracts under longevity risk, *Risks* **9**(1), 20. (Cited on page 69.)

Bacinello, A. R., Millossovich, P., Olivieri, A. and Pitacco, E. (2011), Variable annuities: A unifying valuation approach, *Insurance: Mathematics and Economics* **49**(3), 285–297. (Cited on page 104.)

BaFin (2018), Solvency II: Reviewing the standard formula, https://www.bafin.de/SharedDocs/Veroeffentlichungen/EN/Fachartikel/2018/fa_bj_1807_SCR_SolvencyII_en.html;jsessionid=5333867C0888850CA565A1EE2FC08309.1_cid501. Accessed: 2022-01-13. (Cited on page 13.)

Balter, A. G., Mahayni, A. and Schweizer, N. (2021), Time-Consistency of Optimal Investment Under Smooth Ambiguity, *European Journal of Operational Research* **293**(2), 643–657. (Cited on pages 122, 125, and 132.)

Barbarin, J. and Devolder, P. (2005), Risk measure and fair valuation of an investment guarantee in life insurance, *Insurance: Mathematics and Economics* **37**(2), 297–323. (Cited on page 70.)

Barnett, M., Brock, W. and Hansen, L. P. (2020), Pricing uncertainty induced by climate change, *The Review of Financial Studies* **33**(3), 1024–1066. (Cited on page 2.)

Basak, S. and Chabakauri, G. (2010), Dynamic Mean-Variance Asset Allocation, *The Review of Financial Studies* **23**(8), 2970–3016. (Cited on page 126.)

Basak, S. and Shapiro, A. (2001), Value-at-risk-based risk management: optimal policies and asset prices, *The Review of Financial Studies* **14**(2), 371–405. (Cited on pages 29, 30, 34, 50, 52, 60, 61, 62, 94, and 106.)

Basel Committee on Banking Supervision (1996), Overview of the amendment to the capital accord to incorporate market risks, https://www.bis.org/publ/bcbs23.pdf. Accessed: 2022-01-13. (Cited on page 1.)

Basel Committee on Banking Supervision (2013), Fundamental review of thetradring book: A revised market risk framework, https://www.bis.org/publ/bcbs265.pdf. Accessed: 2022-01-13. (Cited on pages 2 and 14.)

Basel Committee on Banking Supervision (2020), Risk Based Capital Requirements - Calculation of minimum risk-based capital requirements, https://www.bis.org/basel_framework/chapter/RBC/20.htm. Accessed: 2022-01-13. (Cited on page 8.)

Basel Committee on Banking Supervision (2021a), Climate-related financial risks – measurement methodologies, https://www.bis.org/bcbs/publ/d518.pdf. Accessed: 2022-01-13. (Cited on page 167.)

Basel Committee on Banking Supervision (2021b), Climate-related risk drivers and their transmission channels, https://www.bis.org/bcbs/publ/d517.pdf. Accessed: 2022-01-13. (Cited on page 167.)

Bauer, D. and Zanjani, G. (2016), The marginal cost of risk, risk measures, and capital allocation, *Management Science* **62**(5), 1431–1457. (Cited on page 31.)

Bauer, D., Kling, A. and Russ, J. (2008), A universal pricing framework for guaranteed minimum benefits in variable annuities, *ASTIN Bulletin: The Journal of the IAA* **38**(2), 621–651. (Cited on page 104.)

Bauer, H. (2011), *Measure and integration theory*, de Gruyter. (Cited on page 172.)

Bäuerle, N. and Grether, S. (2017), Extremal behavior of long-term investors with power utility, *International Journal of Theoretical and Applied Finance* **20**(5), 1750029. (Cited on page 130.)

Bäuerle, N. and Rieder, U. (2004), Portfolio optimization with Markov-modulated stock prices and interest rates, *IEEE Transactions on Automatic Control* **49**(3), 442–447. (Cited on pages 157 and 159.)

Benninga, S. and Blume, M. (1985), On the optimality of portfolio insurance, *The Journal of Finance* **40**(5), 1341–1352. (Cited on page 34.)

Berdin, E. and Gründl, H. (2015), The effects of a low interest rate environment on life insurers, *The Geneva Papers on Risk and Insurance Issues and Practice* **40**(3), 385–415. (Cited on page 2.)

Berkowitz, J. and O'Brien, J. (2002), How accurate are value-at-risk models at commercial banks?, *The Journal of Finance* **57**(3), 1093–1111. (Cited on page 14.)

Bernard, C., Boyle, P. P. and Vanduffel, S. (2014), Explicit representation of cost-efficient strategies, *Finance* **35**(2), 5–55. (Cited on pages 7 and 23.)

Bernard, C., Cui, Z. and Vanduffel, S. (2017), Impact of flexible periodic premiums on variable annuity guarantees, *North American actuarial journal* **21**(1), 63–86. (Cited on pages 2, 67, and 69.)

Bernard, C., De Staelen, R. H. and Vanduffel, S. (2019), Optimal portfolio choice with benchmarks, *Journal of the Operational Research Society* **70**(10), 1600–1621. (Cited on page 23.)

Bernard, C., Le Courtois, O. and Quittard-Pinon, F. (2005), Market value of life insurance contracts under stochastic interest rates and default risk, *Insurance: Mathematics and Economics* **36**(3), 499–516. (Cited on page 70.)

Bernard, C., Moraux, F., Rüschendorf, L. and Vanduffel, S. (2015), Optimal payoffs under state-dependent preferences, *Quantitative Finance* **15**(7), 1157–1173. (Cited on pages 23 and 29.)

Bertrand, P. and Prigent, J.-L. (2001), Portfolio insurance strategies: OBPI versus CPPI, *Working Paper*. (Cited on page 76.)

Bi, J. and Cai, J. (2019), Optimal investment-reinsurance strategies with state dependent risk aversion and VaR constraints in correlated markets, *Insurance: Mathematics and Economics* **85**, 1–14. (Cited on pages 61 and 64.)

Biagini, S. and Pınar, M. Ç. (2017), The Robust Merton Problem of an Ambiguity Averse Investor, *Mathematics and Financial Economics* **11**(1), 1–24. (Cited on pages 117 and 126.)

Biener, C., Eling, M. and Wirfs, J. H. (2015), Insurability of cyber risk: An empirical analysis, *The Geneva Papers on Risk and Insurance-Issues and Practice* **40**(1), 131–158. (Cited on page 167.)

Björk, T. and Murgoci, A. (2014), A Theory of Markovian Time-Inconsistent Stochastic Control in Discrete Time, *Finance and Stochastics* **18**(3), 545–592. (Cited on page 125.)

Björk, T., Khapko, M. and Murgoci, A. (2017), On Time-Inconsistent Stochastic Control in Continuous Time, *Finance and Stochastics* **21**(2), 331–360. (Cited on page 125.)

Bo, L., Liao, H. and Yu, X. (2019), Risk sensitive portfolio optimization with default contagion and regime-switching, *SIAM Journal on Control and Optimization* **57**(1), 366–401. (Cited on page 160.)

Bollerslev, T., Chou, R. Y. and Kroner, K. F. (1992), ARCH modeling in finance: A review of the theory and empirical evidence, *Journal of econometrics* **52**(1-2), 5–59. (Cited on page 119.)

Boonen, T. J. (2017), Solvency II solvency capital requirement for life insurance companies based on expected shortfall, *European actuarial journal* **7**(2), 405–434. (Cited on page 44.)

Boonen, T. J. (2019), Equilibrium recoveries in insurance markets with limited liability, *Journal of mathematical economics* **85**, 38–45. (Cited on page 167.)

Borgonovo, E. and Marinacci, M. (2015), Decision Analysis Under Ambiguity, *European Journal of Operational Research* **244**(3), 823–836. (Cited on page 126.)

