Modern Policyholder Preferences and Scenario-Based Projections

Ninna Reitzel Jensen

PhD Thesis

This thesis has been submitted to the PhD School of the Faculty of Science, University of Copenhagen

March 31, 2016

Department of Mathematical Sciences
Faculty of Science
University of Copenhagen

Ninna Reitzel Jensen

Department of Mathematical Sciences
University of Copenhagen
Universitetsparken 5
DK-2100 København Ø

ninnareitzel@gmail.com

Assessment Committee:
Professor Jostein Paulsen (chairman),
University of Copenhagen

Professor Frank Thomas Seifried,
University of Trier

Professor Nadine Gatzert,
Friedrich-Alexander University Erlangen-Nürnberg

Supervisor:
Professor Mogens Steffensen,
University of Copenhagen

Co-Supervisor:
Associate Professor Jesper Lund Pedersen,
University of Copenhagen

Submission:
This thesis was submitted to the PhD School of the Faculty of Science,
University of Copenhagen on March 31, 2016

Copyright:
Abstract, Preface, Summary, Chapters 1,4,7: © N.R. Jensen
Chapter 2: © Elsevier B.V.
Chapter 3: © M.A. Fahrenwaldt, N.R. Jensen, M. Steffensen
Chapter 5: © K. Bruhn, N.R. Jensen, M. Steffensen
Chapter 6: © N.R. Jensen, K.J. Schomacker

ISBN:
978-87-7078-944-8

Abstract

This PhD thesis covers aspects of policyholder preferences and life insurance projections based on economic scenarios. Both topics are relevant to policyholders as well as to the life insurance and pension industry—and equally important, the topics give rise to a variety of interesting mathematical problems and industry related considerations.

From a policyholder perspective, being aware of one's own preferences is central for making the best possible financial decisions, with life insurance and pensions playing a major role. From an industry perspective, understanding policyholder preferences is important for designing competitive life insurance and savings products and for providing sound advice to policyholders. Preferences come in many shapes and forms. In this thesis, we focus on separation of risk and time preferences and preferences for smooth investment. The latter is modeled with something as unconventional as explicit preferences for not trading, and if not careful, the former entails time-inconsistency.

From a policyholder and advisory perspective, scenario-based projections allow for tailor-made bonus, benefit, and retirement savings prognoses that illustrate financial riskiness to the policyholder. From an industry and accounting perspective, scenario-based projections allow for valuation of life insurance contracts taking into account both guaranteed and non-guaranteed payments. In this thesis, we focus on economic scenarios because they ensure a low mathematical complexity even for complex financial markets, and we model participating life and unit-linked insurance in the same two-account framework.

Preface

This thesis has been prepared in fulfillment of the requirements for the PhD degree at the Department of Mathematical Sciences (MATH), Faculty of Science, University of Copenhagen. The work has been carried out under the supervision of Professor Mogens Steffensen from MATH in the period from April 1, 2013 to March 31, 2016.

The main body of the thesis consists of an introduction to the overall work and six chapters on different but related topics. The six chapters are written as individual academic papers. They appear as such, and the notation varies slightly from paper to paper. The papers are self-contained and can be read independently with minor overlaps. At the time of submission, two of the papers are published in international peer-review journals, and three of the remaining papers have been submitted for publication [subsequently, the last paper has also been submitted for publication].

Acknowledgments

First and foremost, I would like to thank my supervisor Mogens Steffensen for encouraging me to undertake the PhD study. I am grateful for his ideas, guidance, feedback, co-authoring, visions, and experience. Also, I am grateful for my participation in the Actulus project, and I thank Mogens, Edlund, the IT University of Copenhagen, and MATH for setting up the project and Højteknologifonden (now Innovationsfonden) for supporting it financially (grant number 017-2010-3).

Second, I would very much like to thank Daniel Bauer for facilitating my 4 months visit at J. Mack Robinson College of Business at Georgia State University. I am grateful to him and the entire department of Risk Management & Insurance for a warm welcome and a wonderful, exciting, and educational time in Atlanta.

I would like to thank my co-authors Matthias Albrecht Fahrenwaldt from Leibniz Universität Hannover (now at Heriot-Watt University), Kristian Juul Schomacker from Edlund, and Kenneth Bruhn from PenSam for the good and fruitful collaboration.

I would like to thank my office mates (Maj-Britt, Kamille, Rune, Sima, Henrik, and André) and Nina for many nice trips to the coffee machine, my IE colleagues at MATH for many pleasant lunch talks, my fellow actuarial PhD students for great conference travels, and my colleagues at Edlund for many good Thursdays.

I would like to thank my family and friends for always believing in me, being there, and inspiring me to do better. In particular, I am grateful to my

Dad for introducing me to the beauty of mathematics and to my Mom for teaching me that nothing is impossible. Also, I am grateful to my high school math teacher, Claus Glunk from Helsingør Gymnasium, for enriching my joy of mathematics.

Last, but absolutely not least, I would like to thank my boyfriend Mads Reiter Jespersen for his never ending support and for reminding me of my lovely life outside of the PhD project. I am grateful for his encouragement to follow my dreams and for inspiring me to keep dreaming bigger.

Ninna Reitzel Jensen
Østerbro, March 2016

This final version of my thesis is almost identical to the version submitted for assessment (except from the addition of Appendix A). In Chapters 1, 4 and 5, I have corrected a total of 11 typos. The abstract, summary, and Chapters 2, 3, 6, and 7 remain untouched. After submitting the thesis, I have written down a proof of the verification theorem in Chapter 4 and revised the recapitulation of how to derive bonus sample paths in the first numerical example in Chapter 7, after suggestions from journal referees. Also, I have expanded the paragraph on existing literature in the introduction to Chapter 7, after suggestions from the Assessment Committee and a journal referee. To stay true to the submitted version of the thesis, these improvements are not incorporated in Chapters 4 and 7, but they are presented in Appendix A. As the only part of this thesis, Appendix A has not been assessed by the Assessment Committee.

Ninna Reitzel Jensen
Østerbro, May 2016

Contents

List of Papers

This thesis is based on six papers:

- Ninna Reitzel Jensen & Mogens Steffensen (2015). Personal Finance and Life Insurance under Separation of Risk Aversion and Elasticity of Substitution. *Insurance: Mathematics and Economics 62*, 28–41. doi: 10.1016/j.insmatheco.2015.02.006

- Matthias Albrecht Fahrenwaldt, Ninna Reitzel Jensen, & Mogens Steffensen (2016). Nonrecursive Separation of Risk and Time Preferences. *Submitted for publication.*

- Ninna Reitzel Jensen (2015). Life Insurance Decisions under Recursive Utility. *Submitted for publication.*

- Kenneth Bruhn, Ninna Reitzel Jensen, & Mogens Steffensen (2016). Smooth Investment. *Submitted for publication.*

- Ninna Reitzel Jensen & Kristian Juul Schomacker (2015). A Two-Account Life Insurance Model for Scenario-Based Valuation Including Event Risk. *Risks 3* (2), 183—218. doi: 10.3390/risks3020183

- Ninna Reitzel Jensen (2015). Scenario-based Life Insurance Prognoses in a Multi-State Markov Model. *Submitted for publication.*

Summary

This thesis consists of an introductory chapter and six papers, each of which constitutes a chapter. The first four papers are centered around policyholder preferences, and the last two papers are centered around scenario-based projections of the balance. The papers are related by being of relevance to the life insurance and pension industry. Understanding policyholder preferences is important for the industry in order to design better life insurance and savings products and provide sound advice to policyholders. Scenario-based projections are a valuable tool for the industry in order to price new products, valuate existing products and, again, provide sound advice to policyholders.

In the first paper, *"Personal Finance and Life Insurance under Separation of Risk Aversion and Elasticity of Substitution"*, we study optimal consumption, investment, and life insurance decisions for an policyholder with power utility, an uncertain lifetime, and access to a Black-Scholes market. We separate risk aversion from elasticity of inter-temporal substitution and elasticity of substitution between consumption and bequest by non-linearly aggregating certainty equivalents rather than utility. This leads to time-inconsistency issues which are dealt with using equilibrium theory. We illustrate the equilibrium consumption and bequest in a numerical example, and we establish a connection to recursive utility with Epstein-Zin preferences. The paper builds on my thesis for the Master degree in Actuarial Mathematics which carried the title *"On the Theory of Life Insurance Decisions under Recursive Utility"*.

In the second paper, *"Nonrecursive Separation of Risk and Time Preferences"*, we generalize the separation of preferences from the first paper in terms of utility specification and market modeling. We focus on separation of preferences for time and risk and consider only a certain-lived policyholder, but allow for general utility functions and a general market. We continue separating preferences by aggregating certainty equivalents in a non-linear fashion, and time-inconsistency issues are still dealt with using equilibrium theory. The connection to recursive utility established in the first paper is extended beyond power utility and Epstein-Zin preferences. Furthermore, the connection is extended to an incomplete market setting.

In the third paper, *"Life Insurance Decisions under Recursive Utility"*, we generalize recursive utility to include lifetime uncertainty and utility from bequest. Recursive utility allows for separation of preferences for risk and time, and, with our generalization, also preferences for substitution between bequest and future utility. The latter is our way of formulating preferences for mortality risk. We study optimal consumption, investment, and life insurance choice under recursive utility with generalized Epstein-Zin preferences, and we illustrate the optimal consumption and bequest in a numerical example. The

separation of preferences appears similar to the one in the first paper, but the two ways of separating preferences do not cover the same set of preferences. Only by reducing the number of free preference parameters, thereby giving up on the threefold separability, we identify a coincidence.

In the fourth paper, *"Smooth Investment"*, we solve two portfolio optimization problems with the common feature that non-smooth trading is ruled out. By non-smooth trading we mean diffusive trading and bang-bang investment strategies. Non-smooth trading is avoided by restricting the number of stocks to be differentiable and by punishing trading in the objective function. The latter is equivalent with the policyholder having explicit preferences for not trading. We solve the two portfolio problems semi-explicitly and discuss the structures of the solutions. In numerical examples, we illustrate smooth trading and the resulting stock positions.

In the fifth paper, *"A Two-Account Life Insurance Model for Scenario-Based Valuation Including Event Risk"*, we introduce a two-account model with event risk, such as death and disability, for the purpose of valuating life insurance contracts taking into account both guaranteed and non-guaranteed payments in participating life insurance as well as in unit-linked insurance. To allow for complicated financial markets without dramatically increasing the mathematical complexity, we focus on economic scenarios. We formalize how the bonus schemes "consolidation" and "additional benefits" work and interact in participating life insurance and how guarantees can be implemented in unit-linked insurance. By use of a two-account model, we are able to illustrate general concepts without making the model too abstract, and we provide numerical examples to demonstrate the possible applications of the model.

In the sixth paper, *"Scenario-based Life Insurance Prognoses in a Multi-State Markov Model"*, we introduce economic scenarios in participating life and unit-linked insurance to produce tailor-made bonus, benefit, and retirement savings prognoses that illustrate financial riskiness to the policyholder. In our modeling, we condition on the policyholder starting and staying in a certain state of life, typically "alive and active". We have chosen this fixed path approach to provide policyholders with the best possible economic forecast given that they continue their course of life. We model participating life and unit-linked insurance in the same framework, and for both product types, we provide numerical examples to illustrate the possible applications of our model.

Resumé

(Danish Summary)

Denne afhandling består af et introducerende kapitel og seks artikler, som hver udgør et kapitel. De første fire artikler beskæftiger sig med policetagerpræferencer, mens de to sidste artikler beskæftiger sig med scenariebaserede fremregninger af balancen. Artiklerne har det tilfælles, at de henvender sig til pensions- og livsforsikringsbranchen. Det er vigtigt for industrien at forstå policetagerpræferencer for at kunne designe bedre livsforsikrings- og opsparingsprodukter og for at kunne rådgive policetagere bedst muligt. Scenariebaserede fremregninger er et værdifuldt værktøj for industrien til at prisfastsætte nye produkter, værdifastsætte eksisterende produkter og, igen, kunne rådgive policetagere bedst muligt.

I den første artikel, *"Personal Finance and Life Insurance under Separation of Risk Aversion and Elasticity of Substitution"*, studerer vi optimalt forbrugs-, investerings- og livsforsikringsvalg for en policetager med potensnytte, en usikker levetid og adgang til et Black-Scholes marked. Vi adskiller risikoaversion fra elasticitet af intertemporal substitution og elasticitet af substitution mellem forbrug og arv ved ikke-lineært at summere sikkerhedsækvivalenter frem for nytte. Dette fører til udfordringer med tidsinkonsistens, som vi løser med ligevægtsteori. Vi illustrerer ligevægtsforbrug og -arv i et numerisk eksempel, og vi etablerer en forbindelse til rekursiv nytte med Epstein-Zin præferencer.

I den anden artikel, *"Nonrecursive Separation of Risk and Time Preferences"*, udvider vi adskillelsen af præferencer fra den første artikel med hensyn til nyttespecifikation og markedsmodellering. Vi fokuserer på adskillelse af præferencer for tid og risiko og betragter kun en policetager med en kendt levetid, men tillader generelle nyttefunktioner og et generelt marked. Vi adskiller fortsat præferencer ved ikke-lineært at summere sikkerhedsækvivalenter og håndterer tidsinkonsistens ved hjælp af ligevægtsteori. Forbindelsen til rekursiv nytte, som blev etableret i den første artikel, udvides til at omfatte and og mere end potensnytte og Epstein-Zin præferencer. Forbindelsen udvides desuden til et inkomplet marked.

I den tredje artikel, *"Life Insurance Decisions under Recursive Utility"*, udvider vi rekursiv nytte til at omfatte levetidsusikkerhed og nytte fra arv. Rekursiv nytte tillader adskillelse af præferencer for risiko og tid og, med vores udvidelse, også præferencer for substitution mellem arv og fremtidig nytte. Sidstnævnte er vores måde at formalisere præferencer for dødsrisiko. Vi studerer optimalt forbrugs-, investerings- og livsforsikringsvalg under rekursiv nytte med Epstein-Zin præferencer, og vi illustrerer det optimale forbrug og

den optimale arv i et numerisk eksempel. Adskillelsen af præferencer ligner til forveksling adskillelsen i den første artikel, men de to måder at adskille præferencer dækker ikke den samme mængde af præferencer. Kun ved at reducere antallet af frie parametre, og dermed opgive den trefoldige adskillelse, identificerer vi et sammenfald.

I den fjerde artikel, *"Smooth Investment"*, løser vi to porteføljeproblemer, som har det tilfælles, at ikke-glat handel er udelukket. Ved ikke-glat handel forstår vi diffusiv handel og bang-bang investeringsstrategier. Ikke-glat handel undgås ved at kræve at antallet af aktier er differentiabelt og ved at straffe handel i objektfunktionen. Sidstnævnte er ensbetydende med at policetageren har eksplicitte præferencer for ikke at handle. Vi løser de to porteføljeproblemer semieksplicit og diskuterer løsningernes struktur. I numeriske eksempler illustrerer vi glat handel og den tilsvarende aktiebeholdning.

I den femte artikel, *"A Two-Account Life Insurance Model for Scenario-Based Valuation Including Event Risk"*, introducerer vi en to-konto model med forsikringsrisici for at værdifastsætte livsforsikringskontrakter under hensyntagen til både garanterede og ugaranterede betalinger i gennemsnitsrente såvel som markedsrente. For at tillade komplicerede finansielle markeder uden drastisk at hæve den matematiske kompleksitet fokuserer vi på økonomiske scenarier. Vi formaliserer, hvordan bonussystemerne "opskrivning af ydelser" og "styrkelse" fungerer og interagerer i gennemsnitsrente, og hvordan garantier kan implementeres i markedsrente. Ved brug af en to-konto model er vi i stand til at illustrere generelle koncepter uden at gøre modellen for abstrakt, og vi demonstrerer de mulige anvendelsesmuligheder for modellen i numeriske eksempler.

I den sjette artikel, *"Scenario-based Life Insurance Prognoses in a Multi-State Markov Model"*, introducerer vi økonomiske scenarier i gennemsnitsrente og markedsrente for at udarbejde skræddersyede bonus-, ydelses- og opsparingsprognoser, som illustrerer policetagerens finansielle risiko. I vores modellering betinger vi med, at policetageren starter og forbliver i en bestemt livssituation, typisk "i live og aktiv". Vi har valgt denne betingede fremgangsmåde for at kunne udarbejde de bedst mulige økonomiske prognoser til policetagere givet at de fortsætter deres hidtidige livsforløb. Vi modellerer gennemsnitsrente og markedsrente i samme ramme, og for begge produkttyper demonstrerer vi de mulige anvendelsesmuligheder af vores model i numeriske eksempler.

Chapter 1

Introduction

This introductory chapter provides an overview of the contributions of this thesis. As suggested by the title, the thesis is split in two parts. The first part is made up by four chapters addressing modern policyholder preferences. The second part is made up by two chapters addressing scenario-based projections of the balance. The two parts are primarily related by being of relevance to insurance companies in their designing and valuation of products. Within each of the two parts, the chapters are more closely related, and in this introduction, we explain to which extend. The introduction contains no references to related literature, for these we refer to the introductions of the individual chapters.

1.1 Modern Policyholder Preferences

Classical literature on household decision making deals with maximization of expected time-additive utility. For a policyholder making continuous-time decisions on optimal consumption and investment, the generic problem reads

$$\sup_{c,\pi} \mathbb{E}\left[\int_0^T e^{-\delta s} u\left(c_t\right)\,\mathrm{d}t + e^{-\delta T} U\left(X_T^{c,\pi}\right)\right] , \tag{1.1}$$

where $\delta \geq 0$ is a subjective utility discount rate, u is an instantaneous utility function, and U is a utility function for final wealth. The policyholder's wealth, $X^{c,\pi}$, evolves according to the dynamics

$$\begin{aligned} \mathrm{d}X_t^{c,\pi} &= \left(r + \pi_t \lambda\right) X_t^{c,\pi}\,\mathrm{d}t - c_t\,\mathrm{d}t + \pi_t X_t^{c,\pi} \sigma\,\mathrm{d}W_t \ , \\ X_0^{c,\pi} &= x_0 \ , \end{aligned}$$

where c is the policyholder's consumption rate and π is the proportion invested in a Black-Scholes stock. The utility functions u and U characterize the policyholder's preferences with respect to risk. The problem in Equation (1.1) can be solved by embedding it in an optimal value function given by

$$V\left(t,x\right) = \sup_{c,\pi} \mathbb{E}_{t,x}\left[\int_t^T e^{-\delta(s-t)} u\left(c_s\right)\,\mathrm{d}s + e^{-\delta(T-t)} U\left(X_T^{c,\pi}\right)\right] , \tag{1.2}$$

where $\mathbb{E}_{t,x}$ denotes conditional expectation given $X_t^{c,\pi} = x$. By application of dynamic programming techniques, the value function can be characterized by the Hamilton-Jacobi-Bellman equation, i.e. a partial differential equation containing a local optimization problem at each point (t, x). Using the linearity of the expectation operator and the law of iterated expectation, it can be proven that the solution (c, π) to the continuum of local optimization problems is also a solution to the global optimization problem. Dynamic programming applies because time-additive utility satisfies Bellman's principle of optimality which states that *"an optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision".*

Being an actuary, it is natural to include lifetime uncertainty and life insurance decisions. For a policyholder making continuous-time decisions on optimal consumption, investment and life insurance, the generic problem reads

$$\sup_{c,\pi,d} \mathbb{E}\left[\int_0^T e^{-\delta s}\left\{u\left(c_t\right) I_s \,\mathrm{d}t + b\left(X_t^{c,\pi,d} + d_t\right) \mathrm{d}N_s\right\} + I_T e^{-\delta T} U\left(X_T^{c,\pi,d}\right)\right] , \tag{1.3}$$

where $I = 1 - N$ indicates survival of the policyholder, and b is a utility function for bequest. The policyholder's wealth, $X^{c,\pi,d}$, evolves according to the dynamics

$$\begin{aligned} \mathrm{d}X_t^{c,\pi,d} &= \left(r + \pi_t \lambda\right) X_t^{c,\pi,d} \,\mathrm{d}t - c_t \,\mathrm{d}t - \mu_t^* d_t \,\mathrm{d}t + \pi_t X_t^{c,\pi,d} \sigma \,\mathrm{d}W_t \ , \\ X_0^{c,\pi,d} &= x_0 \ , \end{aligned}$$

where d is a life insurance sum which is priced by a premium intensity μ^*. This consumption-investment-life insurance problem can be solved exactly as the consumption-investment problem using dynamic programming.

We mentioned above that the utility functions u and U characterize the policyholder's preferences with respect to risk, but u also plays a different indirect role in the time-additivity of (1.2). If we, for example, consider the case of power utility,

$$u\left(c\right) = \frac{1}{1-\gamma} c^{1-\gamma} \ ,$$

then the parameter γ does not only represent aversion towards risk, but is also related to the elasticity of inter-temporal substitution (EIS). Whereas risk aversion deals with the willingness to gamble, EIS deals with the willingness to substitute consumption over time.

In Chapters 2 and 3, we formalize a way of separating preferences for risk and time. Here, and in the following, "preferences for time" refer to preferences towards variation in consumption over time, not impatience (as modeled by δ above). Similarly, "preferences for risk" refer to preferences towards variation in consumption over states of the world. The separation of preferences in Chapter 3 is more general in terms of utility specification and market modeling, but only applies to consumption-investment decisions. The separation of preferences in Chapter 2 is restricted to power utility, but applies to consumption-investment-life insurance decisions in a Black-Scholes financial market. The resulting optimization problems do not fit into the framework of problems (1.1) and (1.3), and Bellman's principle of optimality is not satisfied. Consequently, dynamic programming does not apply, and solving the problems requires a special toolbox to avoid time-inconsistent policyholder behavior. This special toolbox is equilibrium theory which arises from a game theoretic approach to stochastic control. We summarize our separated preferences and main insights of the two chapters in Section 1.1.1. We note that in Chapter 3 we often replace the word "separate" by "disentangle".

Chapter 2 builds on my thesis for the Master degree in Actuarial Mathematics which carried the title *"On the Theory of Life Insurance Decisions under Recursive Utility"*. The main theorems have been refined, and the proofs have been sharpened. The central findings have been highlighted, and the set-up has been clarified for non-actuarial readers. Finally, the link to existing literature on time-inconsistency, recursive utility, and hump-shaped consumption has been further developed.

In Chapter 4, we extend an already existing way of separating preferences for risk and time, namely recursive utility, to include lifetime uncertainty and utility from bequest. Our generalization enables us to study optimal consumption-investment-life insurance decisions under separation of preferences for market risk, lifetime uncertainty and time. Again, the resulting optimization problem does not fit into the framework of problem (1.3), but recursive utility is by construction time-consistent, so equilibrium theory does not come into play. We summarize our generalization in Section 1.1.2.

We speak of the policyholder preferences in Chapters 2–4 as modern because they challenge the classical literature on household decision making. We part with expected time-additive utility because it leaves too little room for modeling interesting aspects of preferences for risk and time, and in doing so, we have to give up on dynamic programming.

A very different example of modern policyholder preferences that challenge the classical literature on optimal investment decisions is found in Chapter 5. In this case, dynamic programming applies, but the policyholder has preferences for not trading, e.g. due to transaction costs. These preferences are introduced to obtain smooth investment for a policyholder maximizing, respectively, mean-square and power utility of terminal wealth. We summarize our preferences for not trading in Section 1.1.3.

1.1.1 Separation of Preferences by Aggregation of Certainty-Equivalents

In Chapters 2 and 3, we separate preferences for time and risk by forming certainty equivalents

$$u^{-1}\left(\mathbb{E}_{t,x}\left[u\left(c_s\right)\right]\right) \ , \tag{1.4}$$

where u represents the policyholder's preferences for risk. The entity in Equation (1.4) expresses which certain time-s consumption rate the policyholder requires at time t in order to give up the uncertain time-s consumption rate c_s. Since for all $s \geq t$, the certainty equivalents $u^{-1}\left(\mathbb{E}_{t,x}\left[u\left(c_s\right)\right]\right)$ are known at time t, we are inclined to treat them as deterministic future consumption rates. Now, we let a different function, $\bar{\varphi}$, formalize the policyholder's time preferences with respect to these certainty equivalents. The policyholder's utility from time t and onward is

$$\begin{aligned}
&\int_t^T e^{-\delta(s-t)}\bar{\varphi}\left(u^{-1}\left(\mathbb{E}_{t,x}\left[u\left(c_s\right)\right]\right)\right)\mathrm{d}s + \omega e^{-\delta(T-t)}\bar{\varphi}\left(u^{-1}\left(\mathbb{E}_{t,x}\left[u\left(X_T^{c,\pi}\right)\right]\right)\right) \\
&= \int_t^T e^{-\delta(s-t)}\varphi\left(\mathbb{E}_{t,x}\left[u\left(c_s\right)\right]\right)\mathrm{d}s + \omega e^{-\delta(T-t)}\varphi\left(\mathbb{E}_{t,x}\left[u\left(X_T^{c,\pi}\right)\right]\right)\ , \qquad (1.5)
\end{aligned}$$

where $\varphi = \bar{\varphi} \circ u^{-1}$, and ω is a scaling factor allowing for different weight on utility from consumption and final wealth. For $\bar{\varphi} = u$, corresponding to identical preferences for risk and time, we get back to expected time-additive utility, so our separated preferences are a true generalization of the preferences in Equation (1.1).

Due to the transform φ of the expectation, we cannot exploit the linearity of the expectation operator and the law of iterated expectations. Hence, the problem of maximizing (1.5) goes beyond what can be dealt with by classical dynamic programming. While we are at "destroying" the linearity, we multiply the policyholder's utility with the constant δ and transform it with an increasing function f, yielding the value function

$$V^{c,\pi}(t,x) = f\left(\begin{array}{c} \int_t^T \delta e^{-\delta(s-t)}\varphi\left(\mathbb{E}_{t,x}\left[u\left(c_s\right)\right]\right)\mathrm{d}s \\ +\omega\delta e^{-\delta(T-t)}\varphi\left(\mathbb{E}_{t,x}\left[u\left(X_T^{c,\pi}\right)\right]\right) \end{array}\right) . \qquad (1.6)$$

The function f does not change optimal behavior, and it is convenient for making the mathematical representation of the policyholder's utility as tractable as possible. The choice $f = \varphi^{-1}$ (possibly times a constant) turns out to be particularly convenient. Given this insight, the choice $f = \varphi^{-1} = \mathrm{id}$ (the identity function), corresponding to $\bar{\varphi} = u$, shows why there is no "normalization issue" for time-additive utility.

The problem of maximizing the value function, $V^{c,\pi}$, in Equation (1.6) is absolutely non-standard due to its serial non-linearity, and since dynamic programming does not work, there is no reason to believe that solutions to local and global optimization problems coincide as for time-additive utility. On the contrary, the control resulting from maximizing $V^{c,\pi}$ at time 0 is likely to be inconsistent with the control resulting from maximizing $V^{c,\pi}$ at time $t > 0$. By "inconsistent" we mean that the decision we make at time t based on maximizing $V^{c,\pi}$ is not the same as the decision we plan to make at time t based on maximizing $V^{c,\pi}$ at time 0, for the same realization of the wealth process. We dislike this time-inconsistency, and we do not wish to introduce pre-commitment. Instead, we want to take the policyholder's changing preferences into account. We do this by searching for equilibrium controls rather than classical optimal controls. The equilibrium theory arises from a game theoretic approach to stochastic control. The approach produces a control process that does not maximize the value function over all admissible strategies, but, rather, over "all strategies that one actually intends to follow".

In Chapters 2 and 3, we apply this separation of preferences, but with different extensions. In Chapter 2, we exclude utility from terminal wealth and limit our focus to power utility functions,

$$u(x) = \tfrac{1}{1-\gamma} x^{1-\gamma}, \quad \bar{\varphi}(x) = \tfrac{1}{1-\phi} x^{1-\phi}, \quad \text{and } f(x) = \tfrac{1}{1-\gamma}\left((1-\phi)x\right)^{\theta},$$

with $\gamma, \phi \in \mathbb{R}^+ \setminus \{1\}$ and $\theta = \frac{1-\gamma}{1-\phi} > 0$. In return, we allow for lifetime uncertainty and utility from bequest, and in addition to separating preferences for risk and time, we also separate preferences for substitution between the states "alive" and "dead". This is done by introducing the function $v(x) = \kappa x^{\frac{1}{\kappa}}$ and considering the extended value function

$$\begin{aligned} & V^{c,\pi,d}(t,x) \\ &= f\left(\int_t^T \delta e^{-\delta(s-t)} \varphi \circ v^{-1} \begin{pmatrix} v\left(\mathbb{E}_{t,x}\left[u(c_s) I_s\right]\right) + \\ v\left(\mathbb{E}_{t,x}\left[u\left(X_s^{c,\pi,d} + d_s\right) \frac{\mathrm{d}N_s}{\mathrm{d}s}\right]\right) \end{pmatrix} \mathrm{d}s\right) \qquad (1.7) \\ &= \frac{1}{1-\gamma}\left(\int_t^T \delta e^{-\delta(s-t)} \begin{pmatrix} \left(\mathbb{E}_{t,x}\left[c_s^{1-\gamma} I_s\right]\right)^{\frac{1}{\kappa}} + \\ \left(\mathbb{E}_{t,x}\left[\left(X_s^{c,\pi,d} + d_s\right)^{1-\gamma} \frac{\mathrm{d}N(s)}{\mathrm{d}s}\right]\right)^{\frac{1}{\kappa}} \end{pmatrix}^{\frac{\kappa}{\theta}} \mathrm{d}s\right)^{\theta}. \end{aligned}$$

We present an equilibrium verification theorem for the value function in Equation (1.7), and we derive semi-explicit expressions for the equilibrium control and the corresponding value function. We discover that, in the special case without lifetime uncertainty, our optimization approach results in the same control as recursive utility optimization with Epstein-Zin preferences. We find this interesting since the existing literature on recursive utility optimization with Epstein-Zin preferences does not allow for lifetime uncertainty and utility from bequest. We derive a stochastic differential equation for the equilibrium consumption rate, and in a numerical example, we show that our separation of preferences gives rise to hump-shaped consumption patterns as observed in realized consumption. We note that such hump-shaped consumption patterns cannot be obtained by standard recursive utility or time-additive utility under lifetime uncertainty.

In Chapter 3, we focus on a certain-lived policyholder, but allow for general utility functions and more general wealth dynamics. The wealth of the policyholder evolves according to the dynamics

$$\mathrm{d}X_t^{c,\pi} = \mu^{c,\pi}\left(t, X_t^{c,\pi}, Y_t\right) \mathrm{d}t + \sigma^{c,\pi}\left(t, X_t^{c,\pi}, Y_t\right) \mathrm{d}W_t \ , \ X_0^{c,\pi} = x_0 \ ,$$

where Y is a non-traded state process with the dynamics

$$\mathrm{d}Y_t = \alpha\left(t, Y_t\right)\,\mathrm{d}t + \beta\left(t, Y_t\right)\left(\rho\,\mathrm{d}W_t + \sqrt{1-\rho^2}\,\mathrm{d}\bar{W}_t\right)\ ,\ Y_0 = y_0\ .$$

Here, $\mu^{c,\pi}, \sigma^{c,\pi}, \alpha, \beta$ are sufficiently regular functions, and W and $\bar{W}$ are two independent Brownian motions.

Also, we allow the utility of consumption and terminal wealth to depend on the process Y. More specifically, we replace $u\left(c_s\right)$ by $u\left(Y_s, c_s\right)$. This turns out to be mathematically tractable, and we can, for example, think of Y as an index of purchasing power or a minimum subsistence level, depending on the shape of u. Finally, we introduce separate utility functions for consumption, u_1, and final wealth, u_2. Altogether, we consider the extended value function

$$V^{c,\pi}\left(t,x,y\right) = f\left(\begin{array}{c} \int_t^T \delta e^{-\delta(s-t)}\varphi\left(\mathbb{E}_{t,x,y}\left[u_1\left(Y_s, c_s\right)\right]\right)\mathrm{d}s \\ +\omega\delta e^{-\delta(T-t)}\varphi\left(\mathbb{E}_{t,x,y}\left[u_2\left(Y_T, X_T^{c,\pi}\right)\right]\right)\end{array}\right)\ . \tag{1.8}$$

We present an equilibrium verification theorem for the value function in Equation (1.8) and, for power and exponential utility, we derive semi-explicit expressions for the equilibrium control and the corresponding value function.

By construction, our way of separating preferences is different from that of recursive utility, but for both power and exponential utility functions u_1 and u_2 without Y-dependence, the resulting behavior turns out to coincide with that coming from recursive utility optimization. In particular, we establish a connection to recursive utility optimization that goes beyond power utility and Epstein-Zin preferences. Furthermore, the coincidence is not limited to a Black-Scholes market, but extends to an incomplete market. Hence the connection to recursive utility established in Chapter 2 is extended significantly in Chapter 3.

1.1.2 Separation of Preferences by Recursive Utility

In Chapter 4, we generalize recursive utility to include lifetime uncertainty and utility from bequest. Recursive utility theory deals with the separation of preferences for risk and time through a recursive definition, a certainty equivalent, and a time-aggregator. In discrete time, the present (indirect) utility, $V_{t_k}^c$, is a function of present consumption, c_{t_k}, the time between t_k and t_{k+1}, and the certainty equivalent, $\mathfrak{m}_{t_k}(V_{t_{k+1}}^c)$, of the future utility, $V_{t_{k+1}}^c$. In formulas,

$$V_{t_k}^c = W\left(t_{k+1} - t_k, c_{t_k}, \mathfrak{m}_{t_k}\left(V_{t_{k+1}}^c\right)\right)\ , \tag{1.9}$$

The function W is often referred to as the inter-temporal aggregator because in a set-up without risk, it describes the inter-temporal aggregation of present consumption, c_{t_k}, and the utility of future consumption, $V_{t_{k+1}}^c$. Similarly, the certainty equivalent, $\mathfrak{m}$, is referred to as the risk-aggregator since it describes the risk weighted aggregation of possible future values of $V^c_{t_{k+1}}$.

Recursive utility is widely used to study asset pricing and consumption-portfolio choice in various markets, and it has been used to explore ambiguity aversion and preferences for resolution of uncertainty. Despite a growing literature on recursive utility, there are no attempts to accommodate for lifetime uncertainty and utility from bequest. In particular, the existing literature does not allow for utility from a lump sum at a random point in time, and, therefore, cannot accommodate for utility from bequest.

To introduce lifetime uncertainty and utility from bequest, we replace the backward recursion in Equation (1.9) with the backward recursion

$$V^{c,b}_{t_k} = W\left(t_{k+1} - t_k, c_{t_k}, \mathfrak{m}_{t_k}\left(\mathfrak{n}_{t_k}\left(I_{t_{k+1}} V^{c,b}_{t_{k+1}} + \left(1 - I_{t_{k+1}}\right) b_{t_{k+1}}\right)\right)\right).$$

Here, I_{t_k} indicates survival until time t_k, $b_{t_{k+1}}$ is bequest given death between time t_k and t_{k+1}, and $\mathfrak{n}$ is an additional certainty equivalent that describes the mortality risk weighted aggregation of bequest and future utility. $V^{c,b}_{t_k}$ is the utility given survival up to and including time t_k. The intuition behind the changes to the backward recursion is the following:

- The policyholder is alive at time t_k. At the future time point t_{k+1}, the policyholder is either alive or dead. If the policyholder is alive, the utility is $V^{c,b}_{t_{k+1}}$. If the policyholder is dead, the only utility left is the bequest $b_{t_{k+1}}$.
- The certainty equivalent $\mathfrak{n}$ describes the policyholder's aggregation of bequest and future utility. The aggregation is performed given complete information about the market to focus only on preferences for mortality risk (or, more precisely, preferences for substitution between bequest and future utility).

We extend the generalization to continuous time, and we state a verification theorem with a generalized Hamilton-Jacobi-Bellman equation for the optimal control under recursive utility with lifetime uncertainty.

Recursive utility allows for separation of preferences for risk and time, and, with our generalization, also preferences for substitution between bequest and future utility. The concept of substitution between bequest and future utility is our way of formulating preferences for mortality risk. We study optimal consumption, investment, and life insurance choice under separation of (market) risk aversion, elasticity of inter-temporal substitution, and elasticity of substitution between bequest and future utility. The separation gives rise to hump-shaped consumption patterns as in Chapter 2. We repeat that hump-shaped consumption patterns cannot be obtained by standard recursive utility or time-additive utility under lifetime uncertainty. It is the combination of recursive utility and lifetime uncertainty that enables this feature.

Although different by construction, the established threefold separability appears similar to the one in Chapter 2. We, therefore, explore whether the two ways of separating preferences cover the same set of preferences. In general, the answer is no. The two approaches are different in output as well as in construction. Only if we reduce the number of free preference parameters, thereby giving up on the threefold separability, we identify a coincidence between the preferences covered by the two ways of separating preferences.

1.1.3 Preferences for Not Trading

In Chapter 5, we solve two portfolio optimization problems with the common feature that non-smooth trading is ruled out. By non-smooth trading we mean diffusive trading and bang-bang investment strategies. We work in a classical Black-Scholes financial market, and the ban on non-smooth trading is implemented by restricting the number of stocks to be differentiable and by punishing trading in the objective function. Punishing trading in the objective function is equivalent to the policyholder having explicit preferences for not trading. Diffusive trading arises, e.g., from continuous rebalancing of the constant proportion portfolio in the classical Merton investment problem with power utility. From a practical point of view, diffusive trading is not an option. Apart from it being technically impossible, trading costs prevent such a behavior from being optimal. Trading costs are here thought of as the integrate effects from broker expenses and market impact. If the number of stocks is restricted to be differentiable, diffusive trading is ruled out, but if trading is not punished simultaneously, the result is a bang-bang investment strategy.

There are basically two different ways of punishing trading in the control problem formulation. One way is to implement a cost of trading directly in the wealth process such that trading instantly reduces the policyholder's wealth by a pre-specified amount. An alternative way is to formalize preferences for not trading in the objective function. The latter is, of course, an approximation since, probably, no policyholder has explicit preferences for or against trading, except through its impact on wealth. However, we pursue this indirect formalization of trading costs since it is more mathematically tractable and reveals important insight.

We restrict the trading in two steps: First, we require the number of stocks, N, to be differentiable. This corresponds to requiring $\mathrm{d}N = \tau N \mathrm{d}t$ for some drift coefficient process τ. Second, we punish the trading amount rate, $S\,\mathrm{d}N/\,\mathrm{d}t = \tau A$, for being away from zero. Here, S denotes the stock prices, and $A = NS$ denotes the amount invested in stocks. Due to symmetrization in zero and mathematical tractability, we choose to work with a quadratic utility loss from trading. A quadratic cost function has a convex form that would hold for a trader for whom the effect from market impact dominates that from broker expenses.

We solve two different portfolio optimization problems with a quadratic utility loss from trading. In both problems, the policyholder controls the initial number of stocks, n_0, and the trading speed, τ. First, we consider a policyholder with mean-square utility of terminal wealth. We consider the maximization problem

$$\max_{n_0,(\tau_s)_{0\leq s\leq T}} \mathbb{E}\left[X_T - \gamma X_T^2 - \int_0^T \tfrac{1}{2}\Lambda\left(\tau_s A_s\right)^2 \,\mathrm{d}s\right] .$$

Here, the first two terms are terminal mean-square utility. The third term punishes trading amounts, τA, quadratically, and the parameter $\Lambda \geq 0$ weights the third term against the first two. In this problem formulation, the mean-square utility of terminal wealth matches, mathematically, the quadratic utility loss from trading and gives direct access to a semi-explicit solution, separable in time and wealth. Second, we consider a policyholder with power utility of terminal wealth. Mathematically, the combination of a power utility of terminal wealth and quadratic utility loss from trading is inconvenient. Therefore, we reformulate the power objective by punishing quadratic deviations from the Merton portfolio, π^*. We consider the minimization problem

$$\min_{n_0,(\tau_s)_{0\leq s\leq T}} \mathbb{E}\left[\int_0^T \tfrac{1}{2}\left(\theta(\pi^* X_t - A_t)^2 + (\tau_t A_t)^2\right) \mathrm{d}t\right] .$$

Here, the first term punishes quadratically deviations of the stock holdings from the Merton proportion. The second term punishes trading amounts quadratically. The parameter $\theta \geq 0$ weights the two terms against each other. In both cases, we present semi-explicit solutions and, in numerical examples, we show the impact of trading constraints on the portfolio decision over the investment horizon. In particular, we take the one-liner of "aiming in front of the target" to the level of conventional utility of wealth rather than utility of return on wealth.

1.2 Scenario-Based Projections

In Chapters 6 and 7, we introduce economic scenarios in participating life and unit-linked insurance to allow for market valuation of non-guaranteed payments, pricing and hedging of guarantees, bonus and benefit prognoses, and solvency calculations. We model life insurance contracts using two interacting accounts described by stochastic differential equations. One account measures the assets, and the other account is a technical account. For each economic scenario, the sample paths of the stochastic differential equations are known and can be used to project the two accounts. The scenarios may be worst-case scenarios, scenarios generated via Monte Carlo simulation or best-estimate scenarios. For scenarios generated via Monte Carlo simulation, one obtains a valid approximation of the expected future payments, guaranteed as well as non-guaranteed, by averaging over sufficiently many projections (as is common practice with Monte Carlo simulation). For worst-case or best-estimate scenarios, a single projection is enough to obtain the corresponding worst-case or best-estimate approximation of the future payments.

By use of a two-account model, we are able to illustrate general concepts without making the model too abstract. Also, our two-account model offers a common framework for modeling guaranteed and non-guaranteed payments in participating life and unit-linked insurance which allows us to address similarities and differences between participating life insurance and unit-linked insurance. In participating life insurance, we formalize how the bonus schemes "consolidation" and "additional benefits" work and interact, and in unit-linked insurance, we focus on the implementation of guarantees. For both product types, we provide numerical examples based on Monte Carlo simulation to demonstrate the possible applications of our two-account model.

The topic of scenario-based projection is split in two chapters since there is a fundamental difference between projecting for the purpose of valuation and prognoses. For the purpose of valuation, it is the expected evolution of the policy, both financially and across policyholder states, that is relevant. Hence, the evolution of the policy is considered on an average "portfolio level". Projection on portfolio level is the topic of Chapter 6. The chapter distinguishes itself from the existing literature by taking into account the Markov model for the state of the policyholder, thereby including event risk. For retirement savings, benefit, and bonus prognoses, it is the expected financial evolution of the policy that is relevant. The policyholder needs to know what to expect in a certain state, not the expectation across states of life. Hence, the evolution of the policy is considered on an individual "policy level". Projection on policy level is the topic of Chapter 7. The chapter is, to our knowledge, the first paper to address risk-based prognoses from the policyholder's perspective in participating life and unit-linked insurance in a general financial market. In Chapter 6 we project on portfolio level by taking market expectation across future states of the policy. This corresponds to the policy evolving according to its expectation under the market basis which is exactly what is relevant for valuation. The projection procedure is summarized in Section 1.2.1. In Chapter 7 we project on policy level by conditioning on the policy staying in a certain state, 0. This corresponds to the policyholder continuing his course of life which is exactly what most prognoses from the policyholder's perspective focus on. The projection procedure is summarized in Section 1.2.2.

In unit-linked insurance, projection on portfolio and policy level is almost the same. The only difference is that, on portfolio level, the two accounts of the policy are projected by adding and subtracting the market expected premiums and benefits, whereas, on policy level, the two accounts are projected by adding and subtracting the actual premiums and benefits in state 0 and subtracting the market risk premium associated with jumps out of the state. In both cases, each economic scenario consists of two sample paths: one for the short interest rate and one for the return of the fund that the policyholder invests in. Participating life insurance differs from unit-linked insurance by having collective funds. In particular, the amount of bonus allocated to a policy depends on the evolution of the whole portfolio. As a result, projection on portfolio and policy level differs somewhat in participating insurance. On portfolio level, the two accounts of the policy are projected by adding and subtracting the market expected premiums and benefits, exactly as for unit-linked insurance. On policy level, one needs sample paths for the bonus allocation as stochastic input (instead of sample paths for the short interest rate and the fund return). Given the bonus allocation, the policy can be modeled using just one account, namely the technical account of the policy. Sample paths for the bonus allocation can be obtained by projection on portfolio level. Hence, projection on policy level is a two-step task, but the second step is simple.

In participating life insurance, we consider a policy with guaranteed payments based on a technical basis. The conservative basis gives rise to bonus, and we focus on a bonus scheme consisting of two steps: first, consolidation, and then, when the policy is consolidated on a sufficiently low technical interest rate (if ever), additional benefits. The assets of the policy, X, including its share of the collective bonus potential, are invested in a fund with stochastic return, R_X. The technical reserve, Y, accumulates according to the technical interest rate. The assets and the technical reserve are the backbone of the two-account model. In good times, the return rate on the assets exceeds the technical interest rate. Parts of the excess return are allocated to the policy in terms of bonus, d, which adds to the technical reserve, but parts are saved for times where the return rate on the assets is less favorable. In really bad times, the assets may be insufficient to cover the guaranteed payments of the policy. In that case, the equity holders of the insurance company step in with a capital injection, g, taken from the company's equity. The policyholder pays for the company's risk taking by having a guarantee fee, π_g, deducted from the assets and paid to the equity holders of the insurance company in good times.

In unit-linked insurance, we consider a policy where parts of the benefits are directly linked to the assets of the policy. The policy includes a guaranteed minimum retirement savings amount at the retirement date, based on a guarantee account with a guaranteed interest rate. The assets of the policy, X, are invested in a fund with a stochastic return rate. The guarantee account, Y, accumulates according to the guaranteed interest rate, r^*. The assets and the guarantee account are the backbone of the two-account model. In good times, the return rate on the assets, R_X, exceeds the technical interest rate, and then, the assets outgrow the guarantee account. In that case, the guarantee account is increased with an upgrade, u, according to the terms of the contract. At retirement, the maximum value of the assets and the guarantee account is paid out to the policyholder. In bad times where the guarantee account exceeds the assets at retirement, the equity holders of the insurance company step in with a capital injection, g, taken from the company's equity. The policyholder pays for the company's risk taking by having a guarantee fee, π_g, deducted from the assets and paid to the equity holders of the insurance company.

In participating life insurance, as well as in unit-linked insurance, we consider a policy whose state-wise evolution is governed by a continuous-time Markov process with a finite state space. We assume that the process governing the state of the policy is independent of the financial market. The payments of the policy consist of a state-dependent payment stream

$$B^u + B^f - C \ ,$$

where C is the premium stream ("C" for contributions), B^f is a fixed benefit stream ("B" for benefits, and superscript "f" for fixed), and B^u is a benefit stream that depends on the financial evolution. In participating life insurance, B^u denotes the benefits that are upscaled under the bonus scheme additional benefits (superscript "u" for upscaled). In unit-linked insurance, B^u denotes benefits that are linear in the assets of the policy (superscript "u" for unit-linked). The premium stream consists of state-wise continuous payments, and the benefit streams consist of state-wise continuous payment and lump sum payments upon jumps.

1.2.1 Scenario-Based Projections for Valuation

In Chapter 6, all projection is on portfolio level. In participating life insurance, we project the assets, X, and the technical reserve, Y, using the stochastic differential equations

$$\begin{aligned} \mathrm{d}X(t) &= X(t-)\,\mathrm{d}R_X(t) - \mathrm{d}\beta^f(t) - k^{(\varepsilon(t))}\mathrm{d}\beta^u(t) + \mathrm{d}\varsigma(t) \\ &\quad + [g(t) - \pi_g(t)]\,\mathrm{d}\varepsilon(t) \ , \\ \mathrm{d}Y(t) &= Y(t)\,r^{*(\varepsilon(t))}\mathrm{d}t - \mathrm{d}\beta^f(t) - k^{(\varepsilon(t))}\mathrm{d}\beta^u(t) + \mathrm{d}\varsigma(t) \\ &\quad + d(t)\,\mathrm{d}\varepsilon(t) + \alpha\left(t, r^{*(\varepsilon(t))}, k^{(\varepsilon(t))}\right)\mathrm{d}t \ . \end{aligned} \tag{1.10}$$

Here, R_X, g, π_g, and d are introduced on page 11, and $\varepsilon(t)$ counts the number of updates of guarantee injection, guarantee fee and bonus. In addition,

- $\beta^f(t)$, $\beta^u(t)$, and $\varsigma(t)$ are the market expected benefits and premiums,
- $r^{*(\varepsilon(t))}$ is the technical interest rate (after consolidation), and $k^{(\varepsilon(t))}$ is the upscaling factor (after additional benefits),
- $\alpha\left(t, r^{*(\varepsilon(t))}, k^{(\varepsilon(t))}\right)$ is an adjustment term accounting for the market expected surplus arising from the technical transition intensities.

We formalize a procedure for calculating the consolidated technical interest rate $r^{*(\varepsilon(t))}$ and the upscaling factor $k^{(\varepsilon(t))}$, and we address the difficulties of calculating them on portfolio level.

In unit-linked insurance, we project the assets, X, and the guarantee account, Y, using the stochastic differential equations

$$\begin{aligned} \mathrm{d}X(t) &= X(t-)\,\mathrm{d}R_X(t) - X(t-)\,\mathrm{d}\beta^p(t) - \mathrm{d}\beta^f(t) + \mathrm{d}\varsigma(t) \\ &\quad - \pi_g(t)\,\mathrm{d}\varepsilon(t) + g\,\mathrm{d}\varepsilon_R(t) \\ &\quad - 1_{\{t=R\}}\left(Y(R-) - X(R-)\right)^+ \mathrm{d}\beta^p(t) \ , \\ \mathrm{d}Y(t) &= Y(t)\,r^*(t)\,\mathrm{d}t - X(t-)\,\mathrm{d}\beta^p(t) - \mathrm{d}\beta^f(t) + \mathrm{d}\varsigma(t) \\ &\quad + u(t)\,\mathrm{d}\varepsilon(t) \ . \end{aligned} \tag{1.11}$$

Here, R_X, g, π_g, and u are introduced on page 11, and $\varepsilon(t)$ counts the number of guarantee fee payments and guarantee upgrades. In addition,

- $\beta^p(t)$, $\beta^f(t)$, and $\varsigma(t)$ are the market expected benefits and premiums,
- ε_R marks the exercise of the guarantee at the retirement date,
- The last term in the equation for X ensures that the guarantee injection at time R is included in a possible lump sum payment at time R.

The projection in unit-linked insurance is similar to the projection in participating life insurance. The bonus updates, d, in (1.10) are replaced by guarantee upgrades, u, in (1.11), and the running guarantee, g, in (1.10) is replaced by the final guarantee, g, in (1.11). Also, both unit-linked accounts are based on market transition intensities, so the adjustment term, α, from (1.10) vanishes. Apart from that, the driving stochastic differential equations are the same in participating life insurance and unit-linked insurance. In both cases, the stochastic element, R_X, enters via a sample path for the asset return.

As mentioned earlier in the introduction, for both product types, we project the two accounts by adding and subtraction the market expected premiums and benefits to reflect the expected evolution of the policy across states. The projections are suitable for calculating market cash flows and market values of guaranteed *and* non-guaranteed payments, determining a fair guarantee fee, bonus allocation, and/or guarantee upgrade, and assessing solvency capital requirements.

1.2.2 Scenario-Based Projections for Prognoses

In Chapter 7, all projection is on policy level. In participating life insurance, we project the technical reserve, Y, conditional on continued sojourn in state 0, using the stochastic differential equation

$$\begin{aligned} \mathrm{d}Y(t) &= Y(t)\,r^{*(\varepsilon(t))}(t)\,\mathrm{d}t + \mathrm{d}c_0(t) - k^{(\varepsilon(t))}\mathrm{d}b_0^u(t) - \mathrm{d}b_0^f(t) \\ &\quad - \mathrm{d}s_0^*(t) + d_0(t)\,\mathrm{d}\varepsilon(t) \ . \end{aligned}$$

Here, ε, $r^{*(\varepsilon(t))}(t)$ and $k^{(\varepsilon(t))}$ are as in (1.10), and

- $b_0^f(t)$, $b_0^u(t)$, and $c_0(t)$ are the benefits and premiums in state 0,
- s_0^* is the technical risk premium in state 0,
- d_0 is the stochastic bonus allocation in state 0.

The stochastic element d_0 enters via a sample path for the bonus allocation. We formalize a procedure for calculating $r^{*(\varepsilon(t))}$ and $k^{(\varepsilon(t))}$ on policy level which is conceptually simpler than on portfolio level.

In unit-linked insurance, we project the assets, X, and the guarantee account, Y, conditional on continued sojourn in state 0, using the stochastic differential equations

$$\begin{aligned}
\mathrm{d}X(t) &= X(t-)\,\mathrm{d}R_X(t) + \mathrm{d}c_0(t) - \mathrm{d}b_0^f(t) - X(t-)\,\mathrm{d}b_0^p(t) \\
&\quad - \mathrm{d}s_0^f(t) - \mathrm{d}s_0^u(t) \\
&\quad - \pi_g(t)\,\mathrm{d}\varepsilon(t) + (Y(R-) - X(R-))^+\,\mathrm{d}\varepsilon_R(t)\ , \\
\mathrm{d}Y(t) &= Y(t)\,r^*(t)\,\mathrm{d}t + \mathrm{d}c_0(t) - \mathrm{d}b_0^f(t) - X(t-)\,\mathrm{d}b_0^p(t) \\
&\quad - \mathrm{d}s_0^f(t) - \mathrm{d}s_0^u(t) + u(t)\,\mathrm{d}\varepsilon(t)\ .
\end{aligned}$$

Here, R_X, π_g, u, ε, and ε_R are as in (1.11), and

- $b_0^f(t)$, $b_0^p(t)$, and $c_0(t)$ are the benefits and premiums in state 0,
- s_0^u and s_0^f are the market risk premiums in state 0.

The stochastic element R_X enters via a sample path for the asset return.

As mentioned earlier in the introduction, for both product types, we project the accounts by adding and subtraction the actual premiums and benefits in state 0 and subtracting the market risk premium associated with jumps out of the state to reflect the evolution of the policy in this particular state. The projections are suitable for producing bonus, benefit, and retirement savings prognoses with confidence intervals, given that the policyholder continues his course of life.

Chapter 2

Personal Finance and Life Insurance under Separation of Risk Aversion and Elasticity of Substitution

Ninna Reitzel Jensen & Mogens Steffensen (2015)
Insurance: Mathematics and Economics 62, pp. 28–41.

Abstract: In a classical Black-Scholes market, we establish a connection between two seemingly different approaches to continuous-time utility optimization. We study the optimal consumption, investment, and life insurance decision of an investor with power utility and an uncertain lifetime. To separate risk aversion from elasticity of inter-temporal substitution, we introduce certainty equivalents. We propose a time-inconsistent global optimization problem, and we present a verification theorem for an equilibrium control. In the special case without mortality risk, we discover that our optimization approach is equivalent to recursive utility optimization with Epstein-Zin preferences in the sense that the two approaches lead to the same result. We find this interesting since our optimization problem has an intuitive interpretation as a global maximization of certainty equivalents and since recursive utility, in contrast to our approach, gives rise to severe differentiability problems. Also, our optimization approach can there be seen as a generalization of recursive utility optimization with Epstein-Zin preferences to include mortality risk and life insurance.

Keywords: Recursive utility, lifetime uncertainty, stochastic control, generalized Hamilton-Jacobi-Bellman equation, time-inconsistency, certainty equivalents.

2.1 Introduction

In a classical Black-Scholes market, we establish a connection between two seemingly different approaches to continuous-time utility optimization for a certain-lived investor. One approach is recursive utility optimization with Epstein-Zin preferences, studied in Duffie and Epstein (1992b) and Kraft and Seifried (2010) for general preferences. The other approach is non-linear expected power utility optimization with dynamic updating, studied in this paper for an uncertain-lived investor. This approach is apt for a set-up with mortality risk and utility from inheritance, and because of the established connection for a certain-lived investor, our approach can be seen as a generalization of the recursive utility approach to a set-up with mortality risk and life insurance.

Over time, the optimal consumption and investment decisions of a certain-lived investor have been treated in various papers. An important, early example is Merton (1971) who considers time-additive utility optimization in continuous time. Using dynamic programming techniques, the value function of the time-additive optimization problem can be characterized by a partial differential equation. The equation is called a Hamilton-Jacobi-Bellman equation, and it includes a term $u(c)$ where u is the investor's utility function for consumption and c is the consumption rate.

Richard (1975) generalized the work by Merton (1971) to include mortality risk and life insurance. The value function, V, of the generalized optimization problem is characterized by a partial differential equation similar to the original Hamilton-Jacobi-Bellman equation. The main alteration consists in addition of the term

$$\mu(t)\,\tilde{u}(b+x) - \mu(t)\,V(t,x)\ , \tag{2.1}$$

where μ is the investor's mortality intensity, $\tilde{u}$ is the investor's utility function for inheritance, b is a term insurance sum paid out upon death, and x is wealth. Also, there is an effect on the wealth dynamics due to financing of the term insurance. We note that $\mu(t)\,\tilde{u}(b+x)$ can be interpreted as the investor's probability weighted utility gain associated with death. Similarly, $\mu(t)\,V(t,x)$ can be interpreted as the investor's probability weighted utility loss associated with death. The term in (2.1) is therefore the investor's probability weighted net-gain associated with death.

Unfortunately, time-additive utility has the disadvantage that it mixes preferences for risk and preferences for inter-temporal substitution. The recursive utility approach and our approach both deal with this problem, in two seemingly different ways.

Recursive utility is founded in discrete time, and it allows for separation of preferences for risk and inter-temporal substitution through a recursive definition, a (utility) certainty equivalent and a time-aggregator. In Duffie and Epstein (1992b), recursive utility is extended to continuous time where it is called stochastic differential utility. The link to discrete-time recursive utility is vague though, and in Kraft and Seifried (2010), the extension is refined and called continuous-time recursive utility. In both papers, the optimal consumption and investment decisions of a certain-lived investor are studied. The value function, V, of the recursive optimization problem is characterized by a Hamilton-Jacobi-Bellman equation (in the following 'pseudo-Bellman equation') where the term $u(c)$ is replaced by a term $f(c, V(t,x))$. Here, f is the normalized aggregator representing the investor's preferences. In particular, Epstein-Zin preferences are represented by the aggregator

$$f(c,V) = \theta\delta V\left(\left(\frac{c}{((1-\gamma)V)^{\frac{1}{1-\gamma}}}\right)^{\frac{1-\gamma}{\theta}} - 1\right) .$$

The recursive optimization problem is less intuitive than the time-additive optimization problem, and to our knowledge, the literature contains no attempt to extend the recursive utility problem to a set-up with mortality risk and life insurance. However, inspired by the mortality extension in Richard (1975), it is natural to suggest a pseudo-Bellman equation where we combine $f(c,V)$ defined above with the additional term $\mu(t)\,\tilde{u}(b+x) - \mu(t)\,V(t,x)$.

For Epstein-Zin preferences, we present another suggestion—namely an alteration of the normalized aggregator (and no additional term). The altered aggregator arises from the following optimization approach: we consider an uncertain-lived investor with power utility. To separate preferences for risk and preferences for inter-temporal substitution, we introduce consumption certainty equivalents, and we propose a time-global optimization problem that is about maximizing an infinite sum of infinitesimally small certainty equivalents for future consumption and inheritance. The problem is non-linear in expectation, and consequently it is time-inconsistent in the sense that its solution does not obey Bellman's optimality principle. In other words: if we solve the problem at time 0 and apply the corresponding control up to a future time point $t > 0$, then at this future time point, the control is no longer optimal. For more on time-inconsistency, see e.g. Björk et al. (2014) or Björk and Murgoci (2010). To deal with the time-inconsistency, we search for an equilibrium control instead of a classical optimal control, and we present a verification theorem for a particular equilibrium control. The corresponding value function is characterized by a pseudo-Bellman equation where the term $f(c, V(t,x))$ is replaced by the term $\tilde{f}(t, c, x+b, V(t,x))$. Here, the altered aggregator $\tilde{f}$ is given by

$$\tilde{f}(t,c,y,V) = \theta\delta V\left(\left(\frac{c^{1-\gamma}}{V(1-\gamma)}\right)^{\frac{1}{\kappa}} + \left(\frac{\varepsilon(t)\,\mu(t)\,y^{1-\gamma}}{V(1-\gamma)}\right)^{\frac{1}{\kappa}}\right)^{\frac{\kappa}{\theta}} - (\mu(t) + \theta\delta)\,V .$$

For a certain-lived investor (i.e. $\mu = 0$), the two aggregators f and $\tilde{f}$ coincide, and so our approach leads to the same result as recursive utility optimization with Epstein-Zin preferences, for a certain-lived investor. Because of this equivalence, the aggregator $\tilde{f}$ can be seen as a mortality extension of the normalized Epstein-Zin aggregator—that is, our approach can be seen as a generalization of the recursive utility approach with Epstein-Zin preferences to a set-up with mortality risk and life insurance. This proposal is supported by the fact that our optimization problem has an intuitive interpretation as a global maximization of certainty equivalents, both with and without mortality risk. Furthermore, our approach is a generalization of the time-additive utility optimization in Richard (1975) to time-non-additive power utility.

Recursive utility is considered as a standard way to separate risk aversion from elasticity of inter-temporal substitution. We provide a new way to formalize such a separation where, first, risk aversion forms certainty equivalents and, then, elasticity of substitution forms time-global preferences. Yet, a completely different approach to the separation is suggested in Kihlstrom (2009). In discrete time, he suggests to formalize a separation where, first, elasticity of substitution forms time-global preferences and, then, risk aversion forms one certainty equivalent. Since his formalization is not immediately tractable with our method, future research should address further the relation between Kihlstrom's approach, our approach, and recursive utility.