Bos, K. and Gupta, J. (2019), Stranded assets and stranded resources: Implications for climate change mitigation and global sustainable development, *Energy Research & Social Science* **56**, 101215. (Cited on page 2.)

Boyle, P. and Potapchik, A. (2008), Prices and sensitivities of Asian options: A survey, *Insurance: Mathematics and Economics* **42**(1), 189–211. (Cited on page 82.)

Boyle, P. and Tian, W. (2007), Portfolio management with constraints, *Mathematical Finance* **17**(3), 319–343. (Cited on pages 63, 106, and 108.)

Boyle, P. and Tian, W. (2008), The Design of Equity-Indexed Annuities, *Insurance:Mathematics and Economics* **43**(3), 303–315. (Cited on pages 34, 107, and 108.)

Brandtner, M. (2016a), "Spectral Risk Measures: Properties and Limitations": Comment on Dowd, Cotter, and Sorwar, *Journal of Financial Services Research* **49**(1), 121–131. (Cited on page 22.)

Brandtner, M. (2016b), Spektrale Risikomaße: Konzeption, betriebswirtschaftliche Anwendungen und Fallstricke, *Management Review Quarterly* **66**(2), 75–115. (Cited on page 22.)

Brandtner, M. (2018), Expected Shortfall, spectral risk measures, and the aggravating effect of background risk, or: risk vulnerability and the problem of subadditivity, *Journal of Banking & Finance* **89**, 138–149. (Cited on page 21.)

Brandtner, M. and Kürsten, W. (2014), Solvency II, regulatory capital, and optimal reinsurance: How good are Conditional Value-at-Risk and spectral risk measures?, *Insurance: Mathematics and Economics* **59**, 156–167. (Cited on page 20.)

Brandtner, M. and Kürsten, W. (2015), Decision making with Expected Shortfall and spectral risk measures: The problem of comparative risk aversion, *Journal of Banking & Finance* **58**, 268–280. (Cited on page 21.)

Branger, N. and Larsen, L. S. (2013), Robust portfolio choice with uncertainty about jump and diffusion risk, *Journal of Banking & Finance* **37**(12), 5036–5047. (Cited on pages 115 and 117.)

Branger, N., Becker, L., Mahayni, A. and Offermann, S. (2021), Optimal Asset Allocation, Time-Inconsistency and the Value of Information, *Working Paper*. (Cited on page 123.)

Branger, N., Larsen, L. S. and Munk, C. (2013), Robust portfolio choice with ambiguity and learning about return predictability, *Journal of Banking & Finance* **37**(5), 1397–1411. (Cited on pages 121 and 122.)

Branger, N., Mahayni, A. and Schneider, J. C. (2010), On the optimal design of insurance contracts with guarantees, *Insurance: Mathematics and Economics* **46**(3), 485–492. (Cited on pages 34, 70, 76, 106, and 109.)

Brennan, M. and Schwartz, E. (1976), The Pricing of Equity–Linked Life Insurance Policies with an Asset Value Guarantee, *Journal of Financial Economics* **3**, 195–213. (Cited on pages 33 and 69.)

Brennan, M. J. (1998), The role of learning in dynamic portfolio decisions, *Review of Finance* **1**(3), 295–306. (Cited on pages 119, 120, and 122.)

Brenner, M. and Izhakian, Y. (2018), Asset Pricing and Ambiguity: Empirical Evidence, *Journal of Financial Economics* **130**(3), 503–531. (Cited on page 126.)

Briys, E. and De Varenne, F. (1994), Life insurance in a contingent claim framework: pricing and regulatory implications, *The Geneva Papers on Risk and Insurance Theory* **19**(1), 53–72. (Cited on pages 69 and 70.)

Briys, E. and De Varenne, F. (1997), On the risk of insurance liabilities: debunking some common pitfalls, *Journal of Risk and Insurance* **64**(4), 673–694. (Cited on pages 33, 42, and 69.)

Bugár, G. and Ratting, A. (2016), Revision of the quantification of market risk in the Basel III regulatory framework, *Financial and Economic Review* **15**(1), 33–50. (Cited on page 14.)

Bundesbank, D. (2019), Bonitätsanalyse der Deutschen Bundesbank, https://www.bundesbank.de/resource/blob/602050/7375022c234c5932edac071f26 8f78ee/mL/bonitaetsanalyse-kurzuebersicht-data.pdf. Accessed: 2022-01-13. (Cited on page 15.)

Caldecott, B. (2017), Introduction to special issue: stranded assets and the environment, *Journal of Sustainable Finance & Investment* **7**(1), 1–13. (Cited on page 2.)

Campbell, S. D. (2005), A review of backtesting and backtesting procedures. (Cited on page 15.)

Çanakoğlu, E. and Özekici, S. (2012), HARA frontiers of optimal portfolios in stochastic markets, *European Journal of Operational Research* **221**(1), 129–137. (Cited on page 160.)

Capponi, A. and Figueroa-López, J. E. (2014), Dynamic Portfolio Optimization with a Defaultable Security and Regime-Switching, *Mathematical Finance* **24**(2), 207–249. (Cited on pages 157 and 160.)

Chang, C.-L., Jimeez-Martin, J.-A., Maasoumi, E., McAleer, M. and Pérez-Amaral, T. (2019), Choosing expected shortfall over VaR in Basel III using stochastic dominance, *International Review of Economics & Finance* **60**, 95–113. (Cited on page 14.)

Chen, A. and Hieber, P. (2016), Optimal asset allocation in life insurance: The impact of regulation, *ASTIN Bulletin: The Journal of the IAA* **46**(3), 605–626. (Cited on page 70.)

Chen, A., Hentschel, F. and Klein, J. K. (2015), A utility-and CPT-based comparison of life insurance contracts with guarantees, *Journal of Banking & Finance* **61**, 327–339. (Cited on pages 107 and 109.)

Chen, A., Hieber, P. and Nguyen, T. (2019), Constrained non-concave utility maximization: An application to life insurance contracts with guarantees, *European Journal of Operational Research* **273**(3), 1119–1135. (Cited on pages 34 and 49.)

Chen, A., Nguyen, T. and Rach, M. (2021a), A Collective Investment Problem in a Stochastic Volatility Environment: The Impact of Sharing Rules, *Annals of Operations Research* **302**, 85–109. (Cited on page 125.)

Chen, A., Nguyen, T. and Rach, M. (2021b), Optimal Collective Investment: The Impact of Sharing Rules, Management Fees and Guarantees, *Journal of Banking & Finance* **123**, 106012. (Cited on page 125.)

Chen, A., Nguyen, T. and Stadje, M. (2018a), Optimal investment under VaR-regulation and minimum insurance, *Insurance: Mathematics and Economics* **79**, 194–209. (Cited on pages 34, 56, 61, and 63.)

Chen, A., Nguyen, T. and Stadje, M. (2018b), Risk management with multiple VaR constraints, *Mathematical Methods of Operations Research* **88**(2), 297–337. (Cited on pages 34, 58, 59, 61, and 64.)

Chen, H., Ju, N. and Miao, J. (2014), Dynamic asset allocation with ambiguous return predictability, *Review of Economic Dynamics* **17**(4), 799–823. (Cited on pages 114, 115, 117, and 122.)

Chen, J. M. (2018), On exactitude in financial regulation: Value-at-risk, expected shortfall, and expectiles, *Risks* **6**(2), 61. (Cited on page 17.)

Chen, Z. and Epstein, L. (2002), Ambiguity, risk, and asset returns in continuous time, *Econometrica* **70**(4), 1403–1443. (Cited on pages 114 and 123.)

Cherny, A. S. (2006), Weighted V@ R and its properties, *Finance and Stochastics* **10**(3), 367–393. (Cited on page 19.)

Chi, Y. and Lin, X. S. (2012), Are flexible premium variable annuities underpriced?, *Astin Bulletin* **42**(2), 559–574. (Cited on page 69.)

Chiappori, P.-A. and Paiella, M. (2011), Relative risk aversion is constant: Evidence from panel data, *Journal of the European Economic Association* **9**(6), 1021–1052. (Cited on page 49.)

Chong, W. F. (2019), Pricing and hedging equity-linked life insurance contracts beyond the classical paradigm: The principle of equivalent forward preferences, *Insurance: Mathematics and Economics* **88**, 93–107. (Cited on page 69.)

Coleman, T. F., Li, Y. and Patron, M.-C. (2006), Hedging guarantees in variable annuities under both equity and interest rate risks, *Insurance: Mathematics and Economics* **38**(2), 215–228. (Cited on page 33.)

Coleman, T., Kim, Y., Li, Y. and Patron, M. (2007), Robustly hedging variable annuities with guarantees under jump and volatility risks, *Journal of Risk and Insurance* **74**(2), 347–376. (Cited on page 33.)

Cong, F. and Oosterlee, C. W. (2016), On Pre-Commitment Aspects of a Time-Consistent Strategy for a Mean-Variance Investor, *Journal of Economic Dynamics and Control* **70**, 178–193. (Cited on page 126.)

Costa, O. L. and Araujo, M. V. (2008), A generalized multi-period mean–variance portfolio optimization with Markov switching parameters, *Automatica* **44**(10), 2487–2497. (Cited on pages 158 and 159.)

Cuoco, D., He, H. and Isaenko, S. (2008), Optimal dynamic trading strategies with risk limits, *Operations Research* **56**(2), 358–368. (Cited on pages 60 and 63.)

Cvitanić, J., Lazrak, A., Martellini, L. and Zapatero, F. (2006), Dynamic portfolio choice with parameter uncertainty and the economic value of analysts' recommendations, *The Review of Financial Studies* **19**(4), 1113–1156. (Cited on pages 120 and 122.)

Dai, M., Jin, H., Kou, S. and Xu, Y. (2021), A Dynamic Mean-Variance Analysis for Log Returns, *Management Science* **67**(2), 1093–1108. (Cited on page 126.)

Daníelsson, J., Jorgensen, B. N., de Vries, C. G. and Yang, X. (2008), Optimal portfolio allocation under the probabilistic VaR constraint and incentives for financial innovation, *Annals of finance* **4**(3), 345–367. (Cited on page 63.)

Daníelsson, J., Jorgensen, B. N., Samorodnitsky, G., Sarma, M. and de Vries, C. G. (2013), Fat tails, VaR and subadditivity, *Journal of econometrics* **172**(2), 283–291. (Cited on page 16.)

Danielsson, J., Jorgensen, B. N., Sarma, M. and de Vries, C. G. (2006), Comparing downside risk measures for heavy tailed distributions, *Economics letters* **92**(2), 202–208. (Cited on page 16.)

Deelstra, G., Grasselli, M. and Koehl, P.-F. (2003), Optimal investment strategies in the presence of a minimum guarantee, *Insurance: Mathematics and Economics* **33**(1), 189–207. (Cited on pages 106 and 108.)

Deelstra, G., Grasselli, M. and Koehl, P.-F. (2004), Optimal design of the guarantee for defined contribution funds, *Journal of economic dynamics and control* **28**(11), 2239–2260. (Cited on pages 106 and 108.)

Detemple, J. B. (1991), Further results on asset pricing with incomplete information, *Journal of Economic dynamics and Control* **15**(3), 425–453. (Cited on pages 120 and 122.)

Devolder, P. (2018), Solvency requirement for long term guarantee: risk measure versus probability of ruin, *European Actuarial Journal* **8**(1), 9–24. (Cited on page 70.)

Dhaene, J., Kukush, A., Linders, D. and Tang, Q. (2012), Remarks on quantiles and distortion risk measures, *European Actuarial Journal* **2**(2), 319–328. (Cited on page 19.)

Dhaene, J., Laeven, R. J., Vanduffel, S., Darkiewicz, G. and Goovaerts, M. J. (2008), Can a coherent risk measure be too subadditive?, *Journal of Risk and Insurance* **75**(2), 365–386. (Cited on page 11.)

Di Giacinto, M., Federico, S. and Gozzi, F. (2011), Pension funds with a minimum guarantee: a stochastic control approach, *Finance and Stochastics* **15**(2), 297–342. (Cited on page 109.)

Dichtl, H. and Drobetz, W. (2011), Portfolio insurance and prospect theory investors: Popularity and optimal design of capital protected financial products, *Journal of Banking & Finance* **35**(7), 1683–1697. (Cited on pages 107 and 109.)

Dierkes, M., Erner, C. and Zeisberger, S. (2010), Investment horizon and the attractiveness of investment strategies: A behavioral approach, *Journal of Banking & Finance* **34**(5), 1032–1046. (Cited on page 107.)

Dietz, S., Bowen, A., Dixon, C. and Gradwell, P. (2016), 'Climate value at risk'of global financial assets, *Nature Climate Change* **6**(7), 676–679. (Cited on page 167.)

Dimitriadis, T., Bayer, S. et al. (2019), A joint quantile and expected shortfall regression framework, *Electronic Journal of Statistics* **13**(1), 1823–1871. (Cited on page 15.)

Dimmock, S. G., Kouwenberg, R., Mitchell, O. S. and Peijnenburg, K. (2016), Ambiguity Aversion and Household Portfolio Choice Puzzles: Empirical Evidence, *Journal of Financial Economics* **119**(3), 559–577. (Cited on page 126.)

Doff, R. (2008), A critical analysis of the Solvency II proposals, *The Geneva Papers on Risk and Insurance-Issues and Practice* **33**(2), 193–206. (Cited on page 13.)

Døskeland, T. M. and Nordahl, H. A. (2008), Optimal pension insurance design, *Journal of Banking & Finance* **32**(3), 382–392. (Cited on pages 107 and 108.)

Dow, J. and da Costa Werlang, S. R. (1992), Uncertainty aversion, risk aversion, and the optimal choice of portfolio, *Econometrica: Journal of the Econometric Society* **60**(1), 197–204. (Cited on pages 114 and 116.)

Dowd, K. and Blake, D. (2006), After VaR: the theory, estimation, and insurance applications of quantile-based risk measures, *Journal of Risk and Insurance* **73**(2), 193–229. (Cited on page 17.)

Dowd, K., Cotter, J. and Sorwar, G. (2008), Spectral risk measures: properties and limitations, *Journal of Financial Services Research* **34**(1), 61–75. (Cited on page 22.)

Du, Z. and Escanciano, J. C. (2017), Backtesting expected shortfall: accounting for tail risk, *Management Science* **63**(4), 940–958. (Cited on page 15.)

Dudewicz, E. J. and Mishra, S. (1988), *Modern mathematical statistics*, John Wiley & Sons, Inc. (Cited on page 174.)

Dumas, B. (1989), Two-Person Dynamic Equilibrium in the Capital Market, *The Review of Financial Studies* **2**(2), 157–188. (Cited on page 125.)

Dybvig, P. H. (1988a), Distributional analysis of portfolio choice, *Journal of Business* **61**(3), 369–393. (Cited on pages 6, 7, 23, and 26.)

Dybvig, P. H. (1988b), Inefficient dynamic portfolio strategies or how to throw away a million dollars in the stock market, *The Review of Financial Studies* **1**(1), 67–88. (Cited on pages 6 and 23.)

Easley, D. and O'Hara, M. (2009), Ambiguity and nonparticipation: The role of regulation, *The Review of Financial Studies* **22**(5), 1817–1843. (Cited on pages 114, 115, and 116.)

Ebert, S., Koos, B. and Schneider, J. C. (2012), On the optimal type and level of guarantees for prospect theory investors, *Paris December 2012 Finance Meeting EUROFIDAI-AFFI Paper*. (Cited on pages 34, 104, 107, and 109.)