We emphasize that our optimization problem is not a special case of Björk and Murgoci (2010) as our objective function has a considerably different form. In particular, their result about coincidence of solutions for certain time-consistent and time-inconsistent problems does not explain the equivalence between our approach and recursive utility optimization with Epstein-Zin preferences. Also, we wish to focus on our specific investor problem and not on time-consistency in general, so we do not go into details on the game-theoretic equilibrium approach.

We work in a simple Black-Scholes market because we wish to study the qualitative structures of the solution to our optimization problem. We then avoid drowning our key insights in notation and multidimensionality, and we avoid resorting to numerical optimization. For qualitative insight, sticking to a simple model remains efficient.

Structure of the paper

In Section 2.2, we propose an optimization problem and introduce the concept of equilibrium controls. We present a verification theorem for a particular equilibrium control, and we derive closed-form expressions for the control and the corresponding value function. Finally, we compare our results to Richard (1975). In Section 2.3, we give a short introduction to recursive utility, and we demonstrate the similarity of our pseudo-Bellman equation and the pseudo-Bellman equation in Duffie and Epstein (1992b). Also, we outline perspectives of the established equivalence. In Section 2.4, we derive a stochastic differential equation for the optimal consumption rate from Section 2.2, and we construct numerical examples to illustrate how it differs from the optimal consumption rate from time-additive utility. The numerical examples all arise from the special case without market risk.

2.2 Optimization problem

2.2.1 Set-up

We consider an investor making decisions concerning consumption, investment, and life insurance in continuous time. We adopt the classical survival model, and by N and $I = 1 - N$, we indicate whether the investor is dead or alive at a given point in time (e.g. $N(t) = 1$ if the investor is dead at time t). We treat N and I as stochastic processes on an abstract probability space $(\Omega, \mathcal{F}, P)$, and we model the death of the investor by a mortality intensity μ, i.e.

$$P(I(t) = 1) = P(I(s) = 1 : s \in [0, t]) = e^{-\int_0^t \mu(v)\,\mathrm{d}v} \ , \quad t \geq 0 \ .$$

The investor has access to a classical Black-Scholes market consisting of a bank account, B, with risk free short rate r, and a stock, S, with excess return λ and volatility σ. The asset prices are described by the stochastic differential equations (SDEs)

$$\begin{aligned}
\mathrm{d}B(t) &= B(t)\, r\mathrm{d}t \ , \quad t \geq 0 \ , \quad B(0) = 1 \ , \\
\mathrm{d}S(t) &= S(t)\left[(r + \lambda)\,\mathrm{d}t + \sigma \mathrm{d}W(t)\right] \ , \quad t \geq 0 \ , \quad S(0) = s_0 \ ,
\end{aligned}$$

where $r, \lambda, \sigma > 0$ are constants, and W is a standard Brownian motion on the probability space $(\Omega, \mathcal{F}, P)$.

Also, the investor can trade term insurance contracts with a life insurance company. Note that there is no loss of generality in only considering term insurance, since all available life insurance products are linear combinations of term insurance contracts and a savings plan. A death sum b triggers premiums payments at rate $b\hat{\mu}$. Here, $\hat{\mu}$ is the mortality intensity used by the insurance company for pricing, and it may or may not be equal to μ. The term insurance completes the market. For simplicity, we assume that the insurance company does not pay out bonus. Also, because of the chosen premium structure, the insurance company does not build up reserves, but e.g. fixed premiums in combination with reserve building would not alter our fundamental results. This degree of freedom assumes free access to changing the premium and the death sum in accordance with the equivalence principle at any point in time, though. For detailed discussions and calculations in this direction, see Kraft and Steffensen (2008).

We fix a time-horizon T that we think of as the investor's maximum remaining lifetime. The investor has wealth X and invests a proportion π of X in the stock and a proportion $(1 - \pi)$ of X in the bank account. As long as the investor is alive, she consumes at rate c, earns money at rate w (deterministic), and buys life insurance at premium rate $b\hat{\mu}$. When the investor dies, her inheritors receive the death sum b and the remaining wealth. While the investor is alive, her wealth evolves according to the SDE

$$\begin{aligned} \mathrm{d}X(t) &= X(t)\left[(r+\pi(t)\lambda)\,\mathrm{d}t + \pi(t)\,\sigma \mathrm{d}W(t)\right] \\ &\quad - \left(c(t) + b(t)\,\hat{\mu}(t) - w(t)\right)\mathrm{d}t\ , \qquad t \in [0,T]\ , \\ X(0) &= x_0\ , \end{aligned} \tag{2.2}$$

where x_0 is the initial wealth of the investor, w is a continuous, deterministic function, and c, π, b are stochastic processes, i.e.

$$c, \pi, b : [0,T] \times \Omega \to \mathbb{R}\ . \tag{2.3}$$

In addition to the investor's monetary wealth, we also formalize the investor's human wealth which we denote by L. We do this here because the quantity arises in the solution to problems similar to ours. The investor's human wealth is the financial value of her future labour income, and it is given by

$$L(t) = \int_t^T w(s)\, e^{-\int_t^s (r+\hat{\mu}(v))\,\mathrm{d}v}\,\mathrm{d}s\ , \quad t \in [0,T]\ . \tag{2.4}$$

We note that $\hat{\mu}$ (and not μ) appears in (2.4) because $\hat{\mu}$ is the intensity used for pricing the term insurance, and this asset completes the market.

Since the investor cannot look into the future, it is natural to require that the set of control processes (c, π, b) is adapted to the wealth process X. However, for computational convenience, we go one step further and require that (c, π, b) is of feedback form, i.e.

$$\left(c(t), \pi(t), b(t)\right) = \left(\tilde{c}(t, X(t)), \tilde{\pi}(t, X(t)), \tilde{b}(t, X(t))\right)\ , \quad t \in [0,T]\ ,$$

for deterministic, measurable functions

$$\tilde{c}, \tilde{\pi}, \tilde{b} : [0,T] \times \mathbb{R} \to \mathbb{R}\ . \tag{2.5}$$

For simplicity, we redefine $(c, \pi, b) \equiv \left(\tilde{c}, \tilde{\pi}, \tilde{b}\right)$ and speak of the function (c, π, b) as a control. We thereby leave out the tildes in (2.5) and overtype the processes in (2.3). Now the SDE in (2.2) reads

$$\begin{aligned} \mathrm{d}X(t) &= X(t)\left[(r+\pi(t, X(t))\lambda)\,\mathrm{d}t + \pi(t, X(t))\,\sigma \mathrm{d}W(t)\right] \\ &\quad - \left(c(t, X(t)) + b(t, X(t))\,\hat{\mu}(t) - w(t)\right)\mathrm{d}t\ ,\ t \in [0,T]\ , \\ X(0) &= x_0\ , \end{aligned} \tag{2.6}$$

where c, π, b are deterministic, measurable functions of time and wealth.

Definition 2.1. To ensure that (2.6) makes sense, we only consider controls (c, π, b) for which the SDE in (2.6) has a unique solution. Also, we require that the investor's total wealth $X + L$, consumption rate c, and inheritance

$X + b$ never fall below 0. To ensure this, we only consider controls (c, π, b) for which $(c(t,x), \pi(t,x), b(t,x))$ belongs to the set

$$\Gamma(t,x) \equiv \begin{cases} [0,\infty) \times \mathbb{R} \times [-x,\infty) & \text{if } x + L(t) > 0 \ , \\ \{0\} \times \{0\} \times \{-x\} & \text{if } x + L(t) = 0, L(t) > 0 \ , \\ \{0\} \times \mathbb{R} \times \{0\} & \text{if } x = L(t) = 0 \ . \end{cases}$$

It is easy to verify that this constraint ensures the required non-negativity. We say that a control (c, π, b) is *admissible* if it meets the requirements above, and by $\mathcal{U}$ we denote the set of admissible controls. In Subsection 2.2.3, we impose some additional constraints on the admissible controls.

2.2.2 Formulation

For a moment, we think of the investor as certain-lived, i.e. we let $\mu = \hat{\mu} = 0$ in the set-up from the previous subsection. Then a classical optimization problem for the investor is that of maximizing expected time-additive power utility of consumption, i.e.

$$\sup_{c,\pi} E\left[\int_0^T e^{-\delta t} \frac{1}{1-\gamma} c^{1-\gamma}(t, X(t))\, \mathrm{d}t\right] \ , \tag{2.7}$$

where $\delta \geq 0$ is a subjective utility discount rate, $\gamma > 0, \gamma \neq 1$, is thought of as risk aversion, and (c, π) is chosen among a suitable set of admissible controls. This problem can be dealt with by considering the value function

$$W(t,x) = \sup_{c,\pi} E_{t,x}\left[\int_t^T e^{-\delta s} \frac{1}{1-\gamma} c^{1-\gamma}(s, X(s))\, \mathrm{d}s\right] \ ,$$

where $E_{t,x}$ denotes conditional expectation given $X(t) = x$. By application of dynamic programming techniques, the value function can be characterized by the Hamilton-Jacobi-Bellman equation, i.e. a partial differential equation containing a local optimization problem at each point (t, x). Using the linearity of the expectation operator and the law of iterated expectation, it can be proven that the solution (c, π) to the continuum of local optimization problems is also a solution to the global optimization problem (see e.g. Chapter 19 in Björk (2009)). In the following, the linearity (in expectation) of the optimization problem is disrupted, and then there is no longer coincidence between local and global optimization.

We mentioned that γ is thought of as risk aversion, but γ also plays a role in the time-additivity of (2.7). The parameter γ does not only represent aversion towards risk, it is also related to the Elasticity of Inter-temporal Substitution (EIS). Whereas risk aversion expresses the investor's willingness to gamble, EIS expresses the investor's willingness to substitute consumption over time. To illustrate this, we take away the investor's option to invest in the stock. We are then faced with the deterministic optimization problem

$$\sup_c \int_0^T e^{-\delta t} \frac{1}{1-\gamma} c^{1-\gamma} (t, X(t)) \, \mathrm{d}t \;, \tag{2.8}$$

where X is now a deterministic process. Since there is no risk left in the set-up, the solution to (2.8) should not be related to the investor's aversion towards risk, but the solution does depend on γ. Hence, we have found a way to formalize EIS in the case of no risk, and this motivates our formalization of EIS below, in the presence of risk.

In this paper, we separate risk aversion from EIS by forming certainty equivalents

$$u^{-1} (E [u (c (t, X(t)))]) \;, \tag{2.9}$$

where u is a utility function representing the investor's preferences for risk. We then add certainty equivalents (while taking EIS into account) instead of adding utility. The entity in (2.9) is deterministic and expresses which certain time-t consumption rate the investor requires at time 0 in order to give up the uncertain time-t consumption rate $c(t, X(t))$. In the case of power utility, i.e. $u(c) = \frac{1}{1-\gamma} c^{1-\gamma}$, the certainty equivalent in (2.9) equals

$$\left(E \left[c^{1-\gamma} (t, X(t)) \right] \right)^{\frac{1}{1-\gamma}} \;.$$

For the addition of certainty equivalents, we introduce an EIS-parameter $\phi > 0, \phi \neq 1$, and formalize EIS as in (2.8). This gives us the problem

$$\sup_{c,\pi} \int_0^T e^{-\delta t} \frac{1}{1-\phi} \left(E \left[c^{1-\gamma} (t, X(t)) \right] \right)^{\frac{1}{\theta}} \mathrm{d}t \tag{2.10}$$

with $\theta = \frac{1-\gamma}{1-\phi}$. The special case $\gamma = \phi$ corresponds to the problem in (2.7). Given basic knowledge of dynamic programming, it is clear that the problem in (2.10) cannot be dealt with using classical dynamic programming techniques. This is due to the power $\frac{1}{\theta}$. While we are at spoiling linearity, we make yet another transformation and face the problem

$$\sup_{c,\pi} \frac{1}{1-\gamma} \left(\int_0^T \delta e^{-\delta t} \left(E \left[c^{1-\gamma} (t, X(t)) \right] \right)^{\frac{1}{\theta}} \mathrm{d}t \right)^{\theta} \;. \tag{2.11}$$

This problem is equivalent to the problem in (2.10)—that is, if $\delta > 0$ and $(1-\phi)(1-\gamma) > 0$. By 'equivalent' we mean that the control (c, π) realizing the supremum in (2.10) is identical to the control (c, π) realizing the supremum in (2.11). From now on, we assume that $\delta > 0$ and $(1-\phi)(1-\gamma) > 0$, and it turns out that the problem in (2.11) is more convenient to work with than the problem in (2.10). The constants δ and $\frac{1}{1-\gamma}$ match the powers $-\delta$ and

$1-\gamma$ (which is convenient for differentiation), and in some ways, the power θ offsets the complications from the power $\frac{1}{\theta}$. We note that the factor $\frac{1}{1-\gamma}$ is placed outside the integral (and the parentheses) because the factor can be negative and should therefore not be taken to the power θ or $\frac{1}{\theta}$.

Finally, we go back to the original set-up with mortality risk. We assume that the processes N and I are independent of the process W, and we propose to consider the generalized optimization problem

$$\sup_{(c,\pi,b)\in\mathcal{U}} \frac{1}{1-\gamma} \left(\int_0^T \delta e^{-\delta t} \left(\begin{array}{c} \left(E\left[c^{1-\gamma}\left(t, X^{c,\pi,b}(t)\right) \frac{I(t)\,\mathrm{d}t}{\mathrm{d}t} \right]\right)^{\frac{1}{\kappa}} + \\ \left(E\left[\begin{array}{c} \varepsilon(t) \frac{\mathrm{d}N(t)}{\mathrm{d}t} \times \\ \left(\begin{array}{c} X^{c,\pi,b}(t) + \\ b\left(t, X^{c,\pi,b}(t)\right) \end{array} \right)^{1-\gamma} \end{array} \right]\right)^{\frac{1}{\kappa}} \end{array} \right)^{\frac{\kappa}{\theta}} \mathrm{d}t \right)^{\theta}, \tag{2.12}$$

where the expectation operates on all underlying random variables (i.e. W, N, and I), $\mathcal{U}$ is the set of admissible controls defined in Definition 2.1, and ε is a non-negative, continuous, deterministic weight function. Up to a scaling, the first mean value is the expected utility from consumption, and the second mean value is the expected utility from inheritance. We are aware that the expression $\frac{\mathrm{d}N(t)}{\mathrm{d}t}$ leaps to the eye since the process N is not differentiable. However, we use the expression $E\left[\frac{\mathrm{d}N(t)}{\mathrm{d}t}\right]$ as a heuristic representation of $e^{-\int_0^t \mu(v)\,\mathrm{d}v}\mu(t)$, and $\frac{\mathrm{d}N(t)}{\mathrm{d}t}$ never appears outside an expectation operator. We have included the function ε to allow for a different weight on inheritance than on consumption and to allow for a changing weight on inheritance throughout life. We have introduced the additional parameter $\kappa > 0$ to allow for separation of risk aversion and elasticity of substitution between consumption and inheritance, and we have equipped X with superscript c, π, b to emphasize that it is the wealth process stemming from the control (c, π, b). *Altogether, the generalized problem in* (2.12) *is a question of maximizing an infinite sum of infinitesimal certainty equivalents for future consumption and inheritance.* The problem is complicated, and inspired by dynamic programming, one could try to look at the value function

$$W(t,x) = \sup_{(c,\pi,b)\in\mathcal{U}} Z^{c,\pi,b}(t,x) \tag{2.13}$$

where the objective function $Z^{c,\pi,b} : [0,T] \times \mathbb{R} \to \mathbb{R}$ is given by

$$Z^{c,\pi,b}(t,x) =$$

$$\frac{1}{1-\gamma}\left(\int_t^T \delta e^{-\delta(s-t)} \begin{pmatrix} \left(E^0_{t,x}\left[c^{1-\gamma}\left(s, X^{c,\pi,b}(s)\right)\frac{I(s)\,\mathrm{d}s}{\mathrm{d}s}\right]\right)^{\frac{1}{\kappa}} + \\ \left(E^0_{t,x}\left[\begin{matrix} \varepsilon(s)\frac{\mathrm{d}N(s)}{\mathrm{d}s} \times \\ \begin{pmatrix} X^{c,\pi,b}(s) + \\ b\left(s, X^{c,\pi,b}(s)\right)\end{pmatrix}^{1-\gamma}\end{matrix}\right]\right)^{\frac{1}{\kappa}}\end{pmatrix}^{\frac{\kappa}{\theta}} \mathrm{d}s\right)^{\theta} =$$

$$\frac{1}{1-\gamma}\left(\int_t^T \delta e^{-\delta(s-t)} e^{-\frac{1}{\theta}\int_t^s \mu(v)\,\mathrm{d}v} \begin{pmatrix} \left(m^{c,\pi,b}(t,s,x)\right)^{\frac{1}{\kappa}} + \\ \left(n^{c,\pi,b}(t,s,x)\right)^{\frac{1}{\kappa}}\end{pmatrix}^{\frac{\kappa}{\theta}} \mathrm{d}s\right)^{\theta} \tag{2.14}$$

with

$$\begin{aligned} m^{c,\pi,b}(t,s,x) &= E_{t,x}\left[c^{1-\gamma}\left(s, X^{c,\pi,b}(s)\right)\right] , \\ n^{c,\pi,b}(t,s,x) &= E_{t,x}\left[\varepsilon(s)\,\mu(s)\left(X^{c,\pi,b}(s) + b\left(s, X^{c,\pi,b}(s)\right)\right)^{1-\gamma}\right] . \end{aligned}$$

The operators $E_{t,x}$ and $E^0_{t,x}$ denote conditional expectation given $X^{c,\pi,b}(t) = x$ and $(X^{c,\pi,b}(t), N(t)) = (x, 0)$, respectively. The second equality in (2.14) follows from independence between (N, I) and W. By construction, $(1-\gamma) Z^{c,\pi,b}$ is non-negative, and in general, we assume that $Z^{c,\pi,b}(t,x)$ is non-zero for $x + L(t) > 0$ and $t < n$.

Given the non-linearity (in conditional expectation) of $Z^{c,\pi,b}$, the solution to (2.12) is likely to be inconsistent with the solution to (2.13) for $t > 0$. By 'inconsistent' we mean that the decision we make at time t based on (2.13) is not the same as the decision we plan to make at time t based on (2.12), for the same realization of the wealth process. More formally, if we denote the two solutions by (c^0, π^0, b^0) and (c^t, π^t, b^t), it might be that

$$\begin{aligned} &\left(c^0\left(t, X^{c^0,\pi^0,b^0}(t)\right), \pi^0\left(t, X^{c^0,\pi^0,b^0}(t)\right), c^0\left(t, X^{c^0,\pi^0,b^0}(t)\right)\right) \\ &\quad \neq \left(c^t\left(t, X^{c^0,\pi^0,b^0}(t)\right), \pi^t\left(t, X^{c^0,\pi^0,b^0}(t)\right), c^t\left(t, X^{c^0,\pi^0,b^0}(t)\right)\right) . \end{aligned}$$

We dislike this time-inconsistency, and we do not wish to introduce pre-commitment. Instead, we take inspiration from Björk et al. (2014), discard the optimization problem in (2.12)–(2.13), and search for an equilibrium control for the objective function $Z^{c,\pi,b}$, $(c,\pi,b) \in \mathcal{U}$. The equilibrium formulation arises from a game theoretic approach to stochastic control problems, and rewriting Definition 2.1 in Björk et al. (2014) in the language of this paper, we get the following definition:

Definition 2.2 (Equilibrium)**.** Consider a set of admissible controls $\bar{\mathcal{U}}$ and a control (c^*, π^*, b^*) in $\bar{\mathcal{U}}$ (informally viewed as a candidate equilibrium control).

Choose a fixed control $\left(\bar{c}, \bar{\pi}, \bar{b}\right) \in \bar{\mathcal{U}}$, a real number $h > 0$, and an initial point $(u, y) \in [0, T] \times \mathbb{R}$. Define the control $\left(c^h, \pi^h, b^h\right)$ by

$$\left(c^h, \pi^h, b^h\right)(t, x) = \begin{cases} (\bar{c}, \bar{\pi}, \bar{b})(t, x) \ , & u \leq t < u + h, x \in \mathbb{R} \ , \\ (c^*, \pi^*, b^*)(t, x) \ , & u + h \leq t \leq n, x \in \mathbb{R} \ . \end{cases}$$

If for all controls $\left(\bar{c}, \bar{\pi}, \bar{b}\right) \in \bar{\mathcal{U}}$ and all points $(u, y) \in [0, T] \times \mathbb{R}$

$$\liminf_{h \to 0} \frac{Z^{c^*, \pi^*, b^*}(u, y) - Z^{c^h, \pi^h, b^h}(u, y)}{h} \geq 0 \ , \tag{2.15}$$

we say that (c^*, π^*, b^*) is an *equilibrium control* for the function $Z^{c, \pi, b}$, $(c, \pi, b) \in \bar{\mathcal{U}}$. The corresponding *equilibrium value function* V is given by

$$V(t, x) = Z^{c^*, \pi^*, b^*}(t, x) \ .$$

Remark 2.1. We stress that an equilibrium control is not optimal in the sense that it realizes the supremum in (2.12) (or (2.13) for that matter). However, the control is optimal in the 'intuitive' sense that it maximizes the investor's total utility given that the investor continues to use the control. Therefore, we use the terms equilibrium control and *optimal control* interchangeably. With this convention, there might be several or even no optimal controls because Björk et al. (2014) prove neither existence nor uniqueness of the equilibrium control.

In the next subsection, we present a verification theorem for a particular optimal control and the corresponding equilibrium value function. Furthermore, we present closed form expressions for the control and the corresponding value function. To facilitate the proof, we need to introduce of set of non-standard assumptions, see Assumptions 2.1 in Appendix 2.A. These assumptions serve to prove that the equilibrium condition in (2.15) is satisfied. Also, we need to impose some additional constraints on the set of admissible controls and on the candidate equilibrium control, but these are all standard regularity conditions, see the theorem below. Equation numbers (A.) refer to equations in Appendix 2.A.

2.2.3 Solution

Theorem 2.1 (Verification theorem)**.** *Define the set of admissible controls, $\mathcal{U}^e$, as those controls (c, π, b) in $\mathcal{U}$ (see Definition 2.1) for which the partial differential equations (PDEs) in* (2.28) *have solutions in $\mathcal{C}^{1,0,2}$ and the stochastic integrals in* (2.29)*–*(2.30) *are martingales. Also, define the function*

$f : [0,T] \times (0,\infty)^2 \times \left(1_{\{\gamma<1\}}(0,\infty) \cup 1_{\{\gamma>1\}}(-\infty,0)\right) \to \mathbb{R}$ *by*

$$f(t,c,y,z) = \theta\delta z \left(\left(\frac{c^{1-\gamma}}{z(1-\gamma)} \right)^{\frac{1}{\kappa}} + \left(\frac{\varepsilon(t)\mu(t)y^{1-\gamma}}{z(1-\gamma)} \right)^{\frac{1}{\kappa}} \right)^{\frac{\kappa}{\theta}} - (\mu(t)+\theta\delta)z \ . \tag{2.16}$$

Assume that there exist functions (U, l_1, l_2) *in*

$$\mathcal{C}^{1,2}([0,T]\times\mathbb{R}) \times \mathcal{C}^{1,0,2}\left([0,T]^2\times\mathbb{R}\right) \times \mathcal{C}^{1,0,2}\left([0,T]^2\times\mathbb{R}\right)$$

such that the function U *solves the pseudo-Bellman equation*

$$U_t(t,x) = \inf_{(c,\pi,b)\in\Gamma(x,t)} \left[\begin{array}{c} -f\left(t,c,x+b,K^{l_1,l_2}(t,x)\right) \\ -\left((r+\pi\lambda)x - c - \hat{\mu}(t)b + w(t)\right)U_x(t,x) \\ -\frac{1}{2}\sigma^2\pi^2x^2U_{xx}(t,x) \\ +\frac{1}{2}\pi^2\sigma^2x^2I^{l_1,l_2}(t,x) \end{array} \right] ,$$
$$U(T,x) = 0 \ , \tag{2.17}$$

and such that the functions l_1 *and* l_2*, for each fixed* s*, solve the PDEs*

$$\begin{aligned} (l_i)_t(t,s,x) &= -\left[\begin{array}{c} x(r+\pi^*(t,x)\lambda) - c^*(t,x) \\ -\hat{\mu}(t)b^*(t,x) + w(t) \end{array} \right] \times (l_i)_x(t,s,x) \\ &\quad - \frac{1}{2}(\pi^*(t,x))^2\sigma^2x^2(l_i)_{xx}(t,s,x) , \ i = 1,2 \ , \\ l_1(s,s,x) &= (c^*)^{1-\gamma}(s,x) \ , \\ l_2(s,s,x) &= \varepsilon(s)\mu(s)(x+b^*(s,x))^{1-\gamma} \ , \end{aligned} \tag{2.18}$$

where (c^*, π^*, b^*) *is the function of* (t,x) *that realizes the infimum in* (2.17). *In* (2.17)*, the functions* K^{l_1,l_2} *and* I^{l_1,l_2} *are given by* (2.33) *and* (2.34).

Also, assume that the stochastic integrals in (2.38)–(2.39) *are martingales, that the SDE in* (2.6) *has a unique solution for* (c^*, π^*, b^*)*, and that the stochastic integrals in* (2.29)–(2.30) *are martingales for* (c^*, π^*, b^*)*. Finally, assume that the assumptions in Assumptions 2.1 are satisfied.*

Then (c^*, π^*, b^*) *is a control in* $\mathcal{U}^e$*, and it is an optimal control for the function* $Z^{c,\pi,b}$*,* $(c,\pi,b) \in \mathcal{U}^e$*, defined in* (2.14)*. The corresponding equilibrium value function* V *is given by*

$$V(t,x) = U(t,x) \ ,$$

and it holds that

$$\begin{aligned} m^{c^*,\pi^*,b^*}(t,s,x) &= l_1(t,s,x) \ , \\ n^{c^*,\pi^*,b^*}(t,s,x) &= l_2(t,s,x) \ , \\ U(t,x) &= K^{l_1,l_2}(t,x) \ . \end{aligned}$$

Proof. The proof is presented in Appendix 2.A. □

We note that we have replaced the global optimization problem in (2.13) with the continuum of local optimization problems in (2.17). Also, we recognize f as a generalization of the normalized continuous-time Epstein-Zin aggregator. We comment more on this in Section 2.3. We call the PDE in (2.17) a pseudo-Bellman equation because it bears resemblance to—but is different from—the Hamilton-Jacobi-Bellman equation known from dynamic programming.

Applying the verification theorem, we obtain closed form expression for the optimal control, see the theorem below. In working with the pseudo-Bellman equation, we find that the last term vanishes due to separability. For details, see the proof of Theorem 2.2.

Theorem 2.2 (Optimal control). *Define the function $g : [0,T] \to \mathbb{R}$ by*

$$g(t) = \delta \left(\int_t^T \tilde{\mu}(s)\, e^{-\int_t^s \tilde{r}(v)\,\mathrm{d}v}\,\mathrm{d}s \right)^{\phi} \quad , \quad t \le n \ ,$$

where

$$\tilde{r}(v) = -\frac{1}{\phi}\left[(1-\phi)\left(r + \frac{1}{2}\frac{1}{\gamma}\frac{\lambda^2}{\sigma^2} + \hat{\mu}(v) - \frac{\mu(v)}{1-\gamma} \right) - \delta \right] \ ,$$

$$\tilde{\mu}(s) = \left(1 + \left(\frac{\varepsilon(s)\,\mu(s)}{\hat{\mu}^{1-\gamma}(s)} \right)^{\frac{1}{\gamma+\kappa-1}} \right)^{\frac{(\kappa-1+\gamma)(1-\phi)}{(1-\gamma)\phi}} \ .$$

Moreover, define the functions $h_1, h_2 : [0,T]^2 \to \mathbb{R}$ by

$$h_i(t,s) = b_i(s)\, e^{-\int_t^s a(v)\,\mathrm{d}v} \quad , \quad i = 1,2 \ , \quad t \le s \ ,$$

where

$$\begin{aligned} a(v) = &-(1-\gamma)\left(r + \hat{\mu}(v) + \frac{1}{2}\frac{\lambda^2}{\gamma\sigma^2} \right) \\ &-(1-\gamma)\left(-\delta^{\frac{1}{\phi}} g^{-\frac{1}{\phi}}(v) \left(1 + \left(\frac{\varepsilon(v)\,\mu(v)}{\hat{\mu}^{1-\gamma}(v)} \right)^{\frac{1}{\gamma+\kappa-1}} \right)^{\frac{(\kappa-1+\gamma)(1-\phi)}{(1-\gamma)\phi}} \right) \ , \\ b_1(s) = &\left(\delta^{\frac{1}{\phi}} g^{-\frac{1}{\phi}}(s) \right)^{1-\gamma} \left(1 + \left(\frac{\varepsilon(s)\,\mu(s)}{\hat{\mu}^{1-\gamma}(s)} \right)^{\frac{1}{\gamma+\kappa-1}} \right)^{\frac{\kappa-\phi\kappa-1+\gamma}{\phi}} \ , \\ b_2(s) = &\, b_1(s) \left(\frac{\varepsilon(s)\,\mu(s)}{\hat{\mu}^{1-\gamma}(s)} \right)^{\frac{\kappa}{\gamma+\kappa-1}} \ . \end{aligned}$$

The optimal control from Theorem 2.1 is given by

$$\begin{aligned} c^*(t,x) &= \delta^{\frac{1}{\phi}} g^{-\frac{1}{\phi}}(t) \left(1 + \left(\frac{\varepsilon(t)\,\mu(t)}{\hat{\mu}^{1-\gamma}(t)}\right)^{\frac{1}{\gamma+\kappa-1}}\right)^{\frac{\kappa-\phi\kappa-1+\gamma}{(1-\gamma)\phi}} (x + L(t)) \ , \\ \pi^*(t,x)\,x &= \frac{\lambda}{\gamma\sigma^2}(x + L(t)) \ , \\ b^*(t,x) &= c^*(t,x) \left(\frac{\varepsilon(t)\,\mu(t)}{\hat{\mu}^{\kappa}(t)}\right)^{\frac{1}{\gamma+\kappa-1}} - x \ , \end{aligned} \tag{2.19}$$

and it holds that

$$\begin{aligned} V(t,x) &= \frac{1}{1-\gamma}(x + L(t))^{1-\gamma} g^{\theta}(t) \ , \\ m^{c^*,\pi^*,b^*}(t,s,x) &= (x + L(t))^{1-\gamma} h_1(t,s) \ , \\ n^{c^*,\pi^*,b^*}(t,s,x) &= (x + L(t))^{1-\gamma} h_2(t,s) \ . \end{aligned}$$

Proof. The proof is presented in Appendix 2.B. □

We note that c^*, π^*, and b^*+x are all directly proportional to the investor's total wealth $x+L$. The optimal proportion π^* of wealth to invest in the stock is independent of the elasticity parameters κ and ϕ, and it is the same as in the well-known case of time-additive utility. The expressions for the optimal consumption rate and the optimal inheritance are more complicated, but the optimal consumption is directly proportional to the optimal inheritance, and the optimal consumption rate can be written as

$$c^*(t,x) = \frac{x + \int_t^T w(s)\, e^{-\int_t^s (r+\hat{\mu}(v))\,\mathrm{d}v}\,\mathrm{d}s}{\int_t^T \tilde{\mu}(s)\, e^{-\int_t^s \tilde{r}(v)\,\mathrm{d}v}\,\mathrm{d}s} \tilde{\mu}^{\frac{\kappa-\phi\kappa-1+\gamma}{(\kappa-1+\gamma)(1-\phi)}}(t) \ .$$

The influence of the EIS-parameter ϕ is of special interest since our main innovation is the separation of EIS from relative risk aversion. However, despite the closed form solution, it is unclear how the EIS-parameter drives the optimal solution. In Section 2.4 we comment on the link in numerical examples.

With respect to the optimal inheritance, we note that in the case $\varepsilon = 0$ (i.e. the investor does not care about inheritance, for example because she does not have dependants), it holds that $b^*(\cdot, x) = -x$. This means that the investor continuously *sells* term insurance with a death sum equal to her wealth. Thereby, she jeopardizes her wealth in the case of death, in return for a higher consumption rate while alive. This is the design of a life annuity. The optimality of annuitization is a classical result dating back to Yaari (1965) and questioned ever since by experimentalist in terms of the so-called *annuity puzzle*.

2.2.4 Comparison to Richard (1975)

In this subsection, we consider the special case of time-additive utility, i.e. the case $\phi = \gamma$ and $\kappa = 1$. Letting $\phi = \gamma = K$ and (innocently) dividing by $e^{\delta t}$, the global optimization problem in (2.13) reduces to

$$\sup_{(c,\pi,b)\in\mathcal{U}} E_{t,x}\left[\int_t^T \delta e^{-\delta s} e^{-\int_t^s \mu(v)\,\mathrm{d}v} \begin{pmatrix} \frac{c^{1-K}\left(s,X^{c,\pi,b}(s)\right)}{1-K} + \\ \varepsilon(s)\,\mu(s) \times \\ \frac{\left(X^{c,\pi,b}(s)+b\left(s,X^{c,\pi,b}(s)\right)\right)^{1-K}}{1-K} \end{pmatrix} \mathrm{d}s\right]. \tag{2.20}$$

This simpler problem of maximizing expected time-additive utility for an uncertain-lived investor is treated in Richard (1975) (without an explicit state dependent constraint on the controls in $\mathcal{U}$). Richard (1975) allows for a much broader variety of utility functions than power utility functions, but in Section 4, focus is limited to (weighted) power utility. If we, in Section 4 of Richard (1975), let the constant relative risk aversion be given by $\gamma = 1 - K$, and if we let the weights be given by

$$h(t) = \delta e^{-\delta t}\ , \qquad m(t) = \varepsilon(t)\,\delta e^{-\delta t}\ ,$$

then the optimization problem in Richard (1975) coincides with the optimization problem in (2.20). Due to the time-additivity of the simplified problem, time-inconsistency is no longer an issue, and we wonder how our ‘equilibrium’ optimal control relates to the ‘classical’ optimal control in Richard (1975). With $\phi = \gamma = K$ and $\kappa = 1$, our optimal control (c^*, π^*, b^*) is given by

$$\begin{aligned} \frac{c^*(t,x)}{x+L(t)} &= \delta^{\frac{1}{K}} g^{-\frac{1}{K}}(t)\ , \\ \frac{\pi^*(t,x)\,x}{x+L(t)} &= \frac{\lambda}{K\sigma^2}\ , \\ \frac{b^*(t,x)+x}{x+L(t)} &= \left(\frac{\varepsilon(t)\,\mu(t)}{\hat{\mu}(t)}\right)^{\frac{1}{K}} \delta^{\frac{1}{K}} g^{-\frac{1}{K}}(t)\ , \end{aligned}$$

where

$$L(t) = \int_t^T w(s)\, e^{-\int_t^s (r+\hat{\mu}(v))\,\mathrm{d}v}\,\mathrm{d}s\ ,$$

and

$$g(t) = e^{\delta t}\left(\int_t^T \begin{pmatrix} \left(1 + \varepsilon^{\frac{1}{K}}(s)\,\mu(s)\left(\frac{\mu(s)}{\hat{\mu}(s)}\right)^{\frac{1-K}{K}}\right) \times \\ \left(\delta e^{-\delta s}\right)^{\frac{1}{K}} e^{-\int_t^s \mu(v)\,\mathrm{d}v} \times \\ e^{\frac{1-K}{K}\left(r+\frac{1}{2}\frac{1}{K}\frac{\lambda^2}{\sigma^2}\right)(s-t)+\frac{1-K}{K}\int_t^s(\hat{\mu}(v)-\mu(v))\,\mathrm{d}v} \end{pmatrix} \mathrm{d}s\right)^K .$$

When writing down expressions for the optimal control in Richard (1975), we make use of the following correspondence between our notation and Richard's notation:

Us	λ	b	μ	$\hat{\mu}$	X	π	L	$\hat{\mu}-\mu$	$e^{-\int_0^t \mu(s)\,\mathrm{d}s}$
Richard	$\alpha - r$	$P\mu^{-1}$	λ	μ	W	w	b	η	$G(t)$

With $h(t) = \delta e^{-\delta t}$, $m(t) = \varepsilon(t)\,\delta e^{-\delta t}$, and $\gamma = 1 - K$ in Section 4 of Richard (1975), the 'classical' optimal control $(c^{**}, \pi^{**}, b^{**})$ is given by

$$\begin{aligned}
\frac{c^{**}(t,x)}{x+L(t)} &= \left(\delta e^{-\delta t}\right)^{\frac{1}{K}} a^{-\frac{1}{K}}(t) \ , \\
\frac{\pi^{**}(t,x)\,x}{x+L(t)} &= \frac{\lambda}{K\sigma^2} \ , \\
\frac{b^{**}(t,x)+x}{x+L(t)} &= \left(\frac{\mu(t)}{\hat{\mu}(t)}\right)^{\frac{1}{K}} \left(\varepsilon(t)\,\delta e^{-\delta t}\right)^{\frac{1}{K}} a^{-\frac{1}{K}}(t) \ ,
\end{aligned}$$

where

$$a(t) = \left(\int_t^T \left(\begin{array}{c} \left(\left(\frac{\mu(s)}{\hat{\mu}(s)}\right)^{\frac{1-K}{K}} \mu(s) \left(\varepsilon(s)\,\delta e^{-\delta s}\right)^{\frac{1}{K}} + \left(\delta e^{-\delta s}\right)^{\frac{1}{K}} \right) \times \\ e^{-\int_t^s \mu(v)\,\mathrm{d}v} e^{\frac{1-K}{K}\left(r+\frac{1}{2}\frac{1}{K}\frac{\lambda^2}{\sigma^2}\right)(s-t)+\frac{1-K}{K}\int_t^s (\hat{\mu}(v)-\mu(v))\,\mathrm{d}v} \end{array} \right) \mathrm{d}s \right)^K .$$

Actually, Richard (1975) writes down

$$a(t) = \left(\int_t^T \left(\begin{array}{c} \left(\left(\frac{\mu(s)}{\hat{\mu}(s)}\right)^{\frac{1-K}{K}} \mu(s)^{\mathbf{\frac{1}{K}}} \left(\varepsilon(s)\,\delta e^{-\delta s}\right)^{\frac{1}{K}} + \left(\delta e^{-\delta s}\right)^{\frac{1}{K}} \right) \times \\ e^{-\int_t^s \mu(v)\,\mathrm{d}v} e^{\frac{1-K}{K}\left(r+\frac{1}{2}\frac{1}{K}\frac{\lambda^2}{\sigma^2}\right)(s-t)+\frac{1-K}{K}\int_t^s (\hat{\mu}(v)-\mu(v))\,\mathrm{d}v} \end{array} \right) \mathrm{d}s \right)^K$$

—but from his derivation, it appears that the bold power $\mathbf{\frac{1}{K}}$ must be an error. This is supported by formula (1a) in Kraft and Steffensen (2008).

We see that $g(t) = e^{\delta t} a(t)$. Plugging this into our optimal control, we discover that the two optimal controls match perfectly. We consider this to be an interesting discovery since we have not proven our optimal control to be optimal in the usual sense. It is not surprising, though, since in this special case of time-additivity, we have no time-inconsistency issues to deal with. Our work can be seen as an extension of the utility optimization in Richard (1975) to time-*non*-additive utility, and this is one of our most important insights since the literature, to our knowledge, contains no other attempts in that direction. However, the extension is only for power utility.

2.3 Link to recursive utility

2.3.1 Motivation

In the previous section, we introduced certainty equivalents in order to separate risk aversion from elasticity of inter-temporal substitution. This draws our attention in the direction of recursive utility studied in e.g. Duffie and Epstein (1992b) and Kraft and Seifried (2010). In advance, we have no reason to believe that our optimization approach is equivalent to continuous-time recursive utility optimization, but in the special case of no mortality risk, it turns out that the pseudo-Bellman equation characterizing our equilibrium value function coincides with the pseudo-Bellman equation characterizing the value function of the recursive utility optimization problem in Duffie and Epstein (1992b) for Epstein-Zin preferences. In the following subsections, we give an introduction to recursive utility, demonstrate the similarity of pseudo-Bellman equations, and outline the perspectives of our findings.

2.3.2 A short introduction to recursive utility

Let $(\Omega, \mathcal{F}, P)$ be a probability space endowed with a filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$ satisfying the usual conditions. Fix a set $\mathcal{C} \subset \mathbb{R}^k$ of consumption rates and denote by C a class of predictable $\mathcal{C}$-valued processes with time-horizon $[0, T]$. The backbone of recursive utility is the construction of a mapping $\mathfrak{u} : \mathsf{C} \to \mathbb{R}$ that ranks consumption streams in such a way that $\mathfrak{u}(c) \geq \mathfrak{u}(c')$ if and only if the consumption stream c is weakly preferred to the consumption stream c'. This is done by means of a *utility process* V^c associated to c by setting

$$\mathfrak{u}(c) = V^c(0) \ , \quad c \in \mathsf{C} \ .$$

The utility process is assumed to take values in a subinterval $\mathcal{V} \subset \mathbb{R}$ of the real line, and $\mathfrak{u}$ is referred to as a *recursive utility function.*

2.3.2.1 Discrete-time recursive utility

Recursive utility is first defined in discrete time, and in Section 3 of Kraft and Seifried (2010), we find a brief review of discrete-time recursive utility. Let $\{t_0, t_1, \ldots, t_m\}$ be a partition of $[0, T]$, and let $c = \{c(t_k)\}_{k=1,\ldots,m}$ be a discrete-time consumption stream in C. Then the utility process V^c is defined through the backward recursion

$$\begin{aligned} V^c(t_k) &= W\left(t_{k+1} - t_k, c(t_k), \mathfrak{m}\left(\mathcal{L}\left(V^c(t_{k+1}) | \mathcal{F}_{t_k}\right)\right)\right) \ , \\ &\qquad\qquad k = 0, \ldots, m-1 \ , \\ V^c(t_m) &= 0 \ . \end{aligned} \tag{2.21}$$

Here, $W : [0, \infty) \times \mathcal{C} \times \mathcal{V} \to \mathcal{V}$ is a continuous function with $W(0, c, v) = v$ for $c \in \mathcal{C}, v \in \mathcal{V}$, $\mathcal{L}(V^c(t_{k+1}) | \mathcal{F}_{t_k})$ is the conditional distribution of $V^c(t_{k+1})$

given the information $\mathcal{F}_{t_k}$, and $\mathfrak{m}$ is a certainty equivalent on $\mathcal{V}$. Letting $\mathcal{M}_1(\mathcal{V})$ denote the set of probability measures on $\mathcal{B}(\mathcal{V})$ with moments of all orders, a functional $\mathfrak{m} : \mathcal{M}_1(\mathcal{V}) \to \mathbb{R}$ is called a certainty equivalent on $\mathcal{V}$ if $\mathfrak{m}(\delta_v) = v$ for all $v \in \mathcal{V}$ where δ_v is the Dirac measure at v.

W is often referred to as the time-aggregator because in a set-up without risk (implying $\mathfrak{m}(\mathcal{L}(V^c(t_{k+1})|\mathcal{F}_{t_k})) = V^c(t_{k+1})$), it describes the intertemporal aggregation of present consumption c_{t_k} and the value of future consumption $V^c(t_{k+1})$. Similarly, $\mathfrak{m}$ is referred to as the risk-aggregator since it describes the risk weighted aggregation of possible future values of $V^c(t_{k+1})$. The pair $(W, \mathfrak{m})$ completely describes an investor's preferences for discrete-time stochastic consumption streams, and we call $(W, \mathfrak{m})$ a *discrete-time aggregator*.

A special class of certainty equivalents are those given by

$$\mathfrak{m}(\mu) = h^{-1}\left(\int_{\mathcal{V}} h \,\mathrm{d}\mu\right) \ , \quad \mu \in \mathcal{M}_1(\mathcal{V}) \ ,$$

for a strictly increasing, polynomially bounded $\mathcal{C}^2$-function $h : \mathcal{V} \to \mathbb{R}$. Here, $\mathfrak{m}$ is called an expected utility (EU) certainty equivalent. If h is the identity, then $\mathfrak{m}$ is called risk-neutral.

2.3.2.2 Continuous-time recursive utility

Duffie and Epstein (1992b) denote their approach to recursive utility in continuous time by *stochastic differential utility* (SDU). They start from the discrete-time formulation in (2.21) and use a heuristic limiting argument to motivate their formulation of SDU, but SDU is defined in continuous time and does not rely on the heuristic derivation.

Kraft and Seifried (2010) set the heuristic limiting argument from Duffie and Epstein (1992b) on a rigorous basis and denote their approach to recursive utility in continuous time by *continuous-time recursive utility* (CRU). Thereby, CRU is directly related to discrete-time recursive utility, and CRU is defined in a broader set-up than SDU.

We choose not to write down exactly how SDU and CRU are defined since the general definitions are complicated and since we gain sufficient insight from Lemma 2.2. In both SDU and CRU, the utility process V^c is generated by a *continuous-time aggregator* $(f, \mathfrak{m})$ on $\mathcal{V}$, where $f : \mathcal{C} \times \mathcal{V} \to \mathbb{R}$ is a Borel-measurable function, and $\mathfrak{m}$ is a certainty equivalent on $\mathcal{V}$. Also, both approaches have the disadvantage that they rely on the almost sure differentiability of the function $s \mapsto \mathfrak{m}(\mathcal{L}(V^c_{t+s}|\mathcal{F}_t))$ in $s = 0$.

We end this introduction with two lemmas. The first lemma describes the relation between discrete-time recursive utility and CRU. The second lemma shows that SDU and CRU are equivalent when the certainty equivalent is particularly simple. The lemmas follow from Corollary 6.3 and formulas (7), (19), and (21) in Kraft and Seifried (2010):

Lemma 2.1. *Let* $(W, \mathfrak{m})$ *be a discrete-time aggregator on* $\mathcal{V}$*, assume that* W *is a* $\mathcal{C}^{1,0,1}$*-function, and define* $f : \mathcal{C} \times \mathcal{V} \to \mathbb{R}$ *by*

$$f(c, v) = \frac{\frac{\partial W}{\partial \Delta}(0, c, v)}{\frac{\partial W}{\partial v}(0, c, v)} . \tag{2.22}$$

Then $(f, \mathfrak{m})$ *is the CRU continuous-time aggregator corresponding to* $(W, \mathfrak{m})$*. Note that we cannot be sure that the aggregator* $(f, \mathfrak{m})$ *actually generates a utility function, but if it does, then the discrete-time utility function and the continuous-time utility function represent the same preferences.*

Lemma 2.2. *Let* $(f, \mathfrak{m})$ *be a continuous-time aggregator on* $\mathcal{V} = \mathbb{R}$ *and assume that* $\{\mathcal{F}_t\}_{t\in[0,T]}$ *is generated by a standard Brownian motion, a Poisson random measure and the null sets,* $\mathfrak{m}$ *is the risk-neutral certainty equivalent, and* f *satisfies the Lipschitz and linear growth conditions*

$$\begin{aligned} |f(c, v) - f(c, w)| &\le \alpha |v - w| && \forall c \in \mathcal{C}, v, w \in \mathbb{R} , \\ |f(c, 0)| &\le \beta_1 + \beta_2 |c| && \forall c \in \mathcal{C} , \end{aligned}$$

for some $\alpha, \beta_0, \beta_1 > 0$*. Then SDU and CRU generate the same utility function* $\mathfrak{u} : \mathsf{C} \to \mathbb{R}$*, and it is given by* $\mathfrak{u}(c) = V^c(0)$ *where*

$$V^c(t) = E\left[\int_t^T f(c(s), V^c(s))\, \mathrm{d}s \middle| \mathcal{F}_t\right] \quad a.s .$$

We note that a continuous-time aggregator $(f, \mathfrak{m})$ is called normalized if $\mathfrak{m}$ is the risk-neutral certainty equivalent.

2.3.2.3 Example: Epstein-Zin preferences

An important class of recursive preferences are the Epstein-Zin preferences. In discrete time, these can be represented by a discrete-time aggregator $(W, \mathfrak{m})$ on $\mathcal{V} = (0, \infty)$, where $\mathfrak{m}$ is the risk-neutral certainty equivalent, and W is given by

$$W(\Delta, c, v) = \frac{1}{1-\gamma}\left(\delta \Delta c^{1-\phi} + e^{-\delta\Delta}\left((1-\gamma) v\right)^{\frac{1-\phi}{1-\gamma}}\right)^{\frac{1-\gamma}{1-\phi}}$$

with $\gamma, \phi > 0,\ \gamma, \phi \ne 1$. Here, γ is the relative risk aversion, δ is the rate of time preference, and $\frac{1}{\phi}$ is the constant elasticity of inter-temporal substitution. Using formula (2.22), we find that the normalized continuous-time Epstein-Zin aggregator is given by $(f, \mathfrak{m})$, where

$$f(c, v) = \frac{\frac{\partial W}{\partial \Delta}(0, c, v)}{\frac{\partial W}{\partial v}(0, c, v)} = \frac{1-\gamma}{1-\phi}\delta v\left(\left(\frac{c}{((1-\gamma) v)^{\frac{1}{1-\gamma}}}\right)^{1-\phi} - 1\right) .$$

It is easy to verify that f does not satisfy the Lipschitz and growth conditions of Lemma 2.2 for general ϕ and γ, so a priori we do not know if $(f, \mathfrak{m})$ generates a utility function. However, Duffie and Epstein (1992b) mention in Example 3 that existence and uniqueness can be shown, and Kraft and Seifried (2010) make a similar comment in Remark 6.4.

2.3.3 Similarity of pseudo-Bellman equations

For a while, we think of the investor from Section 2.2 as certain-lived, i.e. we fix $\mu = \hat{\mu} = 0$ in the set-up from Section 2.2. The investor's wealth now evolves according to the SDE

$$\begin{aligned} \mathrm{d}X^{c,\pi}(t) &= X^{c,\pi}(t)\left[(r + \pi(t, X^{c,\pi}(t))\lambda)\,\mathrm{d}t + \pi(t, X^{c,\pi}(t))\,\sigma \mathrm{d}W(t)\right] \\ &\quad - (c(t, X^{c,\pi}(t)) - w(t))\,\mathrm{d}t\ , \\ X^{c,\pi}(0) &= x_0\ , \end{aligned}$$

where x_0 is the investor's initial wealth, w is a continuous, deterministic function of time, $r, \sigma, \lambda > 0$ are constants, and c, π are deterministic, measurable functions of time and wealth. The objective functions reads

$$Z^{c,\pi}(t,x) = \frac{1}{1-\gamma}\left(\int_t^T \delta e^{-\delta(s-t)}\left(E_{t,x}\left[c^{1-\gamma}(s, X^{c,\pi}(s))\right]\right)^{\frac{1}{\theta}}\mathrm{d}s\right)^{\theta}\ ,$$

where the parameters n, δ, γ, and θ are as in Section 2.2. We note that the death sum b has disappeared from both the wealth dynamics and the objective function. This is natural since the term insurance costs nothing (due to $\hat{\mu} = 0$) and pays out nothing (due to $\mu = 0$).

The problem of maximizing $Z^{c,\pi}$ is still time-inconsistent, so again we search for an equilibrium control for the function $Z^{c,\pi}$, $(c, \pi) \in \mathcal{U}_0^e$. Here, subscript 0 indicates that we have plugged in $\mu = \hat{\mu} = 0$ and left out b in the constraints defining $\mathcal{U}^e$. The same applies for $\mathcal{U}_0$ and Γ_0 below. We continue to use the terms optimal control and equilibrium control interchangeably. Plugging $\mu = \hat{\mu} = 0$ into Theorem 2.1 and recalling that the last term in the Bellman equation vanishes due to separability, we get the pseudo-Bellman equation

$$\begin{aligned} U_t(t,x) &= \inf_{(c,\pi)\in\Gamma_0(t,x)}\left[\begin{array}{c} -f(c, U(t,x)) \\ -((r+\pi\lambda)x - c + w(t))\,U_x(t,x) \\ -\frac{1}{2}\sigma^2\pi^2x^2U_{xx}(t,x) \end{array}\right]\ , \\ U(T,x) &= 0\ , \end{aligned} \tag{2.23}$$

where the function $f : (0,\infty) \times (1_{\{\gamma\in(0,1)\}}(0,\infty) \cup 1_{\{\gamma\in(1,\infty)\}}(-\infty,0)) \to \mathbb{R}$ is given by

$$f(c,Z) = \theta\delta Z\left(\left(\frac{c}{((1-\gamma)Z)^{\frac{1}{1-\gamma}}}\right)^{\frac{1-\gamma}{\theta}} - 1\right)\ . \tag{2.24}$$

We recognize equation (2.23) from Proposition 9 in Duffie and Epstein (1992b) as the pseudo-Bellman equation characterizing the value function of the continuous-time recursive utility optimization problem

$$\sup_{(c,\pi)\in\mathcal{D}} \mathfrak{u}\left(c^{c,\pi}\right) ,$$

where $c^{c,\pi} = \{c\left(t, X^{c,\pi}\left(t\right)\right)\}_{t\in[0,T]}$, $\mathcal{D}$ is the set of square-integrable, optional controls in $\mathcal{U}_0$, and $\mathfrak{u}$ is the utility function from Lemma 2.2 generated by the aggregator $(f, \mathfrak{m})$, where f is defined in (2.24), and $\mathfrak{m}$ is the risk-neutral certainty equivalent. In other words, $\mathfrak{u}\left(c^{c,\pi}\right) = V^{c,\pi}\left(0\right)$, where $V^{c,\pi}$ is defined via the backward equation

$$V^{c,\pi}\left(t\right) = E\left[\int_t^T f\left(c\left(s, X^{c,\pi}\left(s\right)\right), V^{c,\pi}\left(s\right)\right) \mathrm{d}s \middle| \mathcal{F}_t\right] .$$

Here, $\mathcal{F}_t$ denotes the augmentation of the σ-algebra generated by the sets $\{W\left(s\right) : 0 \leq s \leq t\}$.

We find the similarity of pseudo-Bellman equations interesting since our optimization problem has an intuitive interpretation as a global maximization of certainty equivalents and since our approach does not give rise to the differentiability problems mentioned in the previous subsection. Moreover, we recognize the aggregator $(f, \mathfrak{m})$ as the normalized continuous-time Epstein-Zin aggregator. This is again interesting since Epstein-Zin preferences are widely used in the literature.

The similarity of pseudo-Bellman equations does not mean that we solve the same problem, but as a consequence of the similarity, we end up with the same optimal control.

Recursive utility optimization for a certain-lived investor with Epstein-Zin preferences and no labour income is studied in Kraft et al. (2013). They allow for a general financial market with the Black-Scholes market as a simple special case. In the Black-Scholes special case, the optimal control in Kraft et al. (2013) coincides with our optimal control (c^*, π^*) for $\mu = \hat{\mu} = w = 0$ (see equation (4.4) in Kraft et al. (2013) and Theorem 2.2). Since we are solving the same problem, this constitutes a nice validation of our solution.

When applying Proposition 9 in Duffie and Epstein (1992b), we stumble on the fact that f does not satisfy certain Lipschitz and growth conditions, but Kraft et al. (2013) show that the proposition remains valid for e.g. $\phi \leq \gamma < 1$ and $\phi \geq \gamma > 1$, and in any case, the similarity of pseudo-Bellman equations is noteworthy.

2.3.4 Perspectives

We have demonstrated that—in the special case without mortality risk—the pseudo-Bellman equation characterizing our equilibrium value function coincides with the pseudo-Bellman equation characterizing the value function of the recursive utility optimization problem in Duffie and Epstein (1992b) for Epstein-Zin preferences. We formulate this by saying that our optimization approach (for a certain-lived investor) is equivalent to recursive utility optimization with Epstein-Zin preferences in a Black-Scholes market.

The equivalence between our optimization approach and recursive utility optimization (that is a well-established approach in diffusive markets) supports the use of our approach, also in cases that are not covered by recursive utility optimization. By 'not covered' we mean that neither SDU optimization nor CRU optimization is apt for an extended set-up with mortality risk and utility from inheritance since neither SDU nor CRU allows for utility from a lump sum at a random point in time. With our approach, we can provide such an extension for Epstein-Zin preferences. That is, our work can be seen as a generalization of recursive utility optimization with Epstein-Zin preferences to include mortality risk and life insurance. To our knowledge, the literature contains no other attempts in that direction.

Partial consistency with recursive utility was also found with the approach taken by Kihlstrom (2009). Kihlstrom (2009) separated risk aversion and EIS in the 'opposite order' in a discrete-time setup by first applying the EIS-function on the consumption stream and then taking expected utility. He found that the optimal consumption-investment strategy differs fundamentally from the optimal consumption-investment strategy arising from discrete-time recursive utility, as introduced by Epstein and Zin (1989) (introducer reference). But if the investment decision is left out by deleting one of two investment alternatives, Kihlstrom (2009) and Epstein and Zin (1989) agree on the consumption pattern. So, it appears that our approach provides an even stronger connection to recursive utility by agreeing on both investment and consumption. However, since Kihlstrom (2009) works exclusively in discrete time and we work exclusively in (a simple Black-Scholes model in) continuous time, the connections are not really comparable. E.g. what is the agreement between our approach and Epstein and Zin (1989) in a general return model in discrete-time as the one studied by Kihlstrom (2009)? It is worth noticing that Kihlstrom (2009) considers his distinction from recursive utility in the case of access to investment decisions as an advantage rather than a disadvantage because of its implications for asset pricing calculations as is his ultimate object of study. It is far beyond the scope of this paper to draw any lines in that direction, but given that our approach does provide, in general, an alternative to recursive utility, the path is definitely worth pursuing.

2.4 The optimal consumption rate

2.4.1 Motivation

With the separation of preferences for risk and inter-temporal substitution, our utility optimization approach gives rise to a broader variety of optimal consumption curves than time-additive power utility optimization. To illustrate this, we derive an SDE for the optimal consumption rate from Section 2.2, consider the special case of no market risk, and go through some numerical examples.

2.4.2 SDE

In the following, we assume that μ is differentiable, ε is constant, and

$$\hat{\mu} = \alpha\mu \quad \text{for some constant } \alpha > 0.$$

The optimal consumption rate is characterized by the SDE

$$\begin{aligned} \frac{\mathrm{d}c^*(t, X^*(t))}{c^*(t, X^*(t))} &= \frac{1}{\phi}\left(r - \delta + \left(\alpha - \frac{1}{\theta}\right)\mu(t) + (1+\phi)\frac{1}{2}\frac{\lambda^2}{\gamma\sigma^2} + \beta(t)\right)\mathrm{d}t \\ &\quad + \frac{\lambda}{\gamma\sigma}\mathrm{d}W(t) , \\ c^*(0, X^*(0)) &= c^*(0, x_0) , \end{aligned} \tag{2.25}$$

where

$$\beta(t) = \frac{\kappa - \phi\kappa - 1 + \gamma}{1-\gamma}\frac{\gamma}{\gamma+\kappa-1}\frac{\mu_t(t)\,\varepsilon^{\frac{1}{\gamma+\kappa-1}}\alpha^{\frac{\gamma-1}{\gamma+\kappa-1}}\mu(t)^{\frac{\gamma}{\gamma+\kappa-1}-1}}{1+\varepsilon^{\frac{1}{\gamma+\kappa-1}}\alpha^{\frac{\gamma-1}{\gamma+\kappa-1}}\mu(t)^{\frac{\gamma}{\gamma+\kappa-1}}} .$$

The derivation is presented in Appendix 2.C.

2.4.3 The special case without market risk

With $\lambda = 0$, there is no investment in the stock, and consequently, there is no market risk. The SDE in (2.25) reduces to the differential equation

$$\frac{\mathrm{d}c^*(t, X^*(t))}{c^*(t, X^*(t))} = \frac{1}{\phi}\left(r - \delta + \left(\alpha - \frac{1}{\theta}\right)\mu(t) + \beta(t)\right)\mathrm{d}t . \tag{2.26}$$

The future optimal consumption rate is deterministic, and the initial value $c^*(0, x_0)$ is given by

$$\begin{aligned} c^*(0, x_0) &= \frac{x_0 + \int_0^T w(s)\, e^{-\int_0^s (r+\alpha\mu(v))\,\mathrm{d}v}\,\mathrm{d}s}{\int_0^T \tilde{\mu}(s)\, e^{-\int_0^s \tilde{r}(v)\,\mathrm{d}v}\,\mathrm{d}s} \\ &\quad \times \left(1+\varepsilon^{\frac{1}{\gamma+\kappa-1}}\alpha^{\frac{\gamma-1}{\gamma+\kappa-1}}\mu(0)^{\frac{\gamma}{\gamma+\kappa-1}}\right)^{\frac{\kappa-\phi\kappa-1+\gamma}{(1-\gamma)\phi}} , \end{aligned} \tag{2.27}$$

where

$$\begin{aligned} \tilde{r}(v) &= -\frac{1}{\phi}\left[(1-\phi)\left(r + \left(\alpha - \frac{1}{1-\gamma}\right)\mu(v)\right) - \delta\right] , \\ \tilde{\mu}(s) &= \left(1+\varepsilon^{\frac{1}{\gamma+\kappa-1}}\alpha^{\frac{\gamma-1}{\gamma+\kappa-1}}\mu(s)^{\frac{\gamma}{\gamma+\kappa-1}}\right)^{\frac{(\kappa-1+\gamma)(1-\phi)}{(1-\gamma)\phi}} . \end{aligned}$$

Since all market risk is eliminated, one might be surprised to see that the risk aversion parameter γ is still present, but this is due to mortality risk.

2.4.4 Numerics

2.4.4.1 Set-up

We consider an investor with the following characteristics:

- The investor is $t_0 = 25$ years old at time 0 and has an initial wealth of $x_0 = 10,000$ USD.
- She starts off with a yearly labour income at rate $20,000$ USD (we do not take taxes into account), and her labour income grows with the risk free short rate until the age of 65 when she retires, i.e.

$$w(t) = 20,000 \cdot e^{rt} \cdot 1_{\{t_0+t\leq 65\}} \ .$$

- Her death is governed by the mortality intensity[1]

$$\mu(t) = 5 \cdot 10^{-4} + 5.3456 \cdot 10^{-5} \cdot e^{0.087498(t_0+t)} \ .$$

We only wish to focus on separation of risk aversion and EIS, so we fix $\alpha = \varepsilon = \kappa = 1$. Also, following Kraft et al. (2013), we fix the risk free short rate at $r = 0.05$ and the risk aversion at $\gamma = 2$. Finally, we fix the time-horizon $n = 85$ since there is very little probability that the investor survives the age of 110 with the chosen mortality. The fixed parameter values are summarized in the following table.

Parameter	α	ε	κ	r	γ	T
Fixed value	1	1	1	0.05	2	85

For a given choice of parameters, we first calculate the initial optimal consumption rate $c^*(0, x_0)$ by approximating the integrals in (2.27) with sums. We then calculate the future optimal consumption rates by approximating (2.26) with a difference equation.

2.4.4.2 Graphs

Fixing the EIS-parameter $\phi = 2$, we are in the time-additive case from Richard (1975), and letting δ vary, we get Figure 2.1. The investor's optimal yearly consumption rate is constant over time when δ is equal to r, and the rate is increasing (decreasing) when δ is smaller (larger) than r. This fits well with the intuition that δ is the investor's utility discount factor: if the investor

[1] For the last three decades, this has served as a standard mortality intensity for adult women in Denmark.

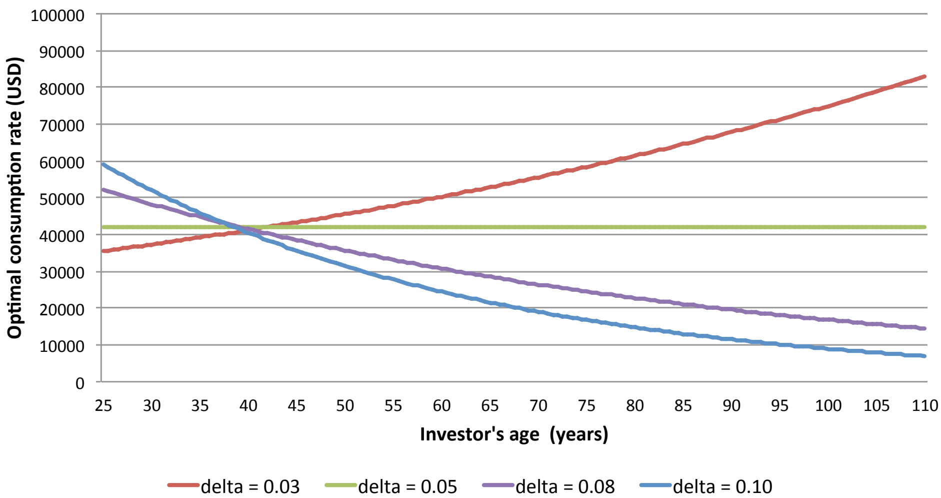

Figure 2.1: The optimal consumption rate as function of δ for fixed $\phi = 2$ $(= \gamma)$.

discounts future consumption with a short rate that is larger than the risk free short rate, then she assigns a higher value to one unit of consumption 'now' than to one unit plus investment returns 'later'. We notice that all the optimal consumption rates seem rather high compared to the investor's initial labour income and wealth. This is because the investor's labour income grows with the risk free short rate, and the plotted optimal consumption curves are expressed in nominal terms.

Fixing $\phi = 1.8$, we enable the separation of risk aversion and EIS that is special for this paper. Letting δ vary, we get Figure 2.2. The investor's optimal yearly consumption rate is increasing for δ smaller than r and non-monotone for δ larger than r. The non-monotone optimal consumption curves are first decreasing and then increasing. For ages below 75 (for small values of δ) or 90 (for larger values of δ), the consumption curves are almost identical to those for $\phi = 2$, but for higher ages, the consumption curves increase rapidly compared to those for $\phi = 2$. This can be explained by the small value of ϕ (corresponding to a higher willingness to substitute consumption over time) combined with the cheap consumption for high ages (because of the low survival probability). The investor's willingness to substitute consumption over time simply allows her to consume more when consumption is cheap.

Fixing $\phi = 3$, we again enable separation of risk aversion and EIS. Letting δ vary, we get Figure 2.3. The investor's optimal yearly consumption rate is decreasing for δ larger than r and non-monotone for δ smaller than r. The non-monotone optimal consumption curves are first increasing and then decreasing. In the literature, this phenomenon is known as hump-shaped consumption. For ages below 50, the consumption curves are similar to those for $\phi = 2$, but for higher ages, the consumption curves decrease compared to

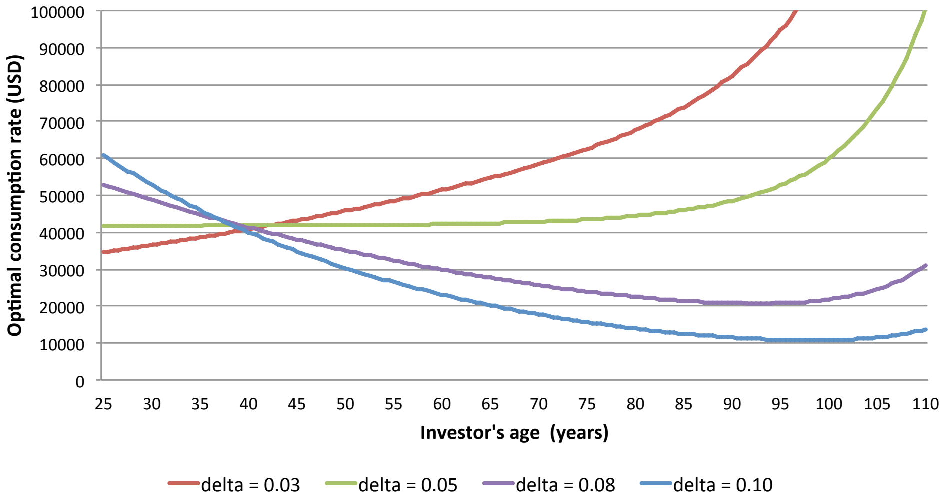

Figure 2.2: The optimal consumption rate as function of δ for fixed $\phi = 1.8$ $(< \gamma)$.

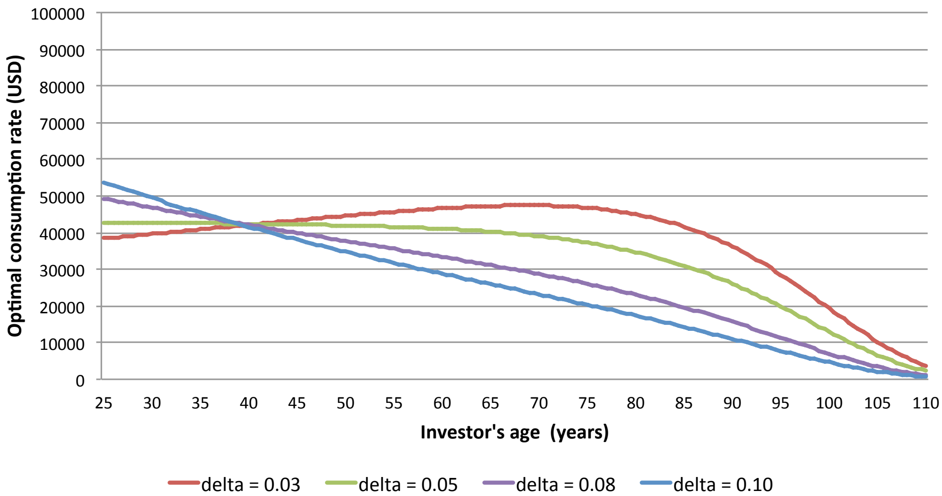

Figure 2.3: The optimal consumption rate as function of δ for fixed $\phi = 3$ $(> \gamma)$.

those for $\phi = 2$. This can be explained by the large value of ϕ (corresponding to a lower willingness to substitute consumption over time) combined with the chosen level of risk aversion and the low survival probability for high ages. The investor is so unwilling to substitute consumption over time that she cannot benefit from the cheap, but uncertain, consumption for high ages.

Hump-shaped consumption is observed in realized consumption, and different articles contain different explanations for this. See e.g. Gourinchas and Parker (2002) who obtain the feature by income uncertainty. They fit to data a hump around age 50. Our hump is not fitted to any data, but the hump around 70 for $\delta = 0.03$ is not necessarily in conflict with their quantities since we illustrate consumption in nominal terms whereas they convert to 1987 dollars. We note that such hump-shaped consumption patterns cannot be obtained by standard recursive utility or time-additive utility under lifetime uncertainty. We do not claim to having found the most important source of hump-shapes, and we do not pursue this particular feature of our approach more for now. Yet, we find it interesting enough to stress that it is the very combination of separation of risk aversion and elasticity of substitution with an uncertain lifetime that takes us to this intriguing feature of realized consumption.

Appendix

2.A Proof of Theorem 2.1

2.A.1 Prerequisites

Fix a control $(c, \pi, b) \in \mathcal{U}^e$. First, we take a look at $m^{c,\pi,b}$ and $n^{c,\pi,b}$. We assume there exist functions $\Lambda_1^{c,\pi,b}$ and $\Lambda_2^{c,\pi,b}$ in $\mathcal{C}^{1,0,2}\left([0,T]^2 \times \mathbb{R}\right)$ such that

$$\begin{aligned}\Lambda_{i,t}^{c,\pi,b}(t,s,x) &= -\left[x\left(r + \pi(t,x)\lambda\right) - c(t,x)\right]\Lambda_{i,x}^{c,\pi,b}(t,s,x)\\ &\quad - \left[-\hat{\mu}(t)\, b(t,x) + w(t)\right]\Lambda_{i,x}^{c,\pi,b}(t,s,x)\\ &\quad - \frac{1}{2}\pi^2(t,x)\,\sigma^2 x^2 \Lambda_{i,xx}^{c,\pi,b}(t,s,x)\ , \quad i = 1,2\ ,\\ \Lambda_1^{c,\pi,b}(s,s,x) &= c^{1-\gamma}(s,x)\ ,\\ \Lambda_2^{c,\pi,b}(s,s,x) &= \varepsilon(s)\,\mu(s)\left(x + b(s,x)\right)^{1-\gamma}\ ,\end{aligned} \tag{2.28}$$

for all $x \in \mathbb{R}$ and $0 \le t \le s \le n$.

Using Itô's formula on $\Lambda_i^{c,\pi,b}\left(t, s, X^{c,\pi,b}(t)\right)$ (for fixed s), plugging in (2.28), and skipping most arguments that are $\left(t, s, X^{c,\pi,b}(t)\right)$, $\left(t, X^{c,\pi,b}(t)\right)$

or t, we get that[2]

$$\begin{aligned} d\Lambda_i^{c,\pi,b}\left(t,s,X^{c,\pi,b}(t)\right) &= \Lambda_{i,t}^{c,\pi,b}\mathrm{d}t + \Lambda_{i,x}^{c,\pi,b}\mathrm{d}X^{c,\pi,b}(t) \\ &\quad + \frac{1}{2}\Lambda_{i,xx}^{c,\pi,b}\mathrm{d}\left[X^{c,\pi,b},X^{c,\pi,b}\right]^c(t) \\ &= \Lambda_{i,x}^{c,\pi,b}X^{c,\pi,b}\pi\sigma\mathrm{d}W(t)\ , \quad i=1,2\ , \quad t\le s\ . \end{aligned}$$

Hence, for $t \le s$, we can write

$$\begin{aligned} &\Lambda_1^{c,\pi,b}\left(t,s,X^{c,\pi,b}(t)\right) \\ &\quad = c^{1-\gamma}\left(s,X^{c,\pi,b}(s)\right) \\ &\qquad - \int_t^s \Lambda_{1,x}^{c,\pi,b}\left(u,s,X^{c,\pi,b}(u)\right)X^{c,\pi,b}(u)\,\pi\left(u,X^{c,\pi,b}(u)\right)\sigma\,\mathrm{d}W(u)\ , \end{aligned} \tag{2.29}$$

$$\begin{aligned} &\Lambda_2^{c,\pi,b}\left(t,s,X^{c,\pi,b}(t)\right) \\ &\quad = \varepsilon(s)\,\mu(s)\left(X^{c,\pi,b}(s)+b\left(s,X^{c,\pi,b}(s)\right)\right)^{1-\gamma} \\ &\qquad - \int_t^s \Lambda_{2,x}^{c,\pi,b}\left(u,s,X^{c,\pi,b}(u)\right)X^{c,\pi,b}(u)\,\pi\left(u,X^{c,\pi,b}(u)\right)\sigma\,\mathrm{d}W(u)\ . \end{aligned} \tag{2.30}$$

We assume that the stochastic integrals in (2.29) and (2.30) are martingales. Taking conditional expectation given $X^{c,\pi,b}(t) = x$ on both sides yields

$$\begin{aligned} \Lambda_1^{c,\pi,b}(t,s,x) &= E_{t,x}\left[c^{1-\gamma}\left(s,X^{c,\pi,b}(s)\right)\right] \\ &= m^{c,\pi,b}(t,s,x)\ , \end{aligned} \tag{2.31}$$

$$\begin{aligned} \Lambda_2^{c,\pi,b}(t,s,x) &= E_{t,x}\left[\varepsilon(s)\,\mu(s)\left(X^{c,\pi,b}(s)+b\left(s,X^{c,\pi,b}(s)\right)\right)^{1-\gamma}\right] \\ &= n^{c,\pi,b}(t,s,x)\ . \end{aligned} \tag{2.32}$$

For strictly positive, sufficiently integrable functions $a, b \in \mathcal{C}^{0,0,1}\left([0,T]^2 \times \mathbb{R}\right)$,

[2] We use Itô's formula as presented in (Protter, 2005, Chapter II, Theorem 33). Several terms are left out or simplified since $X^{c,\pi,b}$ is continuous, the operator $(x,y) \mapsto [x,y]$ is bilinear (see (ibid., p. 66)), $\mathrm{d}[W,W]^c(t) = t$ (see (ibid., p. 67)), and $\mathrm{d}[Id,Id]^c(t) = \mathrm{d}[Id,W]^c(t) = 0$ by (ibid., Theorem 26 and 28) where $Id(t) = t$. The theorems apply since Id is adapted, cadlag, and have path of finite variation on compacts, whereas W is a continuous martingale.

we define the functions $K^{a,b}, I^{a,b} : [0,T] \times \mathbb{R} \to \mathbb{R}$ by

$$K^{a,b}(t,x) = \frac{1}{1-\gamma}\left(\int_t^T \delta e^{-\delta(s-t)} e^{-\frac{1}{\theta}\int_t^s \mu(v)\,\mathrm{d}v}\left(a^{\frac{1}{\kappa}} + b^{\frac{1}{\kappa}}\right)^{\frac{\kappa}{\theta}} \mathrm{d}s\right)^{\theta}, \tag{2.33}$$

$$\begin{aligned}
I^{a,b}(t,x) = {} & \frac{1}{1-\gamma}\left((1-\gamma) K^{a,b}(t,x)\right)^{1-\frac{2}{\theta}} \\
& \times \left(1 - \frac{1}{\theta}\right)\left(\int_t^T \begin{pmatrix} \delta e^{-\delta(s-t)} e^{-\frac{1}{\theta}\int_t^s \mu(v)\,\mathrm{d}v} \times \\ \left(a^{\frac{1}{\kappa}} + b^{\frac{1}{\kappa}}\right)^{\frac{\kappa}{\theta}-1} \times \\ \left(a^{\frac{1}{\kappa}-1} a_x + b^{\frac{1}{\kappa}-1} b_x\right) \end{pmatrix} \mathrm{d}s\right)^2 \\
& + \frac{1}{1-\gamma}\left((1-\gamma) K^{a,b}(t,x)\right)^{1-\frac{1}{\theta}} \\
& \times \left\{\left(\frac{1}{\theta} - \frac{1}{\kappa}\right)\int_t^T \begin{pmatrix} \delta e^{-\delta(s-t)} e^{-\frac{1}{\theta}\int_t^s \mu(v)\,\mathrm{d}v} \times \\ \left(a^{\frac{1}{\kappa}} + b^{\frac{1}{\kappa}}\right)^{\frac{\kappa}{\theta}-2} \\ \times \left(a^{\frac{1}{\kappa}-1} a_x + b^{\frac{1}{\kappa}-1} b_x\right)^2 \end{pmatrix} \mathrm{d}s\right. \\
& \left. + \left(\frac{1}{\kappa} - 1\right)\int_t^T \begin{pmatrix} \delta e^{-\delta(s-t)} e^{-\frac{1}{\theta}\int_t^s \mu(v)\,\mathrm{d}v} \times \\ \left(a^{\frac{1}{\kappa}} + b^{\frac{1}{\kappa}}\right)^{\frac{\kappa}{\theta}-1} \\ \times \left(a^{\frac{1}{\kappa}-2} (a_x)^2 + b^{\frac{1}{\kappa}-2} (b_x)^2\right) \end{pmatrix} \mathrm{d}s\right\}.
\end{aligned} \tag{2.34}$$

Here, we have skipped all arguments (t,s,x) inside the integrals. By (2.31)–(2.32), we can write $Z^{c,\pi,b}(t,x) = K^{\Lambda_1^{c,\pi,b},\Lambda_2^{c,\pi,b}}(t,x)$. Hence, assuming sufficient integrability, applying (2.28), and skipping all arguments that are (t,x) or t, we get the following partial derivative

$$\begin{aligned}
Z_t^{c,\pi,b} = {} & -f\left(t, c, x+b, K^{\Lambda_1^{c,\pi,b},\Lambda_2^{c,\pi,b}}\right) \\
& - \left(x(r+\pi\lambda) - c - \hat{\mu} b + w\right) Z_x^{c,\pi,b} \\
& - \frac{1}{2}\pi^2\sigma^2 x^2 Z_{xx}^{c,\pi,b} + \frac{1}{2}\pi^2\sigma^2 x^2 I^{\Lambda_1^{c,\pi,b},\Lambda_2^{c,\pi,b}}.
\end{aligned} \tag{2.35}$$

Here, f and $I^{\Lambda_1^{c,\pi,b},\Lambda_2^{c,\pi,b}}$ are defined in (2.16) and (2.34). Assuming that $Z^{c,\pi,b}$ is in $\mathcal{C}^{1,2}$, using Itô's formula on $Z^{c,\pi,b}\left(t, X^{c,\pi,b}(t)\right)$, and skipping most arguments that are t or $\left(t, X^{c,\pi,b}(t)\right)$, we get that

$$\begin{aligned}
dZ^{c,\pi,b}\left(t, X^{c,\pi,b}(t)\right) = {} & Z_t^{c,\pi,b}\,\mathrm{d}t + Z_x^{c,\pi,b}\left[X^{c,\pi,b}(r+\pi\lambda) - c - \hat{\mu} b + w\right]\mathrm{d}t \\
& + Z_x^{c,\pi,b} X^{c,\pi,b} \pi\sigma\,\mathrm{d}W(t) + \frac{1}{2} Z_{xx}^{c,\pi,b}\pi^2\sigma^2\left(X^{c,\pi,b}\right)^2 \mathrm{d}t\,.
\end{aligned}$$

Hence, plugging in the partial derivatives of $Z^{c,\pi,b}$ and skipping most arguments that are $\left(u, X^{c,\pi,b}(u)\right)$ or u, we get that

$$Z^{c,\pi,b}\left(t, X^{c,\pi,b}(t)\right) = -\int_t^T Z_x^{c,\pi,b} X^{c,\pi,b} \pi\sigma \, \mathrm{d}W(u) + \int_t^T \left(\begin{array}{c} f\left(u, c, X^{c,\pi,b} + b, K^{\Lambda_1^{c,\pi,b}, \Lambda_2^{c,\pi,b}}\right) \\ -\frac{1}{2}\pi^2\sigma^2 \left(X^{c,\pi,b}\right)^2 I^{\Lambda_1^{c,\pi,b}, \Lambda_2^{c,\pi,b}} \end{array} \right) \mathrm{d}u \; . \tag{2.36}$$

2.A.2 The actual proof

Assume that the functions (U, l_1, l_2) from Theorem 2.1 exist, let (c^*, π^*, b^*) be the function of (t, x) that realizes the infimum in (2.17), and assume that (c^*, π^*, b^*) satisfies the assumptions of Theorem 2.1. Then (c^*, π^*, b^*) is easily seen to be a control in $\mathcal{U}^e$. In the next two subsections, we prove that $U = Z^{c^*,\pi^*,b^*}$, and that (c^*, π^*, b^*) is an equilibrium control for $Z^{c,\pi,b}$.

Proof: $U = Z^{c^*,\pi^*,b^*}$

By assumption, U is in $\mathcal{C}^{1,2}$, so using Itô's formula on $U\left(t, X^{c,\pi,b}(t)\right)$ for some $(c, \pi, b) \in \mathcal{U}^e$, plugging in (2.17), and skipping all arguments that are $\left(u, X^{c,\pi,b}(u)\right)$ or u, we get that

$$U\left(t, X^{c,\pi,b}(t)\right) \geq -\int_t^T U_x X^{c,\pi,b} \pi\sigma \, \mathrm{d}W(u) + \int_t^T \left(\begin{array}{c} f\left(u, c, X^{c,\pi,b} + b, K^{l_1, l_2}\right) \\ -\frac{1}{2}\pi^2\sigma^2 \left(X^{c,\pi,b}\right)^2 I^{l_1, l_2} \end{array} \right) \mathrm{d}u \; . \tag{2.37}$$

We write $Z^* = Z^{c^*,\pi^*,b^*}$, $X^* = X^{c^*,\pi^*,b^*}$, and $\Lambda_i^* = \Lambda_i^{c^*,\pi^*,b^*}$ to simplify notation. To establish the relation $U = Z^*$, we note that $\Lambda_i^* = l_i$, $i = 1, 2$. Plugging this into (2.36) with the control (c^*, π^*, b^*), we get that

$$Z^*(t, X^*(t)) = -\int_t^T Z_x^* X^* \pi^* \sigma \, \mathrm{d}W(u) + \int_t^T \left(\begin{array}{c} f\left(u, c^*, X^* + b^*, K^{l_1, l_2}\right) \\ -\frac{1}{2}(\pi^*)^2 \sigma^2 (X^*)^2 I^{l_1, l_2} \end{array} \right) \mathrm{d}u \; . \tag{2.38}$$

Also, with the control (c^*, π^*, b^*), there is equality in (2.37) (because the infimum in (2.17) is realized), so we get that

$$U(t, X^*(t)) = -\int_t^T U_x X^* \pi^* \sigma \, \mathrm{d}W(u) + \int_t^T \left(\begin{array}{c} f\left(u, c^*, X^* + b^*, K^{l_1, l_2}\right) \\ -\frac{1}{2}(\pi^*)^2 \sigma^2 (X^*)^2 I^{l_1, l_2} \end{array} \right) \mathrm{d}u \; . \tag{2.39}$$

We assume that the stochastic integrals in (2.38) and (2.39) are martingales. Fixing some $(s,y) \in [0,T] \times \mathbb{R}$, subtracting $U(s, X^*(s))$ from $Z^*(s, X^*(s))$, and taking conditional expectation given $X^*(s) = y$, we finally arrive at

$$U(s,y) - Z^*(s,y) = E_{s,y}\left[-\int_s^T (U_x - Z^*_x)\, X^* \pi^* \sigma \, \mathrm{d}W(u)\right] = 0 \ .$$

Since (s,y) were arbitrary, we have proven that $Z^* = U$, and consequently

$$U = K^{\Lambda_1^*, \Lambda_2^*} = K^{l_1, l_2} \ . \tag{2.40}$$

Proof: (c^*, π^*, b^*) **is an equilibrium control**

We fix a control $\left(\bar{c}, \bar{\pi}, \bar{b}\right)$ in $\mathcal{U}^e$, a (small) real number $h > 0$, and an initial point $(u,y) \in [0,T] \times \mathbb{R}$. We then define the control $\left(c^h, \pi^h, b^h\right)$ by

$$\left(c^h, \pi^h, b^h\right)(t,x) = \begin{cases} (\bar{c}, \bar{\pi}, \bar{b})(t,x) \ , & u \le t < u+h, x \in \mathbb{R} \ , \\ (c^*, \pi^*, b^*)(t,x) \ , & u+h \le t \le n, x \in \mathbb{R} \ . \end{cases}$$

Below, we write $Z^h = Z^{c^h, \pi^h, b^h}$. To prove that (c^*, π^*, b^*) is an equilibrium control for $Z^{c,\pi,b}$, we introduce the following non-standard assumptions:

Assumptions 2.1. *We assume that there exist functions Λ_1^h and Λ_2^h that satisfy* (2.28) *for the control $\left(c^h, \pi^h, b^h\right)$ for all $u \le t \le s \le n$ and $x \in \mathbb{R}$. We assume that the functions are sufficiently smooth such that for all $t \in [u,T]$ and $x \in \mathbb{R}$*

$$Z^h(t,x) = K^{\Lambda_1^h, \Lambda_2^h}(t,x) \ . \tag{2.41}$$

Also, we assume that Z^h is twice differentiable in the second argument and once differentiable in the first argument with the t-derivative from (2.35)*. Finally, we assume that the following convergences hold true:*

$$\begin{aligned} Z^h(u,y) &\overset{h\to 0}{\longrightarrow} U(u,y) \ , & Z^h_x(u,y) &\overset{h\to 0}{\longrightarrow} U_x(u,y) \ , \\ Z^h_{xx}(u,y) &\overset{h\to 0}{\longrightarrow} U_{xx}(u,y) \ , & I^{\Lambda_1^h, \Lambda_2^h}(u,y) &\overset{h\to 0}{\longrightarrow} I^{l_1, l_2}(u,y) \ . \end{aligned} \tag{2.42}$$

To prove that (c^*, π^*, b^*) is an equilibrium control in the sense of Definition 2.2, we need to verify that the condition (2.15) is satisfied. We recall that $Z^* = U$. Hence, equation (2.15) reads

$$\liminf_{h\to 0} \frac{U(u,y) - Z^h(u,y)}{h} \ge 0 \ .$$

By construction, we have that $Z^h(t,x) = U(t,x)$ for $t \in [u+h, T]$ and $x \in \mathbb{R}$. Thus, applying Taylor's formula for fixed $x = y$, we get that

$$\begin{aligned} \frac{U(u,y) - Z^h(u,y)}{h} &= \frac{U(u,y) - U(u+h,y) - Z^h(u,y) + Z^h(u+h,y)}{h} \\ &= -U_t(u,y) + Z^h_t(u,y) + o(h) \ . \end{aligned}$$

Hence, what we need to show is that

$$\liminf_{h\to 0} \left[-U_t(u,y) + Z_t^h(u,y)\right] \geq 0 \ . \tag{2.43}$$

Applying (2.35), (2.17), (2.40), and (2.41) and skipping most arguments that are (u, y) or u, we get that

$$\begin{aligned} -U_t + Z_t^h \geq\ & f\left(u, \bar{c}, y + \bar{b}, U\right) - f\left(u, \bar{c}, y + \bar{b}, Z^h\right) \\ & + \left(y\left(r + \bar{\pi}\lambda\right) - \bar{c} - \hat{\mu}\bar{b} + w\right)\left(U_x - Z_x^h\right) \\ & + \frac{1}{2}\bar{\pi}^2\sigma^2 y^2\left(U_{xx} - Z_{xx}^h\right) + \frac{1}{2}\bar{\pi}^2\sigma^2 y^2\left(I^{\Lambda_1^h,\Lambda_2^h} - I^{l_1,l_2}\right) \ . \end{aligned} \tag{2.44}$$

The function f is obviously continuous. Hence, plugging (2.42) into (2.44) as h tends to 0, we see that (2.43) is satisfied. This concludes the proof. □

2.B Proof of Theorem 2.2

We assume that l_1 and l_2 from Theorem 2.1 are separable in the sense that there exist $\mathcal{C}^{1,0}$-functions $h_1, h_2 : [0,T]^2 \to \mathbb{R}$ such that

$$l_i(t,s,x) = h_i(t,s)\left(x + L(t)\right)^{1-\gamma} \ , \quad i = 1,2 \ , \tag{2.45}$$

where L is the investor's human wealth defined in (2.4). Then, by (2.40),

$$U(t,x) = \frac{1}{1-\gamma}\left(x + L(t)\right)^{1-\gamma} g^{\theta}(t) \ ,$$

where the function $g : [0,T] \to \mathbb{R}$ is given by

$$g(t) = \int_t^T \delta e^{-\delta(s-t)} e^{-\frac{1}{\theta}\int_t^s \mu(v)\,\mathrm{d}v}\left(h_1^{\frac{1}{\kappa}}(t,s) + h_2^{\frac{1}{\kappa}}(t,s)\right)^{\frac{\kappa}{\theta}} \mathrm{d}s \ .$$

In the above, we assume that $x + L(t) > 0$ and $t < n$. This can be done without loss of generality because if $x + L(t) = 0$ or $t = n$ then $U(t,x) = 0$. Now, assuming sufficient integrability and skipping all arguments that are (t,s,x), (t,x), (t,s), or t, we get the partial derivatives

$$\begin{aligned} U_x &= (x+L)^{-\gamma} g^{\theta} \ , \\ U_{xx} &= -\gamma (x+L)^{-\gamma-1} g^{\theta} \ , \\ U_t &= \frac{1}{1-\gamma}(x+L)^{1-\gamma}\theta g^{\theta-1} g_t + L_t (x+L)^{-\gamma} g^{\theta} \ , \end{aligned} \tag{2.46}$$

$$\begin{aligned} (l_i)_x &= (1-\gamma) h_i (x+L)^{-\gamma} \ , \\ (l_i)_{xx} &= -(1-\gamma)\gamma h_i (x+L)^{-\gamma-1} \ , \\ (l_i)_t &= (h_i)_t (x+L)^{1-\gamma} + (1-\gamma) h_i (x+L)^{-\gamma} L_t \ , \end{aligned} \tag{2.47}$$

and we easily verify that

$$I^{l_1,l_2}(t,x) = K^{a,b}(t,x) \times \begin{pmatrix} \left(1-\frac{1}{\theta}\right)\left(\frac{1-\gamma}{x+L(t)}\right)^2 + \\ \left(\frac{1}{\theta}-\frac{1}{\kappa}\right)\left(\frac{1-\gamma}{x+L(t)}\right)^2 + \\ \left(\frac{1}{\kappa}-1\right)\left(\frac{1-\gamma}{x+L(t)}\right)^2 \end{pmatrix} = 0\ . \tag{2.48}$$

Plugging (2.48) and (2.40) into (2.17) and skipping all arguments that are (t,x) or t, the differential equation for U reduces to

$$U_t = \inf_{(c,\pi,b)\in\Gamma(x,t)} \begin{bmatrix} -f(t,c,x+b,U) \\ -((r+\pi\lambda)x - c - \hat{\mu}b + w)U_x \\ -\frac{1}{2}\sigma^2\pi^2x^2U_{xx} \end{bmatrix} . \tag{2.49}$$

Also, plugging (2.47) into (2.18), skipping all arguments that are (t,s,x), (t,x), or t, dividing by $(x+L)^{1-\gamma}$, and subtracting $(1-\gamma)h_i(x+L)^{-1}L_t$, we get the following differential equations for h_1 and h_2:

$$\begin{aligned} (h_i)_t &= -\left(r+\hat{\mu}-\frac{c^*}{x+L}-\hat{\mu}\frac{b^*+x}{x+L}+\lambda\frac{\pi^*x}{x+L}-\frac{1}{2}\left(\frac{\pi^*x}{x+L}\right)^2\sigma^2\gamma\right) \\ &\quad\times(1-\gamma)h_i\ ,\quad i=1,2\ , \\ h_1(s,s) &= \left(\frac{c^*(s,x)}{x+L(s)}\right)^{1-\gamma}\ , \\ h_2(s,s) &= \varepsilon(s)\mu(s)\left(\frac{b^*(s,x)+x}{x+L(s)}\right)^{1-\gamma}\ . \end{aligned} \tag{2.50}$$

We need to verify the separability assumption in (2.45). From (2.50) we see that the differential equations for h_1 and h_2 become ordinary (and independent of x), when $\frac{\pi^*x}{x+L}$, $\frac{c^*}{x+L}$, and $\frac{b^*+x}{x+L}$ do not depend on x. Therefore, to verify the assumption (2.45), it suffices to verify that $\frac{\pi^*x}{x+L}$, $\frac{c^*}{x+L}$, and $\frac{b^*+x}{x+L}$ do not depend on x. For the verification, we recall that (c^*,π^*,b^*) solves the continuum of minimization problems in (2.49). Plugging (2.46) into (2.49) and innocently dividing by $(1-\gamma)U$, we face the problem

$$\begin{aligned} &\frac{\theta}{1-\gamma}\frac{g_t}{g}+\frac{L_t}{x+L} \\ &= \inf_{(c,\pi,b)\in\Gamma(x,t)} \begin{bmatrix} -\frac{1}{1-\gamma}\left(\theta\delta\left(\left(\frac{c}{x+L}\right)^{\frac{1-\gamma}{\kappa}}+\varepsilon^{\frac{1}{\kappa}}\mu^{\frac{1}{\kappa}}\left(\frac{b+x}{x+L}\right)^{\frac{1-\gamma}{\kappa}}\right)^{\frac{\kappa}{\theta}}\frac{1}{g}\right) \\ +\frac{1}{1-\gamma}(\mu+\theta\delta)-\left(\frac{(r+\pi\lambda)x}{x+L}-\frac{c+\hat{\mu}b-w}{x+L}-\frac{1}{2}\gamma\sigma^2\left(\frac{\pi x}{x+L}\right)^2\right) \end{bmatrix} . \end{aligned} \tag{2.51}$$

To solve this minimization problem, we differentiate the objective function with respect to each of the (sub)controls and set the partial derivatives equal to zero. Note that we look for an interior solution because of the constraint

$(c, \pi, b) \in \Gamma(x, t)$. We get the solution in (2.19), so we have the crucial independence of x, and it is easily seen that

$$(c^*(t,x), \pi^*(t,x), b^*(t,x)) \in \Gamma(t,x) \ .$$

Hence, we have verified the separability assumption, and we have derived expressions for the optimal control.

Next, we would like to derive closed-form expressions for the functions h_1, h_2, and g. Plugging the optimal control in (2.19) back into (2.51), subtracting $\frac{L_t}{x+L} = \frac{-w+(r+\hat{\mu})L}{x+L}$, and dividing by $\frac{\theta}{(1-\gamma)g}$, we get the following differential equation for g:

$$\begin{aligned} g_t = &-\phi\delta^{\frac{1}{\phi}} \left(1 + \left(\frac{\varepsilon\mu}{\hat{\mu}^{1-\gamma}}\right)^{\frac{1}{\gamma+\kappa-1}}\right)^{\frac{(\kappa-1+\gamma)(1-\phi)}{(1-\gamma)\phi}} g^{1-\frac{1}{\phi}} \\ &- \left((1-\phi)\left(r + \frac{1}{2}\frac{1}{\gamma}\frac{\lambda^2}{\sigma^2} + \hat{\mu} - \frac{\mu}{1-\gamma}\right) - \delta\right) g \ , \\ g(T) = &\ 0 \ . \end{aligned} \tag{2.52}$$

This differential equation has a well-known form, and the solution is given in Theorem 2.2. Moreover, plugging the optimal control in (2.19) back into (2.50), we get the following ordinary differential equations for h_1 and h_2:

$$\begin{aligned} (h_i)_t = &-\left(r + \hat{\mu} - \delta^{\frac{1}{\phi}} g^{-\frac{1}{\phi}} \left(1 + \left(\frac{\varepsilon\mu}{\hat{\mu}^{1-\gamma}}\right)^{\frac{1}{\gamma+\kappa-1}}\right)^{\frac{(\kappa-1+\gamma)(1-\phi)}{(1-\gamma)\phi}} + \frac{1}{2}\frac{\lambda^2}{\gamma\sigma^2}\right) \\ &\times (1-\gamma) h_i \ , \\ h_1(s,s) = &\left(\delta^{\frac{1}{\phi}} g^{-\frac{1}{\phi}}(s)\right)^{1-\gamma} \left(1 + \left(\frac{\varepsilon(s)\mu(s)}{\hat{\mu}^{1-\gamma}(s)}\right)^{\frac{1}{\gamma+\kappa-1}}\right)^{\frac{\kappa-\phi\kappa-1+\gamma}{\phi}} \ , \\ h_2(s,s) = &\ h_1(s,s) \left(\frac{\varepsilon(s)\mu(s)}{\hat{\mu}^{1-\gamma}(s)}\right)^{\frac{\kappa}{\gamma+\kappa-1}} \ . \end{aligned}$$

Again, these differential equations have a well-known form, and the solutions are given in Theorem 2.2. This concludes the proof. □

2.C Derivation of SDE for the optimal consumption

Define the function $v : [0,T] \to \mathbb{R}$ by

$$v(t) = \left(1 + \left(\frac{\varepsilon(t)\mu(t)}{\hat{\mu}^{1-\gamma}(t)}\right)^{\frac{1}{\gamma+\kappa-1}}\right)^{\frac{\kappa-\phi\kappa-1+\gamma}{(1-\gamma)\phi}} \ .$$

Then v is in $\mathcal{C}^1([0,T])$ if the mortality intensities μ, $\hat{\mu}$, and the weight function ε are so, and the optimal consumption rate from Theorem 2.2 can be written as

$$c^*(t,x) = \delta^{\frac{1}{\phi}} g^{-\frac{1}{\phi}}(t)\, v(t)\,(x + L(t)) \ .$$

Assume that ε, μ, and $\hat{\mu}$ are $\mathcal{C}^1$-functions. Since also g and L are $\mathcal{C}^1$-functions, it holds that c^* is in $\mathcal{C}^{1\times 2}$, and we get the partial derivatives

$$\begin{aligned}
c_t^*(t,x) &= \left(-\frac{1}{\phi}\frac{g_t(t)}{g(t)} + \frac{v_t(t)}{v(t)} + \frac{-w(t) + (r + \hat{\mu}(t))\, L(t)}{x + L(t)}\right) c^*(t,x) \ , \\
c_x^*(t,x) &= \delta^{\frac{1}{\phi}} g^{-\frac{1}{\phi}}(t)\, v(t) = \frac{1}{x + L(t)} c^*(t,x) \ , \\
c_{xx}^*(t,x) &= 0 \ .
\end{aligned}$$

Let X^* be the wealth process stemming from the optimal control (c^*, π^*, b^*). Using Itô's formula on $c^*(t, X^*(t))$ (see footnote 2), we get the SDE

$$\begin{aligned}
\frac{\mathrm{d}c^*(t,X^*(t))}{c^*(t,X^*(t))} &= \frac{c_t^*(t,X^*(t))\,\mathrm{d}t + c_x^*(t,X^*(t))\,\mathrm{d}X^*(t)}{c^*(t,X^*(t))} \\
&= \frac{1}{\phi}\left(r + \hat{\mu}(t) - \delta - \frac{1}{\theta}\mu(t) + (1+\phi)\frac{1}{2}\frac{\lambda^2}{\gamma\sigma^2}\right)\mathrm{d}t \\
&\quad + \frac{v_t(t)}{v(t)}\mathrm{d}t + \frac{\lambda}{\gamma\sigma}\mathrm{d}W(t) \ , \\
c^*(0,X^*(0)) &= c^*(0,x_0) \ .
\end{aligned} \tag{2.53}$$

In the calculations, we have used the expressions for the optimal control (c^*, π^*, b^*) from Theorem 2.2. Also, we have plugged in the derivative g_t from (2.52) in Appendix 2.B. In (2.53), the entity $\frac{v_t(t)}{v(t)}$ is rather complicated, but it simplifies if we assume that $\hat{\mu} = \alpha\mu$ for some constant $\alpha > 0$ and that ε is constant. We then get that

$$\begin{aligned}
\frac{v_t(t)}{v(t)} &= \frac{\kappa - \phi\kappa - 1 + \gamma}{(1-\gamma)\,\phi}\left(1 + \varepsilon^{\frac{1}{\gamma+\kappa-1}}\alpha^{\frac{\gamma-1}{\gamma+\kappa-1}}\mu(t)^{\frac{\gamma}{\gamma+\kappa-1}}\right)^{-1} \\
&\quad \times \frac{\gamma}{\gamma+\kappa-1}\mu_t(t)\,\varepsilon^{\frac{1}{\gamma+\kappa-1}}\alpha^{\frac{\gamma-1}{\gamma+\kappa-1}}\mu(t)^{\frac{\gamma}{\gamma+\kappa-1}-1} \ .
\end{aligned}$$

Also, the SDE in (2.53) reduces to

$$\begin{aligned}
\frac{\mathrm{d}c^*(t,X^*(t))}{c^*(t,X^*(t))} &= \frac{1}{\phi}\left(r - \delta + \left(\alpha - \frac{1}{\theta}\right)\mu(t) + (1+\phi)\frac{1}{2}\frac{\lambda^2}{\gamma\sigma^2}\right)\mathrm{d}t \\
&\quad + \frac{v_t(t)}{v(t)}\mathrm{d}t + \frac{\lambda}{\gamma\sigma}\mathrm{d}W(t) \ .
\end{aligned}$$

Chapter 3

Nonrecursive Separation of Risk and Time Preferences

MATTHIAS ALBRECHT FAHRENWALDT, NINNA REITZEL JENSEN, & MOGENS STEFFENSEN (2016)

ABSTRACT: Recursive utility disentangles preferences with respect to time and risk by recursively building up a value function of local increments. This involves certainty equivalents of indirect utility. Instead we disentangle preferences with respect to time and risk by building up a value function as a non-linear aggregation of certainty equivalents of direct utility of consumption. This entails time-consistency issues which are dealt with by looking for an equilibrium control and an equilibrium value function rather than an classically optimal control and a classical optimal value function. We characterize the solution in a general diffusive incomplete market model and find that, in certain special cases of utmost interest, the characterization coincides with what would arise from a recursive utility approach. But also importantly, in other cases, it does not: The two approaches are fundamentally different but match, exclusively but importantly, in the mathematically special case of homogeneity of the value function.

KEYWORDS: Time-consistency, time-global preferences, recursive utility, equilibrium strategies, generalized Hamilton–Jacobi–Bellman equation, continuous time, certainty equivalents.

3.1 Introduction

We formulate a continuous-time dynamic consumption-investment problem where preferences with respect to risk and time variability are disentangled. In contrast to recursive utility which also builds on the idea of disentangling these preferences, our value function is based on a time-global objective with non-time-additive utility. This allows for working with certainty equivalents of direct utility of consumption rather than indirect utility. Time-inconsistency arising from non-time-additivity is dealt with by looking for a subgame perfect equilibrium among a continuum of selves. We consider a general incomplete market with coefficients driven by a non-hedgeable economic state process. In special cases that include the Merton market, we find a resulting behavior that coincides with that coming from recursive utility with Epstein-Zin preferences. Among these special cases, we also find closed-form solutions to new problem formulations beyond standard power utility, including non-hedgeable consumer price indexation and exponential utility. Thus, our contribution to the literature is two-fold: We base—we believe as the first—the disentanglement on a time-global objective for a general financial market. Second we detect new and relevant solveable consumption-investment problems in incomplete markets within our problem formulation where the solution coincides with what would have been obtained by recursive utility.

Recursive utility was developed by Epstein and Zin (1989, 1991) based on work by Kreps and Porteus (1978, 1979). It is celebrated for disentangling preferences with respect to risk and time. Its continuous-time limit, spoken of as stochastic differential utility or, simply, continuous-time recursive utility, developed by Duffie and Epstein (1992a) has the same ability to allow for separate preference functions against variability over risk and (continuous) time. It is widely used to study optimal consumption-portfolio choice in various markets, see e.g. Schroder and Skiadas (1999, 2005); Kraft et al. (2013). Also, recursive utility is used to examine ambiguity aversion and preferences for resolution of uncertainty, see e.g. Chen and Epstein (2002); Skiadas (1998, 2013). Issues with differentiability when going to continuous-time were addressed by Kraft and Seifried (2010, 2014). A particularity of recursive utility is, of course, the definitional recursive building of the value function or indirect utility function. This means that, when locally aggregating present consumption with the utility of future consumption, the latter is represented by its indirect utility. In the recursion appears the certainty equivalent with respect to the representative indirect utility of wealth rather than the underlying future uncertain consumption.

Indirect utility appears to be the right representative of utility of future consumption, given that we start out with a recursive definition. Yet, here we suggest to start out with a time-global objective built up by certainty equivalents with respect to future uncertain consumption. Said differently, we suggest to replace the indirect utility representation of future consumption by the direct utility of future consumption itself. Apart from that, our objective remains the same: To separate preferences for risk and time. Once having formed certainty equivalents of future consumption at different points in time, we think of them as "certain" values attributed to these time points. This allows for a non-linear aggregation of these certainty equivalents which relates to preferences with respect to time only. Our objective becomes non-linear in time which, at first sight, dumps the idea for reasons of time-inconsistency issues that are completely avoided with recursive utility. There, the controls are, definitional from the recursive structure, time-consistent, so why bother with time-inconsistency issues? Because, we find the construction of a time-global objective based on direct utility of future consumption instead of indirect utility appealing and, by now, the complications with time-inconsistency can be overcome. That is, because we should and because we can.

Already in the definition of recursive utility, time-consistency issues are delicately avoided. First the certainty equivalent of the indirect utility is formed. Then this is non-linearly time-aggregated with present consumption. The alternative order is unfriendly: To first non-linearly time-aggregate indirect utility and consumption and then take the expected utility here-of. It is the non-linear time-aggregation under uncertainty that leads to time-inconsistency issues. Although we suggest a completely different formulation, we also have time-inconsistency issues, but for different reasons. We make non-linear time-aggregation of objects we can think of as certain like it is done for recursive utility. But we aggregate over a global time-horizon rather than a local (one-period in discrete-time and infinitesimal in continuous-time) time-horizon, as in the case of recursive utility.

Time-inconsistent behavior was initially formalized by Strotz (1955). Pollak (1968), Goldman (1980), and Laibson (1997) contributed to the understanding of the problem as an intra-personal game and looked for subgame perfect equilibria. Ekeland and Pirvu (2008) defined a continuous-time subgame perfect equilibrium in order to deal with the time-inconsistency arising from replacing exponential discounting of utility by hyperbolic discounting. We follow their definition and derive the equilibrium value and equilibrium strategy when the time-inconsistency arises from the non-linear aggregation of certainty equivalents as explained above.

The idea of summing up certainty equivalents over global time was also pursued by Jensen and Steffensen (2015) [Chapter 2 of this thesis]. They considered a consumption-investment-insurance problem in a Merton market for an investor with an uncertain lifetime and access to life insurance. The disentanglement of preferences for risk and time is, there, a starting point for the idea of also disentangling utility of consumption as alive and inheritors utility of consumption after the death of the investor. Already they show that in the special case of a Merton market the solution to our optimization problem coincides with that of recursive utility with Epstein-Zin preferences. We obtain this coincidence with the Merton market and recursive utility from a different angle.

We start out with a general diffusive, incomplete market with a risky, diffusive asset with price coefficients driven by another diffusive economic state process that cannot be perfectly hedged. Coincidence with the solution for recursive utility with Epstein-Zin preferences in the special case of a Merton market falls out. But we also characterize solutions for much more general markets that have previously been studied under recursive utility. This unveils, in terms of resulting behavior, a fundamental difference between recursive utility and our approach. In general cases, studied under recursive utility by Chacko and Viceira (2005) and Kraft et al. (2013), the generalized Bellman equation that we find to characterize our equilibrium value, contains additional terms compared to the standard recursive utility Bellman-type equation. Only when we have complete separability in time, wealth, and the economic state process, we agree with users of recursive utility on the characterization of the solution. On the other hand, we study in details such special cases leading to linearly homogeneous value functions that, to our knowledge, have not been studied before. They include cases with power utility where we scale consumption by the economic state process, interpreting this process as an only partly hedgeable consumer price index, and cases with exponential utility. We provide explicit solutions in these cases.

The outline of the paper is as follows. In Section 3.2, we present the model for the price and wealth processes. We motivate our problem formulation and relate it to standard recursive utility. In Section 3.3, we define the set of admissible controls and the concept of equilibrium and state our main theorem with sufficient conditions to determine equilibrium controls and the corresponding equilibrium value function. In Section 3.4, we present two nontrivial examples of the framework with incomplete markets. We consider two different choices of the utility functions, namely power utility and exponential utility. We provide explicit solutions, and we establish a connection to recursive utility.

3.2 General Set-Up and Optimization Problem

In this section, we present our optimization problem and its connection to related problems. This section is central because it is our problem formulation, rather than the solution, that is the innovative part of the paper.

We consider an investor making decisions concerning consumption, c, and investment, π, in a Brownian market. The wealth of the investor evolves according to the dynamics

$$\mathrm{d}X_t^{c,\pi} = \mu^{c,\pi}\left(t, X_t^{c,\pi}, Y_t\right)\,\mathrm{d}t + \sigma^{c,\pi}\left(t, X_t^{c,\pi}, Y_t\right)\mathrm{d}W_t\ ,\ X_0^{c,\pi} = x_0\ , \quad (3.1)$$

where Y is a non-traded state process with the dynamics

$$\mathrm{d}Y_t = \alpha\left(t, Y_t\right)\,\mathrm{d}t + \beta\left(t, Y_t\right)\left(\rho\,\mathrm{d}W_t + \sqrt{1-\rho^2}\,\mathrm{d}\bar{W}_t\right)\ ,\ Y_0 = y_0\ . \quad (3.2)$$

Here, $\mu^{c,\pi}, \sigma^{c,\pi}, \alpha, \beta$ are sufficiently regular functions, and W and $\bar{W}$ are two independent Brownian motions. The volatility in $X^{c,\pi}$ arise from investment in a stock. For later use, we introduce the infinitesimal generator $\mathcal{A}^{c,\pi}$ of $(X^{c,\pi}, Y)$ which is given by

$$\mathcal{A}^{c,\pi} = \mu^{c,\pi}\partial_x + \frac{1}{2}\left(\sigma^{c,\pi}\right)^2\partial_x^2 + \alpha\partial_y + \frac{1}{2}\beta^2\partial_y^2 + \rho\beta\sigma^{c,\pi}\partial_{xy}\ .$$

Note that this operator is both time- and space-dependent. In the next section, we make assumptions concerning the coefficients to guarantee the existence and uniqueness of solutions to the Kolmogorov equations $\partial_t l = -\mathcal{A}^{c,\pi} l$ with appropriate terminal condition. Also, our main theorem will assume the existence of such a solution. This has to be checked in concrete situations.

A classical optimization problem formalized for the investor is that of maximizing expected time-additive utility of consumption and final wealth,

$$\sup_{c,\pi} \mathbb{E}\left[\int_0^T e^{-\delta s} u\left(c\left(t, X_t^{c,\pi}, Y_t\right)\right) \mathrm{d}t + e^{-\delta T} \Phi\left(X_T^{c,\pi}\right)\right], \tag{3.3}$$

where $\delta \geq 0$ is a subjective utility discount rate, u is an instantaneous utility function, and Φ is a utility function for final wealth. The utility functions u and Φ characterize the investor's preferences with respect to risk. The problem in (3.3) can be dealt with by embedding it in a value function given by

$$V\left(t,x,y\right) = \sup_{c,\pi} \mathbb{E}_{t,x,y}\left[\begin{array}{c} \int_t^T e^{-\delta(s-t)} u\left(c\left(s, X_s^{c,\pi}, Y_s\right)\right) \mathrm{d}s \\ + e^{-\delta(T-t)} \Phi\left(X_T^{c,\pi}\right) \end{array}\right], \tag{3.4}$$

where $\mathbb{E}_{t,x,y}$ denotes conditional expectation given $X_t^{c,\pi} = x$ and $Y_t = y$. The controls (c, π) are chosen among a set of admissible strategies which essentially means that (3.1) has a solution and that certain integrals with respect to the Brownian motions have expectation zero. By means of dynamic programming techniques, the value function can be characterized by a certain partial differential equation containing a local optimization problem at each point (t, x, y). The solution for (c, π) to the local optimization problem can be proven to also produce the solution for (c, π) to the global optimization problem in (3.3). This is essentially a consequence of the linearity of the expectation operating on an infinitesimal sum of utility of consumption rates. We spell out here that this linearity is essential for the coincidence between local and global optimization since this linearity is serially spoiled below—and, thus, is also the coincidence.

We wrote above that the utility function u characterizes the investor's preferences with respect to risk, but u also plays a different indirect role in the time-additivity of (3.4). Below, we formalize a way to disentangle preferences for risk and time, but, first, we consider briefly how the disentanglement is typically established within the theory of recursive utility. Instead of working with time-global objectives like the one given in (3.4), the standard approach in recursive utility is to study the local discrete-time objective

$$V\left(t, X_t\right) = W\left(c_t, u^{-1}\left(\mathbb{E}_t\left[V\left(t+\Delta, X_{t+\Delta}\right)\right]\right)\right), \tag{3.5}$$

where c_t is the consumption at time t, $u^{-1}\left(\mathbb{E}_t\left[V\left(t+\Delta, X_{t+\Delta}\right)\right]\right)$ is the time t certainty equivalent of having $X_{t+\Delta}$ for consumption from time $t+\Delta$ and onward, and the function W is the so-called time aggregator, aggregating the utility of present consumption c_t and future consumption represented by $u^{-1}\left(\mathbb{E}_t\left[V\left(t+\Delta, X_{t+\Delta}\right)\right]\right)$. An important special case is when $(c,v) \mapsto W\left(c, u^{-1}(v)\right)$ is additive in c and v. Then we obtain time-additive utility, see e.g. Duffie and Epstein (1992a).

The continuous-time equivalent of these patterns of thinking were studied by Duffie and Epstein (1992a,b) under the name stochastic differential utility. The main ingredients are still a certainty equivalent of the value function (indirect utility) and an aggregator. However, taking $\Delta \to 0$ in (3.5) is, in general, a complicated operation that involves differentiability of the certainty equivalent and the aggregator. Kraft and Seifried (2010) propose an alternative notion of differentiability compared to Duffie and Epstein (1992a,b) in order to make the notion of stochastic differential utility more general and robust to e.g. inclusion of non-Brownian markets.

In (3.5), we form a so-called certainty equivalent in terms of

$$u^{-1}\left(\mathbb{E}_t\left[V\left(t+\Delta, X_{t+\Delta}\right)\right]\right) \ ,$$

i.e. in terms of the indirect utility V of wealth rather than the utility function u of consumption. Below we formalize a problem that is based on the certainty equivalence of direct utility of consumption rather than indirect utility. The fundamental idea is to formalize a continuous-time global optimization problem that encompasses both risk and time preferences. This is in sharp contrast to the approach taken by Duffie and Epstein (1992a,b) and Kraft and Seifried (2010) who are challenged by the notion of differentiability when Δ tends to zero in (3.5). We suggest the following approach:

For each future time point s, we form the certainty equivalent of the consumption rate, conditional on $X^{c,\pi}(t) = x$ and $Y(t) = y$,

$$u^{-1}\left(\mathbb{E}_{t,x,y}\left[u\left(c\left(s, X_s^{c,\pi}, Y_s\right)\right)\right]\right) \ .$$

For all $s > t$, these are known at time t, and we are therefore inclined to treat them as deterministic future consumption rates. Now, we let a different function, say $\bar{\varphi}$, formalize the investor's time preferences with respect to these certainty equivalents. The investor's utility from time t and onward is

$$\begin{aligned}
&\int_t^T e^{-\delta(s-t)}\bar{\varphi}\left(u^{-1}\left(\mathbb{E}_{t,x,y}\left[u\left(c\left(s, X_s^{c,\pi}, Y_s\right)\right)\right]\right)\right) \mathrm{d}s \\
&\quad + \omega e^{-\delta(T-t)}\bar{\varphi}\left(u^{-1}\left(\mathbb{E}_{t,x,y}\left[u\left(X_T^{c,\pi}\right)\right]\right)\right) \\
&= \int_t^T e^{-\delta(s-t)}\varphi\left(\mathbb{E}_{t,x,y}\left[u\left(c\left(s, X_s^{c,\pi}, Y_s\right)\right)\right]\right) \mathrm{d}s \\
&\quad + \omega e^{-\delta(T-t)}\varphi\left(\mathbb{E}_{t,x,y}\left[u\left(X_T^{c,\pi}\right)\right]\right) \ ,
\end{aligned}$$

where $u\left(X_T^{c,\pi}\right)$ is utility from final wealth, ω is a scaling factor allowing for different weight on utility from consumption and final wealth, and $\varphi = \bar{\varphi} \circ u^{-1}$.

At this point, it is clear that we have a problem beyond what can be dealt with by classical dynamic programming. Namely, due to the transform φ of the expectation, we cannot exploit the linearity of the expectation operator and interchange expectation and time-addition. Before discussing what we can do "instead of" classical dynamic programming, we twist the problem in three ways. First, we allow the utility of consumption and terminal wealth to depend on the process Y. More specifically, we replace $u\left(c\left(s, X_s^{c,\pi}\right)\right)$ by $u\left(Y_s, c\left(s, X_s^{c,\pi}\right)\right)$. This turns out to be mathematically tractable, and we can, for example, think of Y as an index of purchasing power or a minimum subsistence level, depending on the shape of u. Second, we introduce separate utility functions for consumption, u_1, and final wealth, u_2. Third, while we are at "destroying" the workability of dynamic programming techniques, we multiply the problem with the constant δ and transform it with an increasing function f. Now, the value function reads

$$V^{c,\pi}(t,x,y) = f\left(\begin{array}{c} \int_t^T \delta e^{-\delta(s-t)} \left[\varphi \circ m_1^{c,\pi}\right](t,s,x,y)\, \mathrm{d}s \\ +\omega\delta e^{-\delta(T-t)} \left[\varphi \circ m_2^{c,\pi}\right](t,T,x,y) \end{array} \right), \tag{3.6}$$

where, for $0 \le t \le s \le T$,

$$\begin{aligned} m_1^{c,\pi}(t,s,x,y) &= \mathbb{E}_{t,x,y}\left[u_1\left(Y_s, c\left(s, X_s^{c,\pi}, Y_s\right)\right)\right] \ , \\ m_2^{c,\pi}(t,T,x,y) &= \mathbb{E}_{t,x,y}\left[u_2\left(Y_T, X_T^{c,\pi}\right)\right] \ . \end{aligned}$$

The function f is convenient since it does not hurt optimal behavior such that we can choose f however we want to make the mathematical representation of the value function as attractive as possible. At the moment, the reader may think that $f = \mathrm{id}$ (the identity function) is a particularly attractive choice of f, but this turns out not to be true in general. Rather, one should seek a function f that, in some sense that is made clear in the following section, offsets the complication from the function φ under the integral. In an abstract sense, we seek a non-linearizing function f that offsets the non-linearity stemming from the function φ such that the problem of optimizing (3.6) appears linear.

The choice $f = \varphi^{-1}$ turns out to be particularly convenient, at least in some cases. This choice is motivated by calculations and remarks in the next section. Note that given this insight, the choice $f = \varphi^{-1} = \mathrm{id}$, corresponding to $\bar{\varphi} = u^{-1}$, shows why there is typically no "normalization issue" for time-additive utility. In that case the normalized value function is, indeed, given by (3.4).

The problem of maximizing (3.6) is certainly non-standard due to its serial non-linearity. A seemingly different and new strand of literature on non-linear optimization problems was initiated recently by Basak and Chabakauri (2010). In a Brownian market, they solve the dynamic mean-variance problem without so-called precommitment. The combination of no precommitment and the variance appearing in the objective forms the non-linearity since the variance contains the non-linear square function of the expectation of the wealth. Recent works elaborate on the techniques: Björk et al. (2014) study the mean-variance investment problem in a general Markovian setting; Czichowsky (2013) works with mean-variance problems in a non-Markovian setting by means of quadratic projection methods; Kronborg and Steffensen (2015) study mainly the mean-variance problem in a Black-Scholes setting but include optimization over consumption. It turns out that these techniques are well-suited for approaching non-linear problems like (3.6) in a specific way. The idea of adding up certainty equivalents is already explored in Jensen and Steffensen (2015). There, focus is on disentanglement of risk aversion and EIS for a power maximizing investor with uncertain lifetime and access to life insurance in a Black-Scholes setting.