EC (2009), DIRECTIVE 2009/138/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCILof 25 November 2009 on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II), https://eur-lex.europa.eu/eli/dir/2009/138/2021-06-30. Accessed: 2022-01-13. (Cited on pages 8 and 42.)

Eckert, J., Graf, S. and Kling, A. (2021), A measure to analyse the interaction of contracts in a heterogeneous life insurance portfolio, *European Actuarial Journal* **11**, 87–112. (Cited on page 69.)

EIOPA (2020), 2020 review of Solvency II, https://www.eiopa.europa.eu/browse/solvency-ii/2020-review-of-solvency-ii. Accessed: 2022-01-13. (Cited on pages 2 and 13.)

El Karoui, N., Jeanblanc, M. and Lacoste, V. (2005), Optimal portfolio management with American capital guarantee, *Journal of Economic Dynamics and Control* **29**(3), 449–468. (Cited on pages 34, 55, 106, and 108.)

Eling, M. and Wirfs, J. (2019), What are the actual costs of cyber risk events?, *European Journal of Operational Research* **272**(3), 1109–1119. (Cited on page 167.)

Eling, M., Gatzert, N. and Schmeiser, H. (2008), The Swiss Solvency Test and its market implications, *The Geneva Papers on Risk and Insurance-Issues and Practice* **33**(3), 418–439. (Cited on page 13.)

Eling, M., Schmeiser, H. and Schmit, J. T. (2007), The Solvency II process: Overview and critical analysis, *Risk management and insurance review* **10**(1), 69–85. (Cited on page 13.)

Elliott, R. J. and Miao, H. (2009), VaR and expected shortfall: a non-normal regime switching framework, *Quantitative Finance* **9**(6), 747–755. (Cited on pages 161 and 162.)

Elliott, R. J. and Siu, T. K. (2010), On risk minimizing portfolios under a Markovian regime-switching Black-Scholes economy, *Annals of Operations Research* **176**(1), 271–291. (Cited on pages 161 and 162.)

Elliott, R. J. and Van der Hoek, J. (1997), An application of hidden Markov models to asset allocation problems, *Finance and Stochastics* **1**(3), 229–238. (Cited on pages 156 and 158.)

Elliott, R. J., Siu, T. K. and Badescu, A. (2010), On mean-variance portfolio selection under a hidden Markovian regime-switching model, *Economic Modelling* **27**, 678–686. (Cited on page 159.)

Ellsberg, D. (1961), Risk, ambiguity, and the Savage axioms, *The quarterly journal of economics* **75**(4), 643–669. (Cited on page 112.)

Elstrodt, J. (2018), *Maß-und Integrationstheorie (8te Auflage)*, Springer. (Cited on page 171.)

Embrechts, P. and Hofert, M. (2013), A note on generalized inverses, *Mathematical Methods of Operations Research* **77**(3), 423–432. (Cited on page 11.)

Embrechts, P., McNeil, A. and Straumann, D. (2002), Correlation and dependence in risk management: properties and pitfalls, *Risk management: value at risk and beyond* **1**, 176–223. (Cited on page 16.)

Embrechts, P., Puccetti, G., Rüschendorf, L., Wang, R. and Beleraj, A. (2014), An academic response to Basel 3.5, *Risks* **2**(1), 25–48. (Cited on page 14.)

Emmer, S., Kratz, M. and Tasche, D. (2015), What is the best risk measure in practice? A comparison of standard measures, *Journal of Risk*. (Cited on page 17.)

Epstein, L. G. (2010), A paradox for the "smooth ambiguity" model of preference, *Econometrica* **78**(6), 2085–2099. (Cited on page 114.)

Epstein, L. G. and Schneider, M. (2008), Ambiguity, Information Quality and Asset Pricing, *Journal of Finance* **63**(1), 192–228. (Cited on page 115.)

Epstein, L. G. and Schneider, M. (2010), Ambiguity and Asset Markets, *Annual Review of Financial Economics* **2**(1), 315–346. (Cited on page 112.)

Epstein, L. G. and Wang, T. (2004), *Intertemporal asset pricing under Knightian uncertainty*, Routledge. (Cited on page 114.)

Fang, K., Kotz, S. and Ng, K. (1987), Symmetric Multivariate and Related Distributions Chapman & Hall. (Cited on page 16.)

Filipović, D. and Vogelpoth, N. (2008), A note on the Swiss solvency test risk measure, *Insurance: Mathematics and Economics* **42**(3), 897–902. (Cited on page 13.)

Floreani, A. (2013), Risk measures and capital requirements: A critique of the Solvency II approach, *The Geneva Papers on Risk and Insurance-Issues and Practice* **38**(2), 189–212. (Cited on page 13.)

Föllmer, H. and Leukert, P. (1999), Quantile hedging, *Finance and Stochastics* **3**(3), 251–273. (Cited on pages 34, 56, 168, 175, 176, and 179.)

Föllmer, H. and Schied, A. (2002), Convex measures of risk and trading constraints, *Finance and Stochastics* **6**(4), 429–447. (Cited on pages 7 and 10.)

Föllmer, H. and Schied, A. (2016), *Stochastic Finance (3rd edition)*, De Gruyter. (Cited on pages 10, 13, 19, 171, and 174.)

Fox, C. R. and Tversky, A. (1995), Ambiguity aversion and comparative ignorance, *The quarterly journal of economics* **110**(3), 585–603. (Cited on pages 112 and 113.)

Fox, C. R. and Weber, M. (2002), Ambiguity aversion, comparative ignorance, and decision context, *Organizational behavior and human decision processes* **88**(1), 476–498. (Cited on page 112.)

Frauendorfer, K., Jacoby, U. and Schwendener, A. (2007), Regime switching based portfolio selection for pension funds, *Journal of Banking & Finance* **31**(8), 2265–2280. (Cited on pages 157 and 159.)

Frittelli, M. and Gianin, E. R. (2002), Putting order in risk measures, *Journal of Banking & Finance* **26**(7), 1473–1486. (Cited on pages 7 and 10.)

Fu, J., Wei, J. and Yang, H. (2014), Portfolio optimization in a regime-switching market with derivatives, *European Journal of Operational Research* **233**(1), 184–192. (Cited on pages 157 and 160.)

Gabih, A., Grecksch, W. and Wunderlich, R. (2005), Dynamic portfolio optimization with bounded shortfall risks, *Stochastic analysis and applications* **23**(3), 579–594. (Cited on pages 60 and 62.)

Gao, J., Xiong, Y. and Li, D. (2016), Dynamic mean-risk portfolio selection with multiple risk measures in continuous-time, *European Journal of Operational Research* **249**(2), 647–656. (Cited on pages 34, 61, and 63.)

Garlappi, L., Giammarino, R. and Lazrak, A. (2017), Ambiguity and the Corporation: Group Disagreement and Underinvestment, *Journal of Financial Economics* **125**(3), 417–433. (Cited on page 125.)

Garlappi, L., Uppal, R. and Wang, T. (2007), Portfolio selection with parameter and model uncertainty: A multi-prior approach, *The Review of Financial Studies* **20**(1), 41–81. (Cited on pages 114, 115, and 116.)

Gatzert, N. (2009), Implicit options in life insurance: An overview, *Zeitschrift für die gesamte Versicherungswissenschaft* **98**(2), 141–164. (Cited on page 68.)

Gatzert, N. (2013), On the relevance of premium payment schemes for the performance of mutual funds with investment guarantees, *The Journal of Risk Finance* **14**(5), 436–452. (Cited on pages 69 and 70.)

Gatzert, N. and Kling, A. (2007), Analysis of participating life insurance contracts: a unification approach, *Journal of Risk and Insurance* **74**(3), 547–570. (Cited on page 70.)

Gatzert, N. and Martin, M. (2012), Quantifying credit and market risk under Solvency II: Standard approach versus internal model, *Insurance: Mathematics and Economics* **51**(3), 649–666. (Cited on page 13.)

Gatzert, N. and Wesker, H. (2012), A comparative assessment of Basel II/III and Solvency II, *The Geneva Papers on Risk and Insurance-Issues and Practice* **37**(3), 539–570. (Cited on page 14.)

Gatzert, N. and Wesker, H. (2014), Mortality risk and its effect on shortfall and risk management in life insurance, *Journal of Risk and Insurance* **81**(1), 57–90. (Cited on page 36.)

Gatzert, N., Holzmüller, I. and Schmeiser, H. (2012), Creating customer value in participating life insurance, *Journal of risk and insurance* **79**(3), 645–670. (Cited on pages 33 and 34.)

GDV (2021a), Die deutsche Lebensversicherung in Zahlen 2021, https://www.gdv.de/resource/blob/68738/fc1747f89d09be4b28dd03f831aa6f2e/pdf-data.pdf. Accessed: 2022-01-13. (Cited on page 65.)

GDV (2021b), Kontrollierte Risiken wagen für mehr Chancen in der Altersvorsorge, https://tinyurl.com/5n7mk8aa. Accessed: 2022-01-13. (Cited on page 82.)

Gennotte, G. (1986), Optimal portfolio choice under incomplete information, *The Journal of Finance* **41**(3), 733–746. (Cited on pages 120, 122, and 182.)

Gilboa, I. and Schmeidler, D. (1989), Maxmin expected utility with non-unique prior, *Journal of mathematical economics* **18**(2), 141–153. (Cited on page 114.)

Gneiting, T. (2011), Making and evaluating point forecasts, *Journal of the American Statistical Association* **106**(494), 746–762. (Cited on page 14.)

Gollier, C. (2011), Portfolio choices and asset prices: The comparative statics of ambiguity aversion, *The Review of Economic Studies* **78**(4), 1329–1344. (Cited on pages 112, 115, and 116.)

Grosen, A. and Jørgensen, P. L. (2000), Fair valuation of life insurance liabilities: the impact of interest rate guarantees, surrender options, and bonus policies, *Insurance: Mathematics and Economics* **26**(1), 37–57. (Cited on page 32.)

Grosen, A. and Jørgensen, P. L. (2002), Life insurance liabilities at market value: an analysis of insolvency risk, bonus policy, and regulatory intervention rules in a barrier option framework, *Journal of Risk and Insurance* **69**(1), 63–91. (Cited on pages 32, 33, 46, and 70.)

Guidolin, M. and Rinaldi, F. (2010), A simple model of trading and pricing risky assets under ambiguity: any lessons for policy-makers?, *Applied Financial Economics* **20**(1-2), 105–135. (Cited on pages 115 and 116.)

Guidolin, M. and Rinaldi, F. (2013), Ambiguity in Asset Pricing and Portfolio Choice: A Review of the Literature, *Theory and Decision* **74**(2), 183–217. (Cited on page 126.)

Guidolin, M. and Timmermann, A. (2008), International asset allocation under regime switching, skew, and kurtosis preferences, *The Review of Financial Studies* **21**(2), 889–935. (Cited on page 180.)

Gundel, A. and Weber, S. (2008), Utility maximization under a shortfall risk constraint, *Journal of mathematical Economics* **44**(11), 1126–1151. (Cited on pages 60 and 63.)

Halevy, Y. (2007), Ellsberg revisited: An experimental study, *Econometrica* **75**(2), 503–536. (Cited on page 114.)

Hamilton, J. D. (1989), A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica: Journal of the econometric society* **57**(2), 357–384. (Cited on page 156.)

Handelsblatt (2020), Allianz baut die Lebensversicherung um – 100-Prozent-Beitragsgarantie bei Neuverträgen wird abgeschafft, `https://tinyurl.com/yeyrspsz`. Accessed: 2022-01-13. (Cited on page 2.)

Hansen, L. and Sargent, T. J. (2001), Robust control and model uncertainty, *American Economic Review* **91**(2), 60–66. (Cited on page 114.)

Hardy, M. R. (2001), A regime-switching model of long-term stock returns, *North American Actuarial Journal* **5**(2), 41–53. (Cited on pages 161 and 162.)

He, L., Liang, Z., Liu, Y. and Ma, M. (2020), Weighted utility optimization of the participating endowment contract, *Scandinavian Actuarial Journal* **2020**(7), 577–613. (Cited on page 109.)

He, X. D. and Zhou, X. Y. (2011), Portfolio choice via quantiles, *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics* **21**(2), 203–231. (Cited on page 107.)

He, X. D., Jin, H. and Zhou, X. Y. (2015), Dynamic portfolio choice when risk is measured by weighted VaR, *Mathematics of Operations Research* **40**(3), 773–796. (Cited on pages 19 and 20.)

Hieber, P., Korn, R. and Scherer, M. (2015), Analyzing the effect of low interest rates on the surplus participation of life insurance policies with different annual interest rate guarantees, *European Actuarial Journal* **5**(1), 11–28. (Cited on page 42.)

Hieber, P., Natolski, J. and Werner, R. (2016), Fair Valuation of Cliquet-Style Return Guarantees in (Homogeneous and) Heterogeneous Life Insurance Portfolios, *Working Paper*. (Cited on page 104.)

Hieber, P., Natolski, J. and Werner, R. (2019), Fair valuation of cliquet-style return guarantees in (homogeneous and) heterogeneous life insurance portfolios, *Scandinavian Actuarial Journal* **2019**(6), 478–507. (Cited on pages 33, 69, and 70.)

Hofert, M. and McNeil, A. J. (2015), Subadditivity of Value-at-Risk for Bernoulli random variables, *Statistics & Probability Letters* **98**, 79–88. (Cited on pages 14 and 15.)

Holzmüller, I. (2009), The United States RBC standards, Solvency II and the Swiss solvency test: a comparative assessment, *The Geneva Papers on Risk and Insurance-Issues and Practice* **34**(1), 56–77. (Cited on page 14.)

Honda, T. (2003), Optimal portfolio choice for unobservable and regime-switching mean returns, *Journal of Economic Dynamics and Control* **28**(1), 45–78. (Cited on pages 157, 158, and 182.)

Höring, D. (2013), Will Solvency II market risk requirements bite? The impact of Solvency II on insurers' asset allocation, *The Geneva Papers on Risk and Insurance-Issues and Practice* **38**(2), 250–273. (Cited on page 13.)

Hürlimann, W. (2010), Analytical Pricing of the Unit-linked Endowment with Guarantees and Periodic Premiums, *Astin Bulletin* **40**(2), 631–653. (Cited on page 69.)

Hu, F. and Wang, R. (2017), Optimal investment–consumption strategy with liability and regime switching model under Value-at-Risk constraint, *Applied Mathematics and Computation* **313**, 103–118. (Cited on pages 161 and 163.)

Huang, H., Milevsky, M. A. and Wang, J. (2008), Portfolio choice and life insurance: the CRRA case, *Journal of Risk and Insurance* **75**(4), 847–872. (Cited on pages 34 and 70.)

Hürlimann, W. (2006), A note on generalized distortion risk measures, *Finance Research Letters* **3**(4), 267–272. (Cited on page 19.)

Iwaki, H. and Yumae, S. (2004), An efficient frontier for participating policies in a continuous-time economy, *Insurance: Mathematics and Economics* **35**(3), 611–625. (Cited on page 108.)

Jackson, M. O. and Yariv, L. (2014), Present Bias and Collective Dynamic Choice in the Lab, *American Economic Review* **104**(12), 4184–4204. (Cited on page 125.)

Jang, B.-G., Keun Koo, H., Liu, H. and Loewenstein, M. (2007), Liquidity premia and transaction costs, *The Journal of Finance* **62**(5), 2329–2366. (Cited on page 159.)

Jensen, B. A. and Nielsen, J. A. (2016), How suboptimal are linear sharing rules?, *Annals of Finance* **12**(2), 1–23. (Cited on page 34.)