The problem of maximizing (3.6) is complicated and, since dynamic programming does not work, there is no reason to believe that solutions to local and global optimization problems coincide in the same way as for (3.3) and (3.4). The non-linearity of (3.6) means that the solution at time 0 is likely to be inconsistent with the solution at time $t > 0$ if we search for an optimal control among all the usually admissible ones, namely those for which (3.1) has a solution. By inconsistent we mean that the decision we make at time t is not the same as the decision we plan to make at time t, for the same realization of $(X^{c,\pi}, Y)$. Here, we proceed as in Jensen and Steffensen (2015), take inspiration from Björk et al. (2014), and search for an equilibrium control for the value function $V^{c,\pi}$. The theoretical background for the equilibrium approach is equilibrium theory of continuous-time games. Actually, the resulting strategy is a Nash equilibrium strategy in a game where infinitesimally many so-called multiple selves are competing and where the time t-self knows that the continuum of all “later selves”, i.e. s-selves for $s > t$, face the same game. Therefore, Björk et al. (2014) speak of the resulting strategies as equilibrium strategies rather than optimal strategies. This approach produces an optimal control process that does not solve for the supremum over all usual strategies in a usual sense. Rather, it is the best strategy given that one will later on follow a strategy based on the same objective conditioning on updated information. This conforms with with Basak and Chabakauri (2010) and subsequent papers mentioned above. It should be mentioned that this approach to dynamic decision making dates further back to Strotz (1955) and Pollak (1968).

3.3 Equilibrium and Verification Theorem

In this section we define the set of admissible controls and the concept of equilibrium and state our main theorem. The theorem gives sufficient conditions to determine the equilibrium controls and the equilibrium value function.

In the following, let f and φ be in C^2 and f^{-1} be in C^0. For technical reasons in the proof of Theorem 3.1, we must make a hypothesis that guarantees the existence and uniqueness of $C^{1,2}$-solutions to the terminal value problem

$$\left.\begin{aligned} \partial_t u(t,x,y) &= -\mathcal{A}^{c,\pi} u(t,x,y) \ , \\ u(T,x,y) &= g(x,y) \ . \end{aligned}\right\} \tag{3.7}$$

Recall that we call a function $\mathbb{R}^n \to \mathbb{R}^m$ *slowly increasing* if it is smooth and all derivatives are bounded in norm by a polynomial.

Hypothesis 3.1. The matrix-valued function $[0,T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^2$ given by

$$\begin{pmatrix} \sigma^{c,\pi}(t,x,y) & \rho\beta(t,y)\sigma^{c,\pi}(t,x,y) \\ \rho\beta(t,y)\sigma^{c,\pi}(t,x,y) & \beta(t,y) \end{pmatrix} ,$$

and the vector-valued function $[0,T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^2$ given by $(\mu^{c,\pi}(t,x,y), \alpha(t,y))$ is smooth, have bounded first derivatives, and is slowly increasing. Also, $g \in C^3\left(\mathbb{R}^2\right)$ is polynomially bounded.

This hypothesis is not the most general one can make but is general enough for practical situations. We took these assumptions from Theorem 2.12 of Stroock (1983). The existence and uniqueness of classical solutions of (3.7) is then guaranteed by Theorem 2.21 of Stroock (1983). Given concrete choices of the coefficient functions of $\mathcal{A}^{c,\pi}$ one must check that (3.7) does indeed have a solution for any admissible controls.

Definition 3.1 (Admissible controls)**.** We call a control (c,π) admissible if

(i) $(c,\pi) \in C([0,T] \times \mathbb{R}^2) \times C([0,T] \times \mathbb{R}^2)$,

(ii) Hypothesis 3.1 is satisfied.

Remark 3.1. Note that Hypothesis 3.1 ensures that the system of stochactic differential equations (SDE's) in Equations (3.1)–(3.2) has a unique solution by Theorem 2.12 of Stroock (1983). This solution has the property that $\mathbb{E}\left[||(X_t, Y_t)||^p\right] < \infty$ for all $p \in [2,\infty)$ for given initial conditions (x_0, y_0).

Rewriting Definition 2.1 in Björk et al. (2014) in the language of this paper, we get the following definition of equilibrium:

Definition 3.2. Consider an admissible control (c^*, π^*) (informally viewed as a candidate equilibrium control). Choose a fixed, admissible control $(\bar{c}, \bar{\pi})$, a real number $h > 0$, and an initial point $(u,x,y) \in [0,T] \times \mathbb{R}^2$. Define the control $\left(c^h, \pi^h\right)$ by

$$\left(c^h, \pi^h\right)(t,\bar{x},\bar{y}) = \begin{cases} (\bar{c},\bar{\pi})(t,\bar{x},\bar{y}) \ , & u \le t < u+h, \bar{x}, \bar{y} \in \mathbb{R} \ , \\ (c^*,\pi^*)(t,\bar{x},\bar{y}) \ , & u+h \le t \le n, \bar{x}, \bar{y} \in \mathbb{R} \ . \end{cases}$$

If for all admissible controls $(\bar{c}, \bar{\pi})$ and all points $(u, x, y) \in [0, T] \times \mathbb{R}^2$

$$\liminf_{h \to 0} \frac{V^{c^*, \pi^*}(u, x, y) - V^{c^h, \pi^h}(u, x, y)}{h} \geq 0 \ , \tag{3.8}$$

we say that (c^*, π^*) is an equilibrium control for the function $V^{c,\pi}$. The corresponding equilibrium value function V^* is given by

$$V^*(t, x, y) = V^{c^*, \pi^*}(t, x, y) \ .$$

We stress that an equilibrium control is not optimal in the sense that it maximizes the value function. However, the control is optimal in the "intuitive" sense that it maximizes the investor's utility given that the investor continues to use the control. Björk et al. (2014) prove neither existence nor uniqueness of the equilibrium control, so there might be several or even no equilibrium controls.

We are now ready to state the key result of this section. The proof is in Appendix 3.A.

Theorem 3.1 (Verification theorem)**.** *Assume there exist a pentadruple of functions $(U, l_1, l_2, c^*, \pi^*)$ such that the following holds:*

(i) Regularity:

- *U is in $C^{1,2,2}([0, T] \times \mathbb{R} \times \mathbb{R})$.*
- *(c^*, π^*) are admissible controls.*
- *$l_1(t, s, x, y)$ and $l_2(t, T, x, y)$ are defined for $0 \leq t \leq s \leq T$ and $x, y \in \mathbb{R}$, l_1, l_2 are C^1 in t and C^2 in x, y, and l_1 is jointly continuous in t, s.*

(ii) Equilibrium: *The function U solves the pseudo-Bellman equation*

$$\left.\begin{array}{l} \partial_t U(t, x, y) = \inf\limits_{c,\pi} \left[\begin{array}{c} -F\left(c, y, \bar{U}(t, x, y)\right) - \mathcal{A}^{c,\pi} U(t, x, y) \\ +\frac{1}{2} \sigma^{c,\pi}(t, x, y)^2 R_x(t, x, y) \\ +\frac{1}{2} \beta(t, y)^2 R_y(t, x, y) \\ +\rho \beta(t, y) \sigma^{c,\pi}(t, x, y) R_{xy}(t, x, y) \end{array} \right] , \\ U(T, x, y) = f\left(\omega \delta \left[\varphi \circ u_2\right](y, x)\right) \ , \end{array}\right\} \tag{3.9}$$

where the infimum ranges over all admissible controls, the function F is given by

$$F\left(c, y, \bar{U}\right) = \delta \left[f' \circ f^{-1}\right]\left(\bar{U}\right) \cdot \left(\left[\varphi \circ u_1\right](y, c) - f^{-1}\left(\bar{U}\right)\right) \ , \tag{3.10}$$

and the functions $\bar{U}$, R_x, R_y, and R_{xy} are given in (iv)*. The controls (c^*, π^*) realize the infimum in* (3.9)*.*

(iii) Diffusion equations: *For each fixed* $0 \leq s \leq T$ *the function* l_1 *solves the partial differential equation (PDE)*

$$\left.\begin{aligned} \partial_t l_1 (t,s,x,y) &= -\mathcal{A}^{c^*,\pi^*}(t,s,x,y)\, l_1(t,s,x,y) \ , \\ l_1(s,s,x,y) &= u_1(y, c^*(s,x,y)) \ , \end{aligned}\right\} \tag{3.11}$$

and the function l_2 *solves the PDE*

$$\left.\begin{aligned} \partial_t l_2 (t,T,x,y) &= -\mathcal{A}^{c^*,\pi^*}(t,T,x,y)\, l_2(t,T,x,y) \ , \\ l_2(T,T,x,y) &= u_2(y,x) \ . \end{aligned}\right\} \tag{3.12}$$

(iv) Remainder functions: *Omitting* x, y*-dependence, the function* $\bar{U}$ *is given by*

$$\bar{U}(t,x) = f\left(\begin{array}{c} \int_t^T \delta e^{-\delta(s-t)} \left[\varphi \circ l_1\right](t,s)\, \mathrm{d}s \\ +\omega\delta e^{-\delta(T-t)} \left[\varphi \circ l_2\right](t,T) \end{array} \right),$$

the remainder term R_x *is given by*

$$\begin{aligned} R_x(t) = &\left[f'' \circ f^{-1}\right]\left(\bar{U}(t)\right) \begin{pmatrix} \int_t^T \delta e^{-\delta(s-t)} \left[\varphi' \circ l_1\right](t,s) \partial_x l_1(t,s)\, \mathrm{d}s \\ +\omega\delta e^{-\delta(T-t)} \left[\varphi' \circ l_2\right](t,T) \partial_x l_2(t,T) \end{pmatrix}^2 \\ &+ \left[f' \circ f^{-1}\right]\left(\bar{U}(t)\right) \\ &\cdot \begin{pmatrix} \int_t^T \delta e^{-\delta(s-t)} \left[\varphi'' \circ l_1\right](t,s) \left(\partial_x l_1(t,s)\right)^2 \mathrm{d}s \\ +\omega\delta e^{-\delta(T-t)} \left[\varphi'' \circ l_2\right](t,T) \left(\partial_x l_2(t,T)\right)^2 \end{pmatrix} , \end{aligned}$$

and analogously for R_y*. The cross-term* R_{xy} *is defined as*

$$\begin{aligned} R_{xy}(t) = &\left[f'' \circ f^{-1}\right]\left(\bar{U}(t)\right) \begin{pmatrix} \int_t^T \delta e^{-\delta(s-t)} \left[\varphi' \circ l_1\right](t,s) \partial_x l_1(t,s)\, \mathrm{d}s \\ +\delta e^{-\delta(T-t)} \left[\varphi' \circ l_2\right](t,T) \partial_x l_2(t,T) \end{pmatrix} \\ &\cdot \begin{pmatrix} \int_t^T \delta e^{-\delta(s-t)} \left[\varphi' \circ l_1\right](t,s) \partial_y l_1(t,s)\, \mathrm{d}s \\ +\delta e^{-\delta(T-t)} \left[\varphi' \circ l_2\right](t,T) \partial_y l_2(t,T) \end{pmatrix} \\ &+ \left[f' \circ f^{-1}\right]\left(\bar{U}(t)\right) \\ &\cdot \begin{pmatrix} \int_t^T \delta e^{-\delta(s-t)} \left[\varphi'' \circ l_1\right](t,s) \partial_y l_1(t,s)\, \partial_x l_1(t,s)\, \mathrm{d}s \\ +\delta e^{-\delta(T-t)} \left[\varphi'' \circ l_2\right](t,T) \partial_y l_2(t,T)\, \partial_x l_2(t,T) \end{pmatrix} . \end{aligned}$$

Then the following holds:

(i) The controls (c^*, π^*) *are an equilibrium control for the function* $V^{c,\pi}$ *defined in* (3.6)*.*

(ii) The corresponding equilibrium value function V^{c^*,π^*} *is given by*

$$V^{c^*,\pi^*}(t,x,y) = \bar{U}(t,x,y) = U(t,x,y) \, .$$

(iii) For $0 \le t \le s \le T$*, we have*

$$m_1^{c^*,\pi^*}(t,s,x,y) = l_1(t,s,x,y) \ ,$$
$$m_2^{c^*,\pi^*}(t,T,x,y) = l_2(t,T,x,y) \ .$$

Remark 3.2. Some brief remarks on the pseudo-Bellman equation that allows a comparison with the standard examples.

(i) It may appear unusual to have the $\bar{U}$ as an auxiliary function in the aggregator and indeed one could state the theorem with $\bar{U}$ replaced by U. However, this would lead to a highly nonlinear PDE that would require a more sophisticated mathematical treatment which is beyond the scope of this paper.

(ii) In the pseudo-Bellman equation, the terms have the following meaning:

- The aggregator is given by $F\left(c, y, \bar{U}(t,x,y)\right)$.
- The market dynamics are represented by the operator $\mathcal{A}^{c,\pi}U(t,x,y)$.
- The twisting with f and φ introduces the remainder terms $R_x(t,x,y)$, $R_y(t,x,y)$, and $R_{xy}(t,x,y)$.

(iii) The aggregator can be decomposed in two terms:

- The first term $\delta\left[f' \circ f^{-1}\right]\left(\bar{U}(t,x,y)\right)\left[\varphi \circ u_1\right](y,c)$ is multiplicative in the form: function of $\bar{U}$ times function of c. The twisting yields to a additive perturbation of the standard Bellman equation with explicit remainder terms R_x, R_y, R_{xy}.
- The second term $-\delta\left[f' \circ f^{-1}\right]\left(\bar{U}(t,x,y)\right) f^{-1}\left(\bar{U}(t,x,y)\right)$ is due to the discounting $\delta e^{-\delta(s-t)}$.

(iv) From the above analysis we can expect the same structure of the Bellman equation in higher dimensions, i.e. if we add further diffusion processes.

3.4 Examples for Incomplete Markets

In this section, we present two non-trivial examples of the framework in incomplete markets. We consider an investor making decisions concerning consumption and investment in the incomplete market model formalized by

$$\begin{aligned}
\mathrm{d}B_t &= rB_t\,\mathrm{d}t \ , \\
\mathrm{d}S_t &= S_t\left[(r + \lambda(t,Y_t))\,\mathrm{d}t + \sigma_S(t,Y_t)\,\mathrm{d}W_t\right] \ , \\
\mathrm{d}Y_t &= \mu_Y(t,Y_t)\,\mathrm{d}t + \sigma_Y(t,Y_t)\left(\rho\,\mathrm{d}W_t + \sqrt{1-\rho^2}\,\mathrm{d}\bar{W}_t\right) \ ,
\end{aligned}$$

where $\lambda, \sigma_S, \mu_Y, \sigma_Y$ are regular functions of (t, Y_t), and W and $\bar{W}$ are two independent Brownian motions. The processes B and S represent price processes of a traded bond and stock whereas Y is an additional non-traded state process driving the coefficients of S. The parameter ρ models the level of correlation between S and Y.

We consider an investor investing the proportion π of his wealth in the stock S and the proportion $(1-\pi)$ in the bond B and consuming at rate c. The wealth of the investor evolves according to the dynamics

$$\begin{aligned} \mathrm{d}X_t^{c,\pi} = X_t^{c,\pi} \left(r + \pi(t, X_t^{c,\pi}, Y_t)\lambda(t, Y_t) \right) \mathrm{d}t - c(t, X_t^{c,\pi}, Y_t)\, \mathrm{d}t \\ + X_t^{c,\pi} \pi(t, X_t^{c,\pi}, Y_t) \sigma_S(t, Y_t)\, \mathrm{d}W_t \ . \end{aligned} \tag{3.13}$$

In the notation from the previous section, we have

$$\begin{aligned} \alpha(t,y) &= \mu_Y(t,y) \ , \\ \beta(t,y) &= \sigma_Y(t,y) \ , \\ \mu^{c,\pi}(t,x,y) &= x \left(r + \pi(t,x,y)\lambda(t,y) \right) - c(t,x,y) \ , \\ \sigma^{c,\pi}(t,x,y) &= x\pi(t,x,y)\sigma_S(t,y) \ . \end{aligned}$$

Consumption-investment problems with wealth dynamics given by (3.13) have been studied with various specifications of $\lambda, \sigma_S, \mu_Y, \sigma_Y$, and ρ by a number of authors. They include Wachther (2002) who works in a complete market setting with constant asset volatility and stochastic excess return linear in Y which is modeled by an Ornstein-Uhlenbeck process, i.e. μ_Y affine in Y and σ_Y constant; Chacko and Viceira (2005) who works with constant excess return and stochastic volatility inverse in the square-root of Y which is modelled by a certain square root process, such that μ_Y is affine in Y and σ_Y is linear in its square-root; Liu (2007) who works with an excess return linear in Y and stochastic volatility equal to the square-root of Y which is modeled by a square root process similar to the one used by Chacko and Viceira (2005). Kraft et al. (2013) consider the model in its generality. In order to avoid complicating issues in connection with stochastic interest rates, see Korn and Kraft (2002), it is important that Y governs the coefficients of S only and not the interest rate. Musiela and Zariphopoulou (2010) refer to the model as a Markovian single stochastic factor model.

Below, we consider the optimization problem in Section 3.2 for two different choices of the utility functions, namely power utility and exponential utility. This leads to certain constraints on the coefficient functions $\lambda, \sigma_S, \mu_Y, \sigma_Y$.

3.4.1 Power Utility

To state our result we need the following assumptions on the coefficients driving the financial market dynamics.

Hypothesis 3.2.

(i) Utility functions: let $u_1(y,\xi) = u_2(y,\xi) = y^{\kappa(1-\gamma)}\xi^{1-\gamma}$ for fixed $\kappa \in \mathbb{R}$ and $\gamma \in \mathbb{R}^+, \gamma \neq 1$.

(ii) Elasticity of Inter-temporal Substitution and twisting: let $\varphi(\xi) = \xi^{\frac{1}{\theta}}$ and $f(\xi) = \xi^\theta$ where $\theta = \frac{1-\gamma}{1-\phi}$ for fixed $\phi \in \mathbb{R}^+, \phi \neq 1$.

(iii) Market price of risk: there is a function $h : [0,T] \to \mathbb{R}$ such that $\lambda(t,y) = h(t)\sigma_S(t,y)$.

(iv) Dynamics of Y: there are functions $\alpha, \beta : [0,T] \to \mathbb{R}$ such that $\mu_Y(t,y) = \alpha(t)y$ and $\sigma_Y(t,y) = \beta(t)y$.

To avoid discussions about regularity we assume that all functions are smooth in their arguments. These assumptions ensure that the Verification Theorem is satisfied.

Theorem 3.2 (Equilibrium controls). *Define the function $g : [0,T] \to \mathbb{R}$ by*

$$g(t) = \delta^\theta \left(\int_t^T e^{-\int_t^s \frac{1}{\theta\phi}\tau(v)\,\mathrm{d}v}\,\mathrm{d}s + e^{-\int_t^T \frac{1}{\theta\phi}\tau(v)\,\mathrm{d}v}\omega^{\frac{1}{\phi}} \right)^{\phi\theta} ,$$

where

$$\begin{aligned}\tau(t) = \delta\theta - (1-\gamma)\left[r + \alpha(t)\kappa + \tfrac{1}{2}(\beta(t))^2\kappa((1-\gamma)\kappa - 1)\right] \\ + \left(1 - \frac{1}{\gamma}\right)\left[\rho(1-\gamma)h(t)\beta(t)\kappa + \tfrac{1}{2}h(t)^2 + \tfrac{1}{2}\rho^2\beta(t)^2(1-\gamma)^2\kappa^2\right] .\end{aligned}$$

Also, define the functions $\eta_1, \eta_2 : [0,T]^2 \to \mathbb{R}$ by

$$\begin{aligned}\eta_1(t,s) &= \delta^{\frac{1-\gamma}{\phi}} g(s)^{-\frac{1-\gamma}{\theta\phi}} e^{-\int_t^s \psi(v)\,\mathrm{d}v} , \\ \eta_2(t,T) &= e^{-\int_t^T \psi(v)\,\mathrm{d}v} ,\end{aligned}$$

where

$$\psi(t) = \tau(t) - \delta\theta + (1-\gamma)g(t)^{-\frac{1}{\phi\theta}}\delta^{\frac{1}{\phi}} .$$

Finally, set

$$\begin{aligned}c^*(t,x) &= \delta^{\frac{1}{\phi}} g(t)^{-\frac{1}{\theta\phi}} x , \\ \pi^*(t,y) &= \frac{h(t) + \rho\beta(t)(1-\gamma)\kappa}{\sigma_S(t,y)\gamma} , \\ l_1(t,s,x,y) &= \eta_1(t,s)x^{1-\gamma}y^{\kappa(1-\gamma)} , \\ l_2(t,T,x,y) &= \eta_2(t,T)x^{1-\gamma}y^{\kappa(1-\gamma)} , \\ U(t,x,y) &= x^{1-\gamma}y^{\kappa(1-\gamma)}g(t) ,\end{aligned}$$

where we note that c^* is independent of y and π^* is independent of x. Then Theorem 3.1 is satisfied for $(U, l_1, l_2, c^*, \pi^*)$, i.e. the controls (c^*, π^*) are equilibrium controls and U is the corresponding value function.

We note that the equilibrium controls (c^*, π^*) are admissible, in particular Hypothesis 3.1 is satisfied. The proof is in Appendix 3.B.

We noted in Theorem 3.2 that the equilibrium consumption rate becomes independent of Y. This is due to the specific role and form of Y formalized in Hypothesis 3.2. If we e.g. think of Y as a consumer price index and take $\kappa = -1$, we measure utility of goods bought rather than money spent. In that case, and with homogeneity in the sense of (iii) and (iv), it is reasonable to achieve that the equilibrium consumption is not a function of prices but of wealth (and time) only. Consumption is just a matter of spreading the spending of wealth over time, while hedging away prices changes to the extent possible, and then taking the non-hedgeable part as it comes.

Lemma 3.1. *Under the above assumptions, the dynamics of the equilibrium consumption can be expressed in terms of the SDE*

$$\begin{aligned}\mathrm{d}c^*\left(t, X_t^{c^*,\pi^*}\right) &= c^*\left(t, X_t^{c^*,\pi^*}\right)\left(r + \tfrac{1}{\gamma}(h(t) + \rho(1-\gamma)\kappa\beta(t))h(t) - \tfrac{\tau(t)}{\theta\phi}\right)\mathrm{d}t \\ &\quad + c^*\left(t, X_t^{c^*,\pi^*}\right)\tfrac{1}{\gamma}\left(h(t) + \rho(1-\gamma)\kappa\beta(t)\right)\mathrm{d}W_t\end{aligned}$$

with τ as above.

The proof is in Appendix 3.B.

Remark 3.3. In the introduction we announced a coincidence with the optimal control arising from recursive utility with Epstein-Zin preferences in a Merton market. This coincidence is obtained by setting $\kappa = 0$ and letting λ and σ_S be independent of y such that the state process Y is taken out of the problem. In this case, the equilibrium consumption and investment specified above are exactly those obtained from recursive utility with Epstein-Zin preferences in the same market. It may not be immediately recognizable from the expressions above. Actually, it is easier to recognize from the pseudo-Bellman equation characterizing the solution. The Bellman equation is specialized to power utility in Appendix 3.B. With power utility and $\kappa = 0$, the aggregator in (3.10) becomes

$$F(c, U) = \delta\theta U\left(\left(\frac{c}{U^{\frac{1}{1-\gamma}}}\right)^{1-\phi} - 1\right). \tag{3.14}$$

This is immediately recognized as the Epstein-Zin normalized aggregator, see e.g. Kraft et al. (2013). We highlight this from Appendix 3.B because of the special role of the coincidence. Since further, in this case, the remainder functions R_x, R_y, and R_{xy} become zero (see Appendix 3.B), the pseudo-Bellman equation in (3.9) characterizing the equilibrium value function coincides with the Bellman equation characterizing the value function for recursive utility with Epstein-Zin preferences. Therefore the resulting controls also coincide.

Actually, even including the state process Y in the case where the market price of risk is independent of Y ((iii) and (iv) in Hypothesis 3.2) is a special case that has been commented on by others. Kraft et al. (2013) realize that this is a mathematically tractable case but pay little attention to it. Note however that they only consider the case corresponding to $\kappa = 0$. We think that the case certainly deserves attention, in particular since $\kappa \neq 0$ gives a meaningful interpretation of Y as a non-hedgeable consumer price index.

Note from Lemma 3.1 that the equilibrium consumption rate forms a time-inhomogenous geometric Brownian motion. This form, well-known to arise from recursive utility with Epstein-Zin preferences, in general, and from expected time-additive power utility, in particular, is thus kept under the specifications studied in this section.

3.4.2 Exponential Utility

We define a one-parameter family of functions

$$\varepsilon(t) = \frac{r}{1 + Ce^{rt}} \ ,$$

where C is a constant of our choice. Typically, we choose $C = (r-1)e^{-rT}$ such that $\varepsilon(T) = 1$ and the two utility function below coincide.

To state the result we need the following assumptions on the coefficients driving the financial market dynamics.

Hypothesis 3.3.

(i) Utility functions: let $u_1(y,\xi) = \exp(\gamma\kappa y + \gamma\xi)$ and $u_2(y,\xi) = \exp(\gamma\kappa y + \gamma\varepsilon(T)\xi)$ for fixed $\kappa \in \mathbb{R}$ and $\gamma \in \mathbb{R}^-$.

(ii) Elasticity of Inter-temporal Substitution and twisting: let $\varphi(\xi) = \xi^{\frac{1}{\theta}}$ and $f(\xi) = \xi^\theta$.

(iii) Market price of risk: there is $h : [0,T] \to \mathbb{R}$ with $\lambda(t,y) = h(t)\sigma_S(t,y)$.

(iv) Dynamics of Y: there are $\alpha, \beta : [0,T] \to \mathbb{R}$ with $\mu_Y(t,y) = \alpha(t)$ and $\sigma_Y(t,y) = \beta(t)$.

For simplicity we assume that all functions are smooth. These assumptions ensure that the Verification Theorem is satisfied.

Theorem 3.3 (Equilibrium controls)**.** *Define the function $g : [0,T] \to \mathbb{R}$ by*

$$g(t) = \exp\left(e^{-\int_t^T \varepsilon(v)\,\mathrm{d}v} \theta \log(\delta\omega) - \int_t^T e^{-\int_t^s \varepsilon(v)\,\mathrm{d}v} \tau(s)\,\mathrm{d}s \right) \ ,$$

where

$$\tau(t) = \theta\left(\delta - \varepsilon(t)\right) + \theta\varepsilon(t)\log\left(\frac{\varepsilon(t)}{\delta}\right) + \frac{1}{2}h(t)^2 + \frac{1}{2}\rho^2\beta\left(t\right)^2\gamma^2\kappa^2 \\ - \alpha\left(t\right)\gamma\kappa - \frac{1}{2}\left(\beta\left(t\right)\right)^2\gamma^2\kappa^2 + \rho\beta\left(t\right)\gamma\kappa h(t)\ .$$

Also, define the functions $\eta_1, \eta_2 : [0,T]^2 \to \mathbb{R}$ *by*

$$\eta_1(t,s) = \left(\frac{\varepsilon(s)}{\delta}\right)^{\theta} g\left(s\right) e^{-\int_t^s \psi(v)\,\mathrm{d}v}\ , \\ \eta_2(t,T) = e^{-\int_t^T \psi(v)\,\mathrm{d}v}\ ,$$

where

$$\psi(t) = \tau(t) + \varepsilon(t)\log\left(g\left(t\right)\right) - \theta\left(\delta - \varepsilon(t)\right)\ .$$

Finally, set

$$c^*(t,x) = \varepsilon(t)x + \frac{\theta}{\gamma}\log\left(\frac{\varepsilon(t)}{\delta}\right) + \frac{1}{\gamma}\log\left(g\left(t\right)\right)\ , \\ \pi^*(t,x,y) = -\frac{1}{x}\frac{h(t) + \rho\beta\left(t\right)\gamma\kappa}{\sigma_S(t,y)\varepsilon(t)\gamma}\ , \\ l_1(t,s,x,y) = \eta_1(t,s)\exp(\gamma\kappa y + \gamma\varepsilon(t)x)\ , \\ l_2(t,T,x,y) = \eta_2(t,T)\exp(\gamma\kappa y + \gamma\varepsilon(T)x)\ , \\ U(t,x,y) = g\left(t\right)\exp(\gamma\kappa y + \gamma\varepsilon(t)x)\ ,$$

where we note that c^* *is independent of* y*, and* $x\pi^*$ *is independent of* x*. Then Theorem 3.1 is satisfied for* $(U, l_1, l_2, c^*, \pi^*)$*, i.e. the controls* (c^*, π^*) *are equilibrium controls and* U *is the corresponding value function.*

We note that the equilibrium controls (c^*, π^*) are admissible, in particular Hypothesis 3.1 is satisfied. The proof is in Appendix 3.C.

Lemma 3.2. *Under the above assumptions, the dynamics of the equilibrium consumption can be expressed in term of the SDE*

$$\mathrm{d}c^*\left(t, X_t^{c^*,\pi^*}\right) = \begin{pmatrix} \frac{\theta}{\gamma}(\varepsilon(t) - r) + \frac{1}{\gamma}\tau(t) - \frac{\theta}{\gamma}\varepsilon(t)\log\left(\frac{\varepsilon(t)}{\delta}\right) \\ -\frac{1}{\gamma}(h(t) + \rho\kappa\gamma\beta(t))h(t) \end{pmatrix}\mathrm{d}t \\ - \frac{1}{\gamma}\left(h(t) + \rho\kappa\gamma\beta(t)\right)\mathrm{d}W_t$$

with τ *and* ε *as above.*

The proof is in Appendix 3.C.

Remark 3.4. As for power utility, we obtain a coincidence with known preferences from recursive utility by setting $\kappa = 0$ and making sure that (iii) and (iv) in Hypothesis 3.3 are satisfied such that the remainder functions R_x, R_y, and R_{xy} become zero (for details, see Appendix 3.C). With exponential utility and $\kappa = 0$, the aggregator in (3.10) specializes to

$$F(c, U) = \delta\theta U \left(\frac{\exp\left(\frac{\gamma}{\theta} c\right)}{U^{\frac{1}{\theta}}} - 1 \right) .$$

Up to a constant, this coincides with the aggregator arising from the specification $u(c) = \frac{1}{\gamma}\exp(\gamma c)$ and $g(c) = \frac{\theta}{\gamma}\exp(\frac{\gamma}{\theta} c)$ in Section 6 of Kraft and Seifried (2014) which yields the aggregator

$$f(c, U) = \delta\theta U \left(\frac{\exp(\frac{\gamma}{\theta} c)}{(\gamma U)^{\frac{1}{\theta}}} - 1 \right) .$$

We highlight this because it shows that the coincidence with recursive utility goes beyond power utility and Epstein-Zin preferences.

Note from Lemma 3.2 that the equilibrium consumption rate forms a time-inhomogenous Brownian motion with drift. This is what could be expected from the wealth-non-memorability feature of the exponential utility function. From a decision-making point of view, the exponential utility function thereby, again, proves to be a questionable specification of preferences. However, its usability in indifference pricing is still reason enough to show all the results in detail here, parallel with power utility.

Appendix

3.A Proof of Verification Theorem

The proof of the verification theorem is described in detail in five lemmas.

Proof of Theorem 3.1. We prove the assertions in reverse order.

Assertion (iii) that $m_i^{c^*,\pi^*} = l_i$ for $i = 1, 2$ is in Lemma 3.3.

Assertion (ii) on the characterization of the Value function is split into Corollary 3.1 which says that $V^*(t, x, y) = \bar{U}(t, x, y)$ and Lemma 3.4 giving $V^*(t, x, y) = U(t, x, y)$.

Finally, assertion (i) that (c^*, π^*) are equilibrium controls is proved in Lemma 3.5 □

The first lemma characterizes the functions $m_i^{c^*,\pi^*}$ as PDE solutions.

Lemma 3.3. *Under the assumptions of Theorem 3.1 it holds that*

$$m_1^{c^*,\pi^*}(t,s,x,y) = l_1(t,s,x,y) \ ,$$
$$m_2^{c^*,\pi^*}(t,T,x,y) = l_2(t,T,x,y)$$

for any $0 \le t \le s \le T$.

Proof. Fix an admissible control (c,π). By Definition 3.1 there exist functions $\Lambda_1^{c,\pi}(t,s,x,y)$ and $\Lambda_2^{c,\pi}(t,T,x,y)$, defined for $0 \le t \le s \le T$ and $x,y \in \mathbb{R}$, such that

- $\Lambda_1^{c,\pi}, \Lambda_2^{c,\pi}$ are C^1 in t and C^2 in x,y, and $\Lambda_1^{c,\pi}$ is jointly continuous in t,s.
- For each fixed $0 \le s \le T$, $\Lambda_1^{c,\pi}$ solves the PDE

$$\left.\begin{aligned} \partial_t \Lambda_1^{c,\pi}(t,s,x,y) &= -\mathcal{A}^{c,\pi}\Lambda_1^{c,\pi}(t,s,x,y) \ , \\ \Lambda_1^{c,\pi}(s,s,x,y) &= u_1(y, c(s,x,y)) \ . \end{aligned}\right\} \tag{3.15}$$

- $\Lambda_2^{c,\pi}$ solves the PDE

$$\left.\begin{aligned} \partial_t \Lambda_2^{c,\pi}(t,T,x,y) &= -\mathcal{A}^{c,\pi}\Lambda_2^{c,\pi}(t,T,x,y) \ , \\ \Lambda_2^{c,\pi}(T,T,x,y) &= u_2(y,x) \ . \end{aligned}\right\} \tag{3.16}$$

By the classical Feynman–Kac theorem we have

$$\Lambda_1^{c,\pi}(t,s,x,y) = \mathbb{E}_{t,x,y}\left[u_1\left(Y_s, c\left(s, X_s^{c,\pi}, Y_s\right)\right)\right] = m_1^{c,\pi}(t,s,x,y)\ ,\ t \le s < T,$$

and

$$\Lambda_2^{c,\pi}(t,T,x,y) = \mathbb{E}_{t,x,y}\left[u_2\left(Y_T, X_T^{c,\pi}\right)\right] = m_2^{c,\pi}(t,T,x,y)\ ,\ s < T.$$

Since solutions of the PDEs are unique, we have $\Lambda_i^{c^*,\pi^*} = l_i$ for $i = 1,2$. □

We observe an immediate consequence of the proof.

Corollary 3.1. *The value function* $V^{c,\pi}$ *can be written as*

$$V^{c,\pi}(t,x,y) = f\begin{pmatrix} \int_t^T \delta e^{-\delta(s-t)}\left[\varphi \circ \Lambda_1^{c,\pi}\right](t,s,x,y)\,\mathrm{d}s \\ +\omega\delta e^{-\delta(T-t)}\left[\varphi \circ \Lambda_2^{c,\pi}\right](t,T,x,y) \end{pmatrix}, \tag{3.17}$$

and, in particular, the equilibrium value function satisfies $V^{c^*,\pi^*} = \bar{U}$.

Lemma 3.4. *The equilibrium value function satisfies* $V^{c^*,\pi^*} = U$.

Proof. Since $\Lambda_1^{c,\pi}, \Lambda_2^{c,\pi}$ are C^1 in t and C^2 in x, y, $\Lambda_1^{c,\pi}$ is jointly continuous in t, s, and φ is in C^2, we get from the representation in (3.17) that $V^{c,\pi}$ is in $C^{1,2,2}$. Suppressing x, y-dependence in the Λ-functions, we obtain the partial derivatives (for $i = x, y$)

$$\begin{aligned}
&\partial_t V^{c,\pi}(t,x,y) \\
&= -\delta \left[f' \circ f^{-1}\right] \left(V^{c,\pi}(t,x,y)\right) \left(\left[\varphi \circ u_1\right] (y, c(t,x,y)) - f^{-1}\left(V^{c,\pi}(t,x,y)\right) \right) \\
&\quad + \left[f' \circ f^{-1}\right] \left(V^{c,\pi}(t,x,y)\right) \begin{pmatrix} \int_t^T \delta e^{-\delta(s-t)} \left[\varphi' \circ \Lambda_1^{c,\pi}\right](t,s)\, \partial_t \Lambda_1^{c,\pi}(t,s)\, \mathrm{d}s \\ +\omega \delta e^{-\delta(T-t)} \left[\varphi' \circ \Lambda_2^{c,\pi}\right](t,T)\, \partial_t \Lambda_2^{c,\pi}(t,T) \end{pmatrix},
\end{aligned}$$

$$\begin{aligned}
&\partial_i V^{c,\pi}(t,x,y) \\
&= \left[f' \circ f^{-1}\right] \left(V^{c,\pi}(t,x,y)\right) \begin{pmatrix} \int_t^T \delta e^{-\delta(s-t)} \left[\varphi' \circ \Lambda_1^{c,\pi}\right](t,s)\, \partial_i \Lambda_1^{c,\pi}(t,s)\, \mathrm{d}s \\ +\omega \delta e^{-\delta(T-t)} \left[\varphi' \circ \Lambda_2^{c,\pi}\right](t,T)\, \partial_i \Lambda_2^{c,\pi}(t,T) \end{pmatrix},
\end{aligned}$$

$$\begin{aligned}
&\partial_i^2 V^{c,\pi}(t,x,y) \\
&= \left[f'' \circ f^{-1}\right] \left(V^{c,\pi}(t,x,y)\right) \begin{pmatrix} \int_t^T \delta e^{-\delta(s-t)} \left[\varphi' \circ \Lambda_1^{c,\pi}\right](t,s)\, \partial_i \Lambda_1^{c,\pi}(t,s)\, \mathrm{d}s \\ +\omega \delta e^{-\delta(T-t)} \left[\varphi' \circ \Lambda_2^{c,\pi}\right](t,T)\, \partial_i \Lambda_2^{c,\pi}(t,T) \end{pmatrix}^2 \\
&\quad + \left[f' \circ f^{-1}\right] \left(V^{c,\pi}(t,x,y)\right) \begin{pmatrix} \int_t^T \delta e^{-\delta(s-t)} \left[\varphi'' \circ \Lambda_1^{c,\pi}\right](t,s) \left(\partial_i \Lambda_1^{c,\pi}(t,s)\right)^2 \mathrm{d}s \\ +\omega \delta e^{-\delta(T-t)} \left[\varphi'' \circ \Lambda_2^{c,\pi}\right](t,T) \left(\partial_i \Lambda_2^{c,\pi}(t,T)\right)^2 \end{pmatrix} \\
&\quad + \left[f' \circ f^{-1}\right] \left(V^{c,\pi}(t,x,y)\right) \begin{pmatrix} \int_t^T \delta e^{-\delta(s-t)} \left[\varphi' \circ \Lambda_1^{c,\pi}\right](t,s)\, \partial_i^2 \Lambda_1^{c,\pi}(t,s)\, \mathrm{d}s \\ +\omega \delta e^{-\delta(T-t)} \left[\varphi' \circ \Lambda_2^{c,\pi}\right](t,T)\, \partial_i^2 \Lambda_2^{c,\pi}(t,T) \end{pmatrix},
\end{aligned}$$

$$\begin{aligned}
&\partial_{xy} V^{c,\pi}(t,x,y) \\
&= \left[f' \circ f^{-1}\right] \left(V^{c,\pi}(t,x,y)\right) \\
&\qquad \cdot \begin{pmatrix} \int_t^T \delta e^{-\delta(s-t)} \left[\varphi'' \circ \Lambda_1^{c,\pi}\right](t,s)\, \partial_x \Lambda_1^{c,\pi}(t,s)\, \partial_y \Lambda_1^{c,\pi}(t,s)\, \mathrm{d}s \\ +\omega \delta e^{-\delta(T-t)} \left[\varphi'' \circ \Lambda_2^{c,\pi}\right](t,T)\, \partial_x \Lambda_2^{c,\pi}(t,T)\, \partial_y \Lambda_2^{c,\pi}(t,T) \end{pmatrix} \\
&\quad + \left[f'' \circ f^{-1}\right] \left(V^{c,\pi}(t,x,y)\right) \\
&\qquad \cdot \begin{pmatrix} \int_t^T \delta e^{-\delta(s-t)} \left[\varphi' \circ \Lambda_1^{c,\pi}\right](t,s)\, \partial_x \Lambda_1^{c,\pi}(t,s)\, \mathrm{d}s \\ +\omega \delta e^{-\delta(T-t)} \left[\varphi' \circ \Lambda_2^{c,\pi}\right](t,T)\, \partial_x \Lambda_2^{c,\pi}(t,T) \end{pmatrix} \\
&\qquad \cdot \begin{pmatrix} \int_t^T \delta e^{-\delta(s-t)} \left[\varphi' \circ \Lambda_1^{c,\pi}\right](t,s)\, \partial_y \Lambda_1^{c,\pi}(t,s)\, \mathrm{d}s \\ +\omega \delta e^{-\delta(T-t)} \left[\varphi' \circ \Lambda_2^{c,\pi}\right](t,T)\, \partial_y \Lambda_2^{c,\pi}(t,T) \end{pmatrix} \\
&\quad + \left[f' \circ f^{-1}\right] \left(V^{c,\pi}(t,x,y)\right) \\
&\qquad \cdot \begin{pmatrix} \int_t^T \delta e^{-\delta(s-t)} \left[\varphi' \circ \Lambda_1^{c,\pi}\right](t,s)\, \partial_{xy} \Lambda_1^{c,\pi}(t,s)\, \mathrm{d}s \\ +\omega \delta e^{-\delta(T-t)} \left[\varphi' \circ \Lambda_2^{c,\pi}\right](t,T)\, \partial_{xy} \Lambda_2^{c,\pi}(t,T) \end{pmatrix}.
\end{aligned}$$

Applying (3.15)–(3.16), we obtain the PDE

$$\left.\begin{aligned} \partial_t V^{c,\pi}(t,x,y) &= -F\left(c(t,x,y),y,V^{c,\pi}(t,x,y)\right) - \mathcal{A}^{c,\pi}V^{c,\pi}(t,x,y) \\ &\quad + \tfrac{1}{2}\left(\sigma^{c,\pi}(t,x,y)\right)^2 R_x^{c,\pi}(t,x,y) \\ &\quad + \tfrac{1}{2}\left(\beta(t,y)\right)^2 R_y^{c,\pi}(t,x,y) \\ &\quad + \rho\beta(t,y)\,\sigma^{c,\pi}(t,x,y)\,R_{xy}^{c,\pi}(t,x,y)\ , \\ V^{c,\pi}(T,x,y) &= f\left(\omega\delta\left[\varphi\circ u_2\right](y,x)\right)\ , \end{aligned}\right\} \tag{3.18}$$

where F is given by (3.10) and (for $i = x, y$)

$$\begin{aligned} &R_i^{c,\pi}(t,x,y) \\ &= \left[f''\circ f^{-1}\right]\left(V^{c,\pi}(t,x,y)\right) \\ &\qquad \cdot \begin{pmatrix} \int_t^T \delta e^{-\delta(s-t)}\left[\varphi'\circ\Lambda_1^{c,\pi}\right](t,s)\,\partial_i\Lambda_1^{c,\pi}(t,s)\,\mathrm{d}s \\ +\omega\delta e^{-\delta(T-t)}\left[\varphi'\circ\Lambda_2^{c,\pi}\right](t,T)\,\partial_i\Lambda_2^{c,\pi}(t,T) \end{pmatrix}^2 \\ &\quad + \left[f'\circ f^{-1}\right]\left(V^{c,\pi}(t,x,y)\right) \\ &\qquad \cdot \begin{pmatrix} \int_t^T \delta e^{-\delta(s-t)}\left[\varphi''\circ\Lambda_1^{c,\pi}\right](t,s)\left(\partial_i\Lambda_1^{c,\pi}(t,s)\right)^2\mathrm{d}s \\ +\omega\delta e^{-\delta(T-t)}\left[\varphi''\circ\Lambda_2^{c,\pi}\right](t,T)\left(\partial_i\Lambda_2^{c,\pi}(t,T)\right)^2 \end{pmatrix}\ , \end{aligned}$$

$$\begin{aligned} &R_{xy}^{c,\pi}(t,x,y) \\ &= \left[f'\circ f^{-1}\right]\left(V^{c,\pi}(t,x,y)\right) \\ &\qquad \cdot \begin{pmatrix} \int_t^T \delta e^{-\delta(s-t)}\left[\varphi''\circ\Lambda_1^{c,\pi}\right](t,s)\,\partial_x\Lambda_1^{c,\pi}(t,s)\,\partial_y\Lambda_1^{c,\pi}(t,s)\,\mathrm{d}s \\ +\omega\delta e^{-\delta(T-t)}\left[\varphi''\circ\Lambda_2^{c,\pi}\right](t,T)\,\partial_x\Lambda_2^{c,\pi}(t,T)\,\partial_y\Lambda_2^{c,\pi}(t,T) \end{pmatrix} \\ &\quad + \left[f''\circ f^{-1}\right]\left(V^{c,\pi}(t,x,y)\right) \\ &\qquad \cdot \begin{pmatrix} \int_t^T \delta e^{-\delta(s-t)}\left[\varphi'\circ\Lambda_1^{c,\pi}\right](t,s)\,\partial_x\Lambda_1^{c,\pi}(t,s)\,\mathrm{d}s \\ +\omega\delta e^{-\delta(T-t)}\left[\varphi'\circ\Lambda_2^{c,\pi}\right](t,T)\,\partial_x\Lambda_2^{c,\pi}(t,T) \end{pmatrix} \\ &\qquad \cdot \begin{pmatrix} \int_t^T \delta e^{-\delta(s-t)}\left[\varphi'\circ\Lambda_1^{c,\pi}\right](t,s)\,\partial_y\Lambda_1^{c,\pi}(t,s)\,\mathrm{d}s \\ +\omega\delta e^{-\delta(T-t)}\left[\varphi'\circ\Lambda_2^{c,\pi}\right](t,T)\,\partial_y\Lambda_2^{c,\pi}(t,T) \end{pmatrix}\ . \end{aligned}$$

To establish the relation $U = V^{c^*,\pi^*}$, we recall that $\Lambda_i^{c^*,\pi^*} = l_i$, $i = 1, 2$, and $V^{c^*,\pi^*} = \bar{U}$. This implies that $R_x^{c^*,\pi^*} = R_x$, $R_y^{c^*,\pi^*} = R_y$, and $R_{xy}^{c^*,\pi^*} = R_{xy}$, where R_x, R_Y, and R_{xy} are given in the theorem. Thus, V^{c^*,π^*} solves the PDE

$$\left.\begin{aligned} \partial_t V^{c^*,\pi^*}(t,x,y) &= -F\left(c^*(t,x,y),y,\bar{U}(t,x,y)\right) - \mathcal{A}^{c^*,\pi^*}V^{c^*,\pi^*}(t,x,y) \\ &\quad + \tfrac{1}{2}\sigma^{c^*,\pi^*}(t,x,y)^2 R_x(t,x,y) + \tfrac{1}{2}\beta(t,y)^2 R_y(t,x,y) \\ &\quad + \rho\beta(t,y)\,\sigma^{c^*,\pi^*}(t,x,y)\,R_{xy}(t,x,y)\ , \\ V^{c^*,\pi^*}(T,x,y) &= f\left(\omega\delta\left[\varphi\circ u_2\right](y,x)\right)\ . \end{aligned}\right\}$$

From the Bellman equation, we know that U solves the PDE

$$\left.\begin{aligned} \partial_t U(t,x,y) &= -F\left(c^*(t,x,y),y,\bar{U}(t,x,y)\right) - \mathcal{A}^{c^*,\pi^*}U(t,x,y) \\ &\quad + \frac{1}{2}\sigma^{c^*,\pi^*}(t,x,y)^2 R_x(t,x,y) + \frac{1}{2}\beta(t,y)^2 R_y(t,x,y) \\ &\quad + \rho\beta(t,y)\sigma^{c^*,\pi^*}(t,x,y)R_{xy}(t,x,y)\ , \\ U(T,x,y) &= f\left(\omega\delta\left[\varphi\circ u_2\right](y,x)\right)\ . \end{aligned}\right\}$$

Altogether, the difference $U - V^{c^*,\pi^*}$ solves the PDE

$$\left.\begin{aligned} \partial_t\left(U - V^{c^*,\pi^*}\right)(t,x,y) &= -\mathcal{A}^{c^*,\pi^*}\left(U - V^{c^*,\pi^*}\right)(t,x,y)\ , \\ \left(U - V^{c^*,\pi^*}\right)(T,x,y) &= 0\ . \end{aligned}\right\}$$

Hence, we must have $U = V^{c^*,\pi^*}$. □

Finally, we show that the controls c^* and π^* are indeed equilibrium controls.

Lemma 3.5. *The pair (c^*,π^*) is an equilibrium control.*

Proof. We fix an admissible control $(\bar{c},\bar{\pi})$, a (small) real number $h > 0$, and an initial point $(u,x,y) \in [0,T]\times\mathbb{R}^2$. We then define the control $\left(c^h,\pi^h\right)$ by

$$\left(c^h,\pi^h\right)(t,\bar{x},\bar{y}) = \begin{cases} (\bar{c},\bar{\pi})(t,\bar{x},\bar{y})\ , & u \le t < u+h, \bar{x},\bar{y}\in\mathbb{R}\ , \\ (c^*,\pi^*)(t,\bar{x},\bar{y})\ , & u+h \le t \le T, \bar{x},\bar{y}\in\mathbb{R}\ . \end{cases}$$

Below, we write $V^h = V^{c^h,\pi^h}$, $\Lambda_i^h = \Lambda_i^{c^h,\pi^h}$, and $X^h = X^{c^h,\pi^h}$.

To prove that (c^*,π^*) is an equilibrium control in the sense of Definition 3.2, we need to verify that condition (3.8) is satisfied. Recall that $V^{c^*,\pi^*} = U$. Hence, Equation (3.8) reads

$$\liminf_{h\to 0}\ \frac{U(u,x,y) - V^h(u,x,y)}{h} \ge 0\ .$$

By construction, we have $V^h(t,x,y) = U(t,x,y) = \bar{U}(t,x,y)$ for $t \ge u+h$. Thus, applying Taylor's formula for fixed x,y, we get that

$$\begin{aligned} &\frac{U(u,x,y) - V^h(u,x,y)}{h} \\ &= \frac{U(u,x,y) - U(u+h,x,y) - V^h(u,x,y) + V^h(u+h,x,y)}{h} \\ &= -U_t(u,x,y) + V_t^h(u,x,y) + o(h)\ . \end{aligned}$$

Hence, what we need to show is that

$$\liminf_{h\to 0}\ \left[-U_t(u,x,y) + V_t^h(u,x,y)\right] \ge 0\ .$$

By (3.18) and the Bellman equation, we have

$$\begin{aligned}
&- U_t(u,x,y) + V_t^h(u,x,y) \\
&\geq F\left(\bar{c}(u,x,y), y, \bar{U}(u,x,y)\right) - F\left(\bar{c}(u,x,y), y, V^h(u,x,y)\right) \\
&\quad + \mathcal{A}^{\bar{c},\bar{\pi}}\left(U(u,x,y) - V^h(u,x,y)\right) \\
&\quad + \tfrac{1}{2}\left(\bar{\sigma}(u,x,y)\right)^2\left(R_x^h(u,x,y) - R_x(u,x,y)\right) \\
&\quad + \tfrac{1}{2}\left(\beta(u,y)\right)^2\left(R_y^h(u,x,y) - R_y(u,x,y)\right) \\
&\quad + \rho\beta(u,y)\,\bar{\sigma}(u,x,y)\left(R_{xy}^h(u,x,y) - R_{xy}(u,x,y)\right) \ .
\end{aligned}$$

Hence, it suffices to show that for $h \to 0$

$$F\left(\bar{c}(u,x,y), y, V^h(u,x,y)\right) \to F\left(\bar{c}(u,x,y), y, \bar{U}(u,x,y)\right) \ , \tag{3.19}$$

$$\mathcal{A}^{\bar{c},\bar{\pi}} V^h(u,x,y) \to \mathcal{A}^{\bar{c},\bar{\pi}} U(u,x,y) \ , \tag{3.20}$$

$$R_x^h(u,x,y) \to R_x(u,x,y) \ , \tag{3.21}$$

$$R_y^h(u,x,y) \to R_y(u,x,y) \ , \tag{3.22}$$

$$R_{xy}^h(u,x,y) \to R_{xy}(u,x,y) \ . \tag{3.23}$$

Since V^h and $\bar{U} = U$ are continuous in the first argument, we note that

$$\begin{aligned}
V^h(u,x,y) &= V^h(u,x,y) - V^h(u+h,x,y) + \bar{U}(u+h,x,y) \\
&\to 0 + \bar{U}(u,x,y) \quad \text{as} \quad h \to 0 \ .
\end{aligned}$$

By assumption f and φ are in C^2 and f^{-1} is in C^0. Hence, F is continuous, and (3.19) follows immediately. Furthermore, since V^h and U are in $C^{1,2,2}$, we get (3.20):

$$\begin{aligned}
\mathcal{A}^{\bar{c},\bar{\pi}} V^h(u,x,y) &= \mathcal{A}^{\bar{c},\bar{\pi}} V^h(u,x,y) - \mathcal{A}^{\bar{c},\bar{\pi}} V^h(u+h,x,y) + \mathcal{A}^{\bar{c},\bar{\pi}} U(u+h,x,y) \\
&\to 0 + \mathcal{A}^{\bar{c},\bar{\pi}} U(u,x,y) \quad \text{as} \quad h \to 0 \ .
\end{aligned}$$

Finally, $f'' \circ f^{-1}$ is continuous, so for (3.21) to hold, it suffices to show that

$$\begin{aligned}
&\int_u^T \delta e^{-\delta(v-u)}\left[\varphi' \circ \Lambda_1^h\right](u,v,x,y)\,\partial_x \Lambda_1^h(u,v,x,y)\,\mathrm{d}v \\
&\quad + \omega\delta e^{-\delta(T-u)}\left[\varphi' \circ \Lambda_2^h\right](u,T,x,y)\,\partial_x \Lambda_2^h(u,T,x,y) \\
\to &\int_u^T \delta e^{-\delta(v-u)}\left[\varphi' \circ l_1\right](u,v,x,y)\,\partial_x l_1(u,v,x,y)\,\mathrm{d}v \\
&\quad + \omega\delta e^{-\delta(T-u)}\left[\varphi' \circ l_2\right](u,T,x,y)\,\partial_x l_2(u,T,x,y) \quad \text{as} \quad h \to 0
\end{aligned}$$

and

$$\begin{aligned}&\int_u^T \delta e^{-\delta(v-u)}\left[\varphi''\circ\Lambda_1^h\right](u,v,x,y)\left(\partial_x\Lambda_1^h(u,v,x,y)\right)^2\,\mathrm{d}v\\&\quad+\omega\delta e^{-\delta(T-u)}\left[\varphi''\circ\Lambda_2^h\right](u,T,x,y)\left(\partial_x\Lambda_2^h(u,T,x,y)\right)^2\\&\to\int_u^T \delta e^{-\delta(v-u)}\left[\varphi''\circ l_1\right](u,v,x,y)\left(\partial_x l_1(u,v,x,y)\right)^2\,\mathrm{d}v\\&\quad+\omega\delta e^{-\delta(T-u)}\left[\varphi''\circ l_2\right](u,T,x,y)\left(\partial_x l_2(u,T,x,y)\right)^2\quad\text{as}\quad h\to 0\ .\end{aligned}$$

This is ensured by the fact that φ is in C^2 and Λ_i^h and l_i are in $C^{1,2,2}$. To see this, realize that $\Lambda_i^h(t,s,x,y)=l_i(t,s,x,y)$ for $s\geq t\geq u+h$ by construction and write

$$\begin{aligned}&\int_u^T \delta e^{-\delta(v-u)}\left[\varphi'\circ\Lambda_1^h\right](u,v,x,y)\,\partial_x\Lambda_1^h(u,v,x,y)\,\mathrm{d}v\\&=\int_u^{u+h}\delta e^{-\delta(v-u)}\left[\varphi'\circ\Lambda_1^h\right](u,v,x,y)\,\partial_x\Lambda_1^h(u,v,x,y)\,\mathrm{d}v\\&\quad+\int_{u+h}^T\delta e^{-\delta(v-u)}\left[\varphi'\circ l_1\right](u+h,v,x,y)\,\partial_x l_1(u+h,v,x,y)\,\mathrm{d}v\\&\quad+\int_{u+h}^T\delta e^{-\delta(v-u)}\begin{pmatrix}\left[\varphi'\circ\Lambda_1^h\right](u,v,x,y)\,\partial_x\Lambda_1^h(u,v,x,y)\\-\left[\varphi'\circ\Lambda_1^h\right](u+h,v,x,y)\,\partial_x\Lambda_1^h(u+h,v,x,y)\end{pmatrix}\mathrm{d}v\\&\to 0+\int_u^T\delta e^{-\delta(v-u)}\left[\varphi'\circ l_1\right](u,v,x,y)\,\partial_x l_1(u,v,x,y)\,\mathrm{d}v+0\quad\text{as}\quad h\to 0\ .\end{aligned}$$

To complete the proof, we note that (3.22)–(3.23) follow from the same arguments as (3.21). □

3.B Details for the Power Utility Example

Proof of Theorem 3.2. We proceed in 3 steps.

1. Compute infinitesimal generator and aggregator. With the assumptions on the market dynamics, we get the infinitesimal generator

$$\begin{aligned}\mathcal{A}^{c,\pi}&=\left[x\left(r+\pi(t,x,y)\lambda(t,y)-c(t,x,y)\right]\partial_x+\tfrac{1}{2}\left(x\pi(t,x,y)\sigma_S(t,y)\right)^2\partial_x^2\\&\quad+\alpha(t,y)\,\partial_y+\tfrac{1}{2}\left(\beta(t,y)\right)^2\partial_y^2+\rho\alpha(t,y)\,x\pi(t,x,y)\sigma_S(t,y)\partial_{xy}\\&=\left[x\left(r+\pi(t,x,y)h(t)\sigma_S(t,y)\right)-c(t,x,y)\right]\partial_x\\&\quad+\tfrac{1}{2}\left(x\pi(t,x,y)\sigma_S(t,y)\right)^2\partial_x^2+\alpha(t)\,y\partial_y\\&\quad+\tfrac{1}{2}\left(\beta(t)\,y\right)^2\partial_y^2+\rho\beta(t)\,yx\pi(t,x,y)\sigma_S(t,y)\partial_{xy}\ .\end{aligned}$$

Also, with power utility we get the aggregator

$$F(c,y,U)=\delta\theta\left(U^{\frac{1}{\theta}}\right)^{\theta-1}\left(y^{\frac{\kappa(1-\gamma)}{\theta}}c^{\frac{1-\gamma}{\theta}}-U^{\frac{1}{\theta}}\right)=\delta\theta U\left(\left(\frac{y^\kappa c}{U^{\frac{1}{1-\gamma}}}\right)^{1-\phi}-1\right)\ .$$

Note that for $\kappa = 0$ we get the classical normalized aggregator in (3.14) arising in recursive utility with Epstein-Zin preferences, see also the comments in Remark 3.3.

2. Verify that (c^*, π^*) are equilibrium controls and U solves the Bellman equation. From Theorem 3.2, we have

$$l_i(t,s,x,y) = \eta_i(t,s) x^{1-\gamma} y^{\kappa(1-\gamma)} \ , \quad i = 1,2 \ , \tag{3.24}$$

so the remainder terms vanish, i.e. $R_x = R_y = R_{xy} = 0$. We also have

$$U(t,x,y) = x^{1-\gamma} y^{\kappa(1-\gamma)} g(t) \ .$$

Plugging this and $R_x = R_y = R_{xy} = 0$ into the Bellman equation that U must solve, i.e. Equation (3.9), and dividing by $x^{1-\gamma}y^{\kappa(1-\gamma)}$, we obtain

$$\left.\begin{aligned} g'(t) &= \inf_{c,\pi} \begin{bmatrix} -\delta\theta \left(g(t)\right)^{1-\frac{1}{\theta}} \left(\left(\frac{c}{x}\right)^{\frac{1-\gamma}{\theta}} - g(t)^{\frac{1}{\theta}} \right) \\ -\left((r + \pi h(t) \sigma_S(t,y)) - \frac{c}{x} \right) (1-\gamma) g(t) \\ -\frac{1}{2} \left(\pi\sigma_S(t,y)\right)^2 (1-\gamma)(-\gamma) g(t) \\ -\alpha(t)(1-\gamma)\kappa g(t) \\ -\frac{1}{2}\left(\beta(t)\right)^2 (1-\gamma)\kappa((1-\gamma)\kappa - 1) g(t) \\ -\rho\beta(t)\pi\sigma_S(t,y)(1-\gamma)^2\kappa g(t) \end{bmatrix} , \\ g(T) &= (\delta\omega)^\theta \ . \end{aligned}\right\} \tag{3.25}$$

The first order conditions for c, π read

$$\begin{aligned} 0 &= -\delta\theta g(t)^{1-\frac{1}{\theta}} \frac{1-\gamma}{\theta} c^{\frac{1-\gamma}{\theta}-1} \left(\frac{1}{x}\right)^{\frac{1-\gamma}{\theta}} + \frac{1}{x}(1-\gamma) g(t) \ , \\ 0 &= -h(t)\sigma_S(t,y)(1-\gamma)g(t) - \pi\left(\sigma_S(t,y)\right)^2 (1-\gamma)(-\gamma) g(t) \\ &\quad - \rho\beta(t)\sigma_S(t,y)(1-\gamma)^2\kappa g(t) \ . \end{aligned}$$

These are satisfied by the controls $c^*(t,x,y)$ and $\pi^*(t,x,y)$ from Theorem 3.2. Plugging the controls into (3.25), we obtain

$$\left.\begin{aligned} g'(t) &= -\theta\phi \left(g(t)\right)^{1-\frac{1}{\theta\phi}} \delta^{\frac{1}{\phi}} + \tau(t) g(t) \ , \\ g(T) &= (\delta\omega)^\theta \ , \end{aligned}\right\}$$

where

$$\begin{aligned} \tau(t) = \delta\theta - (1-\gamma) & \left[r + \alpha(t)\kappa + \frac{1}{2}\left(\beta(t)\right)^2 \kappa((1-\gamma)\kappa - 1) \right] \\ + \left(1 - \frac{1}{\gamma}\right) & \left[\rho(1-\gamma)h(t)\beta(t)\kappa + \frac{1}{2}h(t)^2 + \frac{1}{2}\rho^2\beta(t)^2(1-\gamma)^2\kappa^2 \right] \ . \end{aligned}$$

This ordinary differential equation (ODE) is obviously solved by the function g from Theorem 3.2. Hence, the function U from Theorem 3.2 solves the necessary Bellman equation.