Jensen, B. A. and Sørensen, C. (2001), Paying for minimum interest rate guarantees: Who should compensate who?, *European Financial Management* **7**(2), 183–211. (Cited on pages 33, 34, 106, and 108.)

Jeong, D., Kim, H. and Park, J. Y. (2015), Does ambiguity matter? Estimating asset pricing models with a multiple-priors recursive utility, *Journal of Financial Economics* **115**(2), 361–382. (Cited on page 114.)

Jin, H. and Zhou, X. Y. (2015), Continuous-Time Portfolio Selection Under Ambiguity, *Mathematical Control & Related Fields* **5**(3), 475. (Cited on page 126.)

Ju, N. and Miao, J. (2012), Ambiguity, learning, and asset returns, *Econometrica* **80**(2), 559–591. (Cited on page 114.)

Kan, R. and Zhou, G. (2007), Optimal Portfolio Choice with Parameter Uncertainty, *Journal of Financial and Quantitative Analysis* **42**(3), 621–656. (Cited on page 112.)

Kaplanski, G. and Levy, H. (2007), Basel's value-at-risk capital requirement regulation: An efficiency analysis, *Journal of Banking & Finance* **31**(6), 1887–1906. (Cited on page 14.)

Kawata, R. and Kijima, M. (2007), Value-at-risk in a market subject to regime switching, *Quantitative Finance* **7**(6), 609–619. (Cited on pages 161 and 162.)

Kellner, R. and Rösch, D. (2016), Quantifying market risk with Value-at-Risk or Expected Shortfall? Consequences for capital requirements and model risk, *Journal of Economic Dynamics and Control* **68**, 45–63. (Cited on page 14.)

Kijima, M. and Wong, T. (2007), Pricing of ratchet equity-indexed annuities under stochastic interest rates, *Insurance: Mathematics and Economics* **41**(3), 317–338. (Cited on page 105.)

Klibanoff, P., Marinacci, M. and Mukerji, S. (2005), A Smooth Model of Decision Making under Ambiguity, *Econometrica* **73**(6), 1849–1892. (Cited on pages 4, 113, 123, 142, 150, and 165.)

Klibanoff, P., Marinacci, M. and Mukerji, S. (2009), Recursive smooth ambiguity preferences, *Journal of Economic Theory* **144**(3), 930–976. (Cited on page 114.)

Kling, A., Ruez, F. and Russ, J. (2011), The impact of stochastic volatility on pricing, hedging, and hedge efficiency of withdrawal benefit guarantees in variable annuities, *ASTIN Bulletin: The Journal of the IAA* **41**(2), 511–545. (Cited on page 69.)

Korn, R., Siu, T. K. and Zhang, A. (2011), Asset allocation for a DC pension fund under regime switching environment, *European Actuarial Journal* **1**(2), 361–377. (Cited on page 160.)

Kou, S., Peng, X. and Heyde, C. C. (2013), External risk measures and Basel accords, *Mathematics of Operations Research* **38**(3), 393–417. (Cited on pages 11 and 31.)

Kraft, H. and Steffensen, M. (2013), A dynamic programming approach to constrained portfolios, *European Journal of Operational Research* **229**(2), 453–461. (Cited on page 63.)

Kratz, M., Lok, Y. H. and McNeil, A. J. (2018), Multinomial VaR backtests: A simple implicit approach to backtesting expected shortfall, *Journal of Banking & Finance* **88**, 393–407. (Cited on page 15.)

Laas, D. and Siegel, C. F. (2017), Basel III versus Solvency II: An analysis of regulatory consistency under the New Capital Standards, *Journal of Risk and Insurance* **84**(4), 1231–1267. (Cited on page 14.)

Lambert, N. S., Pennock, D. M. and Shoham, Y. (2008), *Eliciting properties of probability distributions*, 129-138. (Cited on page 15.)

Laurent, J.-P., Sestier, M. and Thomas, S. (2016), Trading book and credit risk: How fundamental is the Basel review?, *Journal of Banking & Finance* **73**, 211–223. (Cited on page 14.)

Ledlie, M. C., Corry, D. P., Finkelstein, G. S., Ritchie, A. J., Su, K. and Wilson, D. (2008), Variable annuities, *British Actuarial Journal* **14**(2), 327–389. (Cited on page 104.)

Lee, P. M. (2012), *Bayesian statistics: An Introduction*, 4 edn, Oxford University Press London. (Cited on page 118.)

Leippold, M., Trojani, F. and Vanini, P. (2006), Equilibrium impact of value-at-risk regulation, *Journal of Economic Dynamics and Control* **30**(8), 1277–1313. (Cited on pages 60 and 62.)

Leland, H. E. (1980), Who should buy portfolio insurance?, *The Journal of Finance* **35**(2), 581–594. (Cited on pages 34 and 49.)

Lien, D., Wang, Z. and Yu, X. (2021), Optimal quantile hedging under Markov regime switching, *Empirical Economics* **60**, 2177–2201. (Cited on page 175.)

Liptser, R. S. and Shiriaev, A. N. (1977), *Statistics of random processes: General theory*, Springer. (Cited on page 119.)

Liu, H. (2011), Dynamic portfolio choice under ambiguity and regime switching mean returns, *Journal of Economic Dynamics and Control* **35**(4), 623–640. (Cited on pages 157, 160, and 181.)

Liu, J., Yiu, K.-F. C. and Siu, T. K. (2012), A decomposition method for optimal portfolios with regime-switching and risk constraint, *Risk and Decision Analysis* **3**(4), 269–276. (Cited on pages 161 and 162.)

Liu, J., Yiu, K.-F. C. and Siu, T. K. (2014), Optimal investment of an insurer with regime-switching and risk constraint, *Scandinavian Actuarial Journal* **2014**(7), 583–601. (Cited on pages 161 and 162.)

Lundtofte, F. (2006), The effect of information quality on optimal portfolio choice, *Financial Review* **41**(2), 157–185. (Cited on page 120.)

Maccheroni, F., Marinacci, M. and Ruffino, D. (2013), Alpha as Ambiguity: Robust Mean-Variance Portfolio Analysis, *Econometrica* **81**(3), 1075–1113. (Cited on pages 115, 117, and 126.)

Maenhout, P. J. (2004), Robust portfolio rules and asset pricing, *Review of financial studies* **17**(4), 951–983. (Cited on pages 114 and 116.)

Mahayni, A. and Schlögl, E. (2008), The risk management of minimum return guarantees, *BuR-Business Research* **1**(1), 55–76. (Cited on page 33.)

Mahayni, A. and Schneider, J. C. (2016), Minimum return guarantees, investment caps, and investment flexibility, *Review of Derivatives Research* **19**(2), 85–111. (Cited on page 34.)

Mahayni, A., Lubos, O. and Offermann, S. (2021a), Minimum Return Rate Guarantees Under Default Risk - Optimal Design of Quantile Guarantees, *Review of Managerial Science* **15**, 1821–1848. (Cited on pages 32, 70, 94, 102, 106, and 109.)

Mahayni, A., Lubos, O. and Stein, K. (2021b), Natural Hedging with Fixed and Floating Strike Guarantees, *Working Paper*. (Cited on page 35.)

Mahayni, A., Offermann, S. and Stein, K. (2021c), Participating Life Insurance Contracts with Periodic Premium Payments under Regime Switching, *Working Paper*. (Cited on page 67.)

McCabe, G. M. and Witt, R. C. (1980), Insurance pricing and regulation under uncertainty: A chance-constrained approach, *Journal of Risk and Insurance* **47**(4), 607–635. (Cited on pages 33 and 49.)

McNeil, A. J., Frey, R. and Embrechts, P. (2015), *Quantitative risk management: concepts, techniques and tools - revised edition*, Princeton University Press. (Cited on pages 16 and 22.)