3. Verify that l_1, l_2 solve diffusion equations. Plugging (3.24) and the controls (c^*, π^*) into the diffusion equations that l_1 and l_2 must solve, i.e. Equations (3.11)–(3.12), and dividing by $x^{1-\gamma}y^{\kappa(1-\gamma)}$, we obtain

$$\left.\begin{aligned} \partial_t \eta_1(t,s) &= \psi(t)\eta_1(t,s) \ , \\ \eta_1(s,s) &= \delta^{\frac{1-\gamma}{\phi}} g\left(s\right)^{-\frac{1-\gamma}{\theta\phi}} \ , \end{aligned}\right\}$$

and

$$\left.\begin{aligned} \partial_t \eta_2(t,T) &= \psi(t)\eta_2(t,T) \ , \\ \eta_2(T,T) &= 1 \ , \end{aligned}\right\}$$

where

$$\psi(t) = \tau(t) - \delta\theta + (1-\gamma)g\left(t\right)^{-\frac{1}{\phi\theta}}\delta^{\frac{1}{\phi}} \ .$$

These ODE's are obviously solved by the function η_1, η_2 from Theorem 3.2. Hence, l_1, l_2 from Theorem 3.2 solve the necessary diffusion equations. □

Proof of Lemma 3.1. The idea is to apply Itô's Lemma. Recall that

$$c^*(t,x) = \delta^{\frac{1}{\phi}} g(t)^{-\frac{1}{\phi\theta}} x \ ,$$

where g satisfies the ODE

$$g'(t) = g(t)\left[-\theta\phi\delta^{\frac{1}{\phi}} g(t)^{1-\frac{1}{\phi\theta}} + \tau(t)g(t)\right] .$$

We get the partial derivatives

$$\begin{aligned} \partial_t c^*(t,x) &= \delta^{\frac{1}{\phi}}\left(-\frac{1}{\theta\phi}\right) g(t)^{-\frac{1}{\phi\theta}-1} g'(t)x \\ &= -\frac{\delta^{\frac{1}{\phi}}}{\theta\phi} g(t)^{-\frac{1}{\phi\theta}-1}\left(-\theta\phi\delta^{\frac{1}{\phi}} g(t)^{1-\frac{1}{\phi\theta}} + \tau(t)g(t)\right) x \\ &= \frac{c^*(t,x)^2}{x} - \frac{\tau(t)}{\theta\phi} c^*(t,x) \ , \\ \partial_x c^*(t,x) &= \delta^{\frac{1}{\phi}} g(t)^{-\frac{1}{\phi\theta}} \\ &= \frac{c^*(t,x)}{x} \ , \\ \partial_x^2 c^*(t,x) &= 0 \ . \end{aligned}$$

The drift and diffusion coefficient under the equilibrium controls are given as

$$\begin{aligned} \mu^*(t,x) &= x[r + \pi^*(t,y)\lambda(t,y)] - c^*(t,x) \\ &= x\left[r + \tfrac{1}{\gamma}\left(h(t) + \rho(1-\gamma)\kappa\beta(t)\right)h(t)\right] - c^*(t,x) \ , \\ \sigma^*(t,x) &= x\pi^*(t,y)\sigma_S(t,y) \\ &= x\tfrac{1}{\gamma}\left(h(t) + \rho(1-\gamma)\kappa\beta(t)\right) \ . \end{aligned}$$

Applying Itô's Lemma, we obtain

$$\begin{aligned}
\mathrm{d}c^*\left(t, X_t^{c^*,\pi^*}\right) &= \left(\partial_t c^*(t, X_t^{c^*,\pi^*}) + \mu^*\left(t, X_t^{c^*,\pi^*}\right)\partial_x c^*\left(t, X_t^{c^*,\pi^*}\right)\right)\mathrm{d}t \\
&\quad + \sigma^*\left(t, X_t^{c^*,\pi^*}\right)\partial_x c^*\left(t, X_t^{c^*,\pi^*}\right)\mathrm{d}W_t \\
&= \begin{bmatrix} \frac{c^*(t,X_t^{c^*,\pi^*})^2}{X_t^{c^*,\pi^*}} - \frac{\tau(t)}{\theta\phi}c^*(t, X_t^{c^*,\pi^*}) \\ c^*\left(t, X_t^{c^*,\pi^*}\right)\left[r + \frac{1}{\gamma}(h(t) + \rho(1-\gamma)\kappa\beta(t))h(t)\right] \\ -\frac{c^*(t,X_t^{c^*,\pi^*})^2}{X_t^{c^*,\pi^*}} \end{bmatrix}\mathrm{d}t \\
&\quad + c^*(t, X_t^{c^*,\pi^*})\tfrac{1}{\gamma}(h(t) + \rho(1-\gamma)\kappa\beta(t))\,\mathrm{d}W_t\ ,
\end{aligned}$$

which yields the assertion. □

3.C Details for the Exponential Utility Example

Proof of Theorem 3.3. We proceed in 3 steps.

1. Compute infinitesimal generator and aggregator. With the assumptions on the market dynamics, we get the infinitesimal generator

$$\begin{aligned}
\mathcal{A}^{c,\pi} &= \left[x\left(r + \pi(t,x,y)\lambda(t,y) - c(t,x,y)\right]\partial_x + \tfrac{1}{2}\left(x\pi(t,x,y)\sigma_S(t,y)\right)^2\partial_x^2 \\
&\quad + \alpha\left(t,y\right)\partial_y + \tfrac{1}{2}\left(\beta\left(t,y\right)\right)^2\partial_y^2 + \rho\alpha\left(t,y\right)x\pi(t,x,y)\sigma_S(t,y)\partial_{xy} \\
&= \left[x\left(r + \pi(t,x,y)h(t)\sigma_S(t,y)\right) - c(t,x,y)\right]\partial_x \\
&\quad + \tfrac{1}{2}\left(x\pi(t,x,y)\sigma_S(t,y)\right)^2\partial_x^2 + \alpha\left(t\right)\partial_y \\
&\quad + \tfrac{1}{2}\left(\beta\left(t\right)\right)^2\partial_y^2 + \rho\beta\left(t\right)x\pi(t,x,y)\sigma_S(t,y)\partial_{xy}\ .
\end{aligned}$$

Also, with exponential utility we get the aggregator

$$\begin{aligned}
F\left(c,y,U\right) &= \delta\left[f'\circ f^{-1}\right]\left(U\right)\cdot\left(\left[\varphi\circ u_1\right]\left(y,c\right) - f^{-1}\left(U\right)\right) \\
&= \delta\theta\left(U^{\frac{1}{\theta}}\right)^{\theta-1}\left(\exp\left(\frac{\gamma}{\theta}\kappa y + \frac{\gamma}{\theta}c\right) - U^{\frac{1}{\theta}}\right) \\
&= \delta\theta U\left(\frac{\exp\left(\frac{\gamma}{\theta}\kappa y + \frac{\gamma}{\theta}c\right)}{U^{\frac{1}{\theta}}} - 1\right)\ .
\end{aligned}$$

2. Verify that (c^*, π^*) are equilibrium controls and U solves the Bellman equation. From Theorem 3.3, we have

$$l_i(t,s,x,y) = \eta_i(t,s)\exp\left(\gamma\kappa y + r\varepsilon(t)x\right)\ ,\quad i = 1,2\ , \tag{3.26}$$

so the remainder terms vanish, i.e. $R_x = R_y = R_{xy} = 0$. We also have

$$U\left(t,x,y\right) = \exp\left(\gamma\kappa y + \varepsilon(t)\gamma x\right)g\left(t\right)\ .$$

Plugging this and $R_x = R_y = R_{xy} = 0$ into the Bellman equation that U must solve, i.e. Equation (3.9), and dividing by $\exp(\gamma\kappa y + \varepsilon(t)\gamma x)$, we obtain

$$\left.\begin{aligned} g'(t) + \varepsilon'(t)\gamma x g(t) &= \inf_{c,\pi}\left[\begin{array}{c} -\delta\theta g(t)\left(\frac{\exp\left(\frac{\gamma}{\theta}c - \varepsilon(t)\frac{\gamma}{\theta}x\right)}{g(t)^{\frac{1}{\theta}}} - 1\right) \\ -\left(x\left(r + \pi h(t)\sigma_S(t,y)\right) - c\right)\varepsilon(t)\gamma g(t) \\ -\frac{1}{2}\left(x\pi\sigma_S(t,y)\right)^2\varepsilon(t)^2\gamma^2 g(t) \\ -\alpha(t)\gamma\kappa g(t) - \frac{1}{2}\left(\beta(t)\right)^2\gamma^2\kappa^2 g(t) \\ -\rho\beta(t)\pi x\sigma_S(t,y)\gamma^2\kappa\varepsilon(t)g(t) \end{array}\right], \\ g(T) &= (\omega\delta)^\theta \ , \end{aligned}\right\} \quad (3.27)$$

The first order conditions for c, π read

$$\begin{aligned} 0 &= -\delta\gamma g(t)\frac{\exp\left(\frac{\gamma}{\theta}c - \varepsilon(t)\frac{\gamma}{\theta}x\right)}{g(t)^{\frac{1}{\theta}}} + \varepsilon(t)\gamma g(t) \ , \\ 0 &= -xh(t)\sigma_S(t,y)\varepsilon(t)\gamma g(t) - \pi\left(x\sigma_S(t,y)\right)^2\varepsilon(t)^2\gamma^2 g(t) \\ &\quad - \rho\beta(t)x\sigma_S(t,y)\gamma^2\kappa\varepsilon(t)g(t) \ . \end{aligned}$$

These are satisfied by the controls $c^*(t,x,y)$ and $\pi^*(t,x,y)$ from Theorem 3.3. Plugging the controls into (3.27), we obtain

$$\left.\begin{aligned} g'(t) &= \tau(t)g(t) + \log\left(g(t)\right)\varepsilon(t)g(t) \ , \\ g(T) &= (\delta\omega)^\theta \ , \end{aligned}\right\}$$

where

$$\begin{aligned} \tau(t) &= \theta\left(\delta - \varepsilon(t)\right) + \theta\varepsilon(t)\log\left(\frac{\varepsilon(t)}{\delta}\right) + \frac{1}{2}h(t)^2 + \frac{1}{2}\rho^2\beta(t)^2\gamma^2\kappa^2 \\ &\quad - \alpha(t)\gamma\kappa - \frac{1}{2}\left(\beta(t)\right)^2\gamma^2\kappa^2 + \rho\beta(t)\gamma\kappa h(t) \ . \end{aligned}$$

This ODE is obviously solved by the function g from Theorem 3.3. Hence, the function U from Theorem 3.3 solves the necessary Bellman equation.

3. Verify that l_1, l_2 solve diffusion equations. Plugging (3.26) and the controls (c^*, π^*) into the diffusion equations that l_1 and l_2 must solve, i.e. (3.11)–(3.12), inserting $\varepsilon'(t) = \varepsilon(t)\left(\varepsilon(t) - r\right)$, and dividing by $\exp(\gamma\kappa y + \varepsilon(t)\gamma x)$, we obtain

$$\left.\begin{aligned} \partial_t\eta_1(t,s) &= \psi(t)\eta_1(t,s) \ , \\ \eta_1(s,s) &= \left(\frac{\varepsilon(t)}{\delta}\right)^\theta g(s) \ , \end{aligned}\right\}$$

and

$$\left.\begin{aligned} \partial_t\eta_2(t,T) &= \psi(t)\eta_2(t,T) \ , \\ \eta_2(T,T) &= 1 \ , \end{aligned}\right\}$$

where

$$\psi(t) = \tau(t) + \varepsilon(t)\log\left(g\left(t\right)\right)\right) - \theta\left(\delta - \varepsilon(t)\right) \ .$$

These ODE's are obviously solved by the function η_1, η_2 from Theorem 3.3. Hence, l_1, l_2 from Theorem 3.3 solve the necessary diffusion equations. □

Proof of Lemma 3.2. Again we apply Itô's Lemma. We have

$$c^*(t,x) = \varepsilon(t)x + \tfrac{\theta}{\gamma}\log\left(\frac{\varepsilon(t)}{\delta}\right) + \tfrac{1}{\gamma}\log\left(g\left(t\right)\right) \ ,$$

where

$$\begin{aligned} g'(t) &= \tau(t)g(t) + \log\left(g\left(t\right)\right)\varepsilon(t)g(t) \ , \\ \varepsilon'(t) &= \varepsilon(t)(\varepsilon(t) - r) \ . \end{aligned}$$

We get the partial derivatives

$$\begin{aligned} \partial_t c^*(t,x) &= \varepsilon'(t)x + \varepsilon'(t)\tfrac{\theta}{\gamma}\frac{1}{\varepsilon(t)} + g'(t)\frac{1}{\gamma}\frac{1}{g(t)} \\ &= \varepsilon(t)(\varepsilon(t)-r)x + \tfrac{\theta}{\gamma}(\varepsilon(t)-r) + \frac{1}{\gamma}\tau(t) \\ &\quad + \varepsilon(t)\left(c^*(t,x) - \varepsilon(t)x - \tfrac{\theta}{\gamma}\log\left(\tfrac{\varepsilon(t)}{\delta}\right)\right) \\ &= -r\varepsilon(t)x + \tfrac{\theta}{\gamma}(\varepsilon(t)-r) + \tfrac{1}{\gamma}\tau(t) - \tfrac{\theta}{\gamma}\varepsilon(t)\log\left(\tfrac{\varepsilon(t)}{\delta}\right) + \varepsilon(t)c^*(t,x) \ , \\ \partial_x c^*(t,x) &= \varepsilon(t) \ , \\ \partial_x^2 c^*(t,x) &= 0 \ . \end{aligned}$$

The drift and diffusion coefficient under the equilibrium controls are given as

$$\begin{aligned} \mu^*(t,x) &= rx - \frac{h(t) + \rho\beta\left(t\right)\gamma\kappa}{\varepsilon(t)\gamma}h(t) - c^*(t,x) \ , \\ \sigma^*(t) &= -\frac{h(t) + \rho\gamma\kappa\beta(t)}{\gamma\varepsilon(t)} \ . \end{aligned}$$

Applying Itô's Lemma, we obtain

$$\begin{aligned} \mathrm{d}c^*\left(t, X_t^{c^*,\pi^*}\right) &= \left(\partial_t c^*\left(t, X_t^{c^*,\pi^*}\right) + \mu^*\left(t, X_t^{c^*,\pi^*}\right)\partial_x c^*\left(t, X_t^{c^*,\pi^*}\right)\right)\mathrm{d}t \\ &\quad + \sigma^*\left(t, X_t^{c^*,\pi^*}\right)\partial_x c^*\left(t, X_t^{c^*,\pi^*}\right)\mathrm{d}W_t \\ &= \begin{bmatrix} -\varepsilon(t)rX_t^{c^*,\pi^*} + \tfrac{\theta}{\gamma}(\varepsilon(t)-r) + \tfrac{1}{\gamma}\tau(t) \\ -\tfrac{\theta}{\gamma}\varepsilon(t)\log\left(\tfrac{\varepsilon(t)}{\delta}\right) + \varepsilon(t)c^*(t, X_t^{c^*,\pi^*}) \\ + \left(rX_t^{c^*,\pi^*} - c^*(t, X_t^{c^*,\pi^*}) - \tfrac{h(t)+\rho\kappa\gamma\beta(t)}{\gamma\varepsilon(t)}h(t)\right)\varepsilon(t) \end{bmatrix}\mathrm{d}t \\ &\quad - \frac{h(t) + \rho\gamma\kappa\beta(t)}{\gamma\varepsilon(t)}\varepsilon(t)\,\mathrm{d}W_t \ , \end{aligned}$$

from which the claim follows. □

Chapter 4

Life Insurance Decisions under Recursive Utility

Ninna Reitzel Jensen (2015)

Abstract: In this paper, we generalize recursive utility to include lifetime uncertainty and utility from bequest. The generalization applies to discrete-time as well as continuous-time recursive utility, and it is an important step forward in the development of recursive utility. We formalize the problem of optimal consumption, investment, and life insurance choice under recursive utility, and we state a verification theorem with a generalized Hamilton-Jacobi-Bellman equation. Our generalization of recursive utility allows us to study optimal consumption, investment, and life insurance choice under separation of (market) risk aversion, elasticity of inter-temporal substitution, and elasticity of substitution between bequest and future utility. The separation gives rise to hump-shaped consumption patterns as observed in realized consumption.

Keywords: Recursive utility, lifetime uncertainty, stochastic control, generalized Hamilton-Jacobi-Bellman equation, hump-shaped consumption.

4.1 Introduction

In this paper, we generalize recursive utility to include lifetime uncertainty and utility from bequest. The generalization applies to discrete-time as well as continuous-time recursive utility, and it is a much needed next step in the development of recursive utility. In continuous time, recursive utility is also known as stochastic differential utility. Recursive utility plays an important role in the literature on optimal consumption and investment choice for agents with a certain lifetime, but to the knowledge of the author, it has never before been generalized to agents with an uncertain lifetime, utility from bequest, and, consequently, a need for life insurance. The generalization allows us to study optimal consumption, investment, and life insurance choice under separation of (market) risk aversion, elasticity of inter-temporal substitution, and elasticity of substitution between bequest and future utility. Here, the concept of substitution between bequest and future utility is our way of formulating preferences for mortality risk. A similar separation is obtained in Jensen and Steffensen (2015) [Chapter 2 of this thesis], but, there, the separation entails time-inconsistency, and the resulting control is an equilibrium control and not a classical optimal control. Recursive utility, on the other hand, is by construction time-consistent, therefore the separation of this paper is time-consistent. Interestingly, the separation gives rise to hump-shaped consumption patterns as observed in realized consumption, see e.g. Bullard and Feigenbaum (2007); Feigenbaum (2008); Gourinchas and Parker (2002). In Bullard and Feigenbaum (2007); Feigenbaum (2008); Gourinchas and Parker (2002), hump-shaped consumption patterns are obtained by introducing income uncertainty or utility from leisure, or by excluding access to life insurance. However, without such modifications, hump-shaped consumption patterns cannot be obtained by standard recursive utility or time-additive utility under lifetime uncertainty. It is the very combination of recursive utility and lifetime uncertainty that brings us closer to realized consumption. We note that the separation in Jensen and Steffensen (2015) also gives rise to hump-shaped consumption patterns, and it is natural to ask if the two ways of separating preferences cover the same set of preferences. In general, the answer is no. The two approaches are different in output as well as in construction. Only if we reduce the number of free preference parameters, thereby giving up on the threefold separability, we identify a coincidence between the preferences covered by the two ways of separating utility.

The existing literature on optimal consumption, investment, and life insurance choice focuses on time-additive utility. An important, early continuous-time example is Richard (1975) where the seminal work by Merton (1971) on optimal consumption and investment choice is generalized to include lifetime uncertainty and life insurance. Time-additive utility is tractable in that Bellman's principle of optimality applies. In Bellman (1957), Bellman states that *"an optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision"*. Using dynamic programming techniques, the value function of a time-additive optimization problem can be characterized by a partial differential equation (in continuous time) or a difference equation (in discrete time). Unfortunately, time-additive utility has the disadvantage that it mixes preferences for risk and preferences for inter-temporal substitution. In Epstein and Zin (1991), Hall (1988), and other papers, it is argued that this imposes an undesirable constraint on the agent's risk aversion and elasticity of inter-temporal substitution. Also, the entanglement is proposed as a reason for the so-called equity premium puzzle introduced by Mehra and Prescott (1985).

Recursive utility theory deals with the disentanglement of preferences for risk and preferences for inter-temporal substitution through a recursive definition, a certainty equivalent, and a time-aggregator. Recursive utility was first defined in discrete time, see e.g. Kreps and Porteus (1978); Epstein and Zin (1989). It was extended to continuous time by Duffie and Epstein (1992b) and refined by Kraft and Seifried (2010, 2014). Recursive utility allows for separation of preferences for risk and inter-temporal substitution, and it is widely used to study asset pricing and consumption-portfolio choice in various markets, see e.g. Schroder and Skiadas (1999, 2005); Kraft et al. (2013). Also, recursive utility has been used to explore ambiguity aversion and preferences for resolution of uncertainty, see e.g. Chen and Epstein (2002); Skiadas (1998, 2013). Despite the growing literature on recursive utility, the literature contains no attempt to accommodate for lifetime uncertainty and utility from bequest. The closest attempt is Kraft and Seifried (2010) where the authors allow for Poisson jumps in the wealth process which can to some extend be used to model lifetime uncertainty. However, they do not allow for utility from a lump sum at a random point in time, and, therefore, they cannot accommodate for utility from bequest. In all, our paper is a great advance in the field of recursive utility with lifetime uncertainty.

In Section 4.2, we provide a short introduction to recursive utility in discrete time. In Sections 4.3, we present our generalization of discrete-time recursive utility to include lifetime uncertainty and utility from bequest. In Sections 4.4, we extend the generalization to continuous time. In Section 4.5, we formalize the problem of optimal consumption, investment, and life insurance choice under recursive utility, and we provide a verification theorem with a generalized Hamilton-Jacobi-Bellman equation for its solution. In Section 4.6, we study optimal consumption, investment, and life insurance under generalized Epstein-Zin preferences, and we provide numerical results to exemplify our results and, in particular, our consumption and bequest patterns.

4.2 Discrete-time Recursive Utility

We fix a probability space $(\Omega, \mathcal{G}, P)$ endowed with a filtration $\mathcal{F} = \{\mathcal{F}_t\}_{t\in[0,T]}$ satisfying the usual conditions of completeness and right-continuity. The filtration $\mathcal{F}$ represents all information of relevance to an agent with a certain lifetime, for example information about the surrounding financial market. We fix a set $\mathcal{C} \subset \mathbb{R}^k$ of feasible consumption rates, and we denote by C a class of $\mathcal{F}$-adapted $\mathcal{C}$-valued processes with time-horizon $[0, T]$. The objective of recursive utility theory is the construction of a mapping $\mathfrak{u} : \mathsf{C} \to \mathbb{R}$ that ranks consumption streams in such a way that $\mathfrak{u}(c) \geq \mathfrak{u}(c')$ if and only if the consumption stream c is weakly preferred to the consumption stream c'. This is done by means of a utility process V^c associated to c by setting

$$\mathfrak{u}(c) = V_0^c \ , \quad c \in \mathsf{C} \ .$$

The utility process is assumed to take values in a subinterval $\mathcal{V} \subset \mathbb{R}$ of the real line, and $\mathfrak{u}$ is referred to as a recursive utility function.

Starting in discrete time, we fix a partition $0 = t_0 \leq t_1 \leq \ldots \leq t_m = T$ of $[0, T]$ and consider a discrete-time consumption stream $c = \{c_{t_k}\}_{k=0,\ldots,m-1}$ in C. The utility process V^c is defined through the backward recursion

$$\begin{aligned} V^c_{t_k} &= W\left(t_{k+1} - t_k, c_{t_k}, \mathfrak{m}\left(\mathcal{L}\left(V^c_{t_{k+1}} \middle| \mathcal{F}_{t_k}\right)\right)\right) \ , \quad k = 0, \ldots, m-1 \ , \\ V^c_T &= \xi \ . \end{aligned} \tag{4.1}$$

Here, it holds that

- $c_{t_k} \in \mathcal{C}$ is the consumption rate between time t_k and t_{k+1},
- ξ is terminal utility,
- $W : [0, \infty) \times \mathcal{C} \times \mathcal{V} \to \mathcal{V}$ is a continuous function with $W(0, c, V) = V$ for $c \in \mathcal{C}, V \in \mathcal{V}$,
- $\mathcal{L}\left(V^c_{t_{k+1}} \middle| \mathcal{F}_{t_k}\right)$ is the conditional distribution of $V^c_{t_{k+1}}$ given $\mathcal{F}_{t_k}$,
- $\mathfrak{m}$ is a certainty equivalent on $\mathcal{V}$.

Letting $\mathcal{M}_1(\mathcal{V})$ denote the set of probability measures on $\mathcal{B}(\mathcal{V})$ with moments of all orders, a functional $\mathfrak{m} : \mathcal{M}_1(\mathcal{V}) \to \mathbb{R}$ is called a certainty equivalent on $\mathcal{V}$ if $\mathfrak{m}(\delta_v) = v$ for all $v \in \mathcal{V}$ where δ_v is the Dirac measure at v.

The function W is often referred to as the inter-temporal aggregator because in a set-up without risk, implying $\mathfrak{m}\left(\mathcal{L}\left(V^c_{t_{k+1}}) \middle| \mathcal{F}_{t_k}\right)\right) = V^c_{t_{k+1}}$, it describes the inter-temporal aggregation of present consumption c_{t_k} and the utility of future consumption $V^c_{t_{k+1}}$. Similarly, the certainty equivalent $\mathfrak{m}$ is referred to as the risk-aggregator since it describes the risk weighted aggregation of possible future values of $V^c_{t_{k+1}}$. The pair $(W, \mathfrak{m})$ completely describes a certain-lived agent's preferences for discrete-time stochastic consumption streams, and we call $(W, \mathfrak{m})$ a discrete-time aggregator.

A special class of certainty equivalents are those given by

$$\mathfrak{m}(\mu) = h^{-1}\left(\int_{\mathcal{V}} h \, \mathrm{d}\mu\right) \ , \quad \mu \in \mathcal{M}_1(\mathcal{V}) \ , \tag{4.2}$$

for a strictly increasing, polynomially bounded C^2-function $u : \mathcal{V} \to \mathbb{R}$. Here, $\mathfrak{m}$ is called an expected utility (EU) certainty equivalent or the Kreps-Porteus certainty equivalent induced by the function h. If h is the identity, then $\mathfrak{m}$ denotes expectation and is called risk-neutral. In that case, we speak of the pair $(W, \mathfrak{m})$ as a normalized discrete-time aggregator. When dealing with a normalized aggregator, we generally leave out the certainty equivalent $\mathfrak{m} = \mathbb{E}$ and speak of W as a normalized aggregator instead of writing $(W, \mathbb{E})$.

Utility has an ordinal interpretation rather than a cardinal. Therefore, if $\Phi : \mathcal{V} \to \bar{\mathcal{V}}$ is a strictly increasing function, and if we define the mapping $\bar{\mathfrak{u}} : \bar{\mathsf{C}} \to \mathbb{R}$ by

$$\bar{\mathfrak{u}}\left(c\right) = \Phi\left(\mathfrak{u}\left(c\right)\right) \ ,$$

then $\bar{\mathfrak{u}}$ is a recursive utility function representing the same preferences as $\mathfrak{u}$. Here, barred quantities are interpreted in the same way as their non-barred counterparts. In this case, we say that $\mathfrak{u}$ and $\bar{\mathfrak{u}}$ are equivalent, and the underlying discrete-time aggregators $(W, \mathfrak{m})$ and $\left(\bar{W}, \bar{\mathfrak{m}}\right)$ are said to be ordinally equivalent. In Kraft and Seifried (2010), it is shown that if $\mathfrak{m}$ is given by Equation (4.2), then the normalized aggregator $\bar{W}$ ordinally equivalent to $(W, \mathfrak{m})$ is given by

$$\bar{W}\left(\Delta, c, V\right) = h\left(W\left(\Delta, c, h^{-1}\left(V\right)\right)\right) \ .$$

Example 4.1. A particular class of discrete-time inter-temporal aggregators $W : [0, \infty) \times \mathcal{C} \times \mathcal{C} \mapsto \mathcal{C}$ is given by

$$W\left(\Delta, c, V\right) = g^{-1}\left(\left(1 - \delta\Delta\right) g\left(V\right) + \delta\Delta g\left(c\right)\right)$$

for a strictly increasing function $g : \mathcal{C} \mapsto \mathbb{R}$ and a subjective discount rate $\delta > 0$, see e.g. Kraft and Seifried (2014). Since the expression for the aggregator contains both $g\left(V\right)$ and $g\left(c\right)$, the utility value V and the consumption rate c are measured on the same scale. Also, utility is measured in consumption units thanks to the outer function g^{-1}.

Let $\mathfrak{m}$ be the Kreps-Porteus certainty equivalent induced by a function $u : \mathcal{C} \mapsto \bar{\mathcal{C}}$. Then the normalized discrete-time aggregator $\bar{W} : [0, \infty) \times \mathcal{C} \times \bar{\mathcal{C}} \mapsto \bar{\mathcal{C}}$ corresponding to $(W, \mathfrak{m})$ is given by

$$\begin{aligned} \bar{W}\left(\Delta, c, V\right) &= u\left(g^{-1}\left(\left(1 - \delta\Delta\right) g\left(u^{-1}\left(V\right)\right) + \delta\Delta g\left(c\right)\right)\right) \\ &= \left(u \circ g^{-1}\right)\left(\left(1 - \delta\Delta\right)\left(u \circ g^{-1}\right)^{-1}\left(V\right) + \delta\Delta g\left(c\right)\right) \ . \end{aligned}$$

Utility is now measured in units of felicity.

4.3 The Discrete-Time Aggregator under Lifetime Uncertainty

In the previous section, we only addressed recursive utility for an agent with a certain lifetime. To allow for lifetime uncertainty and utility from bequest, we start by introducing a stochastic process $I = (I_t)_{t \in [0,T]}$ on $(\Omega, \mathcal{G}, \mathbb{P})$ that indicates survival, i.e. $I_{t_k} = 1$ indicates survival up to and including time t_k. We think of T as the agent's maximum remaining lifetime, and we assume that $(I_t)_{t \in [0,T]}$ is independent of the filtration $\mathcal{F}$ that still represents all information of relevance to an agent with a certain lifetime. For convenience, we refer to all non-mortality risk as market risk. With this convention, the filtration $\mathcal{F}$ represents information about developments in the market.

In the following, we limit our focus to inter-temporal aggregators

$$W : [0,\infty) \times \mathcal{C} \times \mathcal{C} \mapsto \mathcal{C}$$

that measures utility in consumption units. This facilitates easy aggregation of bequest and future utility. In addition to the certainty equivalent $\mathfrak{m}$ that describes the risk weighted aggregation of possible future levels of utility for a certain-lived agent, we introduce a certainty equivalent $\mathfrak{n}$ that describes the mortality risk weighted aggregation of bequest and future utility. By mortality risk weighted aggregation, we mean aggregation taking into account the agent's preferences for substitution between bequest and future utility. When applying $\mathfrak{n}$, we condition on $\mathcal{F}_T$ which is seen as complete information about the market. We condition on $\mathcal{F}_T$ to separate lifetime uncertainty from all other uncertainty as represented by $\mathcal{F}_T$. To simplify notation, we write $\mathfrak{m}_{t_k}(\cdot) = \mathfrak{m}\left(\mathcal{L}\left(\cdot|\mathcal{F}_{t_k}\right)\right)$ and $\mathfrak{n}_{t_k}(\cdot) = \mathfrak{n}\left(\mathcal{L}\left(\cdot|I_{t_k}=1,\mathcal{F}_T\right)\right)$.

To introduce lifetime uncertainty and utility from bequest, we replace the backward recursion in Equation (4.1) with the backward recursion

$$\begin{aligned} V_{t_k}^{c,b} &= W\left(t_{k+1}-t_k, c_{t_k}, \mathfrak{m}_{t_k}\left(\mathfrak{n}_{t_k}\left(I_{t_{k+1}}V_{t_{k+1}}^{c,b} + \left(1-I_{t_{k+1}}\right)b_{t_{k+1}}\right)\right)\right), \\ V_T^{c,b} &= \xi \,, \end{aligned} \tag{4.3}$$

where $c_{t_k} \in \mathcal{C}$ is the consumption rate between time t_k and t_{k+1}, $b_{t_{k+1}} \in \mathcal{C}$ is bequest given death between time t_k and t_{k+1}, and $\xi \in \mathcal{C}$ is terminal wealth. Now, $V_{t_k}^{c,b}$ is the utility given survival up to and including time t_k. We assume that the processes $(c_t)_{t\in[0,T]}$ and $(b_t)_{t\in[0,T]}$ are adapted to the filtration $\mathcal{F}$, and that ξ is $\mathcal{F}_T$-measurable. Also, by construction, $\left(V_t^{c,b}\right)_{t\in[0,T]}$ is adapted to the filtration $\mathcal{F}$. The intuition behind the backward recursion in Equation (4.3) is the following:

- The agent is alive at time t_k. At the future time point t_{k+1}, the agent is either alive or dead, as indicated by $I_{t_{k+1}}$. If the agent is alive, the utility is $V_{t_{k+1}}^{c,b}$. If the agent is dead, the only utility left is the bequest $b_{t_{k+1}}$.
- The certainty equivalent $\mathfrak{n}$ describes the agent's mortality risk weighted aggregation of bequest and future utility. The aggregation is performed given the information $\mathcal{F}_T$ to focus only on preferences for mortality risk (or, more precisely, preferences for substitution between bequest and future utility). The result,

$$\mathfrak{n}_{t_k}\left(I_{t_{k+1}}V_{t_{k+1}}^{c,b} + \left(1-I_{t_{k+1}}\right)b_{t_{k+1}}\right) \,,$$

is the agent's mortality risk weighted future utility aggregated across the states dead and alive.

- Due to market risk, the size of this future utility is not known at time t_k. The certainty equivalent $\mathfrak{m}$ describes the market risk weighted aggregation of its possible values.

- Finally, the function W describes the inter-temporal aggregation of present consumption c_{t_k} and the market and mortality risk weighted utility of future consumption and bequest,

$$\mathfrak{m}_{t_k}\left(\mathfrak{n}_{t_k}\left(I_{t_{k+1}}V^{c,b}_{t_{k+1}} + \left(1 - I_{t_{k+1}}\right) b_{t_{k+1}}\right)\right) \ .$$

We mention that we could easily have multiplied the bequest $b_{t_{k+1}}$ by a non-zero weight function. There are three main reasons for including a weight function; 1) it accounts for the fact that agents without heirs might not care much about bequest, 2) it allows for a different weight on bequest throughout life, 3) it accommodates for the fact that bequest is not consumed all at once, but typically over a number of years, thereby raising the utility from bequest. For notational convenience, we have chosen not to include a weight function, but the inclusion of a deterministic and non-zero weight function would not alter our results.

We define $\mathfrak{m}$ and $\mathfrak{n}$ to be Kreps-Porteus certainty equivalents induced by strictly increasing polynomially bounded C^2-functions $u, v : \mathcal{C} \mapsto \mathbb{R}$, i.e.

$$\begin{aligned}
\mathfrak{m}\left(\mu\right) &= u^{-1}\left(\int_{\mathcal{C}} u\left(s\right)\mu\left(\mathrm{d}s\right)\right) \ , \quad \mu \in \mathcal{M}_1\left(\mathcal{C}\right) \ , \\
\mathfrak{n}\left(\mu\right) &= v^{-1}\left(\int_{\mathcal{C}} v\left(s\right)\mu\left(\mathrm{d}s\right)\right) \ , \quad \mu \in \mathcal{M}_1\left(\mathcal{C}\right) \ .
\end{aligned}$$

That is, v determines the agent's preferences for substitution between bequest and future utility, whereas u determines the market risk weighted aggregation of future levels of utility. Letting $p_{t_k,t_{k+1}} = \mathbb{P}\left(I_{t_{k+1}} = 1 \middle| I_{t_k} = 1\right)$ denote the conditional survival probability from time t_k to time t_{k+1}, we get

$$\begin{aligned}
&\mathfrak{n}_{t_k}\left(I_{t_{k+1}}V^{c,b}_{t_{k+1}} + \left(1 - I_{t_{k+1}}\right) b_{t_{k+1}}\right) \\
&= v^{-1}\left(\mathbb{E}\left[v\left(I_{t_{k+1}}V^{c,b}_{t_{k+1}} + \left(1 - I_{t_{k+1}}\right) b_{t_{k+1}}\right)\middle|\mathcal{F}_T, I_{t_k} = 1\right]\right) \\
&= v^{-1}\left(\mathbb{E}\left[I_{t_{k+1}} v\left(V^{c,b}_{t_{k+1}}\right) + \left(1 - I_{t_{k+1}}\right) v\left(b_{t_{k+1}}\right)\middle|\mathcal{F}_T, I_{t_k} = 1\right]\right) \\
&= v^{-1}\left(v\left(V^{c,b}_{t_{k+1}}\right)\mathbb{E}\left[I_{t_{k+1}}\middle|I_{t_k} = 1\right] + v\left(b_{t_{k+1}}\right)\mathbb{E}\left[\left(1 - I_{t_{k+1}}\right)\middle|, I_{t_k} = 1\right]\right) \\
&= v^{-1}\left(p_{t_k,t_{k+1}} v\left(V^{c,b}_{t_{k+1}}\right) + \left(1 - p_{t_k,t_{k+1}}\right) v\left(b_{t_{k+1}}\right)\right) \ .
\end{aligned}$$

At the third equality, we have used that $V^{c,b}_{t_{k+1}}$ and $b_{t_{k+1}}$ are $\mathcal{F}_T$-measurable and that $I_{t_{k+1}}$ is independent of $\mathcal{F}_T$. We obtain

$$\begin{aligned}&\mathfrak{m}_{t_k}\left(\mathfrak{n}_{t_k}\left(I_{t_{k+1}}V^{c,b}_{t_{k+1}}+\left(1-I_{t_{k+1}}\right)b_{t_{k+1}}\right)\right)\\&=u^{-1}\left(\mathbb{E}_{t_k}\left[u\left(v^{-1}\left(p_{t_k,t_{k+1}}v\left(V^{c,b}_{t_{k+1}}\right)+\left(1-p_{t_k,t_{k+1}}\right)v\left(b_{t_{k+1}}\right)\right)\right)\right]\right)\ ,\end{aligned}$$

where we have used the notation $\mathbb{E}_{t_k}[\cdot]=\mathbb{E}[\cdot|\mathcal{F}_{t_k}]$.

We notice that if $u=v$, meaning that the agent's preferences for market and mortality risk are the same, the agent's risk weighted utility of future consumption and bequest is given by

$$\begin{aligned}&\mathfrak{m}_{t_k}\left(\mathfrak{n}_{t_k}\left(I_{t_{k+1}}V^{c,b}_{t_{k+1}}+\left(1-I_{t_{k+1}}\right)b_{t_{k+1}}\right)\right)\\&=u^{-1}\left(\mathbb{E}_{t_k}\left[p_{t_k,t_{k+1}}u\left(V^{c,b}_{t_{k+1}}\right)+\left(1-p_{t_k,t_{k+1}}\right)u\left(b_{t_{k+1}}\right)\right]\right)\ .\end{aligned}$$

Hence, the agent's utility is additive across the states dead and alive. We refer to this as the utility-bequest-additive case. One focal point of this paper is the separation of preferences for market risk and preferences for substitution between bequest and future utility (i.e. preferences for mortality risk).

The backward recursion in Equation (4.3) is intuitive, but not as tractable as a backward recursion with a normalized aggregator. For convenience, assume that the partition $t_0,\dots,t_m$ is equidistant with step size $\Delta=t_{k+1}-t_k$. The transformations

$$\tilde{V}^{c,b}_{t_{k+1}}=u\left(v^{-1}\left(p_{t_k,t_k+\Delta}v\left(V^{c,b}_{t_{k+1}}\right)+\left(1-p_{t_k,t_k+\Delta}\right)v\left(b_{t_{k+1}}\right)\right)\right)$$

and

$$\tilde{\xi}=u\left(v^{-1}\left(p_{T-\Delta,T}v\left(\xi\right)+\left(1-p_{T-\Delta,T}\right)v\left(b_T\right)\right)\right)$$

yield the backward recursion

$$\tilde{V}^{c,b}_{t_k}=\tilde{W}\left(t_k,\Delta,c_{t_k},b_{t_k},\mathbb{E}_{t_k}\left[\tilde{V}^{c,b}_{t_{k+1}}\right]\right)\ ,\ \tilde{V}^{c,b}_T=\tilde{\xi}\ ,$$

where the normalized aggregator $\tilde{W}$ is given by

$$\begin{aligned}&\tilde{W}\left(t,\Delta,c,b,\tilde{V}\right)\\&=u\left(v^{-1}\left(p_{t-\Delta,t}v\left(W\left(\Delta,c,u^{-1}\left(\tilde{V}\right)\right)\right)+\left(1-p_{t-\Delta,t}\right)v\left(b\right)\right)\right)\\&=\left(u\circ v^{-1}\right)\left(p_{t-\Delta,t}\left(u\circ v^{-1}\right)^{-1}\left(\bar{W}\left(\Delta,c,\tilde{V}\right)\right)+\left(1-p_{t-\Delta,t}\right)v\left(b\right)\right)\ .\end{aligned}$$

Here, $\bar{W}$ is the normalized discrete-time aggregator, ordinally equivalent to $(W,\mathfrak{m})$ which is given by

$$\bar{W}\left(\Delta,c,\tilde{V}\right)=u\left(W\left(\Delta,c,u^{-1}\left(\tilde{V}\right)\right)\right)\ .$$

We note that the normalized aggregator $\tilde{W}$ measures utility in units of felicity. We call $\tilde{W}$ a mortality adjusted normalized discrete-time aggregator. Contrary to the aggregator $\bar{W}$, the mortality adjusted aggregator $\tilde{W}$ is time-dependent because the conditional survival probability is time-dependent. This concludes our generalization of discrete-time recursive utility.

Example 4.2. The mortality adjusted normalized discrete-time aggregator corresponding to the discrete-time aggregator $(W, \mathfrak{m})$ from Example 4.1 is given by

$$\begin{aligned}
&\tilde{W}\left(t, \Delta, c, b, \tilde{V}\right) \\
&= u\Bigg(v^{-1}\Big(p_{t-\Delta,t} v\left(g^{-1}\left((1-\delta\Delta)\, g\left(u^{-1}\left(\tilde{V}\right)\right) + \delta\Delta g(c)\right)\right) \\
&\qquad\qquad\qquad\qquad + (1-p_{t-\Delta,t})\, v(b)\Big)\Bigg).
\end{aligned}$$

First, the present consumption and future utility are aggregated, taking into account the agent's preferences for inter-temporal substitution as represented by the function g. Second, the resulting utility is aggregated with the bequest, taking into account the agent's preferences for substitution between bequest and utility as represented by the function v. The two aggregations have exactly the same form; the only difference is that $p_{t-\Delta,t}$ replaces $\delta\Delta$ and v replaces g. We consider this to be an elegant feature of our generalization.

4.4 The Continuous-Time Aggregator under Lifetime Uncertainty

To extend our generalization to continuous time, we model the survival indicator process $(I_t)_{t\in[0,T]}$ by a deterministic, time-dependent mortality intensity $\mu : [0, T] \mapsto \mathbb{R}$ satisfying

$$\frac{\partial}{\partial\Delta} \log\left(p_{t,t+\Delta}\right) = -\mu\left(t+\Delta\right) \ .$$

We assume that μ is bounded on $[0, T]$. With this assumption, T cannot possibly be the agent's maximum remaining lifetime, but we assume that the agent has a very low probability of being alive after time T such that the time horizon $[0, T]$ can be thought of as the agent's remaining lifespan. We have the boundary condition $p_{t,t} = 1$. Hence, the conditional survival probability from time $t - \Delta$ to t is given by

$$p_{t-\Delta,t} = e^{-\int_{t-\Delta}^{t} \mu(v)\,\mathrm{d}v} \ .$$

We assume that the assumptions of Theorem 4.1 or Theorem 6.1 in Kraft and Seifried (2014) are satisfied for the aggregators $\tilde{W}$ and $\bar{W}$ (with the bequest

b playing the same role as the consumption c). Then the continuous-time aggregator corresponding to $\tilde{W}$ is given by

$$f\left(t, c, b, \tilde{V}\right) = \tilde{W}_\Delta\left(t, 0, c, b, \tilde{V}\right) .$$

In the derivation of the continuous-time aggregator in Kraft and Seifried (2014), there are neither time-dependence nor bequest, but in the limiting argument only the step size Δ and the recursively defined utility process are relevant. Hence, the result stands for our time-dependent aggregator with bequest. In Appendix 4.A, we easily derive that

$$\begin{aligned} &f\left(t, c, b, \tilde{V}\right) \\ &= \bar{f}\left(c, \tilde{V}\right) + \mu(t)\left(v(b) - v\left(u^{-1}\left(\tilde{V}\right)\right)\right)\left(u \circ v^{-1}\right)'\left(\left(u \circ v^{-1}\right)^{-1}\left(\tilde{V}\right)\right) . \end{aligned} \tag{4.4}$$

Here, $\bar{f}$ is the "classical" normalized continuous-time aggregator corresponding to the normalized discrete-time aggregator $\bar{W}$. We call f a mortality adjusted normalized continuous-time aggregator. The aggregator is time-dependent exactly as its discrete-time counterpart. As in discrete time, it would be more correct to write $(f, \mathbb{E})$ when speaking of a normalized continuous-time aggregator, but we generally leave out the expectation $\mathbb{E}$.

The utility process $\tilde{V}_t^{c,b}$ associated to the consumption-bequest stream $(c_s, b_s)_{s\in[0,T]}$ and terminal wealth ξ is given by

$$\tilde{V}_t^{c,b} = \mathbb{E}_t\left[\int_t^T f\left(s, c_s, b_s, \tilde{V}_s^{c,b}\right) \mathrm{d}s + u(\xi)\right] .$$

Necessary and sufficient conditions for existence and uniqueness of this utility process cannot be characterized by simple explicit conditions. In recursive utility without lifetime uncertainty, it is sufficient to assume that the aggregator f is measurable, Lipschitz in utility, and satisfies a growth condition in consumption, see e.g. Theorem 1 in Duffie and Epstein (1992b). However, the Lipschitz condition rules out the important class of Epstein-Zin preferences that we focus on in Section 4.6. Following standard practice, we settle for an implicit condition and consider only consumption and bequest processes for which the utility process is well-defined. We denote by $\mathcal{U}$ the class of progressively measurable consumption and bequest processes $(c, b) \in \mathsf{C}^2$ such that $\tilde{V}^{c,b}$ is a uniquely determined semi-martingale. In applications, one then has to verify existence of the utility process corresponding to the optimal solution on a case-by-case basis. For a detailed discussion in this direction, see Kraft et al. (2013).

Example 4.3. The continuous-time aggregator corresponding to the normalized discrete-time aggregator $\bar{W}$ from Example 4.1 is given by

$$\bar{f}\left(c,\tilde{V}\right)=\delta\frac{u'\left(u^{-1}\left(\tilde{V}\right)\right)}{g'\left(u^{-1}\left(\tilde{V}\right)\right)}\left(g\left(c\right)-g\left(u^{-1}\left(\tilde{V}\right)\right)\right)\ .$$

The corresponding mortality adjusted normalized aggregator f is given by

$$\begin{aligned}f\left(t,c,b,\tilde{V}\right)=\mu\left(t\right)&\frac{u'\left(u^{-1}\left(\tilde{V}\right)\right)}{v'\left(u^{-1}\left(\tilde{V}\right)\right)}\left(v\left(b\right)-v\left(u^{-1}\left(\tilde{V}\right)\right)\right)\\&+\delta\frac{u'\left(u^{-1}\left(\tilde{V}\right)\right)}{g'\left(u^{-1}\left(\tilde{V}\right)\right)}\left(g\left(c\right)-g\left(u^{-1}\left(\tilde{V}\right)\right)\right)\ .\end{aligned}$$

For the derivation of both aggregators, see Appendix 4.B. We notice that the aggregation of consumption and future utility and the aggregation of future utility and bequest are now performed simultaneously and not consecutively as in the discrete-time case. Again, the two aggregations have exactly the same form; the only difference is that $\mu\left(t\right)$ replaces δ and v replaces g. As in discrete time, we consider it to be an elegant feature of our generalization that preferences for inter-temporal substitution and preferences for substitution between bequest and utility are treated symmetrically. For concrete choices of the functions u, v, and g, one needs to check up on the assumptions of Theorem 4.1 or Theorem 6.1 in Kraft and Seifried (2014) (with b playing the same role as c).

4.5 Life Insurance Decisions under Recursive Utility

The introduction of lifetime uncertainty and utility from bequest allows us to consider an agent making decisions concerning consumption, investment, and life insurance under continuous-time recursive utility. We model the death of the agent by the survival indicator process I from the previous sections. In particular, the agent's probability of being alive at time t is given by

$$P\left(I_t=1\right)=P\left(I_s=1:s\in[0,t]\right)=e^{-\int_0^t\mu(v)\,\mathrm{d}v}\ ,\quad t\geq 0\ .$$

We assume that the agent has a very low probability of being alive after time T such that the optimization problem below can be thought of as the agent's lifetime consumption-investment-life insurance problem.

The agent has access to a classical Black-Scholes market consisting of a bank account, B, with risk free short rate r, and a stock, S, with excess return

λ and volatility σ. The asset prices evolve according to the dynamics

$$\begin{aligned} \mathrm{d}B_t &= B_t r \mathrm{d}t \ , \quad t \geq 0 \ , \quad B_0 = 1 \ , \\ \mathrm{d}S_t &= S_t \left[(r + \lambda) \, \mathrm{d}t + \sigma \mathrm{d}W_t \right] \ , \quad t \geq 0 \ , \quad S_0 = s_0 \ , \end{aligned}$$

where $r, \lambda, \sigma > 0$ are constants, and W is a standard Brownian motion on the probability space $(\Omega, \mathcal{G}, P)$. We assume that the filtration $\mathcal{F}$ from the previous sections is generated by the Brownian motion W.

The agent can trade term insurance with a life insurance company. A death sum d triggers premiums payments at rate $d\hat{\mu}$. Here, $\hat{\mu}$ is the mortality intensity used by the insurance company for pricing, and it may or may not be equal to μ.

The agent's preferences are described by a mortality adjusted normalized continuous-time aggregator $f\,(t, c, b, v)$ and a utility function u for terminal wealth. The terminal wealth comes into play in the improbable case of survival until time T.

The agent has wealth X and invests an amount π in the stock and the remaining wealth $X - \pi$ in the bank account. As long as the agent is alive, she consumes at rate c, receives labor income at a deterministic rate w, and buys life insurance at premium rate $d\hat{\mu}$. When the agent dies, her inheritors receive the death sum d and the remaining wealth X. While the agent is alive, her wealth evolves according to the dynamics

$$\begin{aligned} \mathrm{d}X_t^{c,\pi,d} &= \left[r X_t^{c,\pi,d} + \pi_t \lambda - c_t - \hat{\mu}\,(t)\, d_t + w\,(t) \right] \mathrm{d}t + \pi_t \sigma \, \mathrm{d}W_t \ , \\ X_0^{c,\pi,d} &= x_0 \ , \end{aligned} \tag{4.5}$$

where x_0 is the initial wealth of the agent, w is a continuous, deterministic function, and c, π, d are stochastic processes, i.e.

$$c, \pi, d : [0, T] \times \Omega \to \mathbb{R} \ .$$

We require that the set of control processes (c, π, d) is adapted to the filtration $\mathcal{F}$ and chosen from the class of admissible controls

$$\begin{aligned} \mathcal{A}\,(x_0) = \Big\{ (c, \pi, d) : \left(c, d + X^{c,\pi,d} \right) \in \mathcal{U} \text{ and} \\ \text{(4.5) has a unique solution } X^{c,\pi,d} \text{ in } \mathbb{R} \text{ with} \\ (c_t, \pi_t, d_t) \in \Gamma \left(t, X_t^{c,\pi,d} \right) \text{ for all } t \in [0, T] \Big\} \, . \end{aligned}$$

Here, the set function $\Gamma : [0, T] \times \mathbb{R}$ models a possible state-dependent constraint on the set of controls.

In addition to the agent's monetary wealth, we also formalize the agent's human wealth which we denote by L. The agent's human wealth is the financial value of her future labour income, and it is given by

$$L\,(t) = \int_t^T w\,(s)\, e^{-\int_t^s (r + \hat{\mu}(v))\, \mathrm{d}v} \, \mathrm{d}s \ , \quad t \in [0, T] \ . \tag{4.6}$$

We speak of the sum $X^{c,\pi,d} + L$ as the agent's total wealth. We note that the mortality intensity $\hat{\mu}$ (and not μ) appears in Equation (4.6) because $\hat{\mu}$ is the intensity used for pricing the term insurance, and this asset completes the market.

The agent wishes to solve the problem

$$\max_{(c,\pi,d)\in\mathcal{A}(x_0)} V_0^{c,\pi,d} \ , \tag{4.7}$$

where $V^{c,\pi,d}$ is the utility process corresponding to the investment strategy $\pi = (\pi_t)_{t\in[0,T]}$, the consumption strategy $c = (c_t)_{t\in[0,T]}$ and the term insurance strategy $d = (d_t)_{t\in[0,T]}$, i.e.

$$V_t^{c,\pi,d} = \mathbb{E}_t\left[\int_t^T f\left(s, c_s, d_s + X_s^{c,\pi,d}, V_s^{c,\pi,d}\right) \mathrm{d}s + u\left(X_T^{c,\pi,d}\right)\right] \ .$$

Repeating the proof of Theorem 3.1 in Kraft et al. (2013) with our mortality adjusted aggregator $f(t,c,b,v)$, we realize that the additional dependence on time and bequest is never an issue. Hence, the theorem and proof in Kraft et al. (2013) allow us to state the following verification theorem.

Theorem 4.1. *Suppose there exists a* $k > 0$ *such that*

$$f(t,c,b,v) - f(t,c,b,w) \leq k(v-w) \tag{4.8}$$

for all $t \in [0,T]$*,* $c, b \in \mathcal{C}$ *and* $v, w \in \mathcal{V}$ *with* $v \geq w$*.*

Assume there exists a function $J \in C^{1,2}([0,T]\times\mathbb{R})$ *that solves the generalized Hamilton-Jacobi-Bellman equation*

$$\begin{aligned} 0 &= \sup_{(c,\pi,d)\in\Gamma(t,x)} \left\{ \begin{array}{c} J_t(t,x) + (rx + \pi\lambda - c - \hat{\mu}(t)d + w(t)) J_x(t,x) \\ +\frac{1}{2}\pi^2\sigma^2 J_{xx}(t,x) \\ +f(t,c,d+x,J(t,x)) \end{array} \right\} \ , \\ J(T,x) &= u(x) \ , \end{aligned} \tag{4.9}$$

and assume that the local martingale

$$\int_0^\cdot J_x\left(t, X_t^{c,\pi,d}\right) \pi_t \sigma \, \mathrm{d}W_t$$

is a true martingale for every $(c,\pi,d) \in \mathcal{A}(x_0)$*.*

If there exists a control $(c^*,\pi^*,d^*) \in \mathcal{A}(x_0)$ *that realizes the supremum in Equation* (4.9)*, then* (c^*,π^*,d^*) *is the optimal control, and* J *is the value function of the problem in Equation* (4.7)*. In particular, it holds that*

$$\max_{(c,\pi,d)\in\mathcal{A}(x_0)} V_0^{c,\pi,d} = V_0^{c^*,\pi^*,d^*} = J(0,x_0) \ .$$

[Post-submission comment: For a later added proof, see Appendix A.]

4.6 Life Insurance Decisions with Epstein-Zin Preferences

We focus on an agent with Epstein-Zin preferences. These preferences arise by setting $\mathcal{C} = (0, \infty)$, $g(x) = \frac{1}{1-\phi} x^{1-\phi}$, and $u(x) = \frac{1}{1-\rho} x^{1-\rho}$ in Examples 4.1–4.3. The "classical" normalized Epstein-Zin continuous-time aggregator reads

$$\bar{f}\left(c, \tilde{V}\right) = \delta \frac{1-\rho}{1-\phi} \tilde{V} \left(\left(\frac{c}{\left((1-\rho)\tilde{V}\right)^{\frac{1}{1-\rho}}} \right)^{1-\phi} - 1 \right).$$

Here, ρ is relative (market) risk aversion, and $\frac{1}{\phi}$ is elasticity of inter-temporal substitution (EIS). Epstein-Zin preferences are also known as CEIS-CRRA preferencence ("constant EIS and constant relative risk aversion").

Letting $\frac{1}{\kappa}$ denote the elasticity of substitution between bequest and future utility, we set $v(x) = \frac{1}{1-\kappa} x^{1-\kappa}$. This gives us the following normalized continuous-time mortality adjusted Epstein-Zin aggregator

$$\begin{aligned} f\left(t, c, b, \tilde{V}\right) = \mu(t) \frac{1-\rho}{1-\kappa} \tilde{V} & \left(\left(\frac{b}{\left((1-\rho)\tilde{V}\right)^{\frac{1}{1-\rho}}} \right)^{1-\kappa} - 1 \right) \\ + \delta \frac{1-\rho}{1-\phi} \tilde{V} & \left(\left(\frac{c}{\left((1-\rho)\tilde{V}\right)^{\frac{1}{1-\rho}}} \right)^{1-\phi} - 1 \right). \end{aligned} \tag{4.10}$$

We notice that the aggregator is a sum of two Epstein-Zin aggregators. Therefore, we can use Proposition 3.2 in Kraft et al. (2013) to conclude that the aggregator satisfies the regularity condition in Equation (4.8) in each of the following four cases:

1. $\rho > 1$ and $\phi, \kappa < 1$,
2. $\rho > 1$ and $\phi, \kappa > 1$ with $\rho \leq \phi, \kappa$,
3. $\rho < 1$ and $\phi, \kappa > 1$,
4. $\rho < 1$ and $\phi, \kappa < 1$ with $\rho \geq \phi, \kappa$.

In the following, we only consider these cases.

We mention that Epstein-Zin preferences do not necessarily satisfy Theorem 4.1 or Theorem 6.1 in Kraft and Seifried (2014). However, Epstein-Zin preferences are widely used in the literature, so we continue working with the aggreagtor in Equation (4.10).

In Appendix 4.C, we show that the optimal controls for the problem in Equation (4.7) with the aggregator in Equation (4.10) are given by

$$\begin{aligned}
\pi^*(t,x) &= \frac{\lambda}{\rho\sigma^2}\left(x + L(t)\right) \ , \\
c^*(t,x) &= \delta^{\frac{1}{\phi}} \left(g(t)\right)^{-\frac{q(1-\phi)}{\phi(1-\rho)}} \left(x + L(t)\right) \ , \\
d^*(t,x) &= \left(\frac{\mu(t)}{\hat{\mu}(t)}\right)^{\frac{1}{\kappa}} \left(g(t)\right)^{-\frac{q(1-\kappa)}{\kappa(1-\rho)}} \left(x + L(t)\right) - x \ ,
\end{aligned}$$

where L is the agent's human wealth, q is a free parameter that can be chosen based on convenience, and g is the solution to the ordinary differential equation (ODE)

$$\begin{aligned}
g_t(t) = &-\frac{1}{q}(1-\rho)\left[r + \frac{\lambda^2}{2\rho\sigma^2} + \hat{\mu}(t) - \frac{\delta}{1-\phi} - \frac{\mu(t)}{1-\kappa}\right] g(t) \\
&- \frac{1}{q}\frac{\phi(1-\rho)}{1-\phi}\delta^{\frac{1}{\phi}}\left(g(t)\right)^{-\frac{q(1-\phi)}{\phi(1-\rho)}+1} \\
&- \frac{1}{q}\frac{\kappa(1-\rho)}{1-\kappa}\left(\hat{\mu}(t)\right)^{1-\frac{1}{\kappa}}\left(\mu(t)\right)^{\frac{1}{\kappa}}\left(g(t)\right)^{-\frac{q(1-\kappa)}{\kappa(1-\rho)}+1} \ , \\
g(T) = &\ 1 \ .
\end{aligned} \tag{4.11}$$

We note that c^*, π^*, and $d^* + x$ are all directly proportional to the agent's total wealth $x + L$. The optimal proportion π^* of wealth to invest in the stock is independent of the elasticity parameters κ and ϕ, and it is the same as in the well-known case of time-additive utility. The ODE for g is non-linear for any choice of q, except in the time-additive and utility-bequest-additive case $\rho = \phi = \kappa$ where it is linear for $q = \rho$. In general, an explicit solution is therefore not available for the optimal consumption rate c^* or bequest $d^* + x$.

4.6.1 The optimal consumption rate and bequest

In this subsection, we study the agent's optimal consumption rate and bequest by deriving the dynamics of the optimal consumption rate. We start by fixing $q = \frac{\kappa(1-\rho)}{1-\kappa}$ to eliminate one of the non-linearities in Equation (4.11). We then get the ODE

$$\begin{aligned}
g_t(t) = &-\frac{1-\kappa}{\kappa}\left[r + \frac{\lambda^2}{2\rho\sigma^2} + \hat{\mu}(t) - \frac{\delta}{1-\phi} - \frac{\mu(t)}{1-\kappa}\right] g(t) \\
&- \frac{\phi(1-\kappa)}{\kappa(1-\phi)}\delta^{\frac{1}{\phi}}\left(g(t)\right)^{-\frac{\kappa(1-\phi)}{\phi(1-\kappa)}+1} - \left(\hat{\mu}(t)\right)^{1-\frac{1}{\kappa}}\left(\mu(t)\right)^{\frac{1}{\kappa}} \ , \\
g(T) = &\ 1 \ .
\end{aligned} \tag{4.12}$$

Also, the optimal controls read

$$\begin{aligned} \pi^*(t,x) &= \frac{\lambda}{\rho\sigma^2}\left(x+L(t)\right) \ , \\ c^*(t,x) &= \delta^{\frac{1}{\phi}}\left(g(t)\right)^{-\frac{\kappa(1-\phi)}{\phi(1-\kappa)}}\left(x+L(t)\right) \ , \\ d^*(t,x) &= \left(\frac{\mu(t)}{\hat{\mu}(t)}\right)^{\frac{1}{\kappa}}\left(g(t)\right)^{-1}\left(x+L(t)\right)-x \\ &= c^*(t,x)\,\delta^{-\frac{1}{\phi}}\left(\frac{\mu(t)}{\hat{\mu}(t)}\right)^{\frac{1}{\kappa}}\left(g(t)\right)^{\frac{\kappa(1-\phi)}{\phi(1-\kappa)}-1}-x \ . \end{aligned} \tag{4.13}$$

We denote the optimal bequest by b^*, i.e. $b^*(t,x)=x+d^*(t,x)$. Since the optimal bequest can be expressed in terms of the optimal consumption rate c^*, we start by focusing on c^*. In Appendix 4.D, we show that the optimal consumption rate has the dynamics

$$\begin{aligned} \frac{\mathrm{d}c^*(t,X_t^*)}{c^*(t,X_t^*)} &= \frac{1}{\phi}\left[r+(1+\phi)\frac{\lambda^2}{2\rho\sigma^2}+\hat{\mu}(t)-\delta-\mu(t)\frac{1-\phi}{1-\kappa}\right]\mathrm{d}t \\ &\quad+\left(\frac{\kappa(1-\phi)}{\phi(1-\kappa)}-1\right)\left(\hat{\mu}(t)\right)^{1-\frac{1}{\kappa}}\left(\mu(t)\right)^{\frac{1}{\kappa}}\left(g(t)\right)^{-1}\mathrm{d}t+\frac{\lambda}{\rho\sigma}\mathrm{d}W_t \ , \\ c^*(0,X_0) &= \left(x_0+L(0)\right)\delta^{\frac{1}{\phi}}\left(g(0)\right)^{-\frac{\kappa(1-\phi)}{\phi(1-\kappa)}} \ . \end{aligned}$$

The dynamics allow us to simulate the optimal consumption rate and thereby also the optimal bequest using the relation

$$b^*(t,X_t^*) = c^*(t,X_t^*)\,\delta^{-\frac{1}{\phi}}\left(\frac{\mu(t)}{\hat{\mu}(t)}\right)^{\frac{1}{\kappa}}\left(g(t)\right)^{\frac{\kappa(1-\phi)}{\phi(1-\kappa)}-1} \ .$$

Simulating the optimal consumption rate (obviously) results in a wide range of optimal consumption paths for each choice of parameters. This makes interpretation difficult, and instead we focus on the expected optimal consumption rate $\hat{c}^*$ which is given by the ODE

$$\begin{aligned} \frac{\hat{c}_t^*(t)}{\hat{c}^*(t)} &= \frac{1}{\phi}\left[r+(1+\phi)\frac{\lambda^2}{2\rho\sigma^2}+\hat{\mu}(t)-\delta-\mu(t)\frac{1-\phi}{1-\kappa}\right] \\ &\quad+\left(\frac{\kappa(1-\phi)}{\phi(1-\kappa)}-1\right)\left(\hat{\mu}(t)\right)^{1-\frac{1}{\kappa}}\left(\mu(t)\right)^{\frac{1}{\kappa}}\left(g(t)\right)^{-1} \ , \\ \hat{c}^*(0) &= \left(x_0+L(0)\right)\delta^{\frac{1}{\phi}}\left(g(0)\right)^{-\frac{\kappa(1-\phi)}{\phi(1-\kappa)}} \ . \end{aligned} \tag{4.14}$$

The expected optimal bequest $\hat{b}^*$ is given by

$$\hat{b}^*(t) = \hat{c}^*(t)\,\delta^{-\frac{1}{\phi}}\left(\frac{\mu(t)}{\hat{\mu}(t)}\right)^{\frac{1}{\kappa}}\left(g(t)\right)^{\frac{\kappa(1-\phi)}{\phi(1-\kappa)}-1} \ . \tag{4.15}$$

In the next subsection, we comment on the size and shape of $\hat{c}^*$ and $\hat{b}^*$ in a numerical example.

4.6.2 Numerical example

We consider a female agent who is 30 years old at time 0 and has an initial wealth of $x_0 = 10,000$ USD. She starts off with a yearly labour income at rate $20,000$ USD, and her labour income grows with 2% each year due to inflation. We do not take taxes into account. She retires at age 65. Altogether, her labour income rate is given by

$$w(t) = 20,000 \cdot e^{0.02t} \cdot 1_{\{30+t \leq 65\}} \ .$$

Following Kraft et al. (2013), we choose the following market and preference parameters values:

r	σ	λ	δ	ρ	ϕ
0.05	0.20	0.07	0.08	2	8

This choice of ρ and ϕ places us in the second of the four cases on page 94. Hence, we only consider $\kappa \geq 2$. We set both μ and $\hat{\mu}$ equal to the G82 mortality intensity for a female aged 30 at time 0, i.e.

$$\hat{\mu}(t) = \mu(t) = 5 \cdot 10^{-4} + 5.3456 \cdot 10^{-5} \cdot e^{0.087498(30+t)} \ .$$

For the last three decades, the gender specific G82 mortality intensities have served as standard mortality intensities for adults in Denmark. We fix the time-horizon $T = 80$ since the probability of surviving the age of 110 is very small with the G82 mortality.

We study the agent's optimal consumption rate $\hat{c}^*$ and bequest $\hat{b}^*$ numerically with particular focus on the elasticity of substitution between bequest and future utility, $\frac{1}{\kappa}$. We solve the ODEs in Equations (4.12) and (4.14) numerically by use of a simple Euler scheme. In Figures 4.1–4.2, we plot the expected optimal consumption rate and bequest for different values of κ. The optimal bequest is computed using the relation in Equation (4.15).

We notice that the expected optimal consumption rate in Figure 4.1 starts out the same for all values of κ whereas the expected optimal bequest in Figure 4.2 starts out high for low values of κ and vice versa. For low values of κ, the expected optimal consumption rate and bequest are humped-shaped, meaning that they increase and then decrease. For high values of κ, the expected optimal consumption rate and bequest are strictly increasing over time. To explain the differences, we remember the following rules of thumb: For low ages, life insurance is cheap because of the low mortality. For high ages, consumption is cheap because of the low survival probability. The agent's utility from consumption is decreasing in mortality since the agent only consumes while alive. The agent's utility from bequest is increasing in mortality since the bequest only comes into play when the agent dies.

A high value of κ corresponds to a low elasticity of substitution between bequest and future utility. Hence, an agent with a high value of κ hardly

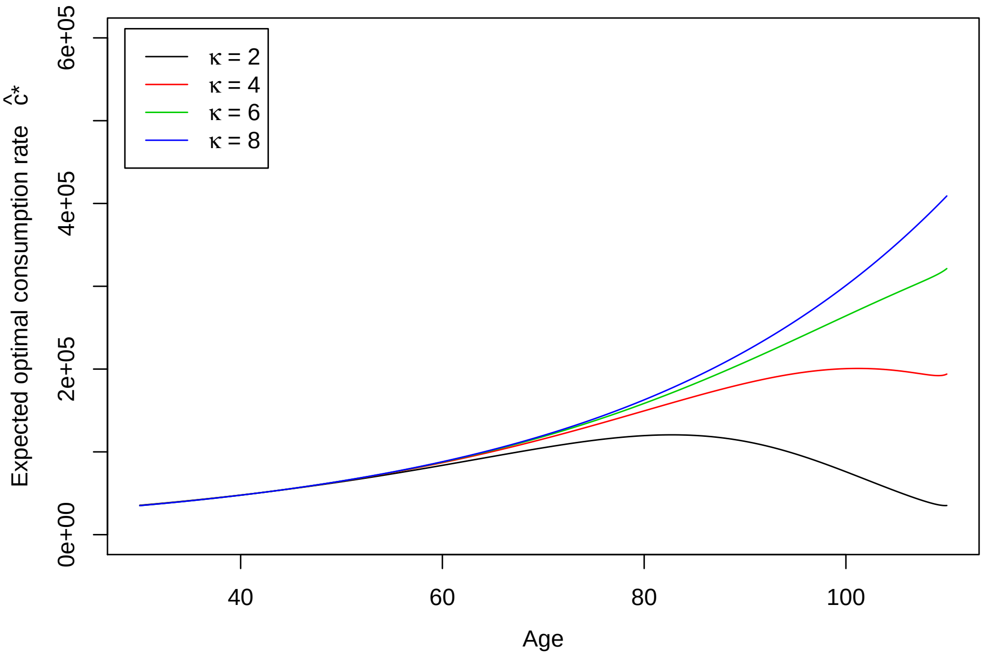

Figure 4.1: Expected optimal consumption rate as a function of age.

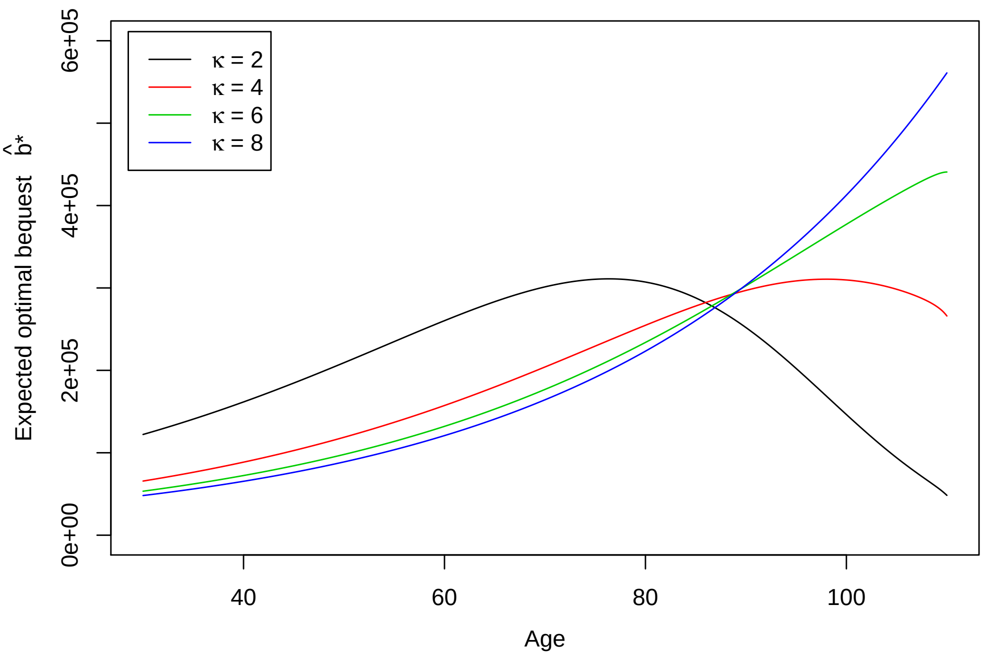

Figure 4.2: Expected optimal bequest as a function of age.

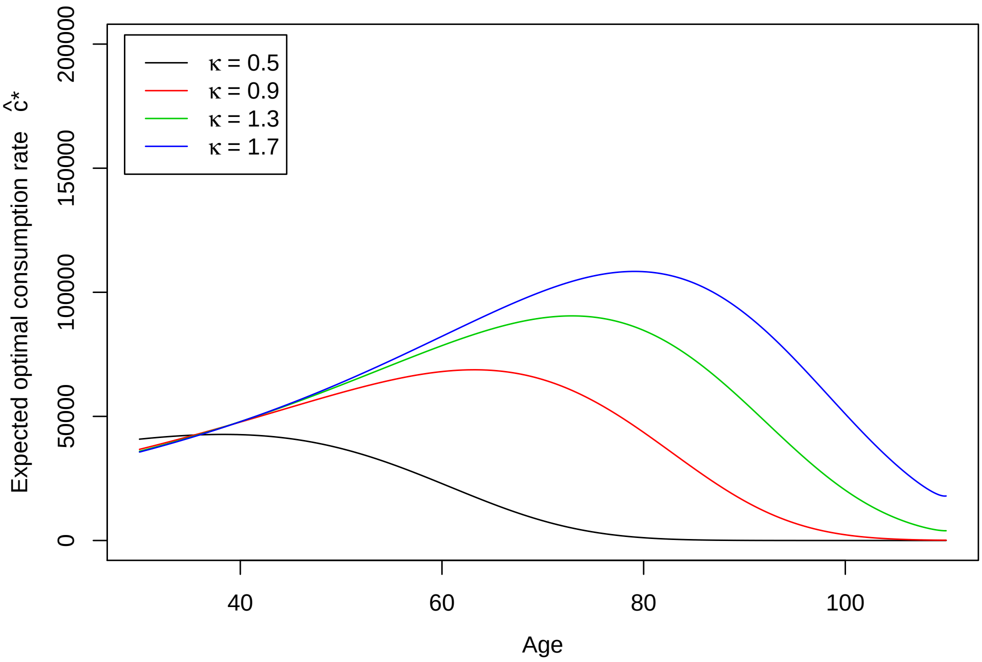

Figure 4.3: Expected optimal consumption rate as a function of age for low values of κ.

benefits from bequest in the early years when the probability of dying (and leaving bequest) is low. This means that the agent cannot benefit from the cheap life insurance early in life. Instead, the agent has a high consumption rate and bequest late in life when consumption is cheap and when she benefits more from bequest because of the increased mortality.

A low value of κ corresponds to a high elasticity of substitution between bequest and future utility. Hence, an agent with a low value of κ has more freedom to exploit the cheap life insurance early in life. The agent's hump-shaped bequest can be explained by the following two competing effects: On one hand, increased mortality means that the agents benefits more from bequest. On the other hand, increased mortality means more expensive life insurance. The agent's bequest increases as long as the first effect is stronger. The value of κ decides when the second effect takes over and the bequest starts to decrease. The agent's hump-shaped consumption rate can be explained in much the same way: On one hand, increased mortality means cheaper consumption. On the other hand, increased mortality means that the agents benefits less from consumption. The agent's consumption rate increases as long as the first effect is stronger. Again, the value of κ decides when the second effect takes over and the consumption rate starts to decrease.

Values of κ below 2 are not covered by the four cases on page 94. However, if we consider values of κ between 0 and 2, we get Figures 4.3 and 4.4 (beware

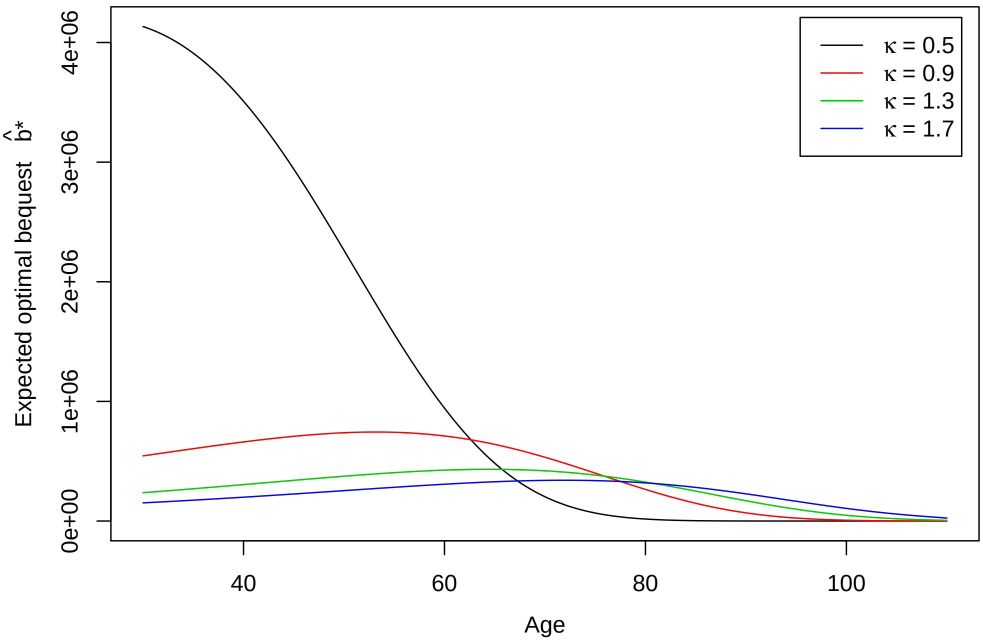

Figure 4.4: Expected optimal bequest as a function of age for low values of κ.

of the different scaling of the value axes). We notice that for κ tending to zero, the agent buys more and more life insurance, and she does so earlier and earlier in life–while consuming less and less. This is because she has a high elasticity of substitution between bequest and future utility which allows her to exploit the cheap life insurance early in life.

Hump-shaped consumption patterns are observed in realized consumption, see e.g. Bullard and Feigenbaum (2007); Feigenbaum (2008); Gourinchas and Parker (2002). Hump-shaped consumption patterns cannot be obtained by standard recursive utility or time-additive utility under lifetime uncertainty. It is the combination of recursive utility and lifetime uncertainty that enables this interesting feature. In Feigenbaum (2008), hump-shaped consumption patterns are obtained in a general time-additive equilibrium model with mortality risk, but only by excluding access to life insurance which is no innocuous assumption. In Gourinchas and Parker (2002); Bullard and Feigenbaum (2007), hump-shaped consumption patterns are explained by income uncertainty and utility from leisure, but not without a significant increase in complexity of the model. Our model, on the other hand, offers an explanation for hump-shaped consumption patterns in a simple model.

4.6.3 Comparison to Jensen and Steffensen (2015)

In the previous section, we studied optimal consumption, investment, and life insurance choice under separation of (market) risk aversion, elasticity of inter-temporal substitution, and elasticity of substitution between bequest and future utility. A similar separation is obtained in Jensen and Steffensen (2015), and, there, the separation also gives rise to hump-shaped consumption patterns. It is natural to consider if the two approaches cover the same set of preferences. By construction the two ways of separating preferences are very different. Recursive utility separates preferences by recursively building up a value function of local certainty equivalents of future (indirect) utility. Jensen and Steffensen (2015) separate preferences by building up a value function as a non-linear global aggregation of certainty equivalents of future consumption and bequest. The latter leads to time-consistency issues which are overcome using equilibrium theory.

In Jensen and Steffensen (2015), the agent's preferences are described by the aggregator

$$f\left(t,c,b,v\right)=\theta\delta v\left(\left(\frac{c^{1-\gamma}}{v\left(1-\gamma\right)}\right)^{\frac{1}{\kappa}}+\left(\frac{\varepsilon\left(t\right)\mu\left(t\right)b^{1-\gamma}}{v\left(1-\gamma\right)}\right)^{\frac{1}{\kappa}}\right)^{\frac{\kappa}{\theta}}-\left(\mu\left(t\right)+\theta\delta\right)v,$$

where γ models risk aversion, $\frac{1}{\phi}$ models elasticity of inter-temporal substitution, κ is a parameter in the modeling of substitution between consumption and bequest, and ε is a deterministic weight function allowing for a different weight on inheritance than on consumption and for a changing weight on inheritance throughout life. In the previous section, the agent's preferences are described by the aggregator

$$\begin{aligned} f\left(t,c,b,v\right)=\mu\left(t\right)\frac{1-\rho}{1-\kappa}v&\left(\frac{b^{1-\kappa}}{\left(\left(1-\rho\right)v\right)^{\frac{1-\kappa}{1-\rho}}}-1\right)\\ +\delta\frac{1-\rho}{1-\phi}v&\left(\frac{c^{1-\phi}}{\left(\left(1-\rho\right)v\right)^{\frac{1-\phi}{1-\rho}}}-1\right)\ . \end{aligned}$$

It is clear that the two aggregators cannot be reparameterized to coincide without reducing the number of free parameters. Hence, the aggregators do, in general, not cover the same set of preferences. Consumption and bequest are somehow more entangled in the aggregator from Jensen and Steffensen (2015), and, from the general aggregator in Equation (4.4), it is apparent that we cannot replicate this entanglement with a different specification of u, v, and $\bar{f}$. This means that the two ways of separating preferences are not only different in their construction, but also in output.

However, if we reduce the number of free parameters, the two aggregators can be reparameterized to almost coincide. Letting $\kappa=\theta=\frac{1-\gamma}{1-\phi}$ and $\varepsilon\left(t\right)=\mu\left(t\right)^{-1+\theta}$ in the aggregator from Jensen and Steffensen (2015), we obtain

$$f\left(t,c,b,v\right)=\theta v\left(\delta\frac{c^{1-\phi}}{\left(v\left(1-\gamma\right)\right)^{\frac{1}{\theta}}}+\delta\mu\left(t\right)\frac{b^{1-\phi}}{\left(v\left(1-\gamma\right)\right)^{\frac{1}{\theta}}}\right)-\left(\tfrac{1}{\theta}\mu\left(t\right)+\delta\right)\theta v\ .$$

Similarly, letting $\kappa = \phi$ and $\bar{\theta} = \frac{1-\rho}{1-\phi}$ in the aggregator from this paper, we obtain

$$f(t,c,b,v) = \bar{\theta} v \left(\delta \frac{c^{1-\phi}}{((1-\rho)v)^{\frac{1}{\bar{\theta}}}} + \mu(t) \frac{b^{1-\phi}}{((1-\rho)v)^{\frac{1}{\bar{\theta}}}} \right) - (\mu(t) + \delta) \bar{\theta} v \ .$$

Now, the aggregators are seen to coincide up to scaling of the mortality intensity. Hence, if we give up on the threefold separability, there is some coincidence between the two ways of separating preferences, but not for general preferences.

4.7 Conclusion

Recursive utility plays an important role in the literature on optimal consumption and investment choice for agents with a certain lifetime, but to the knowledge of the author, it has never before been generalized to agents with an uncertain lifetime. We generalize recursive utility to include lifetime uncertainty and utility from bequest. Recursive utility allows for separation of preferences for risk and inter-temporal substitution, and, with our generalization, also preferences for substitution between bequest and future utility. The concept of substitution between bequest and future utility is our way of formulating preferences for mortality risk. We state a verification theorem with a generalized Hamilton-Jacobi-Bellman equation for optimal control under recursive utility with lifetime uncertainty. We study optimal consumption, investment, and life insurance choice under separation of (market) risk aversion, elasticity of inter-temporal substitution, and elasticity of substitution between bequest and future utility. The separation gives rise to hump-shaped consumption patterns as observed in realized consumption. The hump-shaped consumption is a result of the following two non-linear effects of increased mortality; cheaper consumption and lower utility from consumption. The consumption rate increases as long as the first effect is stronger and then starts to decrease.

Appendix

4.A Derivation of Continuous-Time Aggregator

Differentiating the normalized discrete-time aggregator $\tilde{W}$, we get

$$\begin{aligned}
&\tilde{W}_\Delta\left(t,\Delta,c,b,\tilde{V}\right)\\
&=\left(u\circ v^{-1}\right)'\left(p_{t-\Delta,t}\left(u\circ v^{-1}\right)^{-1}\left(\bar{W}\left(\Delta,c,\tilde{V}\right)\right)+\left(1-p_{t-\Delta,t}\right)v\left(b\right)\right)\\
&\quad\times p_{t-\Delta,t}\Bigg(\left(\left(u\circ v^{-1}\right)^{-1}\right)'\left(\bar{W}\left(\Delta,c,\tilde{V}\right)\right)\bar{W}_\Delta\left(\Delta,c,\tilde{V}\right)\\
&\qquad\qquad+\mu\left(t-\Delta\right)\left(v\left(b\right)-\left(u\circ v^{-1}\right)^{-1}\left(\bar{W}\left(\Delta,c,\tilde{V}\right)\right)\right)\Bigg)\,.
\end{aligned}$$

Using $\left(\left(u\circ v^{-1}\right)^{-1}\right)'=\left(\left(u\circ v^{-1}\right)'\circ\left(u\circ v^{-1}\right)^{-1}\right)^{-1}$, $\bar{W}\left(0,c,\tilde{V}\right)=\tilde{V}$, and $p_{t,t+0}=1$, we obtain

$$\begin{aligned}
&\tilde{W}_\Delta\left(t,0,c,b,\tilde{V}\right)\\
&=\left(u\circ v^{-1}\right)'\left(\left(u\circ v^{-1}\right)^{-1}\left(\tilde{V}\right)\right)\\
&\quad\times\left(\left(\left(u\circ v^{-1}\right)^{-1}\right)'\left(\tilde{V}\right)\bar{W}_\Delta\left(0,c,\tilde{V}\right)+\mu\left(t\right)\left(v\left(b\right)-\left(u\circ v^{-1}\right)^{-1}\left(\tilde{V}\right)\right)\right)\\
&=\left(u\circ v^{-1}\right)'\left(\left(u\circ v^{-1}\right)^{-1}\left(\tilde{V}\right)\right)\mu\left(t\right)\left(v\left(b\right)-v\left(u^{-1}\left(\tilde{V}\right)\right)\right)+\bar{W}_\Delta\left(0,c,\tilde{V}\right)\,.
\end{aligned}$$

Hence, we get

$$\begin{aligned}
f\left(t,c,b,\tilde{V}\right)&=\mu\left(t\right)\left(v\left(b\right)-v\left(u^{-1}\left(\tilde{V}\right)\right)\right)\left(u\circ v^{-1}\right)'\left(\left(u\circ v^{-1}\right)^{-1}\left(\tilde{V}\right)\right)\\
&\quad+\bar{f}\left(c,\tilde{V}\right)\,,
\end{aligned}$$

where $\bar{f}$ is the "classical" normalized continuous-time aggregator corresponding to $\bar{W}$.

4.B Derivation of Aggregators in Example 4.3

For the normalized discrete-time aggregator $\bar{W}$ in Example 4.1, we get

$$\begin{aligned}
\bar{W}_\Delta\left(\Delta,c,\tilde{V}\right)&=\left(u\circ g^{-1}\right)'\left(\left(1-\delta\Delta\right)\left(u\circ g^{-1}\right)^{-1}\left(\tilde{V}\right)+\delta\Delta g\left(c\right)\right)\\
&\quad\times\left(-\delta\left(u\circ g^{-1}\right)^{-1}\left(\tilde{V}\right)+\delta g\left(c\right)\right)\,.
\end{aligned}$$

Using $\left(\left(u\circ g^{-1}\right)^{-1}\right)' = \left(\left(u\circ g^{-1}\right)'\circ\left(u\circ g^{-1}\right)^{-1}\right)^{-1}$, we obtain

$$\bar{W}_\Delta\left(0,c,\tilde{V}\right) = \delta\left(u\circ g^{-1}\right)'\left(\left(u\circ g^{-1}\right)^{-1}\left(\tilde{V}\right)\right)\left(g\left(c\right) - g\left(u^{-1}\left(\tilde{V}\right)\right)\right) .$$

Therefore, the continuous-time aggregator corresponding to the normalized discrete-time aggregator $\bar{W}$ from Example 4.1 is given by

$$\begin{aligned}\bar{f}\left(c,\tilde{V}\right) &= \bar{W}_\Delta\left(0,c,\tilde{V}\right)\\ &= \delta\left(u\circ g^{-1}\right)'\left(\left(u\circ g^{-1}\right)^{-1}\left(\tilde{V}\right)\right)\left(g\left(c\right) - g\left(u^{-1}\left(\tilde{V}\right)\right)\right)\\ &= \delta\frac{u'\left(u^{-1}\left(\tilde{V}\right)\right)}{g'\left(u^{-1}\left(\tilde{V}\right)\right)}\left(g\left(c\right) - g\left(u^{-1}\left(\tilde{V}\right)\right)\right) .\end{aligned}$$

The corresponding mortality adjusted normalized aggregator f is given by

$$\begin{aligned}f\left(t,c,b,\tilde{V}\right) &= \mu\left(t\right)\left(v\left(b\right) - v\left(u^{-1}\left(\tilde{V}\right)\right)\right)\left(u\circ v^{-1}\right)'\left(\left(u\circ v^{-1}\right)^{-1}\left(\tilde{V}\right)\right)\\ &\quad + \delta\frac{u'\left(u^{-1}\left(\tilde{V}\right)\right)}{g'\left(u^{-1}\left(\tilde{V}\right)\right)}\left(g\left(c\right) - g\left(u^{-1}\left(\tilde{V}\right)\right)\right)\\ &= \mu\left(t\right)\frac{u'\left(u^{-1}\left(\tilde{V}\right)\right)}{v'\left(u^{-1}\left(\tilde{V}\right)\right)}\left(v\left(b\right) - v\left(u^{-1}\left(\tilde{V}\right)\right)\right)\\ &\quad + \delta\frac{u'\left(u^{-1}\left(\tilde{V}\right)\right)}{g'\left(u^{-1}\left(\tilde{V}\right)\right)}\left(g\left(c\right) - g\left(u^{-1}\left(\tilde{V}\right)\right)\right) .\end{aligned}$$

4.C Solution of Hamilton-Jacobi-Bellman equation

For the Hamilton-Jacobi-Bellman equation in Equation (4.9) with Epstein-Zin preferences, we conjecture that

$$J\left(t,x\right) = \frac{1}{1-\rho}\left(x + L\left(t\right)\right)^{1-\rho}\left(g\left(t\right)\right)^{q} ,$$

where q is a non-zero constant, $g:[0,T]\mapsto\mathbb{R}$ is a C^1-function with $g\left(T\right)=1$, and L is given by Equation (4.6). We recall that $L\left(T\right)=0$. We get the partial derivatives

$$J_x(t,x) = (x+L(t))^{-\rho}(g(t))^q = \frac{1-\rho}{x+L(t)}J(t,x) \ ,$$
$$J_{xx}(t,x) = -\rho(x+L(t))^{-\rho-1}(g(t))^q = -\frac{\rho(1-\rho)}{(x+L(t))^2}J(t,x) \ ,$$
$$\begin{aligned} J_t(t,x) &= \frac{q}{1-\rho}g_t(t)(x+L(t))^{1-\rho}(g(t))^{q-1} + (x+L(t))^{-\rho}(g(t))^q L_t(t) \\ &= \frac{g_t(t)\,q}{g(t)}J(t,x) + \frac{(1-\rho)L_t(t)}{x+L(t)}J(t,x) \ , \end{aligned}$$

where

$$L_t(t) = -w(t) + (r+\hat{\mu}(t))L(t) \ .$$

We note that the boundary condition is satisfied since

$$J(T,x) = \frac{1}{1-\rho}(x+0)^{1-\rho}(1)^q = \frac{1}{1-\rho}x^{1-\rho} = u(x) \ .$$

Furthermore, we have

$$\begin{aligned} &f(t,c,d+x,J(t,x)) \\ &= \mu(t)\frac{1-\rho}{1-\kappa}J(t,x)\left(\left(\frac{d+x}{((1-\rho)J(t,x))^{\frac{1}{1-\rho}}}\right)^{1-\kappa} - 1\right) \\ &\quad + \delta\frac{1-\rho}{1-\phi}J(t,x)\left(\left(\frac{c}{((1-\rho)J(t,x))^{\frac{1}{1-\rho}}}\right)^{1-\phi} - 1\right) \\ &= \mu(t)\frac{1-\rho}{1-\kappa}J(t,x)\left(\left(\frac{d+x}{\left((1-\rho)\frac{1}{1-\rho}(x+L(t))^{1-\rho}(g(t))^q\right)^{\frac{1}{1-\rho}}}\right)^{1-\kappa} - 1\right) \\ &\quad + \delta\frac{1-\rho}{1-\phi}J(t,x)\left(\left(\frac{c}{\left((1-\rho)\frac{1}{1-\rho}(x+L(t))^{1-\rho}(g(t))^q\right)^{\frac{1}{1-\rho}}}\right)^{1-\phi} - 1\right) \\ &= \mu(t)\frac{1-\rho}{1-\kappa}J(t,x)\left(\left(\frac{d+x}{(x+L(t))(g(t))^{\frac{q}{1-\rho}}}\right)^{1-\kappa} - 1\right) \\ &\quad + \delta\frac{1-\rho}{1-\phi}J(t,x)\left(\left(\frac{c}{(x+L(t))(g(t))^{\frac{q}{1-\rho}}}\right)^{1-\phi} - 1\right) \ . \end{aligned}$$

Plugging the above into Equation (4.9), we get the reduced Hamilton-Jacobi-Bellman equation

$$0=\sup_{c,\pi,d} J(t,x)\left\{\begin{array}{c}\frac{g_t(t)q}{g(t)}+\frac{(1-\rho)L_t(t)}{x+L(t)}\\ +\frac{1-\rho}{x+L(t)}\left[rx+\pi\lambda-c-\hat{\mu}(t)\,d+w(t)\right]\\ -\frac{1}{2}\frac{\rho(1-\rho)}{(x+L(t))^2}\pi^2\sigma^2\\ +\mu(t)\,\frac{1-\rho}{1-\kappa}\left(\left(\frac{d+x}{(x+L(t))(g(t))^{\frac{q}{1-\rho}}}\right)^{1-\kappa}-1\right)\\ +\delta\frac{1-\rho}{1-\phi}\left(\left(\frac{c}{(x+L(t))(g(t))^{\frac{q}{1-\rho}}}\right)^{1-\phi}-1\right)\end{array}\right\}. \quad (4.16)$$

The first-order conditions for the supremum read

$$\begin{aligned}0&=\frac{1-\rho}{x+L(t)}\lambda-\frac{\rho(1-\rho)}{(x+L(t))^2}\pi\sigma^2\ ,\\ 0&=-\frac{1-\rho}{x+L(t)}+\delta(1-\rho)(x+L(t))^{\phi-1}(g(t))^{-\frac{q(1-\phi)}{1-\rho}}c^{-\phi}\ ,\\ 0&=-\hat{\mu}(t)\frac{1-\rho}{x+L(t)}+\mu(t)(1-\rho)(x+L(t))^{\kappa-1}(g(t))^{-\frac{q(1-\kappa)}{1-\rho}}(d+x)^{-\kappa}\ .\end{aligned}$$

This gives us the candidate solutions

$$\begin{aligned}\pi^*(t)&=\frac{\lambda}{\rho\sigma^2}(x+L(t))\ ,\\ c^*(t)&=\delta^{\frac{1}{\phi}}(g(t))^{-\frac{q(1-\phi)}{\phi(1-\rho)}}(x+L(t))\ ,\\ d^*(t)&=\left(\frac{\mu(t)}{\hat{\mu}(t)}\right)^{\frac{1}{\kappa}}(g(t))^{-\frac{q(1-\kappa)}{\kappa(1-\rho)}}(x+L(t))-x\ .\end{aligned}$$

Plugging the candidate solutions back into Equation (4.16) and dividing by J, we obtain

$$
\begin{aligned}
0 = {} & \frac{g_t(t)\,q}{g(t)} + (1-\rho)\frac{-w(t) + (r+\hat{\mu}(t))\,L(t)}{x+L(t)} \\
& + \frac{1-\rho}{x+L(t)}\left[rx + \frac{\lambda}{\rho\sigma^2}(x+L(t))\,\lambda - \delta^{\frac{1}{\phi}}(g(t))^{-\frac{q(1-\phi)}{\phi(1-\rho)}}(x+L(t))\right] \\
& + \frac{1-\rho}{x+L(t)}\left[w(t) - \hat{\mu}(t)\left(\left(\frac{\mu(t)}{\hat{\mu}(t)}\right)^{\frac{1}{\kappa}}(g(t))^{-\frac{q(1-\kappa)}{\kappa(1-\rho)}}(x+L(t)) - x\right)\right] \\
& - \frac{1}{2}\frac{\rho(1-\rho)}{(x+L(t))^2}\left(\frac{\lambda}{\rho\sigma^2}(x+L(t))\right)^2\sigma^2 \\
& + \mu(t)\frac{1-\rho}{1-\kappa}\left(\left(\frac{\left(\frac{\mu(t)}{\hat{\mu}(t)}\right)^{\frac{1}{\kappa}}(g(t))^{-\frac{q(1-\kappa)}{\kappa(1-\rho)}}(x+L(t))}{(x+L(t))\,(g(t))^{\frac{q}{1-\rho}}}\right)^{1-\kappa} - 1\right) \\
& + \delta\frac{1-\rho}{1-\phi}\left(\left(\frac{\delta^{\frac{1}{\phi}}(g(t))^{-\frac{q(1-\phi)}{\phi(1-\rho)}}(x+L(t))}{(x+L(t))\,(g(t))^{\frac{q}{1-\rho}}}\right)^{1-\phi} - 1\right) \\
= {} & \frac{g_t(t)\,q}{g(t)} + \frac{(1-\rho)\,\lambda^2}{\rho\sigma^2} - (1-\rho)\left[\delta^{\frac{1}{\phi}}(g(t))^{-\frac{q(1-\phi)}{\phi(1-\rho)}} + (\hat{\mu}(t))^{1-\frac{1}{\kappa}}(\mu(t))^{\frac{1}{\kappa}}\right] \\
& + \frac{1-\rho}{x+L(t)}\left[rx + (r+\hat{\mu}(t))\,L(t) + \hat{\mu}(t)\,x\right] \\
& - \frac{1-\rho}{2}\frac{\lambda^2}{\rho\sigma^2} + \frac{1-\rho}{1-\kappa}(\hat{\mu}(t))^{1-\frac{1}{\kappa}}(\mu(t))^{\frac{1}{\kappa}}(g(t))^{-\frac{q(1-\kappa)}{\kappa(1-\rho)}} - \mu(t)\frac{1-\rho}{1-\kappa} \\
& + \frac{1-\rho}{1-\phi}\delta^{\frac{1}{\phi}}(g(t))^{-\frac{q(1-\phi)}{\phi(1-\rho)}} - \delta\frac{1-\rho}{1-\phi} \\
= {} & \frac{g_t(t)\,q}{g(t)} + (1-\rho)\left[r + \frac{\lambda^2}{2\rho\sigma^2} + \hat{\mu}(t) - \frac{\delta}{1-\phi} - \frac{\mu(t)}{1-\kappa}\right] \\
& + \frac{\phi(1-\rho)}{1-\phi}\delta^{\frac{1}{\phi}}(g(t))^{-\frac{q(1-\phi)}{\phi(1-\rho)}} \\
& + \frac{\kappa(1-\rho)}{1-\kappa}(\hat{\mu}(t))^{1-\frac{1}{\kappa}}(\mu(t))^{\frac{1}{\kappa}}(g(t))^{-\frac{q(1-\kappa)}{\kappa(1-\rho)}} \quad .
\end{aligned}
$$

Finally, dividing by $\frac{q}{g}$ and including the boundary condition, we arrive at

$$\begin{aligned} g_t(t) = &-\frac{1}{q}(1-\rho)\left[r + \frac{\lambda^2}{2\rho\sigma^2} + \hat{\mu}(t) - \frac{\delta}{1-\phi} - \frac{\mu(t)}{1-\kappa}\right] g(t) \\ &- \frac{1}{q}\frac{\phi(1-\rho)}{1-\phi}\delta^{\frac{1}{\phi}}(g(t))^{-\frac{q(1-\phi)}{\phi(1-\rho)}+1} \\ &- \frac{1}{q}\frac{\kappa(1-\rho)}{1-\kappa}(\hat{\mu}(t))^{1-\frac{1}{\kappa}}(\mu(t))^{\frac{1}{\kappa}}(g(t))^{-\frac{q(1-\kappa)}{\kappa(1-\rho)}+1} \ , \\ g(T) = 1 \ . \end{aligned}$$

4.D Derivation of Dynamics for the Optimal Consumption Rate

Differentiating the optimal consumption rate in Equation (4.13), we obtain the partial derivatives

$$\begin{aligned} c_t^*(t,x) &= -\frac{\kappa(1-\phi)}{\phi(1-\kappa)}(x+L(t))\,\delta^{\frac{1}{\phi}}(g(t))^{-\frac{\kappa(1-\phi)}{\phi(1-\kappa)}-1} g_t(t) \\ &\quad + \delta^{\frac{1}{\phi}}(g(t))^{-\frac{\kappa(1-\phi)}{\phi(1-\kappa)}} L_t(t) \ , \\ c_x^*(t,x) &= \delta^{\frac{1}{\phi}}(g(t))^{-\frac{\kappa(1-\phi)}{\phi(1-\kappa)}} \ , \\ c_{xx}^*(t,x) &= 0 \ . \end{aligned}$$

Let X^* be the wealth process stemming from the optimal control (c^*, π^*, d^*) in Equation (4.13). Using Itô's formula on $c^*(t, X_t^*)$, we get the dynamics

$$\begin{aligned} \frac{\mathrm{d}c^*(t,X_t^*)}{c^*(t,X_t^*)} &= \frac{c_t^*(t,X_t^*)\,\mathrm{d}t + c_x^*(t,X_t^*)\,\mathrm{d}X_t^*}{c^*(t,X_t^*)} \\ &= \frac{-\frac{\kappa(1-\phi)}{\phi(1-\kappa)}(X_t^*+L(t))\,\delta^{\frac{1}{\phi}} g_t(t)\,(g(t))^{-\frac{\kappa(1-\phi)}{\phi(1-\kappa)}-1}\,\mathrm{d}t}{(X_t^*+L(t))\,\delta^{\frac{1}{\phi}}(g(t))^{-\frac{\kappa(1-\phi)}{\phi(1-\kappa)}}} \\ &\quad + \frac{\delta^{\frac{1}{\phi}}(g(t))^{-\frac{\kappa(1-\phi)}{\phi(1-\kappa)}} L_t(t)\,\mathrm{d}t}{(X_t^*+L(t))\,\delta^{\frac{1}{\phi}}(g(t))^{-\frac{\kappa(1-\phi)}{\phi(1-\kappa)}}} \\ &\quad + \frac{\delta^{\frac{1}{\phi}}(g(t))^{-\frac{\kappa(1-\phi)}{\phi(1-\kappa)}}\,\mathrm{d}X_t^*}{(X_t^*+L(t))\,\delta^{\frac{1}{\phi}}(g(t))^{-\frac{\kappa(1-\phi)}{\phi(1-\kappa)}}} \\ &= -\frac{\kappa(1-\phi)}{\phi(1-\kappa)}\frac{g_t(t)}{g(t)}\mathrm{d}t + \frac{L_t(t)\,\mathrm{d}t + \mathrm{d}X_t^*}{X_t^*+L(t)} \ . \end{aligned}$$

It holds that

$$
\begin{aligned}
-\frac{\kappa(1-\phi)}{\phi(1-\kappa)}\frac{g_t(t)}{g(t)} &= \frac{1-\phi}{\phi}\left[r+\frac{\lambda^2}{2\rho\sigma^2}+\hat{\mu}(t)-\frac{\delta}{1-\phi}-\frac{\mu(t)}{1-\kappa}\right] \\
&\quad+\delta^{\frac{1}{\phi}}(g(t))^{-\frac{\kappa(1-\phi)}{\phi(1-\kappa)}} \\
&\quad+\frac{\kappa(1-\phi)}{\phi(1-\kappa)}(\hat{\mu}(t))^{1-\frac{1}{\kappa}}(\mu(t))^{\frac{1}{\kappa}}(g(t))^{-1}
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{L_t(t)\,\mathrm{d}t+\mathrm{d}X_t^*}{X_t^*+L(t)} &= \frac{-w(t)+(r+\hat{\mu}(t))L(t)}{X_t^*+L(t)}\mathrm{d}t+\frac{\pi_t^*\sigma}{X_t^*+L(t)}\mathrm{d}W_t \\
&\quad+\frac{rX_t^*+\pi_t^*\lambda-c_t^*-\hat{\mu}(t)d_t^*+w(t)}{X_t^*+L(t)}\mathrm{d}t \\
&= \left(r+\frac{\lambda^2}{\rho\sigma^2}-\delta^{\frac{1}{\phi}}(g(t))^{-\frac{\kappa(1-\phi)}{\phi(1-\kappa)}}\right)\mathrm{d}t \\
&\quad+\left(\hat{\mu}(t)-(\hat{\mu}(t))^{1-\frac{1}{\kappa}}(\mu(t))^{\frac{1}{\kappa}}(g(t))^{-1}\right)\mathrm{d}t+\frac{\lambda}{\rho\sigma}\mathrm{d}W_t\ .
\end{aligned}
$$

Hence, we arrive at the dynamics

$$
\begin{aligned}
\frac{\mathrm{d}c^*(t,X_t^*)}{c^*(t,X_t^*)} &= \frac{1-\phi}{\phi}\left[r+\frac{\lambda^2}{2\rho\sigma^2}+\hat{\mu}(t)-\frac{\delta}{1-\phi}-\frac{\mu(t)}{1-\kappa}\right]\mathrm{d}t \\
&\quad+\delta^{\frac{1}{\phi}}(g(t))^{-\frac{\kappa(1-\phi)}{\phi(1-\kappa)}}\mathrm{d}t \\
&\quad+\frac{\kappa(1-\phi)}{\phi(1-\kappa)}(\hat{\mu}(t))^{1-\frac{1}{\kappa}}(\mu(t))^{\frac{1}{\kappa}}(g(t))^{-1}\mathrm{d}t \\
&\quad\left(r+\frac{\lambda^2}{\rho\sigma^2}-\delta^{\frac{1}{\phi}}(g(t))^{-\frac{\kappa(1-\phi)}{\phi(1-\kappa)}}\right)\mathrm{d}t \\
&\quad+\left(\hat{\mu}(t)-(\hat{\mu}(t))^{1-\frac{1}{\kappa}}(\mu(t))^{\frac{1}{\kappa}}(g(t))^{-1}\right)\mathrm{d}t+\frac{\lambda}{\rho\sigma}\mathrm{d}W_t \\
&= \frac{1}{\phi}\left[r+(1+\phi)\frac{\lambda^2}{2\rho\sigma^2}+\hat{\mu}(t)-\delta-\mu(t)\frac{1-\phi}{1-\kappa}\right]\mathrm{d}t \\
&\quad+\left(\frac{\kappa(1-\phi)}{\phi(1-\kappa)}-1\right)(\hat{\mu}(t))^{1-\frac{1}{\kappa}}(\mu(t))^{\frac{1}{\kappa}}(g(t))^{-1}\mathrm{d}t \\
&\quad+\frac{\lambda}{\rho\sigma}\mathrm{d}W_t\ , \\
c^*(0,X_0) &= (x_0+L(0))\,\delta^{\frac{1}{\phi}}(g(0))^{-\frac{\kappa(1-\phi)}{\phi(1-\kappa)}}\ .
\end{aligned}
$$

Chapter 5

Smooth Investment

Kenneth Bruhn, Ninna Reitzel Jensen, & Mogens Steffensen
(2016)

Abstract: In the classical portfolio optimization problem considered by Merton, the resulting constant proportion investment plan requires a diffusive trading strategy. This means that, within any arbitrarily small time interval, the investor has to both buy and sell stocks. From a practical point of view, this can serve as nothing else than an idealistic benchmark. We study the problems of a mean-square and a power utility investor for whom the trading strategy is constrained to be smooth, i.e. non-diffusive. In particular, this means that, over sufficiently small time intervals, the investor is either a seller or a buyer of stocks. The mathematical framework is built around quadratic objectives such that trading activity is punished quadratically. Mean-square utility is in itself quadratic, and power utility is covered by quadratic punishment of distance to Merton's power utility portfolio. We present semi-explicit solutions and, in a series of numerical illustrations, we show the impact of trading constraints on the portfolio decision over the investment horizon.

Keywords: Smooth investment, diffusive trading, trading costs, power utility, mean-square utility.

5.1 Introduction

We solve two portfolio optimization problems with the common feature that the number of stocks is restricted to be differentiable such that diffusive trading is avoided. Diffusive trading, e.g. resulting from continuous rebalancing of the constant proportion portfolio in a Merton problem, is here disallowed by punishing, in the utility function, quadratic deviations of the trading rate away from zero. The trading objective is economically and mathematically motivated here in the introduction. We solve the problems semi-explicitly for both a mean-square investor and a tailor-made version of a power utility investor. A mean-square investor is known to be equivalent to a mean-variance investor with precommitment, see e.g. Zhou and Li (2000). We discuss the structures of the solutions and illustrate trading and stock positions in a series of numerical examples. Our contribution is to cover how far we can get under the restriction of smooth stock positions within a mathematically tractable and economical meaningful framework. Further, we provide structures of portfolios with trading costs mainly stemming from market impact. This takes the oneliner by Gârleanu and Pedersen (2013) of “aiming in front of the target” to the level of conventional utility of wealth rather than utility of return on wealth.

The canonical approach to continuous-time dynamic portfolio choice was proposed by Merton (1971) who optimized the proportion of diffusively risky stocks in a portfolio of an investor maximizing utility of terminal wealth and/or consumption. With π, X, N, and S denoting the stock proportion, the wealth, the number of stocks, and the stock price, respectively, we have that the amount invested in stocks A has the following different representations,

$$A = \pi X = NS \ .$$

Whether the problem is formulated as a control problem over A, π, or N makes no difference for the resulting stock position. Yet, most often the problem is formulated with π as the control process, mainly due to mathematical tractability in markets without frictions. The well-known result for power utility is that π is constant over time and, in that case, for diffusive modelling of the stock price, also N in general becomes diffusive. A diffusive N is not limited to the case of a power utility investor, but it is immediate to conclude in that case due to the simple form of π. Only in the singular cases $\pi = 0$ (buy and hold no stocks) and $\pi = 1$ (buy and hold stocks for all your money), a diffusive N is avoided.

Diffusive stock numbers are not for real. Apart from it being technically impossible to trade diffusively, also trading costs prevent such a behavior from being optimal. Trading costs are here thought of as the integrate effects from broker expenses and market impact. For both effects, the intuition is that trading should be limited in order to limit the aggregate trading costs.

There are essentially two different ways to punish trading in the control problem formulation. One way is to implement a cost of trading directly in the wealth process such that a specific amount is withdrawn from the wealth if a trade is made. This has been studied in many different works on trading costs. When trading costs are broker expenses paid by a marginal investor with no market impact, this formalization has been studied by e.g. Magill and Constantinides (1976); Davis and Norman (1990); Oksendal and Sulem (2002); Chellathurai and Draviam (2007). When trading costs arise from market impact for a large investor with negligible broker expenses, this formalization has been studied by e.g. Bank and Baum (2004); Bertsimas and Lo (1998); Moazeni et al. (2010); Soner and Vukelja (2013). The formalization explained in this paragraph specifies a direct impact of trading costs on wealth such that the wealth is, at any time, reduced with the nominal cost of trading. At a terminal time point, utility of wealth, including an aggregate loss from trading, can be measured.

An alternative way of punishing trading in the control problem is to formalize preferences for not trading in the objective function. This is, of course, an indirect way of specifying trading costs since, probably, no investor has really explicit preferences for or against trading, only through its direct impact on wealth. The wealth is not reduced by trading costs, but terminal utility of wealth is measured along with a punishment for the aggregate trading made during the course of investment. Still, it may be meaningful to work with this indirect formalization of trading costs, e.g. due to mathematical tractability. This is the way we explore throughout this paper.

A different way of combining preferences for wealth and trading costs was proposed by Gârleanu and Pedersen (2013, 2014). Their idea is to specify preferences for return on wealth rather than on wealth itself. In particular, their objective corresponds to a mean-variance preference for return on wealth. This approach has some advantages since, in combination with quadratic utility of return on wealth, the preferences for trading can be, consistently, interpreted as direct trading costs: The current trading costs instantly reduce the return on wealth such that direct measurement of quadratic trading costs in the wealth dynamics and indirect measurement of trading costs in the objective function coincide. However, this consistency is obtained only because the objective is formalized in terms of return on wealth and not in terms of wealth itself. Utility from the return of wealth rather than utility of wealth appears in itself to be an indirect approach. Gârleanu and Pedersen (2013, 2014) point out that *"[...]The objective can be shown to approximate a standard utility function [...]"*. To us, the distinction between utility of return on wealth versus wealth appears non-bridgeable in a world with trading costs, though. It is by no means clear that the short cut they make via measuring utility from return on wealth is economically better than the short cut we make by specifying no direct impact of trading on wealth but an indirect punishment via preferences for no trading in combination with classical utility of wealth. Both short cuts serve the mathematical tractability of the problem.

Given the approach to punish trading amounts in the value function of the control problem, we need to formalize this punishment. First we choose as control process the change in N rather than N itself. This does not, upfront, make any difference since specifying the two processes are equivalent. However, we now restrict the trading $\mathrm{d}N$ in two steps. First, we rule out diffusive trading by saying that N has to be differentiable, i.e. $\mathrm{d}N$ has to be of order $\mathrm{d}t$. This, at least in a world of infinite divisibility, appears much more natural than a diffusive N. Second, we punish the trading amount rate, $S\,\mathrm{d}N/\,\mathrm{d}t$, for being away from zero. The first step is identical to Longstaff (2001), but, there, trading is not punished, and the result is a so-called bang-bang investment strategy.

Given the approach to punish the trading amount rate, we need to choose a cost function. If the only objective of this cost function were to account for indirect broker expenses paid by a marginal investor, one would probably prefer a function which is zero at zero, concave on the positive half line, and symmetric in zero. Concavity reflects some kind of “quantity discount” on trading. Adding e.g. fixed costs to proportional costs would correspond to “quantity discount” reflected in a concave cost functional on the positive half line. Due to symmetrization in zero, the function would obviously not be concave across zero but only piecewise concave on the positive and negative half-lines. However, in this paper, costs also account for market impact. Here an effect opposite to “quantity discount” kicks in. A large investor experiences a convex cost of trading since limit order books produce an impact on prices that, even marginally, increases with the volume of a trade. The cost functional in the value function is supposed to account, in a stylized way, for the aggregate impact of broker expenses and market impact. Here, we work with a quadratic cost function. A quadratic function as cost functional is a particularly tractable cost function with the desired convex form that would hold for a trader for whom the effect from market impact dominates that from broker expenses.

We solve two different problems with a quadratic utility loss from trading. First, we consider an investor with mean-square utility of terminal wealth. In that problem formulation, the mean-square utility of terminal wealth matches, mathematically, the quadratic utility loss from trading and gives direct access to a semi-explicit solution, separable in time and wealth. Second, we consider an investor with power utility of terminal wealth. Mathematically, it is inconvenient with the combination of a power utility of terminal wealth and quadratic utility loss from trading. Therefore, we reformulate the power objective by punishing quadratic deviations from the power utility optimal stock position. In the limit where there is no punishment from trading and infinite punishment from deviating from the power portfolio, we get, of course, the power utility optimal stock position, equivalent to the Merton portfolio. Other weights on the two separate objectives balances off the two considerations such that the Merton stock position is smoothed with consideration to the utility loss from trading. The transformation of the power utility objective into a quadratic objective is somewhat tailor-made for the purpose, but given the Merton portfolio as the limiting solution, we find it to have considerable backing. It is also striking that Liu and Zheng (2016) worked with the same construction, even without benefiting mathematically from it to the same extent as we do. Other related works includes Pliska and Suzuki (2004) who minimize deviations from a target portfolio in the light of proportional transaction costs.

In the title and this introduction, we refer to smoothing of investments. This refers to ruling out diffusive trading by requiring the number of stocks to be differentiable. In mathematics, smoothing has different connotations.

Some use the word smooth about a function being (at least) once continuously differentiable. For others, the function needs to be infinitely continuously differentiable. Mathematically, we here use the word in the first sense. The differential quotient of the number of stocks in our optimal portfolio will, indeed, be diffusive. Also, economically the word smoothing is used for different purposes. One speaks, e.g., about a smoothing mechanism in certain pension savings products. This means that the diffusive market return on investments of a portfolio of pension savers is smoothed over time before being distributed to individual saving accounts, see e.g. Guillén et al. (2006). In more classical finance, one has worked with consumption smoothing as the general idea of redistributing limited years of labor income to a life-long consumption plan. But the term is also used for investment and consumption in a mathematical way, similar to our use. Longstaff (2001) considered the Merton problem requiring the number of stocks to be differentiable and obtaining a bang-bang investment strategy. Bruhn and Steffensen (2013) studied differentiable consumption streams, but worked in a mathematical framework similar to ours, where quadratic functions arrange for mathematical tractability. In many ways, our paper takes the patterns of thinking in Bruhn and Steffensen (2013) as far as possible in the investment dimension rather than the consumption dimension.

The outline of the paper is as follows: In Section 5.2, we illustrate how diffusive trading arises in a Black-Scholes financial market and introduce a trading constraint to avoid it. In Section 5.3, we show how the trading constraint leads to a bang-bang investment strategy if the constraint is not accompanied by a utility loss from trading. In Section 5.4, we derive a smooth investment strategy for a mean-square utility investor and illustrate the strategy and the resulting trading in a numerical example based on simulation. In Section 5.5, we derive a smooth investment strategy for a power utility investor who minimizes quadratic deviations from the Merton investment proportion. Again, we illustrate the strategy and the resulting trading in a numerical example.

5.2 Model and Motivation

We consider an investor making decisions about investment in continuous time. The investor has access to a classical Black-Scholes market consisting of a bank account, B, with risk free short rate r, and a stock, S, with excess return λ and volatility σ. The asset prices are described by the stochastic differential equations (SDEs)

$$
\begin{aligned}
\mathrm{d}B_t &= rB_t\mathrm{d}t \ , \quad t \geq 0 \ , \quad B_0 = 1 \ , \\
\mathrm{d}S_t &= S_t\left[(r+\lambda)\,\mathrm{d}t + \sigma\mathrm{d}W_t\right] \ , \quad t \geq 0 \ , \quad S_0 = s_0 \ ,
\end{aligned}
$$

where $r, \lambda, \sigma > 0$ are constants, and W is a standard Brownian motion on the probability space $(\Omega, \mathcal{F}, P)$.

We fix a time-horizon T. The investor has wealth X and invests a proportion π of X in the stock and a proportion $(1-\pi)$ of X in the bank account. The investor's wealth evolves according to the SDE

$$\begin{aligned} \mathrm{d}X_t &= X_t\left[(r+\pi_t\lambda)\,\mathrm{d}t + \pi_t\sigma\mathrm{d}W_t\right] \ , \qquad t\in[0,T] \ , \\ X_0 &= x_0 \ , \end{aligned}$$

where x_0 is the initial wealth of the investor.

The traditional Merton investment problem considers expected utility maximization for an investor with constant relative risk aversion γ and a fixed time-horizon T, i.e.

$$\max_{(\pi_s)_{0\le s\le T}} \mathbb{E}\left[\frac{1}{1-\gamma}X_T^{1-\gamma}\right] .$$

The optimal investment proportion for this problem is

$$\pi_t^* = \pi^* \ ,$$

where π^* is the so-called Merton proportion, see Merton (1971), which is given by

$$\pi^* = \frac{\lambda}{\sigma^2\gamma} \ .$$

Keeping a constant investment proportion means selling when prices go up and buying when prices go down (in short "sell high, buy low"). The optimal amount invested in stocks, $A^*=\pi^*X^*$, follows the dynamics

$$\begin{aligned} \mathrm{d}A^* &= \pi^*X_t^*\left[(r+\pi^*\lambda)\,\mathrm{d}t + \pi^*\sigma\mathrm{d}W_t\right] \\ &= \frac{A_t^*}{S_t}\mathrm{d}S_t + (\pi^*-1)\,A_t^*\left[\lambda\mathrm{d}t+\sigma\mathrm{d}W_t\right] \ . \end{aligned}$$

Here, the first term is the market return on A^* whereas the second term is the necessary stock trading in order to keep a constant investment proportion. For reasonable parameters, we have $\pi^*<1$, and so the drift and volatility of the trading amount are negative. Consequently, the investor on average trades a negative amount in stocks (i.e. sells). This is an intuitive behavior since the drift of the optimal wealth, X^*, is smaller than that of the optimal amount invested in stocks, A^*.

On the other hand, the optimal number of stocks held by the investor, $N^*=\frac{A^*}{S}$, follows the dynamics

$$\begin{aligned} \mathrm{d}N_t^* &= \pi^*\left(\frac{1}{S_t}\,\mathrm{d}X_t^* - \frac{X_t^*}{S_t^2}\,\mathrm{d}S_t - \frac{1}{S_t^2}\,\mathrm{d}X_t^*\,\mathrm{d}S_t + \frac{X_t^*}{S_t^3}\,(\,\mathrm{d}S_t)^2\right) \\ &= \ N_t^*\,(1-\pi^*)\left(\left(\sigma^2-\lambda\right)\,\mathrm{d}t - \sigma\,\mathrm{d}W_t\right) . \end{aligned}$$

Notice that the number of stocks has a positive drift for reasonable parameters with $\sigma^2>\lambda$ and $\pi^*<1$. Hence, the investor on average trades a positive number of stocks (i.e. buys). Altogether, the investor trades a negative amount (i.e. sells) but a positive number (i.e. buys) when stock volatility is sufficiently high ($\sigma^2 > \lambda$). This counter-intuitive fact is explained by investor's "sell high, buy low" strategy; on average, he buys a number of stocks, but he buys them cheap and sells them expensive, resulting in negative trading amounts. We are going to refer to this peculiarity several times later on when a similar behavior is observed also for smooth investment.