Melnikov, A. and Smirnov, I. (2012), Dynamic hedging of conditional value-at-risk, *Insurance: Mathematics and Economics* **51**(1), 182–190. (Cited on page 175.)

Melnikov, A. and Tong, S. (2014), Quantile hedging on equity-linked life insurance contracts with transaction costs, *Insurance: Mathematics and Economics* **58**, 77–88. (Cited on page 175.)

Merton, R. (1971), Optimal Portfolio and Consumption Rules in a Continuous-Time Model, *Journal of Economic Theory* **3**(4), 373–413. (Cited on pages 6, 28, 34, 49, 50, 60, 62, 70, 94, 125, 153, and 156.)

Merton, R. C. (1969), Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case, *The Review of Economics and Statistics* **51**(3), 247–257. (Cited on pages 34, 60, and 156.)

Merton, R. C. (1980), On estimating the expected return on the market: An exploratory investigation, *Journal of financial economics* **8**(4), 323–361. (Cited on page 119.)

Miao, J. (2009), Ambiguity, Risk and Portfolio Choice under Incomplete Information., *Annals of Economics & Finance*. (Cited on page 122.)

Milevsky, M. A. and Young, V. R. (2007), Annuitization and asset allocation, *Journal of Economic Dynamics and Control* **31**(9), 3138–3177. (Cited on pages 36 and 107.)

Milevsky, M. and Kyrychenko, V. (2008), Portfolio Choice with Puts: Evidence from Variable Annuities, *Financial Analysts Journal* **64**(3), 80–95. (Cited on page 34.)

Moldenhauer, F. and Pitera, M. (2019), Backtesting Expected Shortfall: a simple recipe?, *Journal of Risk*. (Cited on page 15.)

Mukhopadhyay, A., Chatterjee, S., Saha, D., Mahanti, A. and Sadhukhan, S. K. (2013), Cyber-risk decision models: To insure IT or not?, *Decision Support Systems* **56**, 11–26. (Cited on page 167.)

Nagai, H. and Runggaldier, W. J. (2007), PDE approach to utility maximization for market models with hidden Markov factors, *Progress in Probability* **59**, 493–506. (Cited on page 159.)

Neyman, J. and Pearson, E. S. (1933), IX. On the problem of the most efficient tests of statistical hypotheses, *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* **231**(694-706), 289–337. (Cited on pages 168 and 172.)

Nguyen, T. and Stadje, M. (2020), Nonconcave optimal investment with value-at-risk constraint: An application to life insurance contracts, *SIAM journal on control and optimization* **58**(2), 895–936. (Cited on pages 61 and 64.)

Nielsen, J. A. and Sandmann, K. (1995), Equity-linked life insurance: A model with stochastic interest rates, *Insurance: Mathematics and Economics* **16**(3), 225–253. (Cited on pages 69 and 74.)

Nielsen, J. A. and Sandmann, K. (1996), Uniqueness of the fair premium for equity-linked life insurance contracts, *The Geneva Papers on Risk and Insurance Theory* **21**(1), 65–102. (Cited on pages 69 and 71.)

Nielsen, J. A. and Sandmann, K. (2003), Pricing bounds on Asian options, *Journal of Financial and Quantitative Analysis* **38**(2), 449–473. (Cited on page 81.)

Nielsen, J. A., Sandmann, K. and Schlögl, E. (2011), Equity-linked pension schemes with guarantees, *Insurance: Mathematics and Economics* **49**(3), 547–564. (Cited on pages 33, 35, and 36.)

Nieto, M. R. and Ruiz, E. (2016), Frontiers in VaR forecasting and backtesting, *International Journal of Forecasting* **32**(2), 475–501. (Cited on page 15.)

O'Brien, J. and Szerszen, P. (2017), An evaluation of bank measures for market risk before, during and after the financial crisis, *Journal of Banking & Finance* **80**, 215–234. (Cited on page 14.)

Ocejo, A. (2018), Explicit solutions to utility maximization problems in a regime-switching market model via Laplace transforms, *Nonlinear Analysis: Hybrid Systems* **30**, 45–57. (Cited on pages 153, 157, and 160.)

Offermann, S. and Stein, K. (2021), The Impact of Guarantee Schemes in Insurance Contracts with Periodic Premium Payments: A Utility-Based Approach, *Working Paper*. (Cited on page 105.)

Orozco-Garcia, C. and Schmeiser, H. (2019), Is fair pricing possible? An analysis of participating life insurance portfolios, *Journal of Risk and Insurance* **86**(2), 521–560. (Cited on page 69.)

Osmundsen, K. K. (2018), Using expected shortfall for credit risk regulation, *Journal of International Financial Markets, Institutions and Money* **57**, 80–93. (Cited on page 14.)

Owen, J. and Rabinovitch, R. (1983), On the class of elliptical distributions and their applications to the theory of portfolio choice, *The Journal of Finance* **38**(3), 745–752. (Cited on page 16.)

Pastor, L. and Veronesi, P. (2009), Learning in financial markets, *Annu. Rev. Financ. Econ.* **1**(1), 361–381. (Cited on page 121.)

Pedersen, J. L. and Peskir, G. (2017), Optimal mean-variance portfolio selection, *Mathematics and Financial Economics* **11**(2), 137–160. (Cited on page 126.)

Peijnenburg, K. (2018), Life-cycle asset allocation with ambiguity aversion and learning, *Journal of Financial and Quantitative Analysis* **53**(5), 1963–1994. (Cited on pages 121 and 122.)

Pflug, G. and Wozabal, D. (2007), Ambiguity in Portfolio Selection, *Quantitative Finance* **7**(4), 435–442. (Cited on pages 116 and 126.)

Pınar, M. Ç. (2014), Equilibrium in an Ambiguity-Averse Mean-Variance Investors Market, *European Journal of Operational Research* **237**(3), 957–965. (Cited on pages 117 and 126.)

Pirvu, T. A. (2007), Portfolio optimization under the value-at-risk constraint, *Quantitative Finance* **7**(2), 125–136. (Cited on pages 60 and 62.)

Rieder, U. and Bäuerle, N. (2005), Portfolio optimization with unobservable Markov-modulated drift process, *Journal of Applied Probability* **42**(2), 362–378. (Cited on pages 157 and 159.)

Rockafellar, R. T. and Uryasev, S. (2000), Optimization of conditional value-at-risk, *Journal of Risk* **2**, 21–42. (Cited on page 13.)

Rogers, L. (2009), Optimal and robust contracts for a risk-constrained principal, *Mathematics and Financial Economics* **2**(3), 151–171. (Cited on pages 60 and 63.)

Rogers, L. C. G. (2001), The relaxed investor and parameter uncertainty, *Finance and Stochastics* **5**(2), 131–154. (Cited on pages 70, 112, 119, 120, and 122.)

Rootzen, H. and Klüppelberg, C. (1999), A single number can't hedge against economic catastrophes, *AMBIO: A Journal of the Human Environment*. (Cited on page 11.)

Ruß, J. and Schelling, S. (2018), Multi cumulative prospect theory and the demand for cliquet-style guarantees, *Journal of Risk and Insurance* **85**(4), 1103–1125. (Cited on pages 34, 104, 105, 107, and 109.)

Sass, J. and Haussmann, U. G. (2004), Optimizing the terminal wealth under partial information: The drift process as a continuous time Markov chain, *Finance and Stochastics* **8**(4), 553–577. (Cited on page 159.)

Schmeiser, H. and Wagner, J. (2015), A proposal on how the regulator should set minimum interest rate guarantees in participating life insurance contracts, *Journal of Risk and Insurance* **82**(3), 659–686. (Cited on pages 33, 34, 35, 42, 49, 51, 70, 106, and 109.)

Setyani, F., Novita, M. and Malik, M. (2018), Optimal portfolio selection with regime-switching Hamilton-Jacobi-Bellman (HJB) equation and maximum value-at-risk (MVaR) constraint, **1108**(1), 1–6. (Cited on page 163.)