The optimal number of stocks held by the investor follows a process of unbounded variation. This phenomenon is not isolated to power utility. Consider for example expected mean-square utility maximization,

$$\max_{(\pi_s)_{0\leq s\leq T}} \mathbb{E}\left[X_T - \gamma X_T^2\right] . \tag{5.1}$$

In Appendix 5.A, we show that the optimal investment proportion for this problem is

$$\pi_t^* = \frac{\lambda}{\sigma^2}\left(\frac{1}{2X_t^*\gamma e^{r(T-t)}} - 1\right) , \tag{5.2}$$

and that the optimal investment proportion follows the dynamics

$$\mathrm{d}\pi_t^* = \left(\pi_t^* + \frac{\lambda}{\sigma^2}\right)\left(\left((\pi_t^*)^2\,\sigma^2 - \pi_t^*\lambda\right)\mathrm{d}t - \pi_t^*\sigma\mathrm{d}W_t\right) .$$

Since the optimal investment proportion is decreasing in wealth, the mean-square investor sells when prices go up and buys when prices go down—and he does so more rapidly than the power investor. The investment proportion is seen to have a negative drift for $\pi_t^* \in \left(0, \frac{\lambda}{\sigma^2}\right)$ which is the case for reasonable parameters.

In Appendix 5.A, we also show that the optimal number of stocks held by the investor, $N^* = \pi^*\frac{X^*}{S}$, follows the dynamics

$$\mathrm{d}N_t^* = N_t^*\left(\left(1 + \frac{1-r}{\pi_t^*} - \frac{\lambda^2}{\sigma^2} - r + \sigma^2\right)\mathrm{d}t - \left(\frac{\lambda}{\sigma} + \sigma\right)\mathrm{d}W_t\right) ,$$

and that the optimal amount invested in stocks, $A^* = \pi^* X^*$, follows the dynamics

$$\mathrm{d}A^* = \frac{A_t^*}{S_t}\mathrm{d}S_t + (\pi_t^* - 1)\,A_t^*\,[\lambda\mathrm{d}t + \sigma\mathrm{d}W_t] + X_t^*\,\mathrm{d}\pi_t^* .$$

Notice that the optimal number of stocks has a positive drift for

$$\pi_t^* > \frac{1-r}{\lambda^2/\sigma^2 + r - \sigma^2 - 1} ,$$

which holds for reasonable parameters. The second and third term in the dynamics of the optimal amount invested in stocks are the necessary stock trading in order to obtain the desired investment proportion. For $0 < \pi_t^* < \min\left\{\frac{\lambda}{\sigma^2}, 1\right\}$, the drift of the trading amount is negative . Hence, for reasonable parameters, the mean-square investor's "sell high, buy low" strategy leads to the peculiarity of trading a negative amount (i.e. selling) but a positive number (i.e. buying)—exactly as for the power investor.

Like in the case of power utility, the optimal number of stocks follows a process of unbounded variation. A typical investor trades a lot more smoothly. We constrain the number of stocks held by the investor, N, to follow the dynamics

$$\mathrm{d}N_t = \tau_t N_t \,\mathrm{d}t \ , \ N_0 = n_0 \ , \tag{5.3}$$

where the investor controls τ_t and n_0. Thereby, the amount invested in stocks, $A_t = N_t S_t$, follows the dynamics

$$\mathrm{d}A_t = A_t\left(\left(r + \lambda + \tau_t\right)\mathrm{d}t + \sigma\,\mathrm{d}W_t\right) \ , \ A_0 = s_0 n_0 \ ,$$

and the investor's wealth follows the dynamics

$$\mathrm{d}X_t = rX_t\,\mathrm{d}t + A_t\left(\lambda\,\mathrm{d}t + \sigma\,\mathrm{d}W_t\right) \ , \ X_0 = x_0 \ .$$

Throughout the paper, we refer to τ as the trading rate, τA as the trading amount rate, and τN as the trading number rate.

5.3 Standard Utility Optimization with Bounded Variation

The two utility maximization problems in the previous section can be written as

$$\max_{(\pi_s)_{0\le s\le T}} \mathbb{E}\left[u\left(X_T\right)\right]$$

for a suitable utility function u. If we consider this general problem with the trading constraint in Equation (5.3), we get the Hamilton-Jacobi-Bellman equation

$$\begin{aligned} J_t = -\max_{\tau}\Big\{ &\left(\lambda a + rx\right)J_x + a\left(\lambda + r + \tau\right)J_a \\ &+ \frac{1}{2}\sigma^2 a^2\left(J_{xx} + J_{aa} + 2J_{ax}\right)\Big\} \ , \\ J\left(T, x, a\right) = u\left(x\right) \ . \end{aligned}$$

If we allow the trading rate τ to vary in the interval $\left[\tau^-, \tau^+\right]$, we see that the optimal τ is given by

$$\tau^* = \begin{cases} \tau^- & \text{if } J_a < 0 \ , \\ \tau^+ & \text{if } J_a \ge 0 \ . \end{cases}$$

Hence, the result is a bang-bang strategy which is not really as smooth as we think it should be. For details on the bang-bang strategy, see Longstaff

(2001). In the following sections, we propose two different approaches to obtain optimal smooth investment strategies—one for mean-square utility and one for power utility.

5.4 Smooth Investment with Mean-Square Utility

To obtain smooth investment for an expected mean-square utility maximizing agent, we propose the optimization criteria

$$\max_{n_0,(\tau_s)_{0\leq s\leq T}} \mathbb{E}\left[X_T - \gamma X_T^2 - \int_0^T \frac{1}{2}\Lambda\left(\tau_s A_s\right)^2 \,\mathrm{d}s\right] . \tag{5.4}$$

Here, the parameter $\Lambda \geq 0$ weights the third term against the first two. The first two terms are terminal mean-square utility. The third term punishes trading amounts, τA, quadratically. Hence, we have simply subtracted the expected utility loss from trading from the investor's original optimization criteria.

In Appendix 5.B, we show that the optimal trading amount rate is

$$\tau_t^* A_t^* = \frac{2}{\Lambda} f_3\left(t\right) A_t^* + \frac{1}{\Lambda} f_4\left(t\right) + \frac{1}{\Lambda} f_5\left(t\right) X_t^* , \tag{5.5}$$

where f_3, f_4, and f_5 are solutions to the following system of ordinary differential equations (ODEs):

$$\begin{aligned}
f_1'\left(t\right) &= -2rf_1\left(t\right) - \frac{1}{2\Lambda} f_5^2\left(t\right) \ , \quad f_1\left(T\right) = -\gamma \ , \\
f_2'\left(t\right) &= -rf_2\left(t\right) - \frac{1}{\Lambda} f_4\left(t\right) f_5\left(t\right) \ , \quad f_2\left(T\right) = 1 \ , \\
f_3'\left(t\right) &= -\left(\lambda + \sigma^2\right) f_5\left(t\right) - \frac{2}{\Lambda} f_3^2\left(t\right) - \sigma^2 f_1\left(t\right) - \left(2\lambda + 2r + \sigma^2\right) f_3\left(t\right) \ , \\
&\qquad f_3\left(T\right) = 0 \ , \\
f_4'\left(t\right) &= -\lambda f_2\left(t\right) - \frac{2}{\Lambda} f_3\left(t\right) f_4\left(t\right) - \left(\lambda + r\right) f_4\left(t\right) \ , \ f_4\left(T\right) = 0 \ , \\
f_5'\left(t\right) &= -2\lambda f_1\left(t\right) - \frac{2}{\Lambda} f_3\left(t\right) f_5\left(t\right) - \left(\lambda + 2r\right) f_5\left(t\right) \ , \quad f_5\left(T\right) = 0 \ , \\
f_6'\left(t\right) &= -\frac{1}{2\Lambda} f_4^2\left(t\right) \ , \quad f_6\left(T\right) = 0 \ .
\end{aligned} \tag{5.6}$$

Also, the investor's optimal number of stocks at time zero is given by

$$n_0^* = -\frac{f_4\left(0\right) + f_5\left(0\right) x_0}{2 f_3\left(0\right) s_0} . \tag{5.7}$$

Notice that this choice of n_0^* results in $\tau_0^* = 0$, meaning that the investor starts out at his target.

Remark 5.1 (Establishment and liquidation punishment)**.** Since we are punishing continuous trading, it is also natural to punish trading upon establishment and liquidation of the portfolio. This results in the optimization criteria

$$\max_{n_0,(\tau_s)_{0\le s\le T}} \mathbb{E}\left[X_T - \gamma X_T^2 - \frac{1}{2}\Lambda_1 a_0^2 - \int_0^T \frac{1}{2}\Lambda\left(\tau_s A_s\right)^2 \,\mathrm{d}s - \frac{1}{2}\Lambda_2 A_T^2\right] \ ,$$

where $\Lambda_1, \Lambda_2 \ge 0$ weight the punishment terms against each other. The only thing that changes in the solution is the boundary conditions for f_3. With establishment and liquidation punishment, we have

$$\begin{aligned} f_3\left(T\right) &= -\frac{1}{2}\Lambda_2 \ , \\ f_3'\left(t\right) &= -\left(\lambda+\sigma^2\right) f_5\left(t\right) - \frac{2}{\Lambda} f_3^2\left(t\right) - \sigma^2 f_1\left(t\right) - \left(2\lambda + 2r + \sigma^2\right) f_3\left(t\right) \ , \\ & \qquad t \in \left(0, T\right] \ , \\ f_3\left(0\right) &= -\frac{1}{2}\Lambda_1 + f_3\left(0+\right) \ . \end{aligned}$$

5.4.1 Numerics

To study the impact of the trading constraint in Equation (5.3) for an expected mean-square utility maximizing investor, we consider the investor's optimal wealth, X^*, and optimal amount invested in stocks, A^*. They evolve according to the SDE's

$$\begin{aligned} \mathrm{d}X_t^* &= rX_t^* \,\mathrm{d}t + A_t^*\left(\lambda \,\mathrm{d}t + \sigma \,\mathrm{d}W_t\right) \ , \quad X_0^* = x_0 \ , \\ \mathrm{d}A_t^* &= A_t^*\left(\left(r + \lambda + \tau_t^*\right) \mathrm{d}t + \sigma \,\mathrm{d}W_t\right) \ , \quad A_0^* = a_0 \ , \end{aligned} \tag{5.8}$$

where the optimal trading rate, τ^*, is given by Equation (5.5). In general, there is no explicit solution to the system of ODE's in Equation (5.6), and the distribution of X^* and A^* is unknown. Therefore, to study the evolution of τ^*, X^*, and A^*, we solve the system of ODE's in (5.6) numerically for selected parameter values and simulate X^* and A^* using standard Monte-Carlo methods.

We consider a time horizon of $T = 30$ years and fix the investor's initial wealth at $x_0 = 100$ and the initial value of the stock at $s_0 = 1$. We fix $\gamma = \frac{0.08}{x_0}$ since this level of risk aversion results in an expected final wealth of the unconstrained mean-square investor that matches the expected final wealth of the unconstrained power investor in Section 5.5.1. The division by x_0 is inspired by Björk et al. (2014). We choose the following market parameters values:

r	σ	λ
0.04	0.20	0.03

We solve the system of ODE's in Equation (5.6) numerically for five different values of Λ, namely $\Lambda = 10^{-5}, 0.01, 0.1, 1, 10$. The larger the value of Λ,

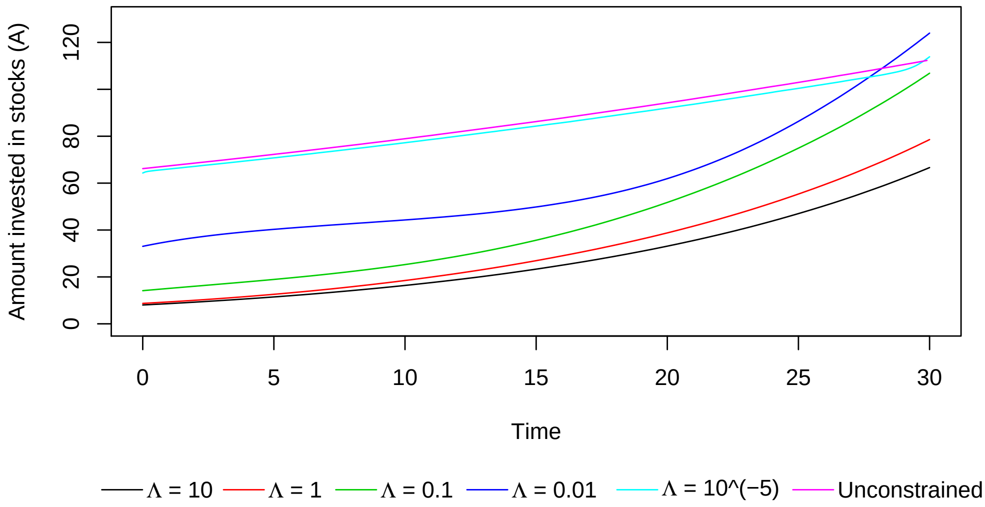

Figure 5.1: Average amount A^* invested in stocks as a function of time.

the larger the utility loss associated with trading. Applying the formulas in Equations (5.2) and (5.7), we calculate the investor's optimal number of stocks at time zero, n_0^*, for each value of Λ and for the unconstrained mean-square investor:

	Unconstrained	$\Lambda = 10^{-5}$	$\Lambda = 0.01$	$\Lambda = 0.1$	$\Lambda = 1$	$\Lambda = 10$
n_0^*	66.2	64.3	33.0	14.1	8.7	8.0

For each value of Λ, we simulate 1 million sample paths for the investor's optimal wealth, X^*, and amount invested in stocks, A^*, using discretized versions of the SDE's in Equation (5.8). We also simulate 1 million sample paths for the investor's optimal wealth and stock investment without the constraint, i.e. with the proportion π^* from Equation (5.2) invested in stocks.

In Figure 5.1, we plot the investor's average amount, A^*, invested in stocks. Recall that we optimize over both n_0 and τ for each value of Λ. For all values of Λ, the amount in stocks starts out lower than in the unconstrained case. This is to avoid having to sell too large amounts later on; the investor anticipates his future desired investment proportion and seeks to slide towards it, without sliding to far. Thereby, the investor "aims in front of the target" as seen in Gârleanu and Pedersen (2013). The lower the value of Λ, the closer the amount in stocks gets to the unconstrained investment since trading is punished less. Notice that the case $\Lambda = 10^{-5}$ is very close to the unconstrained case, also in terms of certainty equivalents as is shown below.

In Figure 5.2, we plot the average trading amount rate, $\tau^* A^*$, and in Figure 5.3, we plot the average trading number rate, $\tau^* N^*$. As for the unconstrained mean-square investor, the constrained investor, on average, trades a negative amount (i.e. sells) but a positive number (i.e. buys), mimicking the

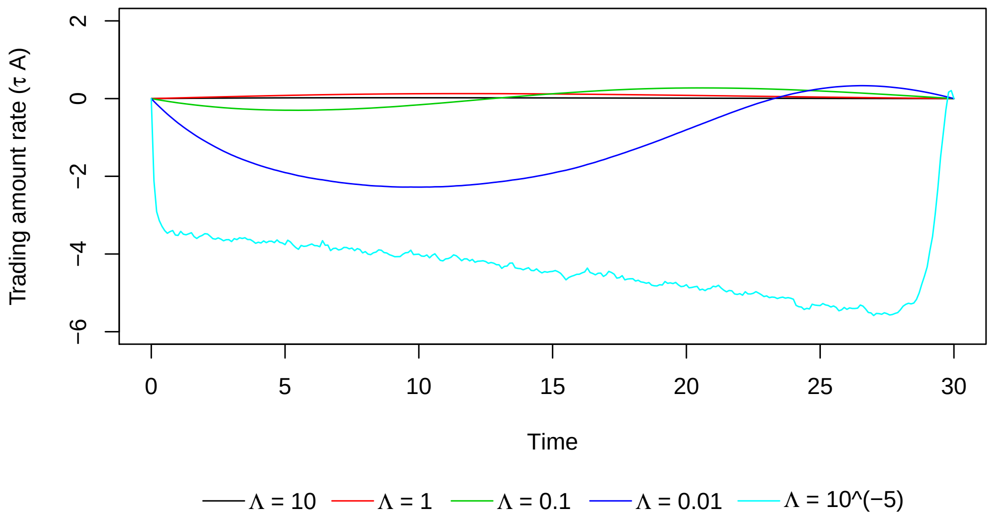

Figure 5.2: Average trading amount rate $\tau^* A^*$ as a function of time.

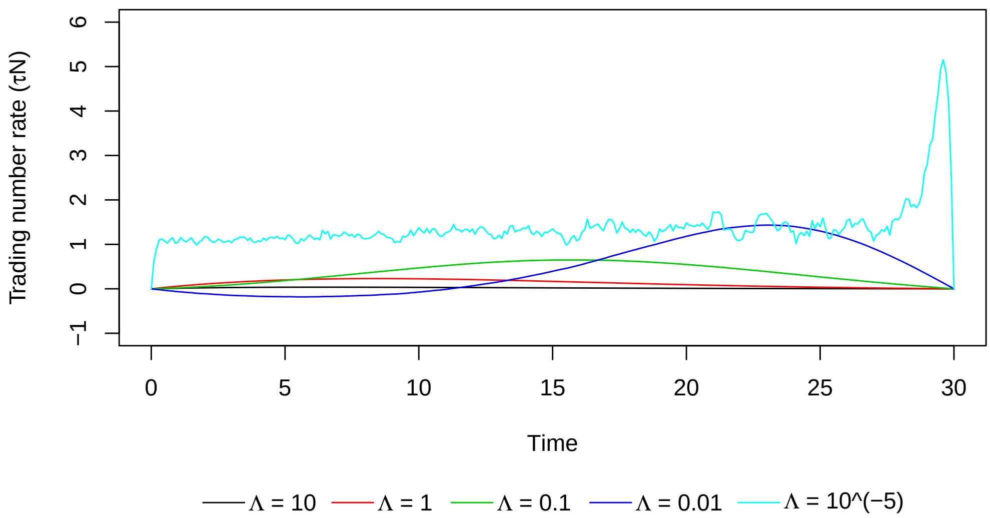

Figure 5.3: Average trading number rate $\tau^* N^*$ as a function of time.

unconstrained investor's "sell high, buy low" strategy. Investors with high values of Λ tend to be "buy-and-hold" investors since trading is punished harder. Investors with low values of Λ tend to trade more; they trade a larger negative amount and a larger positive number. Investors with a moderate value of Λ first trade positive amounts and then trade negative amounts which is caused by the trade-off between punishment for trading and mean-square maximization. All investors, except the investor with $\Lambda = 0.01$, on average trade a positive number of stocks, corresponding to a positive drift of N^* as for the unconstrained investor. The investor with $\Lambda = 0.01$ starts out by trading a small negative number of stocks, but then switches to trading a positive number like the other investors.

To study the investor's expected utility loss from the trading constraint in Equation (5.3), we use the simulated sample paths to approximate certainty equivalents for the investor's optimal final wealth. We calculate the expected utility

$$\mathrm{EU} = \mathbb{E}\left[X_T - \gamma X_T^2\right] \;,$$

and we determine the so-called certainty equivalent as the solution CE to the equation

$$\mathrm{EU} = \mathrm{CE} - \gamma \mathrm{CE}^2 \;.$$

The certainty equivalent expresses which certain amount the investor requires at time T in order to give up his uncertain wealth X_T. It is more meaningful to compare certainty equivalents than utility since utility has an ordinal interpretation. The equation has two solutions, and we pick the smallest one. We approximate the mean $\mathbb{E}\left[X_T - \gamma X_T^2\right]$ by the average of the simulations. We apply the procedure for the unconstrained wealth and for the constrained wealth for each value of Λ. We get the following table:

	No constraint	$\Lambda = 10^{-5}$	$\Lambda = 0.01$	$\Lambda = 0.1$	$\Lambda = 1$	$\Lambda = 10$
EU	277	277	268	261	255	253
CE	414	414	389	371	357	352

The investor's expected utility loss is significant for high values of Λ. This is consistent with the unconstrained mean-square investor being an active investor. A high value of Λ corresponds to a hard punishment for trading and this forces the investor to trade less and deviate from his optimal investment strategy, resulting in an expected utility loss.

5.5 Smooth Investment with Power Utility

Inspired by the previous section, it is natural to propose the following optimization criteria to obtain smooth investment for an expected power utility maximizing agent;

$$\max_{n_0,(\tau_s)_{0\leq s\leq T}} \mathbb{E}\left[\frac{1}{1-\gamma}X_T^{1-\gamma} - \int_0^T \frac{1}{2}\Lambda\left(\tau_s A_s\right)^2 \,\mathrm{d}s\right] .$$

However, if we proceed along the lines of the previous section, we cannot obtain a trading solution given by ODEs. We are stuck with a solution given by a partial differential equation (PDE) that does not provide us with useful insight on smooth investment. Therefore, we take a different road to investment smoothing for the expected power utility maximizing agent.

Based on the Merton result that optimal investment is constant proportional to wealth and with the trading constraint in Equation (5.3), we propose the optimization criteria

$$\min_{n_0,(\tau_s)_{0\leq s\leq T}} \mathbb{E}\left[\int_0^T \frac{1}{2}\left(\theta(\pi^* X_t - A_t)^2 + (\tau_t A_t)^2\right) \mathrm{d}t\right] . \tag{5.9}$$

Here, the parameter $\theta \geq 0$ weights the two terms against each other. The first term punishes quadratically if the actual stock holdings of the investor deviate from the Merton proportion π^* of the investor's wealth. The second term punishes trading amounts quadratically.

In Appendix 5.C, we show that the optimal trading amount rate is

$$\tau_t^* A_t^* = h\left(t\right) X_t^* - g\left(t\right) A_t^* = g\left(t\right)\left(\frac{h\left(t\right)}{g\left(t\right)} X_t^* - A_t^*\right) , \tag{5.10}$$

where f, g, and h are solutions to the following system of ODEs

$$\begin{aligned} f'\left(t\right) &= -\theta\left(\pi^*\right)^2 + \left(h\left(t\right)\right)^2 - 2rf\left(t\right) \ , \quad f\left(T\right) = 0 \ , \qquad (5.11)\\ g'\left(t\right) &= -\theta - \sigma^2 f\left(t\right) + 2\left(\lambda + \sigma^2\right) h\left(t\right) - \left(2r + 2\lambda + \sigma^2\right) g\left(t\right) + \left(g\left(t\right)\right)^2 ,\\ &\qquad g\left(T\right) = 0 \ ,\\ h'\left(t\right) &= -\theta\pi^* + \lambda f\left(t\right) - \left(2r + \lambda\right) h\left(t\right) + g\left(t\right) h\left(t\right) \ , \quad h\left(T\right) = 0 \ . \end{aligned}$$

Also, the investor's optimal number of stocks at time zero is given by

$$n_0^* = \frac{h\left(0\right) x_0}{g\left(0\right) s_0} . \tag{5.12}$$

From the expression in Equation (5.10), we see that the investor trades towards a target stock proportion of $\frac{h}{g}$, with the speed g. By trading towards this target instead of the Merton proportion, π^*, the investor anticipates his future desired investment proportion and seeks to slide towards it, thereby “aiming in front of his target” as seen in Gârleanu and Pedersen (2013). Notice that the choice of n_0^* results in $\tau_0^* = 0$, meaning that the investor starts out at his target.

Remark 5.2 (Initial and terminal punishment)**.** Since we are punishing continuous trading and deviations from the Merton proportion, it is also natural to punish trading upon establishment and liquidation and to punish initial and terminal deviations from the Merton proportion. This results in the optimization criteria

$$\min_{n_0,(\tau_s)_{0\le s\le T}} \mathbb{E}\Bigg[\frac{1}{2}\Lambda_1 a_0^2 + \frac{1}{2}\Theta_1(\pi^* X_0 - a_0)^2 + \int_0^T \frac{1}{2}\Big(\theta(\pi^* X_t - A_t)^2 + (\tau_t A_t)^2\Big)\,\mathrm{d}t + \frac{1}{2}\Theta_2(\pi^* X_T - A_T)^2 + \frac{1}{2}\Lambda_2 A_T^2\Bigg],$$

where $\Theta_1, \Theta_2, \Lambda_1, \Lambda_2 \ge 0$ weight the punishment terms against each other. The only thing that changes is the boundary conditions for f, g, and h. With initial and terminal punishment, we have

$$\begin{aligned}
f(T) &= \Theta_2 (\pi^*)^2 \ , \\
f'(t) &= -\theta(\pi^*)^2 + (h(t))^2 - 2rf(t) \ , \ t\in(0,T] \ , \\
f(0) &= \Theta_1(\pi^*)^2 + f(0+) \ , \\
g(T) &= \Lambda_2 + \Theta_2 \ , \\
g'(t) &= -\theta - \sigma^2 f(t) + 2\left(\lambda+\sigma^2\right)h(t) - \left(2r + 2\lambda + \sigma^2\right)g(t) + (g(t))^2 \ , \\
&\qquad t\in(0,T] \ , \\
g(0) &= \Lambda_1 + \Theta_1 + g(0+) \ , \\
h(T) &= \Theta_2 \pi^* \ , \\
h'(t) &= -\theta\pi^* + \lambda f(t) - (2r+\lambda)h(t) + g(t)h(t) \ , \ t\in(0,T] \ , \\
h(0) &= \Theta_1\pi^* + h(0+) \ .
\end{aligned}$$

5.5.1 Numerics

To study the impact of the trading constraint in Equation (5.3) for an expected power utility maximizing investor, we consider the investor's optimal wealth, X^*, and optimal amount invested in stocks, A^*. They evolve according to the SDE's

$$\begin{aligned}
\mathrm{d}X_t^* &= rX_t^*\,\mathrm{d}t + A_t^*\,(\lambda\,\mathrm{d}t + \sigma\,\mathrm{d}W_t) \ , \quad X_0^* = x_0 \ , \\
\mathrm{d}A_t^* &= A_t^*\,((r+\lambda+\tau_t^*)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t) \ , \quad A_0^* = a_0 \ ,
\end{aligned} \tag{5.13}$$

where the optimal trading rate τ_t^* is given by Equation (5.10). In general, there is no explicit solution to the system of ODE's in Equation (5.11), and the distribution of X^* and A^* is unknown. Therefore, to study the evolution of τ^*,

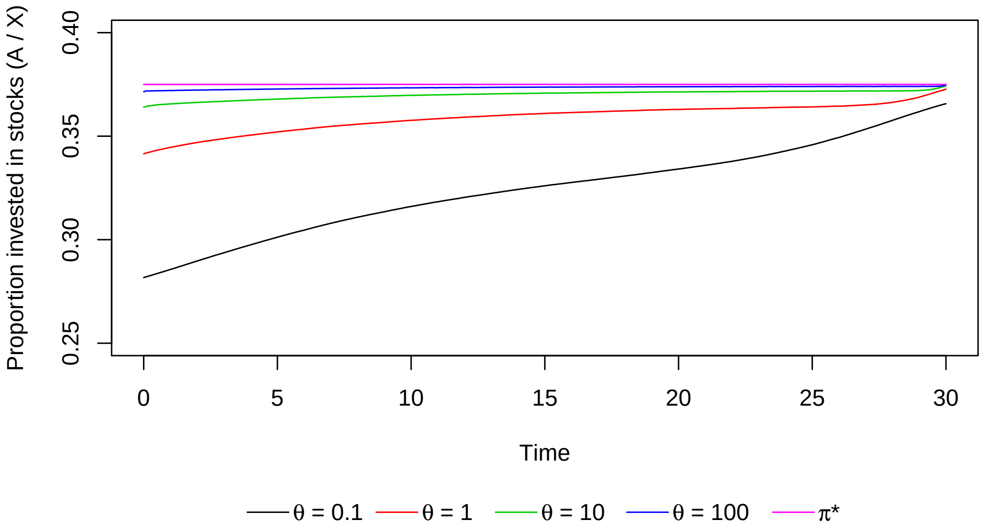

Figure 5.4: Average proportion $\frac{A^*}{X^*}$ invested in stocks as a function of time.

X^*, and A^*, we solve the system of ODE's numerically for selected parameter values and simulate X^* and A^* using standard Monte-Carlo methods.

We consider a time horizon of $T = 30$ years and fix the investor initial wealth at $x_0 = 100$. We choose the following market and preference parameters values:

r	σ	λ	γ
0.04	0.20	0.03	2

The Merton proportion yields $\pi^* = 0.375$. We fix the initial value of the stock at $s_0 = 1$.

We solve the system of ODE's in Equation (5.11) numerically for four different values of θ, namely $\theta = 0.1, 1, 10, 100$. The larger the value of θ, the larger the weight on stock holdings being close to the Merton proportion π^* of the wealth.

Applying the formula in Equation (5.12), we calculate the investor's optimal number of stocks at time zero for each value of Θ:

	No constraint	$\theta = 100$	$\theta = 10$	$\theta = 1$	$\theta = 0.1$
n_0^*	37.5	37.1	36.4	34.1	27.2

For each value of θ, we simulate 1 million sample paths for the investor's optimal wealth, X^*, and amount invested in stocks, A^*, using discretized versions of the SDE's in Equation (5.13). We also simulate 1 million sample paths for the investor's optimal wealth and stock investment without the constraint, i.e. with the Merton proportion, π^*, invested in stocks.

In Figure 5.4, we plot the investor's proportion of wealth, $\frac{A^*}{X^*}$, invested in stocks. For all four values of θ, the proportion starts out lower than the Merton

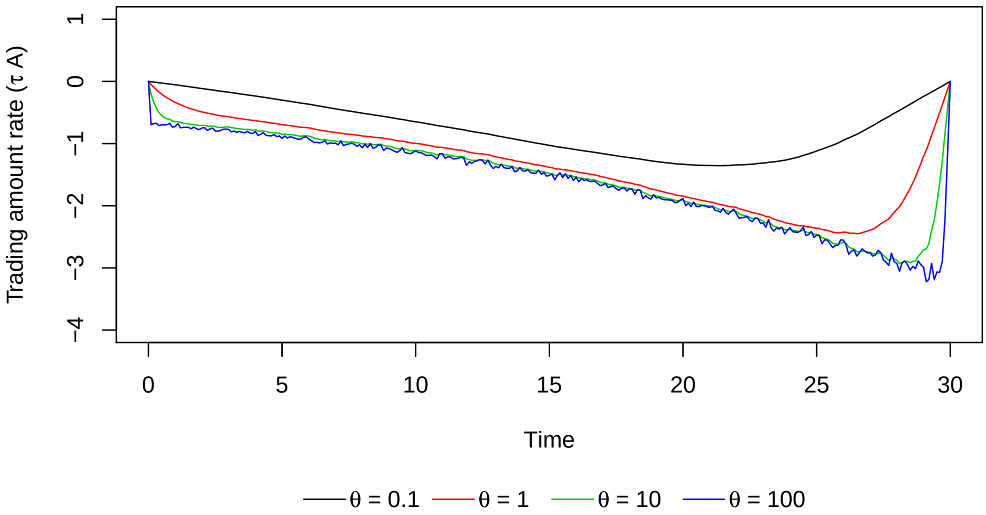

Figure 5.5: Average trading amount rate $\tau^* A^*$ as a function of time.

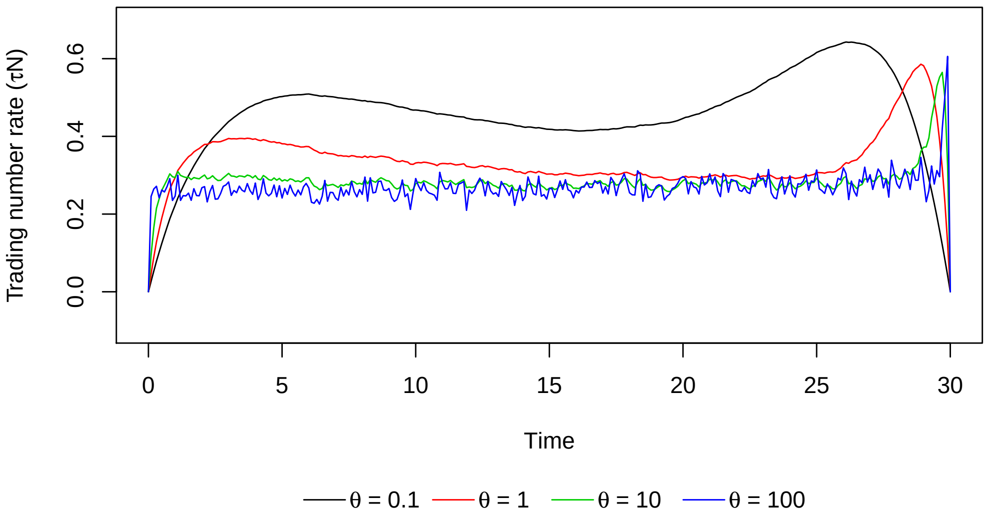

Figure 5.6: Average trading number rate $\tau^* N^*$ as a function of time.

proportion π^*. As for the mean-square investor, this is to avoid having to sell too large amounts later on, by "aiming in front of the target". An investor with a high value of θ stays closer to the Merton proportion which agrees with the intuition about θ. We notice that the case $\theta = 100$ is very close to the unconstrained case.

In Figure 5.5, we plot the average trading amount rate, $\tau^* A^*$, and in Figure 5.6, we plot the average trading number rate, $\tau^* N^*$. Since the investor mimics the unconstrained power investor's "sell high, buy low" strategy, the investor trades a negative amount (i.e. sells) but a positive number (i.e. buys), on average. An investor with a high value of θ trades larger amounts in order to stay close to the Merton proportion. An investor with a low value of θ trades smaller amounts because a low value of θ corresponds to harder punishment for trading. All investors on average trade a positive number of stocks, corresponding to a positive drift of N^* as for the unconstrained investor. An investor with a low value of θ trades a larger number of stocks, but this is due to the investor's utility function punishing trading amounts and not trading numbers.

To study the investor's expected utility loss from the trading constraint in Equation (5.3), we use the simulated sample paths to approximate certainty equivalents for the investor's optimal final wealth. We calculate the certainty equivalent

$$\mathrm{CE} = \left(\mathbb{E}\left[X_T^{1-\gamma}\right]\right)^{\frac{1}{1-\gamma}} .$$

Again, the certainty equivalent expresses which certain amount the investor requires at time T in order to give up his uncertain wealth X_T. We approximate the mean $\mathbb{E}\left[X_T^{1-\gamma}\right]$ by the average of the simulations. We apply the procedure for the unconstrained wealth and for the constrained wealth for each value of θ. We get the following table:

	No constraint	$\theta = 100$	$\theta = 10$	$\theta = 1$	$\theta = 0.1$
CE	392.0	392.0	391.9	391.2	389.1

The investor's expected utility loss is insignificant. The utility loss is larger for low value of θ. This agrees with the intuition about θ since a low value of θ corresponds to a harder punishment for trading. Notice that the expected utility loss in general is much smaller for the power investor than for the mean-square investor which conforms with the unconstrained power investor being a less active investor.

Appendix

5.A Calculations for Mean-Square Utility Optimization

We embed the problem in Equation (5.1) in an optimal value function J given by

$$J(t,x) = \max_{(\pi_s)_{t\le s\le T}} \mathbb{E}\left[X_T - \gamma X_T^2 \middle| X_t = x\right] ,$$

The Hamilton-Jacobi-Bellman equation is

$$\begin{aligned} J_t &= -\max_{\pi} \left\{ (\pi\lambda + r)\, x J_x + \frac{1}{2}\sigma^2\pi^2 x^2 J_{xx} \right\} , \\ J(T,x) &= x - \gamma x^2 . \end{aligned}$$

The first order condition reads

$$\pi = -\frac{\lambda}{\sigma^2}\frac{J_x}{xJ_{xx}} \ .$$

Inserting this in the Hamilton-Jacobi-Bellman equation, we obtain the PDE

$$0 = J_t + \frac{1}{2}\frac{\lambda^2 J_x^2}{\sigma^2 J_{xx}} + rxJ_x \ ,$$
$$J(T,x) = x - \gamma x^2 \ .$$

We make the ansatz

$$J(t,x) = f(t)\,x + g(t)\,x^2 + h(t) \ ,$$

where f, g, and h are continuously differentiable functions satisfying the terminal conditions $f(T) = 1$, $g(T) = -\gamma$, and $h(T) = 0$. Plugging in, we obtain

$$\begin{aligned}
0 &= h'(t) + f'(t)\,x + g'(t)\,x^2 + \frac{1}{2}\frac{\lambda^2\left(f(t) + 2g(t)\,x\right)^2}{2\sigma^2 g(t)} \\
&\quad + rx\left(f(t) + 2g(t)\,x\right) \\
&= h'(t) + f'(t)\,x + g'(t)\,x^2 + \frac{\lambda^2}{4\sigma^2}\frac{\left(f(t)\right)^2}{g(t)} \\
&\quad + \frac{\lambda^2}{\sigma^2}g(t)\,x^2 + \frac{\lambda^2}{\sigma^2}f(t)\,x + rf(t)\,x + 2rg(t)\,x^2 \\
&= \left[h'(t) + \frac{\lambda^2}{4\sigma^2}\frac{\left(f(t)\right)^2}{g(t)}\right] + \left[f'(t) + \frac{\lambda^2}{\sigma^2}f(t) + rf(t)\right]x \\
&\quad + \left[g'(t) + \frac{\lambda^2}{\sigma^2}g(t) + 2rg(t)\right]x^2 \ .
\end{aligned}$$

Hence, we get the following system of ODEs:

$$\begin{aligned}
f'(t) &= -\left(\frac{\lambda^2}{\sigma^2} + r\right)f(t) \ , \quad f(T) = 1 \ , \\
g'(t) &= -\left(\frac{\lambda^2}{\sigma^2} + 2r\right)g(t) \ , \quad g(T) = -\gamma \ , \\
h'(t) &= -\frac{\lambda^2}{4\sigma^2}\frac{\left(f(t)\right)^2}{g(t)} \ , \quad h(T) = 0 \ .
\end{aligned}$$

The solutions are

$$\begin{aligned}
f(t) &= e^{\left(\frac{\lambda^2}{\sigma^2} + r\right)(T-t)} \ , \\
g(t) &= -\gamma e^{\left(\frac{\lambda^2}{\sigma^2} + 2r\right)(T-t)} \ , \\
h(t) &= \int_t^T -\frac{\lambda^2}{4\sigma^2\gamma}e^{\frac{\lambda^2}{\sigma^2}(T-s)}\,\mathrm{d}s = \left[\frac{1}{4\gamma}e^{\frac{\lambda^2}{\sigma^2}(T-s)}\right]_t^T = \frac{1}{4\gamma} - \frac{1}{4\gamma}e^{\frac{\lambda^2}{\sigma^2}(T-t)} \ .
\end{aligned}$$

The optimal investment reads

$$\pi_t^* = -\frac{\lambda}{\sigma^2}\frac{f(t) + 2g(t)\,X_t^*}{2X_t^* g(t)} = \frac{\lambda}{\sigma^2}\left(\frac{1}{2X_t^*\gamma e^{r(T-t)}} - 1\right) .$$

The optimal number of stocks held by the investor, $N^* = \frac{\pi^* X^*}{S}$, follows the dynamics

$$\begin{aligned}
\mathrm{d}N_t^* &= \mathrm{d}\left(\frac{\frac{\lambda}{\sigma^2}\left(\frac{1}{2\gamma e^{r(T-t)}} - X_t^*\right)}{S_t}\right) \\
&= \frac{\lambda}{\sigma^2}\Bigg(\frac{\frac{r}{2\gamma e^{r(T-t)}}}{S_t}\,\mathrm{d}t - \frac{1}{S_t}\,\mathrm{d}X_t^* - \frac{\frac{1}{2\gamma e^{r(T-t)}} - X_t^*}{S_t^2}\,\mathrm{d}S_t \\
&\qquad + \frac{1}{S_t^2}\,\mathrm{d}X_t^*\,\mathrm{d}S_t + \frac{\frac{1}{2\gamma e^{r(T-t)}} - X_t^*}{S_t^3}\,(\mathrm{d}S_t)^2\Bigg) \\
&= \frac{\lambda}{\sigma^2}\Bigg(\frac{\frac{r}{2\gamma e^{r(T-t)}}}{S_t}\,\mathrm{d}t - \frac{1}{S_t}X_t^*\left((r + \pi_t^*\lambda)\,\mathrm{d}t + \pi_t^*\sigma\,\mathrm{d}W_t\right) \\
&\qquad - \frac{\frac{1}{2\gamma e^{r(T-t)}} - X_t^*}{S_t}\left((r+\lambda)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t\right) \\
&\qquad + \frac{1}{S_t}X_t^*\pi_t^*\sigma^2\,\mathrm{d}t + \frac{\frac{1}{2\gamma e^{r(T-t)}} - X_t^*}{S_t}\sigma^2\,\mathrm{d}t\Bigg) \\
&= N_t^*\left(\left(1 + \frac{1-r}{\pi_t^*} - \frac{\lambda^2}{\sigma^2} - r + \sigma^2\right)\mathrm{d}t - \left(\frac{\lambda}{\sigma} + \sigma\right)\mathrm{d}W_t\right) .
\end{aligned}$$

The optimal investment proportion π^* follows the dynamics

$$\begin{aligned}
\mathrm{d}\pi_t^* &= r\left(\pi_t^* + \frac{\lambda}{\sigma^2}\right)\mathrm{d}t - \left(\pi_t^* + \frac{\lambda}{\sigma^2}\right)\left((r + \pi_t^*\lambda)\,\mathrm{d}t + \pi_t^*\sigma\mathrm{d}W_t\right) \\
&\quad + \left(\pi_t^* + \frac{\lambda}{\sigma^2}\right)(\pi_t^*)^2\,\sigma^2\,\mathrm{d}t \\
&= \left(\pi_t^* + \frac{\lambda}{\sigma^2}\right)\left(\left((\pi_t^*)^2\,\sigma^2 - \pi_t^*\lambda\right)\mathrm{d}t - \pi_t^*\sigma\mathrm{d}W_t\right) .
\end{aligned}$$

The optimal amount invested in stocks, $A^* = \pi^* X^*$, follows the dynamics

$$\begin{aligned}
\mathrm{d}A^* &= \pi_t^* X_t^*\left[(r + \pi^*\lambda)\,\mathrm{d}t + \pi^*\sigma\mathrm{d}W_t\right] + X_t^*\,\mathrm{d}\pi_t^* \\
&= \frac{A_t^*}{S_t}\mathrm{d}S_t + (\pi_t^* - 1)\,A_t^*\left[\lambda\mathrm{d}t + \sigma\mathrm{d}W_t\right] + X_t^*\,\mathrm{d}\pi_t^* .
\end{aligned}$$

5.B Calculations for Smooth Investment with Mean-Square Utility

We embed the problem in Equation (5.4) in an optimal value function J given by

$$J(t,x,a) = \max_{(\tau_s)_{t\le s\le T}} \mathbb{E}\left[X_T - \gamma X_T^2 - \int_t^T \frac{1}{2}\Lambda\left(\tau_s A_s\right)^2 \, \mathrm{d}s \,\middle|\, X_t = x, A_t = a \right] \ ,$$

where $\mathbb{E}_{t,x,a}$ denotes conditional expectation given $X_t = x$ and $A_t = a$. The problem in Equation (5.4) then reads

$$\max_{n_0} J\left(0, x_0, s_0 n_0\right) \ .$$

The Hamilton-Jacobi-Bellman equation for the function J is

$$\begin{aligned} J_t &= -\max_{\tau} \Big\{ -\frac{1}{2}\Lambda\tau^2 a^2 + (\lambda a + rx) J_x \\ &\qquad + a(\lambda + r + \tau) J_a + \frac{1}{2}\sigma^2 a^2 \left(J_{xx} + J_{aa} + 2J_{ax}\right) \Big\} \ , \\ J(T,x,a) &= x - \gamma x^2 \ . \end{aligned}$$

The first order condition for τ implies

$$\tau^* = \frac{J_a}{a\Lambda} \ .$$

Inserting this in the Hamilton-Jacobi-Bellman equation, we obtain the PDE

$$\begin{aligned} 0 &= J_t - \frac{1}{2}\Lambda\left(\frac{J_a}{a\Lambda}\right)^2 a^2 + (\lambda a + rx) J_x \\ &\qquad + a\left(\lambda + r + \frac{J_a}{a\Lambda}\right) J_a + \frac{1}{2}\sigma^2 a^2 \left(J_{xx} + J_{aa} + 2J_{ax}\right) \\ &= J_t + (\lambda a + rx) J_x + \frac{1}{2\Lambda} J_a^2 + a(\lambda + r) J_a \\ &\qquad + \frac{1}{2}\sigma^2 a^2 \left(J_{xx} + J_{aa} + 2J_{ax}\right) \ , \\ J(T,x,a) &= x - \gamma x^2 \ . \end{aligned}$$

We make the ansatz

$$J(t,x,a) = f_1(t)\, x^2 + f_2(t)\, x + f_3(t)\, a^2 + f_4(t)\, a + f_5(t)\, ax + f_6(t) \ ,$$

where $f_1, \ldots, f_4$ are continuously differentiable functions satisfying the terminal conditions $f_1(T) = -\gamma$, $f_2(T) = 1$, $f_3(T) = f_4(T) = f_5(T) = f_6(T) = 0$.

Inserting in the PDE, we get

$$
\begin{aligned}
0 &= f_1'(t)\,x^2 + f_2'(t)\,x + f_3'(t)\,a^2 + f_4'(t)\,a + f_5'(t)\,ax + f_6'(t)\\
&\quad + (\lambda a + rx)\left(2f_1(t)\,x + f_2(t) + f_5(t)\,a\right)\\
&\quad + \tfrac{1}{2\Lambda}\left(2f_3(t)\,a + f_4(t) + f_5(t)\,x\right)^2\\
&\quad + a(\lambda + r)\left(2f_3(t)\,a + f_4(t) + f_5(t)\,x\right)\\
&\quad + \tfrac{1}{2}\sigma^2 a^2\left(2f_1(t) + 2f_3(t) + 2f_5(t)\right)\\
&= x^2\left(f_1'(t) + 2rf_1(t) + \tfrac{1}{2\Lambda}f_5^2(t)\right)\\
&\quad + x\left(f_2'(t) + rf_2(t) + \tfrac{1}{\Lambda}f_4(t)\,f_5(t)\right)\\
&\quad + a^2\Big(f_3'(t) + \left(\lambda + \sigma^2\right)f_5(t) + \tfrac{2}{\Lambda}f_3^2(t) + \sigma^2 f_1(t)\\
&\qquad\qquad + \left(2\lambda + 2r + \sigma^2\right)f_3(t)\Big)\\
&\quad + a\left(f_4'(t) + \lambda f_2(t) + \tfrac{2}{\Lambda}f_3(t)\,f_4(t) + (\lambda + r)\,f_4(t)\right)\\
&\quad + ax\left(f_5'(t) + 2\lambda f_1(t) + \tfrac{2}{\Lambda}f_3(t)\,f_5(t) + (\lambda + 2r)\,f_5(t)\right)\\
&\quad + f_6'(t) + \tfrac{1}{2\Lambda}f_4^2(t)\ .
\end{aligned}
$$

Hence, we get the following system of ODEs:

$$
\begin{aligned}
f_1'(t) &= -2rf_1(t) - \frac{1}{2\Lambda}f_5^2(t)\ ,\quad f_1(T) = -\gamma\ ,\\
f_2'(t) &= -rf_2(t) - \frac{1}{\Lambda}f_4(t)\,f_5(t)\ ,\quad f_2(T) = 1\ ,\\
f_3'(t) &= -\left(\lambda + \sigma^2\right)f_5(t) - \frac{2}{\Lambda}f_3^2(t) - \sigma^2 f_1(t) - \left(2\lambda + 2r + \sigma^2\right)f_3(t)\ ,\\
&\qquad\qquad f_3(T) = 0\ ,\\
f_4'(t) &= -\lambda f_2(t) - \frac{2}{\Lambda}f_3(t)\,f_4(t) - (\lambda + r)\,f_4(t)\ ,\quad f_4(T) = 0\ ,\\
f_5'(t) &= -2\lambda f_1(t) - \frac{2}{\Lambda}f_3(t)\,f_5(t) - (\lambda + 2r)\,f_5(t)\ ,\quad f_5(T) = 0\ ,\\
f_6'(t) &= -\frac{1}{2\Lambda}f_4^2(t)\ ,\quad f_6(T) = 0\ .
\end{aligned}
$$

We notice that the system of ODEs for f_1, f_3, and f_5 can be solved separately. The optimal trading amount rate is

$$\tau_t^* A_t = \tfrac{2}{\Lambda}f_3(t)\,A_t^* + \tfrac{1}{\Lambda}f_4(t) + \tfrac{1}{\Lambda}f_5(t)\,X_t^*\ .$$

We have that

$$
\begin{aligned}
J(0, x_0, s_0 n_0) &= f_1(0)\,x_0^2 + f_2(0)\,x_0 + f_3(0)\,n_0^2 s_0^2 + f_4(0)\,n_0 s_0\\
&\quad + f_5(0)\,n_0 s_0 x_0 + f_6(0)\,.
\end{aligned}
$$

The first order condition for optimality of n_0 reads

$$2f_3(0)\, n_0 s_0^2 + f_4(0)\, s_0 + f_5(0)\, s_0 x_0 = 0 \ .$$

Hence, the investor's optimal number of stocks at time zero is given by

$$n_0^* = -\frac{f_4(0) + f_5(0)\, x_0}{2f_3(0)\, s_0} \ .$$

5.C Calculations for Smooth Investment with Power Utility

We embed the problem in Equation (5.9) in an optimal value function V given by

$$V(t,x,a) = \min_{(\tau_s)_{t\le s\le T}} \mathbb{E}_{t,x,a}\left[\int_t^T \tfrac{1}{2}\left(\theta(\pi X_s - A_s)^2 + (\tau_s A_s)^2\right) \mathrm{d}s\right] ,$$

where $\mathbb{E}_{t,x,a}$ denotes conditional expectation given $X_t = x$ and $A_t = a$. The problem in Equation (5.9) then reads

$$\max_{n_0} V(0, x_0, s_0 n_0) \ .$$

The Hamilton-Jacobi-Bellman equation for the function V is

$$\begin{aligned} V_t = -\min_{\tau} \Big\{ &\tfrac{1}{2}\left[\theta(\pi^* x - a)^2 + (\tau a)^2\right] + (r + \lambda + \tau)\, a V_a \\ &+ (rx + \lambda a)\, V_x + \tfrac{1}{2}\sigma^2 a^2 (V_{aa} + V_{xx} + 2V_{xa}) \Big\} , \\ V(T,x,a) = 0 \ . & \end{aligned}$$

The first order condition for τ implies

$$\tau^* = -\frac{V_a}{a} \ .$$

Inserting this in the Hamilton-Jacobi-Bellman equation, we obtain the PDE

$$\begin{aligned} 0 = V_t + \tfrac{1}{2}\theta(\pi^* x - a)^2 - \tfrac{1}{2}(V_a)^2 + (r+\lambda)\, a V_a & \\ + (rx + \lambda a)\, V_x + \tfrac{1}{2}\sigma^2 a^2 (V_{aa} + V_{xx} + 2V_{xa}) & \ , \\ V(T,x,a) = 0 \ . & \end{aligned}$$

We make the ansatz

$$V(t,x,a) = \tfrac{1}{2} f(t)\, x^2 + \tfrac{1}{2} g(t)\, a^2 - h(t)\, ax \ ,$$

where f, g, and h are continuously differentiable functions satisfying the terminal conditions $f(T) = g(T) = h(T) = 0$. Inserting in the PDE, we get

$$\begin{aligned}
0 &= \tfrac{1}{2} f'(t)\, x^2 + \tfrac{1}{2} g'(t)\, a^2 - h'(t)\, ax + \tfrac{\theta}{2} (\pi^* x - a)^2 - \tfrac{1}{2} (g(t)\, a - h(t)\, x)^2 \\
&\quad + (r+\lambda)\, a\, (g(t)\, a - h(t)\, x) + (rx + \lambda a)\, (f(t)\, x - h(t)\, a) \\
&\quad + \tfrac{1}{2} \sigma^2 a^2 (g(t) + f(t) - 2h(t)) \\
&= \tfrac{1}{2} f'(t)\, x^2 + \tfrac{1}{2} g'(t)\, a^2 - h'(t)\, ax + \tfrac{\theta}{2} (\pi^*)^2 x^2 + \tfrac{\theta}{2} a^2 - \theta \pi^* xa \\
&\quad - \tfrac{1}{2} (g(t))^2 a^2 - \tfrac{1}{2} (h(t))^2 x^2 + g(t)\, ah(t)\, x \\
&\quad + (r+\lambda)\, g(t)\, a^2 - (r+\lambda)\, h(t)\, xa + rf(t)\, x^2 - rh(t)\, ax \\
&\quad + \lambda f(t)\, xa - \lambda h(t)\, a^2 + \tfrac{1}{2} \sigma^2 a^2 (g(t) + f(t) - 2h(t)) \\
&= \tfrac{1}{2} x^2 \left[f'(t) + \theta (\pi^*)^2 - (h(t))^2 + 2rf(t) \right] \\
&\quad + \tfrac{1}{2} a^2 \Big[g'(t) + \theta + \sigma^2 f(t) - 2 \left(\lambda + \sigma^2 \right) h(t) - (g(t))^2 \\
&\qquad\qquad\qquad\qquad + \left(2r + 2\lambda + \sigma^2 \right) g(t) \Big] \\
&\quad - ax \Big[h'(t) + \theta \pi^* - \lambda f(t) + (2r + \lambda)\, h(t) - g(t)\, h(t) \Big] \ .
\end{aligned}$$

Hence, we get the following system of ODEs:

$$\begin{aligned}
f'(t) &= -\theta (\pi^*)^2 + (h(t))^2 - 2rf(t) \ , \quad f(T) = 0 \ , \\
g'(t) &= -\theta - \sigma^2 f(t) + 2 \left(\lambda + \sigma^2 \right) h(t) - \left(2r + 2\lambda + \sigma^2 \right) g(t) + (g(t))^2 \ , \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad g(T) = 0 \ , \\
h'(t) &= -\theta \pi^* + \lambda f(t) - (2r + \lambda)\, h(t) + g(t)\, h(t) \ , \quad h(T) = 0 \ .
\end{aligned}$$

The optimal trading amount rate is

$$\tau_t^* A_t^* = -V_a (t, X_t^*, A_t^*) = h(t)\, X_t^* - g(t)\, A_t^* \ .$$

We have that

$$V(0, x_0, s_0 n_0) = \tfrac{1}{2} f(0)\, x_0^2 + \tfrac{1}{2} g(0)\, n_0^2 s_0^2 - h(0)\, n_0 s_0 x_0 \ .$$

The first order condition for optimality of n_0 reads

$$g(0)\, n_0 s_0^2 - h(0)\, s_0 x_0 = 0 \ .$$

Hence, the investor's optimal number of stocks at time zero is given by

$$n_0^* = \frac{h(0)\, x_0}{g(0)\, s_0} \ .$$

Chapter 6

A Two-Account Life Insurance Model for Scenario-Based Valuation Including Event Risk

Ninna Reitzel Jensen & Kristian Juul Schomacker (2015)
Risks 3 (2), pp. 183–218.

Abstract: Using a two-account model with event risk, we model life insurance contracts taking into account both guaranteed and non-guaranteed payments in participating life insurance as well as in unit-linked insurance. Here, event risk is used as a generic term for life insurance events, such as death, disability, *etc.* In our treatment of participating life insurance, we have special focus on the bonus schemes "consolidation" and "additional benefits", and one goal is to formalize how these work and interact. Another goal is to describe similarities and differences between participating life insurance and unit-linked insurance. By use of a two-account model, we are able to illustrate general concepts without making the model too abstract. To allow for complicated financial markets without dramatically increasing the mathematical complexity, we focus on economic scenarios. We illustrate the use of our model by conducting scenario analysis based on Monte Carlo simulation, but the model applies to scenarios in general and to worst-case and best-estimate scenarios in particular. In addition to easy computations, our model offers a common framework for the valuation of life insurance payments across product types. This enables comparison of participating life insurance products and unit-linked insurance products, thus building a bridge between the two different ways of formalizing life insurance products. Finally, our model distinguishes itself from the existing literature by taking into account the Markov model for the state of the policyholder and, hereby, facilitating event risk.

Keywords: Two-account model, economic scenarios, participating life insurance, unit-linked insurance, stochastic differential equations, guarantees, bonus, fairness, market valuation.

6.1 Introduction

Classical life insurance mathematics deals with the computation of reserves and cash flows for guaranteed payments in participating life insurance. Non-guaranteed payments in participating life insurance and guaranteed and non-guaranteed payments in unit-linked insurance depend on the evolution of the financial market, and this makes them difficult to model, in particular on top of the state model of the policyholder. Note that by non-guaranteed payments, we mean all future payments that are not guaranteed, with bonus in participating life insurance as the leading example. The paper Møller and Steffensen (2007) offers examples of this advanced combined modeling in the case of a Black–Scholes financial market. To lower the mathematical complexity and allow for more complicated financial markets while maintaining a general biometric state model, we focus on economic scenarios. An economic scenario could, for example, consist of a sample path for the short interest rate and/or a stock index. The scenarios may be worst-case scenarios, stress scenarios from Solvency II, scenarios generated via Monte Carlo simulation or best-estimate scenarios. For a given scenario, the balance of the policy is projected into the future. For scenarios generated via Monte Carlo simulation, one obtains a valid approximation of the expected future payments, guaranteed as well as non-guaranteed, by averaging over sufficiently many projections (as is common practice with Monte Carlo simulation). For worst-case or best-estimate scenarios, a single projection is enough to obtain the corresponding worst-case or best-estimate approximation of the future payments. We will not go into details about the generation of stochastic scenarios; we simply take them as financial input to our model. For Monte Carlo simulation, we refer to Glasserman (2004). For the generation of worst-case scenarios, we refer to Christiansen et al. (2014). Scenario-based calculations have the advantage that the overall projection approach does not change with the financial model, because the stochastic scenarios are the only financial input. Economic scenarios are widely used in the insurance industry; see, for example, Insurance Regulation Committee of the International Actuarial Association (2013) and Chapter 10 "*A Simulation-Based ALM Model in Practical Use by a Norwegian Life Insurance Company*" in Silvestrov and Martin-Löf (2014).

Considering non-guaranteed payments when valuing the liabilities has many applications, such as risk management, product development and solvency. In Solvency II Directive (2009), Paragraph 79, it is stated that the value of financial guarantees and contractual options should take into account non-guaranteed as well as guaranteed payments. Scenario-based calculations allow for market valuation, solvency calculations, hedging and pricing of guaranteed and non-guaranteed payments in participating life and unit-linked insurance. Hence, scenario-based calculations, which is exactly what we propose, are useful for complying with current Solvency II regulation.

We model life insurance contracts using two interacting accounts described by stochastic differential equations. One account measures the assets, and the other account is a technical account. For each scenario, the stochastic differential equations simplify to deterministic differential equations that can be solved numerically. A numerical solution can, for example, be obtained by applying a simple numerical discretization. Thereby, our model is simple to implement. Furthermore, our model allows us to model participating life and unit-linked insurance in the same framework. By doing so, we are able to compare the two. In their nature, unit-linked and participating life insurance seem different, but they are really not. The products may vary in riskiness, but projection-wise, they are almost the same. The main difference lies in the specification of how non-guaranteed payments arise, stated in the contract from the beginning (unit-linked) or determined fairly by the company along the way (participating life).

By use of a two-account model, we are able to illustrate general concepts without making the model too abstract. Our two-account model is based on the two-account model in Steffensen and Waldstrøm (2009), and both models offer a common framework for the valuation of guaranteed and non-guaranteed payments in participating life and unit-linked insurance. Our model distinguishes itself from the model in Steffensen and Waldstrøm (2009) by taking into account the Markov model for the state of the policyholder, thereby including event risk. Here, event risk is used as a generic term for life insurance events, such as death, disability, *etc.* The existing literature considers the valuation of guaranteed and non-guaranteed payments in participating life insurance or unit-linked insurance without event risk, whereas a common framework and inclusion of event risk are rare. The papers Bauer et al. (2006); Zaglauer and Bauer (2008); Bauer et al. (2010); Bohnert and Gatzert (2012) are examples of recent literature that considers valuation in participating life insurance without (or with very limited) event risk. The papers Gatzert and Kling (2007); Graf et al. (2011); Kling et al. (2007) are other examples within the same area, but their focus is more risk-related. On the other hand, Norberg (2001) models participating life insurance, taking into account the Markov model for the state of the policyholder, but the model is only tractable for a very simple financial environment, and it does not apply to unit-linked insurance. The work in Møller and Steffensen (2007) and, more specifically, Steffensen (2006) cover participating life and unit-linked insurance with broad event risk, but only in a Black–Scholes market, and the results involve non-trivial partial differential equations.

In our treatment of participating life insurance, we have special focus on bonus allocation and on the bonus schemes "consolidation" and "additional benefits". These bonus schemes are the most common in the Danish life insurance and pensions industry, but to our knowledge, consolidation is barely mentioned in the literature. An important goal of this paper is to formalize exactly how these bonus schemes work and interact. Other papers with a similar focus on bonus include Grosen and Jørgensen (2000); Jensen et al. (2001);

Hansen and Miltersen (2002); Miltersen and Persson (2003), but again, none of them include event risk. In our treatment of unit-linked insurance, we have special focus on the implementation of guarantees and on the similarities and differences in relation to participating life insurance. For both product types, we include numerical examples to demonstrate the possible applications of our two-account model.