Shen, Y. and Siu, T. K. (2012), Asset allocation under stochastic interest rate with regime switching, *Economic Modelling* **29**(4), 1126–1136. (Cited on page 160.)

Shi, Z. and Werker, B. J. (2012), Short-horizon regulation for long-term investors, *Journal of Banking & Finance* **36**(12), 3227–3238. (Cited on pages 60 and 63.)

Sotomayor, L. R. and Cadenillas, A. (2009), Explicit solutions of consumption-investment problems in financial markets with regime switching, *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics* **19**(2), 251–279. (Cited on pages 153, 157, 159, and 160.)

Strotz, R. H. (1955), Myopia and Inconsistency in Dynamic Utility Maximization, *The Review of Economic Studies* **23**(3), 165–180. (Cited on page 125.)

Swiss Financial Market Supervisory Authority (2006), Technical dokument on the Swiss Solvency Test, https://www.finma.ch/FinmaArchiv/bpv/download/e/SST_techDok_061002_E_wo_Li_20070118.pdf. Accessed: 2022-01-13. (Cited on page 8.)

Taboga, M. (2005), Portfolio selection with two-stage preferences, *Finance Research Letters* **2**(3), 152–164. (Cited on pages 115 and 116.)

Taksar, M. and Zeng, X. (2007), Optimal terminal wealth under partial information: Both the drift and the volatility driven by a discrete-time Markov chain, *SIAM journal on Control and Optimization* **46**(4), 1461–1482. (Cited on page 159.)

Tasche, D. (2002), Expected shortfall and beyond, *Journal of Banking & Finance* **26**(7), 1519–1533. (Cited on page 22.)

Tepla, L. (2001), Optimal investment with minimum performance constraints, *Journal of Economic Dynamics and Control* **25**(10), 1629–1645. (Cited on pages 106 and 108.)

Trautmann, S. T., Vieider, F. M. and Wakker, P. P. (2011), Preference reversals for ambiguity aversion, *Management Science* **57**(7), 1320–1333. (Cited on page 112.)

Tversky, A. and Kahneman, D. (1992), Advances in prospect theory: Cumulative representation of uncertainty, *Journal of Risk and uncertainty* **5**(4), 297–323. (Cited on page 107.)

van Staden, P. M., Dang, D.-M. and Forsyth, P. A. (2021), The surprising robustness of dynamic Mean-Variance portfolio optimization to model misspecification errors, *European Journal of Operational Research* **289**(2), 774–792. (Cited on page 126.)

Vigna, E. (2020), On Time Consistency for Mean-Variance Portfolio Selection, *International Journal of Theoretical and Applied Finance* **23**(06), 1–22. (Cited on page 126.)

Vorst, T. (1992), Prices and hedge ratios of average exchange rate options, *International Review of Financial Analysis* **1**(3), 179–193. (Cited on page 74.)

Wang, R. and Zitikis, R. (2021), An axiomatic foundation for the Expected Shortfall, *Management Science* **67**(3), 1413–1429. (Cited on page 13.)

Wang, S. S. (2000), A class of distortion operators for pricing financial and insurance risks, *Journal of risk and insurance* **67**(1), 15–36. (Cited on page 20.)

Wang, S. S., Young, V. R. and Panjer, H. H. (1997), Axiomatic characterization of insurance prices, *Insurance: Mathematics and Economics* **21**(2), 173–183. (Cited on page 17.)

Wang, Y. (2009), Quantile hedging for guaranteed minimum death benefits, *Insurance: Mathematics and Economics* **45**(3), 449–458. (Cited on page 175.)

Wei, P. (2018), Risk management with weighted VaR, *Mathematical Finance* **28**(4), 1020–1060. (Cited on pages 7, 23, 30, 31, 61, and 64.)

Wei, P. (2021), Risk management with expected shortfall, *Mathematics and Financial Economics* **15**, 847–883. (Cited on pages 61 and 64.)

Williams, J. T. (1977), Capital asset prices with heterogeneous beliefs, *Journal of Financial Economics* **5**(2), 219–239. (Cited on page 120.)

Winkler, R. L., Roodman, G. M. and Britney, R. R. (1972), The determination of partial moments, *Management Science* **19**(3), 290–296. (Cited on page 33.)

Witting, H. (1985), *Mathematische Statistik I: Parametrische Verfahren bei festem Stichprobenumfang*, Springer-Verlag. (Cited on pages 169, 170, 171, 173, and 174.)

Xia, Y. (2001), Learning about predictability: The effects of parameter uncertainty on dynamic asset allocation, *The Journal of Finance* **56**(1), 205–246. (Cited on pages 112, 121, and 122.)

Xie, S. (2009), Continuous-time mean–variance portfolio selection with liability and regime switching, *Insurance: Mathematics and Economics* **45**(1), 148–155. (Cited on pages 158 and 159.)

Yan, M., Yang, H., Zhang, L. and Zhang, S. (2020), Optimal investment-reinsurance policy with regime switching and value-at-risk constraint, *Journal of Industrial & Management Optimization* **16**(5), 2195–2211. (Cited on pages 161 and 163.)

Yin, G. and Zhou, X. Y. (2004), Markowitz's mean-variance portfolio selection with regime switching: From discrete-time models to their continuous-time limits, *IEEE Transactions on automatic control* **49**(3), 349–360. (Cited on page 158.)

Yiu, K.-F. C. (2004), Optimal portfolios under a value-at-risk constraint, *Journal of Economic Dynamics and Control* **28**(7), 1317–1334. (Cited on pages 34, 60, and 62.)

Yiu, K.-F. C., Liu, J., Siu, T. K. and Ching, W.-K. (2010), Optimal portfolios with regime switching and value-at-risk constraint, *Automatica* **46**(6), 979–989. (Cited on pages 161 and 162.)

Zariphopoulou, T. (1992), Investment-consumption models with transaction fees and Markov-chain parameters, *SIAM Journal on Control and Optimization* **30**(3), 613–636. (Cited on pages 156 and 158.)

Zhang, J., Jin, Z. and An, Y. (2017), Dynamic portfolio optimization with ambiguity aversion, *Journal of Banking & Finance* **79**, 95–109. (Cited on pages 115 and 117.)

Zhang, Q. and Gao, Y. (2016), Optimal consumption—portfolio problem with CVaR constraints, *Chaos, Solitons & Fractals* **91**, 516–521. (Cited on page 63.)

Zhang, X., Siu, T. K. and Meng, Q. (2010), Portfolio selection in the enlarged Markovian regime-switching market, *SIAM Journal on Control and Optimization* **48**(5), 3368–3388. (Cited on pages 157 and 160.)

Zhao, P. and Xiao, Q. (2016), Portfolio selection problem with Value-at-Risk constraints under non-extensive statistical mechanics, *Journal of computational and applied mathematics* **298**, 64–71. (Cited on page 63.)

Zhou, X. Y. and Yin, G. (2003), Markowitz's Mean-Variance Portfolio Selection with Regime Switching: A Continuous-Time Model, *SIAM Journal on Control and Optimization* **42**(4), 1466–1482. (Cited on pages 156 and 158.)

Zhu, D.-M., Xie, Y., Ching, W.-K. and Siu, T.-K. (2016), Optimal portfolios with maximum Value-at-Risk constraint under a hidden Markovian regime-switching model, *Automatica* **74**, 194–205. (Cited on pages 161 and 163.)

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich die vorliegende Dissertation selbständig und ohne unerlaubte Hilfe angefertigt und andere als die in der Dissertation angegebenen Hilfsmittel nicht benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus anderen Schriften entnommen sind, habe ich als solche kenntlich gemacht.

Ort, Datum

Unterschrift

DOI: 10.17185/duepublico/76967
URN: urn:nbn:de:hbz:465-20221024-151913-5