In Section 6.2, we discuss scenario-based projection in general. Our main focus is on projection level and which measure to project under (physical or pricing measure). In Section 6.3, we discuss valuation bases in life insurance and formalize a common model for the state-wise evolution of the policies under consideration. In Section 6.4, we consider participating life insurance. We briefly touch upon different bonus schemes, and we present our two-account model for a general participating life insurance policy, although not allowing for policyholder behavior options. We include simple survival model examples to illustrate formulas and provide intuition. We end the section with a numerical example building on the survival model. The example illustrates a fair bonus strategy and the risk of unfair redistribution between policies in a portfolio. It also highlights some of the many possible applications of scenario-based calculations. In Section 6.5, we consider unit-linked insurance. We touch upon different aspects of unit-linked insurance, and we present our two-account model for a general unit-linked insurance policy. Again, we include simple and illustrative survival model examples. We end the section with a numerical example that is a unit-linked version of the numerical example in the previous section. The example illustrates a fair guarantee fee strategy, and we compare the unit-linked insurance policy to its participating life insurance counterpart, making good use of our common modeling framework.

6.2 Projection in General

In participating life insurance, we introduce stochastic scenarios to allow for market valuation of non-guaranteed payments, pricing and hedging of guarantees, bonus and benefit prognoses and solvency calculations. In unit-linked insurance, we introduce stochastic scenarios to utilize retirement savings and benefit prognoses, solvency calculations and hedging and pricing of unit-linked guarantees. In both cases, each scenario consists of two sample paths: one for the short interest rate, r, and one for the return on the fund that the policyholder and/or the insurance company has chosen to invest in, R_X. We assume that the stochastic scenarios arise from a financial model equipped with a physical measure P and a risk-neutral pricing measure Q. For each scenario, we project the accounts that determine the financial progress of a given policy. When making the projections in participating life insurance, as well as in unit-linked insurance, it is important to bear in mind the requested outcome.

For pricing, hedging, market valuation and solvency assessments of guaranteed and non-guaranteed payments (in participating life) or of a provided guarantee (in unit-linked) or for examining a bonus allocation strategy (in participating life), it is the expected evolution of the policy, both financially and across states, that is relevant. Hence, the evolution of the policy is considered on an average "portfolio level". For pricing, hedging and market valuation, the projections are carried out under the pricing measure (Q), since the focus is on pricing and valuation. For solvency assessments, the projections are carried out under the physical measure (P) up to some relevant time point, and from then on, they are carried out under the pricing measure (Q). For examining a bonus allocation strategy and quantifying the total expected future bonus, the projections are carried out under the physical measure (P), since the focus is on the actual outcome.

For retirement savings, benefit and bonus prognoses, it is the expected financial evolution of the policy that is relevant. The policyholder wants to know what to expect in each state, not the average expectation. Hence, the evolution of the policy is considered on an individual "policy level". However, in participating life insurance, the amount of bonus allocated to the policy depends on the financial evolution and the average evolution of the policy. Hence, for the purpose of prognoses in participating life insurance, the assets and the reserves must, first, be projected on portfolio level to produce a sample path for the bonus allocation. Second, the sample path for the bonus allocation is used to project the reserves on an individual path-wise "policy level". In either case, the projections are carried out under the physical measure (P), since the focus is on the actual bonus, retirement savings and benefits.

In this paper, we limit our focus to projection on portfolio level and leave projection on policy level for future research.

6.3 Valuation Bases and Insurance Model

A cornerstone in life insurance mathematics is the principle of equivalence, which states that the expected present values of premiums and benefits should be equal. The principle relies on the law of large numbers that will, then, on average, make premiums and benefits balance in a large insurance portfolio. To apply the equivalence principle, one needs assumptions about interest, mortality and other relevant economic-demographic elements. The uncertain development of these elements subjects the insurance company to a risk that is independent of the size of the portfolio. In participating life insurance, the insurance company can neither raise the premiums nor reduce the benefits along the way, so the only way for the insurance company to mitigate this risk is to build a safety loading into the premiums. This is done by performing the equivalence principle under conservative assumptions about interest, mortality, *etc.* These assumptions make up the so-called technical basis, and it represents a provisional worst-case scenario for its elements. Below, we mark elements of the technical basis by superscript "$*$". For market-consistent valuation of future payments, the technical basis does not apply due to its worst-case nature. Instead, valuation is performed under the so-called market basis, which is made up of best-estimate assumptions about the various elements. Below, we mark elements of the market basis by superscript "m". In unit-linked insurance, the benefits are typically allowed to fluctuate with the market, hereby making the technical basis superfluous.

In participating life insurance, as well as in unit-linked insurance, we consider a policy whose state-wise evolution is governed by a continuous-time Markov process Z with a finite state space $\mathcal{J}$, starting at zero. For a detailed description of the Markov model, see Norberg (1991) or Norberg (1999). For $k, j \in \mathcal{J}, j \neq k$, we define the counting process N_{jk} and the indicator process I_k by

$$\begin{aligned} N_{jk}(t) &= \#\{s \leq t : Z(s-) = j, Z(s) = k\} \ , \\ I_k(t) &= 1_{\{Z(t)=k\}} \ . \end{aligned}$$

With this definition, $N_{jk}(t)$ counts the number of jumps from state j to state k until time t, and $I_k(t)$ indicates a sojourn in state k at time t. Under the technical basis, we model the evolution of Z by the transition intensities $t \mapsto \mu^*_{jk}(t)$, $j, k \in \mathcal{J}, j \neq k$, and under the market basis, we model the evolution of Z by the transition intensities $t \mapsto \mu^m_{jk}(t)$, $j, k \in \mathcal{J}, j \neq k$. The corresponding technical and market transition probabilities from state j to state k over the time interval $[t, s]$ are denoted by $p^*_{jk}(t, s)$ and $p^m_{jk}(t, s)$, and with $\circ = *, m$ indicating the basis, we have

$$\mu^\circ_{jk}(t) = \lim_{h \downarrow 0} \frac{p^\circ_{jk}(t, t+h)}{h} \ .$$

The transition probabilities can be calculated numerically from the transition intensities by use of the Kolmogorov equations; see, for example, Norberg (1991). We assume that the process Z governing the state of the policy is independent of the financial market, and under both P and Q, the evolution of Z is described by the transition intensities from the market basis.

In addition to the market transition intensities, the market basis consists of a market interest rate. The market basis has no more elements, as we do not take expenses or any other economic-demographic elements into account. Similarly, the technical basis consists of the technical transition intensities and a technical interest rate. In unit-linked insurance, only the market basis comes into play.

6.4 Participating Life Insurance

In this section, we consider participating life insurance. We touch upon different bonus schemes, and we present our two-account model for a general participating life insurance policy. We include simple survival model examples to illustrate formulas and provide intuition. We end the section with a numerical example building on the survival model.

6.4.1 Non-Guaranteed Payments (Bonus)

In participating life insurance, the guaranteed payments are based on the technical basis. The conservative technical basis gives rise to a systematic surplus that is to be paid back to the policyholders in terms of bonus. There are many possible ways to do so. For a short survey, see Møller and Steffensen (2007). We consider a bonus scheme consisting of two steps: first, consolidation, and then, when the policy is consolidated on a sufficiently low technical interest rate (if ever), additional benefits. The bonus scheme consolidation (in Danish "*styrkelse*") is much used in the Danish market, but it can easily be skipped below, heading straight for the bonus scheme additional benefits.

The bonus scheme consolidation is primarily used for policies with a technical interest rate that is "too high" compared to the market interest rate. Bonus is used to consolidate the policy on a lower technical interest rate. By consolidate, we mean that the technical interest rate is lowered without changing the guaranteed payments. This may seem to be less favorable for the policyholder, but since the guaranteed payments are not changed, the policyholder is not worse off. When a sufficiently low technical interest rate has been reached, the remaining bonus is used for additional benefits. Hence, consolidation does not benefit the policyholder in terms of more favorable payments immediately after bonus payments, but it helps to ensure that the liabilities of the policy can be met. Furthermore, the lower technical interest rate gives rise to a higher systematic surplus in the future, which will eventually be redistributed and reflected in the payments.

The bonus scheme additional benefits is primarily used for policies with a low technical interest rate compared to the market interest rate. Bonus is used to increase parts of the guaranteed benefits proportionally, whereas the remaining benefits, the premiums and the technical interest rate are maintained. It is usually the retirement part of the benefits (such as a pure endowment or a life annuity) that is increased and the insurance part of the benefits (such as a term insurance or disability coverage) that is not. There is good reason to increase the retirement part of the benefits instead of decreasing the premiums or increasing all of the benefits, since the retirement benefits are typically set according to which premiums the policyholder can afford and which insurance coverage he/she needs, and not the other way around. Furthermore, there is good reason to increase the retirement benefits proportionally, as the benefit profile reflects the policyholder's preferences.

6.4.2 Product Specification

We consider a participating life insurance policy with guaranteed payments based on a technical basis whose elements are marked by superscript "$*$". The state-wise evolution of the policy is described in Section 6.3. We let $r^{*(0)}$ denote the technical interest rate at time 0. By B^u, B^f and C, we denote the guaranteed payment streams at time 0. Here, C is the premium stream ("C" for contributions), B^u is the benefit stream for the benefits that are increased ("B" for benefits and superscript "u" for upscaled) and B^f is the benefit stream for the benefits that are kept fixed (superscript "f" for fixed). The payments streams are given by

$$\begin{aligned}
\mathrm{d}C &= \sum_{j \in \mathcal{J}} I_j \, \mathrm{d}c_j \ , \\
\mathrm{d}B^i &= \sum_{j \in \mathcal{J}} I_j \, \mathrm{d}b_j^i + \sum_{j,k \in \mathcal{J}: k \neq j} b_{jk}^i \, \mathrm{d}N_{jk} \ , \quad i = f, u \ ,
\end{aligned}$$

where c_j, b_j^f and b_j^u are deterministic, state-wise payment streams and b_{jk}^f and b_{jk}^u are deterministic lump sum payments upon jumps. We note that we, hereby, exclude policyholder behavior options, such as surrender and free policy, since they imply non-deterministic payments. However, for surrender modeling, see the remark on Page 152. Examples of deterministic lump sum payments upon jumps include insurance coverage, such as a death sum, a disability sum or a sum upon critical illness. The policy terminates at time T. Thereafter, there are no payments.

6.4.3 Two-Account Model

We denote by X the assets of the policy, including its share of the collective bonus potential, and by Y, we denote the market expected technical reserve for the policy. By market expected technical reserve, we mean the expectation of future state-wise technical reserves where the expectation across states is taken under the market basis. Thus, the market expected technical reserve is not a state-wise reserve, but a probability weighted sum of state-wise reserves. The accounts X and Y are the backbone of our two-account model. The policy is issued before or at time 0, and the two accounts amount to $X(0-) = x_0$ and $Y(0-) = y_0$ just before time 0. For a policy issued at time 0, y_0 and x_0 are both zero. For a policy issued before time 0, y_0 is equal to the technical reserve for the policy just before time 0, and x_0 is equal to the assets of the policy just before time 0. Both are assumed to be known when initiating the projection.

The assets X are invested in a fund with stochastic return R_X. The market expected technical reserve Y accumulates according to the technical interest rate. In good times, the return rate on the assets exceeds the technical interest rate. Parts of the excess return are allocated to the policy in terms of bonus, which adds to the market expected technical reserve, but parts are saved for times where the return rate on the assets is less favorable. In really bad times, the assets may be insufficient to cover the guaranteed payments of the policy. In that case, the equity holders of the insurance company step in with a capital injection taken from the company's equity. We speak of the possible capital injection as a guarantee injection, and its role is to raise the assets in case of unfavorable developments in the financial market. The policyholder pays for the company's risk taking by having a guarantee fee deducted from the assets and paid to the equity holders of the insurance company in good times. We assume that the insurance company's equity is always sufficient to cover the guarantee injections and that all guarantee injections and guarantee fees are settled via the equity. In Denmark, the guarantee fee used to be known as the "*driftsherretillæg*" (translates to "technical yield"). All of the above does not happen continuously, but at pre-specified, deterministic time points $0 < t_1 < \ldots < t_n = T$ (for example, once a year) where the two accounts X and Y are updated. We let

$$\varepsilon(t) = \#\{i = 1, \ldots, n : t_i \leq t\}$$

count the number of updates prior to time t. The updates consist of bonus allocation d (if funds are sufficient), guarantee injection g from the equity holders of the insurance company (if needed) and deduction of the guarantee fee π_g in return for the possible guarantee injection. All three are non-negative. For technical convenience, we assume that the stochastic return on the assets, R_X, does not jump at the time points $0 < t_1 < \ldots < t_n = T$ with account updates. Furthermore, for all t with $\mathrm{d}\varepsilon(t) = 1$, *i.e.*, for all time points with an account update, we assume that $d(t)$ and $\pi_g(t)$ are known at time $t-$ and that $g(t)$ is calculated at time $t-$. This is to ensure predictability and, thereby, stochastic integrability.

6.4.4 Bonus Mechanisms

As mentioned, we consider a bonus scheme where bonus allocated to the policy is, first, used to lower the technical interest rate until it hits a pre-described level r^*. Typically, this level coincides with the technical interest rate for new policies. Thereafter, bonus is used to increase the benefits B^u. The additional benefits are priced using the technical transition intensities μ^*_{jk} and the technical interest rate r^*. This means that the minimum technical interest rate for consolidation and the pricing interest rate for additional benefits are assumed to coincide. One could have chosen another technical interest rate for the pricing of additional benefits, but that would require a division of the technical reserve on two different technical bases, so we insist on using r^*. We let $r^{*(n)}$ denote the technical interest rate after the n-th bonus accrual and $k^{(n)}$ denote the upscaling of the benefits B^u after the n-th bonus accrual. We note that the upscaling factor starts at one, *i.e.*, $k^{(0)} = 1$. After the n-th bonus accrual, the guaranteed benefit stream for the policy is given by

$$B^{(n)} = k^{(n)} B^u + B^f \ .$$

We point out that $r^{*(n)}$ and $k^{(n)}$ depend on the development of the financial market and are therefore stochastic. However, for each economic scenario, we have a procedure for calculating them according to the equivalence principle. The procedure is presented in Section 6.4.8. We note that we have $k^{(n)} = 1$ for all n with $r^{*(n)} > r^*$, and if $k^{(n)} > 1$, then necessarily $r^{*(n)} = r^*$. This is because we do not increase the guaranteed benefits until the technical interest rate has been lowered to r^*. Finally, we note that the technical interest rate and the upscaling factor amount to $r^{*(\varepsilon(t))}$ and $k^{(\varepsilon(t))}$ at time t, since there has been $\varepsilon(t)$ account updates at time t.

For all t with $\mathrm{d}\varepsilon(t) = 1$, that is for all time points with an account update, we assume that the technical interest rate $r^{*(\varepsilon(t))}$ and the upscaling factor $k^{(\varepsilon(t))}$ are calculated at time $t-$. Again, this is to ensure predictability. Furthermore, additional benefits are in effect from time $t-$, such that benefits paid out at time t include the upscaling $k^{(\varepsilon(t))}$. The latter ensures that a policyholder with a final lump sum payment actually benefits from the last bonus update.

6.4.5 Technical Reserves

We denote by $V_j^{f,*,+}(\cdot,\rho)$ and $V_j^{u,*,+}(\cdot,\rho)$ the state-wise technical benefit reserves for the benefit streams B^f and B^u given that the policy is in state j and that the technical interest rate is ρ. Similarly, we denote by $V_j^{*,-}(\cdot,\rho)$ the state-wise technical premium reserves for the premium stream C. Note that we use superscript "+" to indicate the benefit reserves and superscript "−" to indicate the premium reserve. Furthermore, we use superscript "∗" to indicate that the reserve is evaluated under the technical basis. Finally, we use the generic constant ρ in place of the technical interest rate, because we need to evaluate the technical reserves for different technical interest rates in connection with the bonus scheme consolidation. We have

$$\begin{aligned} V_j^{*,-}(t,\rho) &= \mathbb{E}^*\left[\int_t^T e^{-\rho(s-t)}\,\mathrm{d}C(s)\,\middle|\,Z(t)=j\right] \\ &= \int_t^T e^{-\rho(s-t)} \sum_{l\in\mathcal{J}} p_{jl}^*(t,s)\,\mathrm{d}c_l(s) \ , \\ V_j^{i,*,+}(t,\rho) &= \mathbb{E}^*\left[\int_t^T e^{-\rho(s-t)}\,\mathrm{d}B^i(s)\,\middle|\,Z(t)=j\right] \\ &= \int_t^T e^{-\rho(s-t)} \sum_{l\in\mathcal{J}} p_{jl}^*(t,s) \left\{\mathrm{d}b_l^i(s) + \sum_{k\in\mathcal{J}:k\neq l} \mu_{lk}^*(s)\, b_{lk}^i(s)\,\mathrm{d}s\right\}, \end{aligned} \tag{6.1}$$

for $i = f, u$. Here, $\mathbb{E}^*$ denotes technical expectation and p^*_{jl} is the technical probability of transition from state 0 to j. Both are determined by the transition intensities from the technical basis. The state-wise technical reserves can be calculated numerically by use of Thiele's differential equations; see Hoem (1969).

We denote by $V_i^*(\cdot, \rho, k)$ the state-wise technical reserve for the (partly upscaled by k) payment stream $B^f + kB^u - C$, given that the policy is in state i and that the technical interest rate is ρ, *i.e.*,

$$V_i^*(t, \rho, k) = kV_i^{u,*,+}(t, \rho) + V_i^{f,*,+}(t, \rho) - V_i^{*,-}(t, \rho) \ , \quad i \in \mathcal{J} \ . \quad (6.2)$$

Here, $V_j^{f,*,+}$, $V_j^{u,*,+}$, and $V_j^{*,-}$ are the state-wise technical benefit and premium reserves defined in Equation (6.1). With the introduction of V_i^*, we can write the market expected technical reserve as

$$Y(t) = \sum_{j \in \mathcal{J}} p^m_{0j}(0, t) V_j^*\left(t, r^{*(\varepsilon(t))}, k^{(\varepsilon(t))}\right) \ .$$

We recall that p^m_{0j} is the market probability of transition from state 0 to j, which is determined by the transition intensities from the market basis. We note that the stochasticity in $Y(t)$ comes from the stochastic development of the technical interest rate $r^{*(\varepsilon(t))}$ and the upscaling factor $k^{(\varepsilon(t))}$. However, for each t, the technical interest rate $r^{*(\varepsilon(t))}$ is determined as a constant interest rate over $[t, T]$, so we never plug a non-constant technical interest rate into the reserves in Equation (6.1) when calculating $V_j^*\left(t, r^{*(\varepsilon(t))}, k^{(\varepsilon(t))}\right)$.

Example 6.1 (Survival model). We consider a simple example that provides the basis for numerical illustrations later on. The state of the policy is described by the classical survival model with two states, 0 (alive) and 1 (dead). The payments of the policy consist of a constant continuous premium payment π while alive, a term insurance sum b^{ad} upon death before expiration T and a pure endowment sum b^a upon survival until expiration T. Under the bonus scheme "additional benefits", bonus is used to increase the endowment sum. There are no payments in the death state. For simplicity, we write $I = I_1$, $N = N_{01}$, $\mu^\circ = \mu^\circ_{01}$ and $p^\circ = p^\circ_{00}$ for $\circ = *, m$, and we have

$$p^\circ(s, t) = e^{-\int_s^t \mu^\circ(v)\, \mathrm{d}v} \ , \quad s \le t \ .$$

The payment streams of the policy read

$$\begin{aligned} \mathrm{d}C(t) &= \pi I \, \mathrm{d}t \ , \quad t \le T \ , \\ \mathrm{d}B^f(t) &= b^{ad} \, \mathrm{d}N(t) \ , \quad t \le T \ , \\ \mathrm{d}B^u(t) &= b^a I(t) \, \mathrm{d}\varepsilon_T(t) \ , \quad t \le T \ , \end{aligned}$$

where ε_T is the Dirac measure in T, *i.e.*, for a measurable set $A \subseteq \mathbb{R}$

$$\varepsilon_T(A) = 1_{\{T\}}(A) = \begin{cases} 1 & \text{for } T \in A, \\ 0 & \text{for } T \notin A. \end{cases}$$

We note that

$$\int_0^T \mathrm{d}B^u(t) = b^a I(T) \ .$$

The technical premium and benefit reserves are zero in the state "dead", and in the state "alive", they read

$$\begin{aligned} V^{*,-}(t,\rho) &= \int_t^T e^{-\rho(s-t)} p^*(t,s)\,\pi\,\mathrm{d}s \\ &= \pi \int_t^T e^{-\rho(s-t)} e^{-\int_t^s \mu^*(v)\,\mathrm{d}v}\,\mathrm{d}s\ , \quad t \le T\ , \\ V^{f,*,+}(t,\rho) &= \int_t^T e^{-\rho(s-t)} p^*(t,s)\, b^{ad} \mu^*(s)\,\mathrm{d}s \\ &= b^{ad} \int_t^T e^{-\rho(s-t)} e^{-\int_t^s \mu^*(v)\,\mathrm{d}v} \mu^*(s)\,\mathrm{d}s\ , \quad t \le T\ , \\ V^{u,*,+}(t,\rho) &= \int_t^T e^{-\rho(s-t)} p^*(t,s)\, b^a\,\mathrm{d}\varepsilon_T(s) \\ &= b^a e^{-\rho(T-t)} e^{-\int_t^T \mu^*(v)\,\mathrm{d}v}\ , \quad t \le T\ . \end{aligned}$$

6.4.6 Cash Flows

For projection on portfolio level, it it useful to consider market cash flows of the policy. Here, we use the term market cash flows for the expectation of the stochastic payment streams taken under the market basis. By ς, β^f and β^u, we denote the time 0 market cash flows for the premium stream C and the benefit streams B^f and B^u, *i.e.*,

$$\begin{aligned} \varsigma(t) &= \mathbb{E}^m\left[\int_0^t \mathrm{d}C(s)\right] \\ &= \sum_{j\in\mathcal{J}} \int_0^t p_{0j}^m(0,s)\,\mathrm{d}c_j(s)\ , \\ \beta^i(t) &= \mathbb{E}^m\left[\int_0^t \mathrm{d}B^i(s)\right] \\ &= \sum_{j\in\mathcal{J}} \int_0^t p_{0j}^m(0,s) \left\{ \mathrm{d}b_j^i(s) + \sum_{k\in\mathcal{J}:k\neq j} \mu_{jk}^m(s)\, b_{jk}^i(s)\,\mathrm{d}s \right\}\ , \quad i = f,u\ , \end{aligned}$$

where the expectation $\mathbb{E}^m$ is taken under the market basis. Furthermore, we need the market expected market reserve. By $V(t)$, we denote the market expected market reserve at time t for the most recently guaranteed payment stream $B^{(\varepsilon(t))} - C = B^f + k^{(\varepsilon(t))}B^u - C$, *i.e.*,

$$\begin{aligned} V(t) &= \mathbb{E}^m_{k^{(\varepsilon(t))}}\left[\mathbb{E}^m_{k^{(\varepsilon(t))}}\left[\int_t^T e^{-\int_t^s r(v)\,\mathrm{d}v}\,\mathrm{d}\left(B^{(\varepsilon(t))} - C\right)(s)\middle| Z(t)\right]\right] \\ &= \int_t^T e^{-\int_t^s r_t(v)\,\mathrm{d}v}\left(k^{(\varepsilon(t))}\,\mathrm{d}\beta^u(s) + \mathrm{d}\beta^f(s) - \mathrm{d}\varsigma(s)\right) . \end{aligned} \tag{6.3}$$

Here, $\mathbb{E}^m_{k^{(\varepsilon(t))}}$ denotes market expectation given $k^{(\varepsilon(t))}$, r is the stochastic short interest rate and r_t is the yield curve seen from time t. Similar to the market expected technical reserve, the market expected market reserve is not a state-wise reserve, but a market probability weighted sum of state-wise market reserves. This is not evident from the formula above, since the reserve simplifies due to the tower property. We emphasize that only additional benefits, and not consolidation, raise the guarantee. However, consolidation has an effect on the non-guaranteed benefits as the technical reserve increases.

Example 6.2 (Survival model continued)**.** For the simple policy in Example 6.1, the time 0 market premium and benefit cash flows read

$$\begin{aligned} \varsigma(t) &= \int_0^t p^m(0,s)\,\pi\,\mathrm{d}s = \pi\int_0^t e^{-\int_0^s \mu^m(v)\,\mathrm{d}v}\,\mathrm{d}s\ ,\ t \le T\ , \\ \beta^f(t) &= \int_0^t p^m(0,s)\,b^{ad}\mu^m(s)\,\mathrm{d}s = b^{ad}\int_0^t e^{-\int_0^s \mu^m(v)\,\mathrm{d}v}\mu^m(s)\,\mathrm{d}s\ ,\ t \le T\ , \\ \beta^u(t) &= \int_0^t p^m(0,s)\,b^a\,\mathrm{d}\varepsilon_T(s) = b^a e^{-\int_0^T \mu^m(v)\,\mathrm{d}v} I_{\{t \ge T\}}\ ,\ t \le T\ . \end{aligned}$$

The market expected market reserve V reads

$$\begin{aligned} V(t) &= e^{-\int_t^T r_t(v)\,\mathrm{d}v} e^{-\int_0^T \mu^m(v)\,\mathrm{d}v} k^{(\varepsilon(t))} b^a \\ &\quad + \int_t^T e^{-\int_t^s r_t(v)\,\mathrm{d}v} e^{-\int_0^s \mu^m(v)\,\mathrm{d}v}\left(b^{ad}\mu^m(s) - \pi\right)\mathrm{d}s\ . \end{aligned}$$

6.4.7 Two-Account Projection

On portfolio level, the assets X and the market expected technical reserve Y of the policy evolve according to the stochastic differential equations (SDEs)

$$\begin{aligned} \mathrm{d}X(t) &= X(t-)\,\mathrm{d}R_X(t) - \mathrm{d}\beta^f(t) - k^{(\varepsilon(t))}\mathrm{d}\beta^u(t) + \mathrm{d}\varsigma(t) \\ &\qquad + \left[g(t) - \pi_g(t)\right]\mathrm{d}\varepsilon(t)\ \ ,\ \ t \le T\ , \\ X(0-) &= x_0\ , \\ \mathrm{d}Y(t) &= Y(t)\,r^{*(\varepsilon(t))}\mathrm{d}t - \mathrm{d}\beta^f(t) - k^{(\varepsilon(t))}\mathrm{d}\beta^u(t) + \mathrm{d}\varsigma(t) \\ &\qquad + d(t)\,\mathrm{d}\varepsilon(t) + \alpha\left(t, r^{*(\varepsilon(t))}, k^{(\varepsilon(t))}\right)\mathrm{d}t\ ,\ \ t \le T\ , \\ Y(0-) &= y_0\ . \end{aligned} \tag{6.4}$$

Here, α is an adjustment term given by

$$\alpha(t,\rho,k) = \sum_{j\in\mathcal{J}} \sum_{l\in\mathcal{J}:l\neq j} p^m_{0j}(0,t)\left(\mu^*_{jl}(t) - \mu^m_{jl}(t)\right) \times \left(V^*_j(t,\rho,k) - kb^u_{jl}(t) - b^f_{jl}(t) - V^*_l(t,\rho,k)\right) ,$$

where $V^*_i(t,\rho,k)$ are the state-wise technical reserves defined in Equation (6.2). The adjustment term accounts for the market expected surplus arising from the conservative technical transition intensities. See, for example, Norberg (2001).

We recall that R_X is the stochastic return on the assets, g is the guarantee injection provided by the equity holders of the insurance company, π_g is the guarantee fee deducted from the assets and paid to the equity holders, d is the allocated bonus and ε counts the number of updates of guarantee injection, guarantee fee and bonus (typically annual). The bonus d and the guarantee fee π_g are specified by the company, whereas the guarantee injection g is designed to ensure that the assets are at least equal to the guaranteed liabilities. We define the guaranteed liabilities L as the maximum of the market expected market reserve and the market expected technical reserve for the guaranteed payments, *i.e.*,

$$L(t) = \max\{V(t), Y(t)\} . \tag{6.5}$$

This definition has been common practice in Denmark since the introduction of market values. However, the guaranteed liabilities can easily be defined differently, for example $L = V$. The guarantee injection $g(t)$ is calculated according to the formula

$$g(t) = (L(t-) - (X(t-) - \pi_g(t)))^+ . \tag{6.6}$$

This guarantee design ensures that the assets X are sufficient to cover the guaranteed liabilities L after the guarantee fee π_g has been paid to the equity holders of the insurance company. The guaranteed liabilities L represent the lowest amount that the insurance company can set aside for the guaranteed payments. Hence, the assets should always exceed the guaranteed liabilities, and by design of the guarantee injection, this will always be the case after adding the guarantee injection. The inclusion of the guarantee fee is a technicality that ensures that the assets are not drained by guarantee fee payments to the equity holders of the insurance company in bad times where the liabilities exceed the assets. By the design of the guarantee injection, no guarantee fee is deducted from the assets in those times. In Section 6.4.9, we get into details about how the bonus allocation and guarantee fee are determined.

The stochastic element R_X enters via a sample path for the asset returns. Furthermore, the size of the guarantee injection g depends on the sample path for the short interest rate. In practice, one will often work with a discretized version of the stochastic differential equations in Equation (6.4). For an example, see Section 6.4.11.

Example 6.3 (Survival model continued)**.** For the simple policy in Example 6.1–6.2, the adjustment term α reads

$$\alpha\left(t,\rho,k\right)=p^{m}\left(0,t\right)\left(\mu^{*}\left(t\right)-\mu^{m}\left(t\right)\right)\left(V^{*}\left(t,\rho,k\right)-b^{ad}\right)\ ,$$

where the total technical reserve V^{*} in the state "alive" is given by

$$\begin{aligned}V^{*}\left(t,\rho,k\right)&=kV^{u,*,+}\left(t,\rho\right)+V^{f,*,+}\left(t,\rho\right)-V^{*,-}\left(t,\rho\right)\\&=kb^{a}e^{-\rho(T-t)}e^{-\int_{t}^{T}\mu^{*}(v)\,\mathrm{d}v}\\&\quad+\int_{t}^{T}e^{-\rho(s-t)}e^{-\int_{t}^{s}\mu^{*}(v)\,\mathrm{d}v}\left(b^{ad}\mu^{*}\left(s\right)-\pi\right)\mathrm{d}s\ ,\quad t\leq T\ .\end{aligned}$$

6.4.8 Procedure for Determining the Technical Interest Rate and the Upscaling Factor

Assume that $\mathrm{d}\varepsilon\left(t\right)=1$, meaning that there is an update at time t. In determining the technical interest rate $r^{*(\varepsilon(t))}$ and the upscaling factor $k^{(\varepsilon(t))}$, the distribution of the policy across states at time t enters. The distribution depends on the choice of basis; in our case, the technical basis or the market basis. The market basis reflects the true distribution of the policy across states. Therefore, we strongly suggest to work under the market basis. Working under the technical basis has the advantage that the tower property applies (see below), which limits the number of computations. However, taking the short cut and using the artificial technical basis leads to a twisted picture of the evolution of the policy, so we discourage it. For completeness, we include both options and model them by $\circ$ below.

Assume that $r^{*(\varepsilon(t-))}>r^{*}$, so that the policy is still in the consolidation phase of the bonus scheme. Then, necessarily, $k^{(\varepsilon(t-))}=1$ (since we consolidate first), and the technical interest rate $r^{*(\varepsilon(t))}$ is determined as the solution to the equation

$$Y\left(t-\right)+d\left(t\right)=V^{*,\circ}\left(t-,r^{*(\varepsilon(t))}\right)\ ,\tag{6.7}$$

where $V^{*,\circ}\left(\cdot,\rho\right)$ is the market or technical (indicated by the $\circ$) expected technical reserve for the payment stream $B^{(0)}-C=B^{f}+B^{u}-C$, given that the technical interest rate is ρ. That is

$$\begin{aligned}V^{*,\circ}\left(t,\rho\right)&=\mathbb{E}^{\circ}\left[\mathbb{E}^{*}\left[\int_{t}^{T}e^{-\rho(s-t)}\,\mathrm{d}\left(B^{f}+B^{u}-C\right)\left(s\right)\middle|Z\left(t\right)\right]\right]\\&=\mathbb{E}^{\circ}\left[V_{Z(t)}^{*}\left(t,\rho,1\right)\right]\ ,\end{aligned}$$

where the state-wise technical reserves V_{j}^{*}, $j\in\mathcal{J}$, are given in Equation (6.2), and $\mathbb{E}^{\circ}$ denotes market or technical expectation. Hence, $r^{*(\varepsilon(t))}$ is the technical interest rate that complies with the equivalence principle on portfolio level.

Under the technical basis, the tower property applies, and the reserve simplifies to

$$\begin{aligned} V^{*,*}(t,\rho) &= \sum_{j\in\mathcal{J}} p^{*}_{0j}(0,t)\, V^{*}_{j}(t,\rho,1) \\ &= \int_t^T e^{-\rho(s-t)}\,\mathrm{d}\left(\beta^{f,*}+\beta^{u,*}-\varsigma^{*}\right)(s)\,, \end{aligned}$$

where ς^*, $\beta^{f,*}$ and $\beta^{u,*}$ are the time 0 technical cash flows for the premium stream C and the benefit streams B^f and B^u. This means that the reserve can be calculated using only the technical cash flows. Under the market basis, the reserve reads

$$V^{*,m}(t,\rho) = \sum_{j\in\mathcal{J}} p^{m}_{0j}(0,t)\, V^{*}_{j}(t,\rho,1) \ .$$

Hence, using the market basis, both transition probabilities and state-wise technical reserves are needed in order to solve Equation (6.7). This is a drawback, but in our opinion, it is not enough to switch to the artificial technical basis. If the solution $r^{*(\varepsilon(t))}$ is strictly smaller than r^*, then $r^{*(\varepsilon(t))}$ is set to r^*, and the remaining bonus

$$Y(t-) + d(t) - V^{*,\circ}(t-,r^*)$$

is used to raise the upscaling factor $k^{(\varepsilon(t))}$ as below. Otherwise, we set $k^{(\varepsilon(t))} = 1$.

Now, assume that $r^{*(\varepsilon(t-))} = r^*$. Then, the policy is in the additional benefits phase of the bonus scheme, and we set $r^{*(\varepsilon(t))} = r^*$. The upscaling factor $k^{(\varepsilon(t))}$ is determined as the solution to the equation

$$d(t) = \left(k^{(\varepsilon(t))} - k^{(\varepsilon(t-))}\right) V^{u,*,\circ,+}(t-) \ ,$$

i.e.,

$$k^{(\varepsilon(t))} = k^{(\varepsilon(t-))} + \frac{d(t)}{V^{u,*,\circ,+}(t-)} \ .$$

Here, $V^{u,*,\circ,+}$ is the market or technical (indicated by the $\circ$) expected technical reserve for the benefit stream B^u, given that the interest rate is r^*, *i.e.*,

$$\begin{aligned} V^{u,*,\circ,+}(t) &= \mathbb{E}^{\circ}\left[\mathbb{E}^{*}\left[\int_t^T e^{-r^*(s-t)}\,\mathrm{d}B^u(s)\middle| Z(t)\right]\right] \\ &= \mathbb{E}^{\circ}\left[V^{u,*,+}_{Z(t)}(t,r^*)\right] \ , \end{aligned}$$

where the state-wise technical benefit reserves $V^{u,*,+}_j$, $j\in\mathcal{J}$, are given in Equation (6.1), and $\mathbb{E}^{\circ}$ denotes market or technical expectation. Hence, $k^{(\varepsilon(t))}$ is the upscaling factor that satisfies the equivalence principle on portfolio level.

Under the technical basis, the tower property applies, and the reserve simplifies to

$$V^{u,*,*,+}(t) = \sum_{j \in \mathcal{J}} p^{*}_{0j}(0,t)\, V^{u,*,+}_{j}(t, r^{*}) = \int_{t}^{T} e^{-r^{*}(s-t)}\, \mathrm{d}\beta^{u,*}(s) \ .$$

Under the market basis, it reads

$$V^{u,*,m,+}(t) = \sum_{j \in \mathcal{J}} p^{m}_{0j}(0,t)\, V^{u,*,+}_{j}(t, r^{*}) \ .$$

Again, we see that, using the market basis, both transition probabilities and state-wise technical reserves are needed.

We emphasize that there is no reason to consider the case $r^{*(\varepsilon(t-))} < r^{*}$. For the explanation, recall that consolidation serves to lower the technical interest rate, so if the technical interest rate is already low, there is no need for consolidation. If the initial technical interest rate $r^{*(\varepsilon(0))}$ is high compared to the pre-described level r^{*}, the allocated bonus is used for consolidation until $r^{*(\varepsilon(t))} = r^{*}$ for some t. Thereafter, the bonus is used for additional benefits, and the technical interest rate is kept fixed. If the initial technical interest rate $r^{*(\varepsilon(0))}$ is equal to r^{*}, the consolidation phase is skipped, the allocated bonus is used for additional benefits and the technical interest rate is kept fixed from the beginning. In neither case, we arrive at $r^{*(\varepsilon(t-))} < r^{*}$. In the third and last case where the initial technical interest rate $r^{*(\varepsilon(0))}$ is low compared to r^{*}, there is clearly no need for consolidation. Now, one has two options. Either, one can lower r^{*} to $r^{*(\varepsilon(0))}$ and proceed as in the case where $r^{*(\varepsilon(0))}$ is equal to r^{*}; or, one can raise $r^{*(\varepsilon(0))}$ to r^{*}, use the decline in the technical reserve for additional benefits and then proceed as in the case where $r^{*(\varepsilon(0))}$ is equal to r^{*}. Both solutions will avert the case $r^{*(\varepsilon(t-))} < r^{*}$. We note that the case $r^{*(\varepsilon(0))} < r^{*}$ represents a situation with increasing technical interest rate. This has not been observed in Denmark in recent years, which is why we exclude the case from our paper. However, as argued above, our model can easily handle the case.

Example 6.4 (Survival model continued)**.** For the simple policy in Example 6.1–6.3, the expected technical reserve $V^{*,\circ}(\cdot, \rho)$ reads

$$V^{*,\circ}(t,\rho) = e^{-\int_0^t \mu^{\circ}(v)\,\mathrm{d}v} \left(\int_t^T e^{-\rho(s-t)} e^{-\int_t^s \mu^{*}(v)\,\mathrm{d}v} \left(b^{ad}\mu^{*}(s) - \pi \right) \mathrm{d}s + b^{a} e^{-\rho(T-t)} e^{-\int_t^T \mu^{*}(v)\,\mathrm{d}v} \right) \ .$$

The expected technical benefit reserve $V^{u,*,\circ,+}(\cdot)$ reads

$$V^{u,*,\circ,+}(t) = b^{a} e^{-\int_0^t \mu^{\circ}(v)\,\mathrm{d}v} e^{-r^{*}(T-t)} e^{-\int_t^T \mu^{*}(v)\,\mathrm{d}v} \ .$$

Remark 6.1. In Section 6.4.2, we mentioned that our setup does not allow for policyholder behavior options, such as surrender or free policy. However, it is not particularly complicated to include surrender, since it is an absorbing state. For the sake of clarity, we will not go into details on how. We just mention that, under the bonus scheme "additional benefits", the surrender cash flow needs to be split into an upscaled and non-upscaled part. Furthermore, under the bonus scheme "consolidation", the bonus suddenly raises the guarantee through a higher surrender value (typically equal to the technical reserve), and the market cash flows need to be recalculated every time the policy is consolidated to account for the higher surrender value.

6.4.9 Bonus Allocation and Guarantee Fee

In Section 6.4.7, we took the bonus d and the guarantee fee π_g as exogenously given. This is imprecise for at least three reasons. Firstly, the total bonus allocated to the policies in a (homogeneous) portfolio typically depends on the collective bonus potential of the portfolio. The collective bonus potential K is defined as the maximum of zero and assets less guaranteed liabilities, *i.e.*,

$$K(t) = (X(t) - L(t))^{+} \ ,$$

where the assets X and the guaranteed liabilities L are calculated on portfolio level. With this definition, the balance sheet can be represented as in Figure 6.1. The collective bonus potential is a result of the systematic surplus to which the conservative technical basis gives rise. The systematic surplus of the policy is to be paid back to the policyholder in terms of bonus, but the collective bonus potential serves as a buffer in years with poor financial returns and/or poor risk results, so most often, the systematic surplus is not paid out right away. Therefore, to avoid redistribution across policies via the collective bonus potential, the portfolio must be homogeneous with respect to interest rate and risk (and costs, but in this paper, we leave that out). Furthermore, in order to avoid redistribution across generations, the systematic surplus should be paid out as soon as possible.

Secondly, the policy's share of the portfolio bonus depends on how much the policy has contributed to the portfolio's systematic surplus. As mentioned, the adjustment term α in the projection SDEs in Equation (6.4) is the market expected surplus arising from the conservative technical transition intensities. We choose to pay out the adjustment term immediately as risk bonus, such that the collective bonus potential collects surplus from capital gains only. The surplus collected in the collective bonus potential is then paid out, but not immediately, as interest rate bonus, *i.e.*, proportional to the market expected technical reserve Y. If the technical transition intensities are not chosen carefully enough (which can be difficult for varying products), the adjustment term can be negative for some ages. In that case, no risk bonus is paid out.

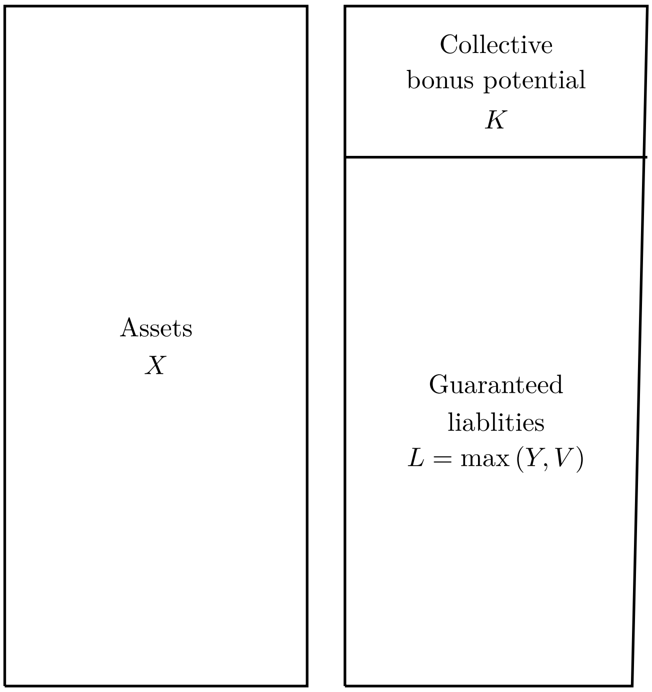

Figure 6.1: Portfolio balance sheet.

Thirdly, for the contract to be fair, the bonus d and the guarantee fee π_g must be chosen in such a way that the equivalence principle is satisfied for the total payments under the market basis, *i.e.*,

$$\mathbb{E}^Q \left[\int_0^T e^{-\int_0^s r(v)\,\mathrm{d}v} \,\mathrm{d} \left(B^{(\varepsilon(s))} - C \right)(s) \right] = x_0 \ . \tag{6.8}$$

In a multi-policy portfolio, the fairness constraint can be difficult to honor. It is possible to have fairness on portfolio level, but not on policy level, implying an unfair redistribution of systematic surplus across policies.

Often, the guarantee fee is a fraction of either the assets or the asset returns. The bonus allocation takes on more forms, but is ultimately a function of the collective bonus potential, the market reserve and the technical reserve. In Section 6.4.11, we present a numerical example with a one-policy and a two-policy portfolio. We show how to find a fair bonus and guarantee fee strategy, and we exemplify the challenges of fairness in a two-policy portfolio.

6.4.10 Application of Projections

We recall that the state process Z is independent of the financial market, and that, under both P and Q, the evolution of Z is described by the market basis. Most importantly, the projections of X and Y can be used to calculate the total time 0 market cash flow CF and market value W for the guaranteed and non-guaranteed payments, *i.e.*, to calculate

$$\begin{aligned} \mathrm{d}CF(t) &= \mathbb{E}^Q\left[k^{(\varepsilon(t))}\,\mathrm{d}B^u(t) + \mathrm{d}B^f(t) - \mathrm{d}C(t)\right] \\ &= \mathbb{E}^Q\left[k^{(\varepsilon(t))}\right]\mathrm{d}\beta^u(t) + \mathrm{d}\beta^f(t) - \mathrm{d}\varsigma(t) \end{aligned} \tag{6.9}$$

and

$$\begin{aligned} W(0) &= \mathbb{E}^Q\left[\int_0^T e^{-\int_0^s r(v)\,\mathrm{d}v}\,\mathrm{d}\left(B^{(\varepsilon(s))} - C\right)(s)\right] \\ &= \mathbb{E}^Q\left[\int_0^T e^{-\int_0^s r(v)\,\mathrm{d}v}\left(k^{(\varepsilon(s))}\mathrm{d}\beta^u(s) + \mathrm{d}\beta^f(s) - \mathrm{d}\varsigma(s)\right)\right] . \end{aligned} \tag{6.10}$$

Here, r is the stochastic short interest rate. We emphasize that the cash flow and market value distinguish themselves from the usual cash flows and market values by including non-guaranteed payments as well as guaranteed payments. In particular, we have

$$W(0) = V(0) + \mathbb{E}^Q\left[\int_0^T e^{-\int_0^s r(v)\,\mathrm{d}v}\left(k^{(\varepsilon(s))} - 1\right)\mathrm{d}\beta^u(s)\right] ,$$

where V is the usual market value from Equation (6.3). The additional term is the market value of the non-guaranteed benefits. If the projections are based on scenarios generated via Monte Carlo simulation, then for each t, the expectation $\mathbb{E}^Q\left[k^{(\varepsilon(t))}\right]$ in Equation (6.9) is approximated by averaging over a sufficient number of Q-projections up to time t. If, instead, the projections are of the worst-case or best-estimate type (and, hence, singular), then $\mathbb{E}^Q\left[k^{(\varepsilon(t))}\right]$ is approximated by the single projected value. If the short interest rate is deterministic, then Equation (6.10) simplifies to

$$W(0) = \int_0^T e^{-\int_0^s r(v)\,\mathrm{d}v}\,\mathrm{d}CF(s) . \tag{6.11}$$

Otherwise, Equation (6.10) is approximated by averaging over a sufficient number of integrated sample paths $t \mapsto k^{(\varepsilon(t))}\,\mathrm{d}\beta^u(t) + \mathrm{d}\beta^f(t) - \mathrm{d}\varsigma(t)$, discounted by the short interest rate. The market value is useful for determining the bonus allocation d and guarantee fee π_g according to the fairness criterion in Equation (6.8), which can be written as

$$W(0) = x_0 .$$

So far, we have suppressed the influence of the investment strategy, but it enters through the stochastic return on the assets. Hence, the task of determining d and π_g is the classical trade-off between the aggressiveness of dividend allocation (expressed by d) and the option price (expressed by π_g)

given the aggressiveness of the investment strategy (typically expressed by the volatility).

The projections of X and Y are also useful for calculating the time 0 P-expected cash flow for the guaranteed and non-guaranteed payments, *i.e.*, for calculating

$$\begin{aligned}\mathrm{d}CF^P(t) &= \mathbb{E}^P\left[k^{(\varepsilon(t))}\,\mathrm{d}B^u(t) + \mathrm{d}B^f(t) - \mathrm{d}C(t)\right]\\ &= \mathbb{E}^P\left[k^{(\varepsilon(t))}\right]\mathrm{d}\beta^u(t) + \mathrm{d}\beta^f(t) - \mathrm{d}\varsigma(t)\ .\end{aligned}$$

If the projections are based on scenarios generated via Monte Carlo simulation, then for each t, the expectation $\mathbb{E}^P\left[k^{(\varepsilon(t))}\right]$ is approximated by averaging over a sufficient number of P-projections up to time t. If instead, the projections are of the worst-case or best-estimate type (and, hence, singular), then $\mathbb{E}^P\left[k^{(\varepsilon(t))}\right]$ is approximated by the single projected value. The P-expected cash flow is an estimate of the money out flow from the insurance company at future time points, and it is, therefore, useful for liquidity considerations. Again, the cash flow distinguishes itself by including non-guaranteed payments as well as guaranteed payments, thereby, providing a more complete picture.

Finally, for solvency purposes, one can use scenarios generated via Monte Carlo simulation to calculate P-quantiles for the capital requirement at time T_1

$$\begin{aligned}&\mathbb{E}^Q\left[\int_0^{T\vee T_1} e^{-\int_{T_1}^s r(v)\,\mathrm{d}v}\,\mathrm{d}\left(B^{(\varepsilon(s))} - C\right)(s)\middle|\left(k^{(\varepsilon(s))}, r(s)\right)_{s\le T_1}\right] =\\ &\int_0^{T_1} e^{-\int_{T_1}^s r(v)\,\mathrm{d}v}\left(k^{(\varepsilon(s))}\mathrm{d}\beta^u(s) + \mathrm{d}\beta^f(s) - \mathrm{d}\varsigma(s)\right) +\\ &\mathbb{E}^Q\left[\int_{T_1}^{T\vee T_1} e^{-\int_{T_1}^s r(v)\,\mathrm{d}v}\left(k^{(\varepsilon(s))}\mathrm{d}\beta^u(s) + \mathrm{d}\beta^f(s) - \mathrm{d}\varsigma(s)\right)\middle| k^{(\varepsilon(T_1))}, r(T_1)\right].\end{aligned}$$

The capital requirement is expressed in terms of the capital needed up to time T_1 plus the market value of future liabilities at time T_1. The conditional Q-expectation appears in the capital requirement, because the capital requirement concerns future payments and balance sheets. The quantiles are obtained by projecting up to time T_1 under the physical measure P. However, for each projection, the Q-expectation is approximated by projecting from time T_1 to time T under the pricing measure Q. Hence, if N sample paths are needed for approximating cash flows and market values, then N^2 paths are needed for the quantiles. The quantiles can be used for solvency assessments of the provided guarantee.

6.4.11 Numerical Examples

In this section, we go through two numerical examples with a one-policy portfolio and a two-policy portfolio. A larger portfolio would, of course, be more realistic, but a large number of policies could easily drown the key insights from the examples. Going from one to two policies is by far the biggest step, and conceptually, there is no impediment to extending the theory to larger portfolios. Working in a discrete projection setup, we show how to find a fair bonus and guarantee fee strategy for the one-policy portfolio, and we exemplify the fairness challenges in a two-policy portfolio. The examples are based on 5000 scenarios generated via Monte Carlo simulation. We have made sure that the number of simulated scenarios is sufficiently high for our numerical results and graphs not to change between simulations, but we do not go into details about the robustness of the simulations, since the examples only serve to demonstrate the possible applications of our model.

6.4.11.1 One-Policy Portfolio

We consider a portfolio consisting of a single policy. The policy is the one from Examples 6.1–6.4. The policyholder is a female aged 25 at time 0, where the policy is issued. We fix $r^* = 0.02$, and we assume that $r^{*(0)} = r^*$, which is natural for a newly-issued policy. Thereby, we only consider the bonus scheme "additional benefits". We recall that bonus is used to increase the endowment sum and not the term insurance sum. The death of the policyholder is governed by the technical mortality intensity

$$\mu^* (t) = 5 \cdot 10^{-4} + 5.3456 \cdot 10^{-5} \cdot e^{0.087498(25+t)} \ .$$

For the last three decades, this has served as a standard mortality intensity for adult women in Denmark. It is part of the so-called G82 technical basis that was set forth as a Danish industry standard in 1982. The market mortality intensity is given by

$$\mu^m (t) = 0.8\mu^* (t) \ .$$

With this choice of mortality intensities and with the product choices below, the technical basis is on the safe side, except for low ages, where the death sum exceeds the savings, resulting in a negative contribution from mortality risk. However, due to the low mortality for low ages, the negative contribution is insignificantly small.

The policy expires at time $T = 40$ when the policyholder is 65. We fix the term insurance sum at $b^{ad} = 1$ and the pure endowment sum at $b^a = 3$. The equivalence premium is determined via the equivalence relation

$$V^{*,-} (0, r^*) = V^{f,*,+} (0, r^*) + V^{u,*,+} (0, r^*) \ ,$$

i.e.,

$$\pi = \frac{b^a e^{-\int_0^T (r^* + \mu^*(v))\,\mathrm{d}v} + b^{ad} \int_0^T e^{-\int_0^s (r^* + \mu^*(v))\,\mathrm{d}v} \mu^* (s)\,\mathrm{d}s}{\int_0^T e^{-\int_0^s (r^* + \mu^*(v))\,\mathrm{d}v}\,\mathrm{d}s} \ .$$

Using numerical methods, we obtain $\pi = 0.04614$. The bonus d is allocated and the guarantee fee π_g is paid once a year. Hence, we have

$$\varepsilon(t) = \#\{i = 1, \ldots, 40 : i \leq t\} \ .$$

We note that $\varepsilon(t) = t$ for $t = 1, \ldots, 40$. We project the two accounts X and Y using steps of a size of one year by applying a discretized version of the stochastic differential equations for X and Y. For the discretization, we recall from Example 6.2 that β^u is a pure jump function and that ς and β^f are continuous functions. Hence, we get the stochastic difference equations

$$\begin{aligned}
X(t-) &= X(t-1)(1 + R_X(t)) - \int_{(t-1,t)} \left(\mathrm{d}\beta^f(s) - \mathrm{d}\varsigma(s)\right) \ , \\
X(t) &= X(t-) - k^{(t)}\Delta\beta^u(t) + g(t) - \pi_g(t) \ , \qquad t = 1, \ldots, 40 \ , \\
X(0) &= 0 \ , \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (6.12) \\
Y(t-) &= Y(t-1)e^{r^*} - \int_{(t-1,t)} \left(\mathrm{d}\beta^f(s) - \mathrm{d}\varsigma(s)\right) + \alpha\left(t, r^*, k^{(t-1)}\right) \ , \\
Y(t) &= Y(t-) - k^{(t)}\Delta\beta^u(t) + d(t) \ , \qquad t = 1, \ldots, 40 \ , \\
Y(0) &= 0 \ .
\end{aligned}$$

We assume a deterministic market interest rate $r = 0.04$, and the assets of the portfolio (in this case, the assets of the policy) are invested in a fund with log-normal returns that are paid out once a year, *i.e.*,

$$R_X(t) = \frac{S(t) - S(t-1)}{S(t-1)} \ , \quad t = 1, \ldots, 40 \ ,$$

where S is a geometric Brownian motion. We basically consider a simple Black–Scholes financial market. We assume that the fund size S has drift 0.07 and volatility 0.2 under the physical measure P (and, consequently, drift $r = 0.04$ and volatility 0.2 under the pricing measure Q).

The bonus d is determined as a fraction θ_1 of the excess collective bonus potential K just before the bonus allocation over a threshold $\bar{K}$ if this fraction exceeds the positive part of the natural risk bonus α (see Section 6.4.9 for more on risk bonus), *i.e.*,

$$d(t) = \max\left\{\left(\alpha\left(t, r^*, k^{(\varepsilon(t-))}\right)\right)^+, \theta_1\left(K(t-) - \bar{K}(t-)\right)^+\right\} \ , \qquad t = 1, \ldots, 40 \ ,$$

where K and $\bar{K}$ are given by

$$\begin{aligned}
K(t) &= (X(t) - L(t))^+ \ , \\
\bar{K}(t) &= \theta_2 L(t) \ ,
\end{aligned}$$

with the guaranteed liabilities L defined in Equation (6.5). The threshold $\bar{K}$ can be seen as a preferred minimum collective bonus potential. We fix $\theta_1 = 0.2$ and $\theta_2 = 0.1$. As mentioned, the chosen technical transition intensity is not on the safe side for low ages. Therefore, we need to take the positive part of α in the expression above to exclude negative risk bonus. Finally, we choose the guarantee fee π_g to be a fraction θ_3 of the positive part of the returns on the assets, *i.e.*,

$$\pi_g(t) = \theta_3 \left(R_X(t) X(t-1)\right)^+ \ .$$

In addition to the yearly guarantee fee, the equity holders of the insurance company receives the remaining collective bonus potential at expiration as part of the final guarantee fee. We determine the fraction θ_3 according to the fairness criterion in Equation (6.8). Furthermore, using this guarantee fee, we consider:

- the expected evolution of the upscaling factor $t \mapsto \mathbb{E}^Q\left[k^{(\varepsilon(t))}\right]$,
- the expected evolution of the assets $t \mapsto \mathbb{E}^Q[X(t)]$, market reserve $t \mapsto \mathbb{E}^Q[V(t)]$, technical reserve $t \mapsto \mathbb{E}^Q[Y(t)]$ and collective bonus potential $t \mapsto \mathbb{E}^Q[K(t)]$,
- the expected level for the guarantee injections $\mathbb{E}^Q[g(t)]$ and guarantee fees $\mathbb{E}^Q[\pi_g(t)]$, $t = 1, \ldots, T$.

Using standard Monte Carlo methods, we simulate 5000 sample paths for the asset returns R_X under the measure Q, and for each sample path, we project X and Y for different values of θ_3, using the difference equations in Equation (6.12). More specifically, we look for a θ_3, such that we get zero when approximating the time 0 market value $W(0)$ from Equation (6.11). We recall that W is the market value of the guaranteed and non-guaranteed payments. We arrive at the fair guarantee fee fraction $\theta_3 = 0.31$.

In Figures 6.2–6.4, we plot the average evolution of the upscaling factor, the average level for the guarantee injection and guarantee fee, and the average evolution of the assets, reserves and collective bonus potential. From Figure 6.2, we see that more than 60% of the final endowment sum comes from bonus. This is primarily due to the technical interest rate being only half the size of the market interest rate. From Figure 6.3, we observe that the guarantee fees and the guarantee injections follow each other closely, implying that the structure of the guarantee fee is reasonable. The final guarantee fee includes the remaining collective bonus potential at expiration and is, consequently, much higher than the other guarantee fees. We note that a considerable part of the guarantee injections are paid for by giving up the remaining collective bonus potential at expiration. In a multi-generation portfolio, this is of course transferred to the other policies, leading to a higher guarantee fee throughout the period. From Figure 6.4, we see how the different parts of the balance in Figure 6.1 evolve in expectation.

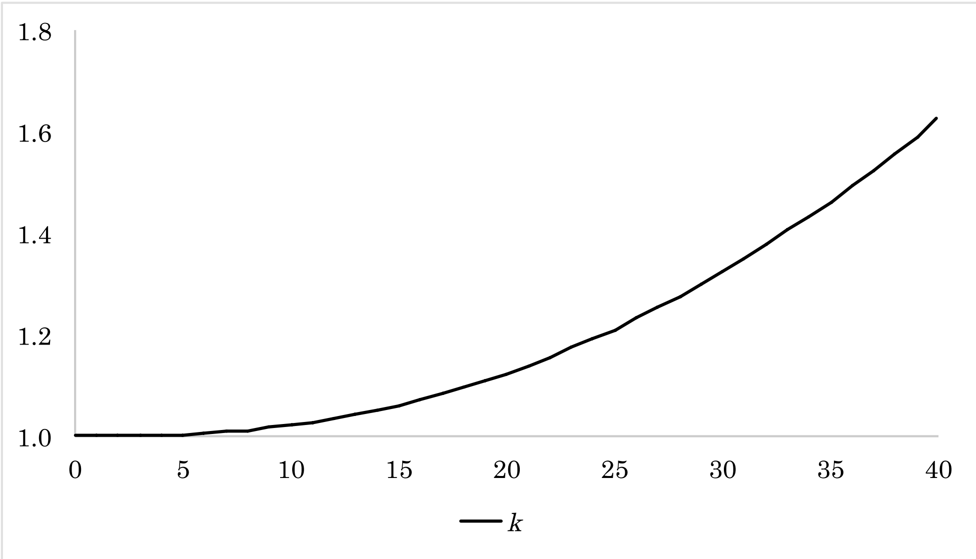

Figure 6.2: Approximated expected upscaling factor k as a function of time.

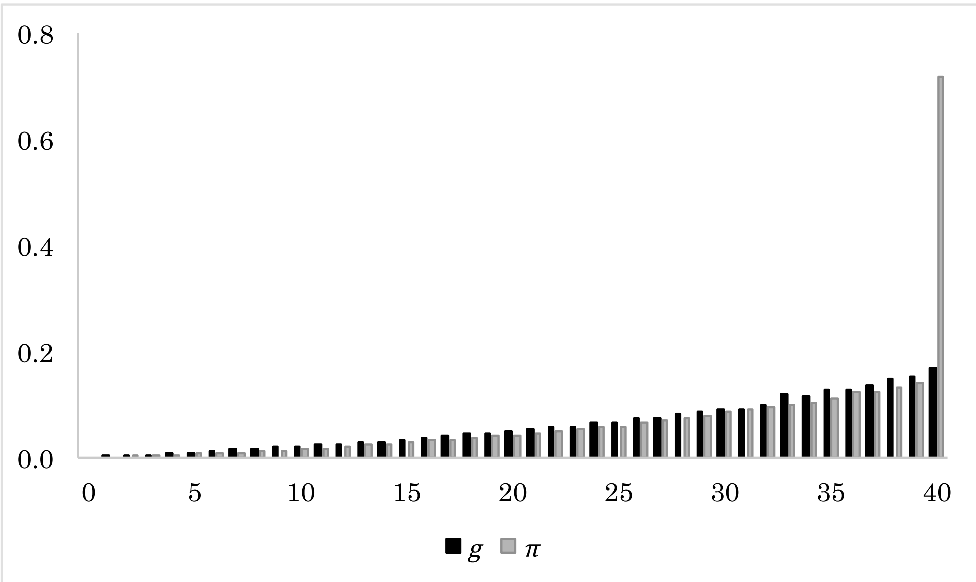

Figure 6.3: Approximated expected guarantee injection g and guarantee fee π_g as a function of time.

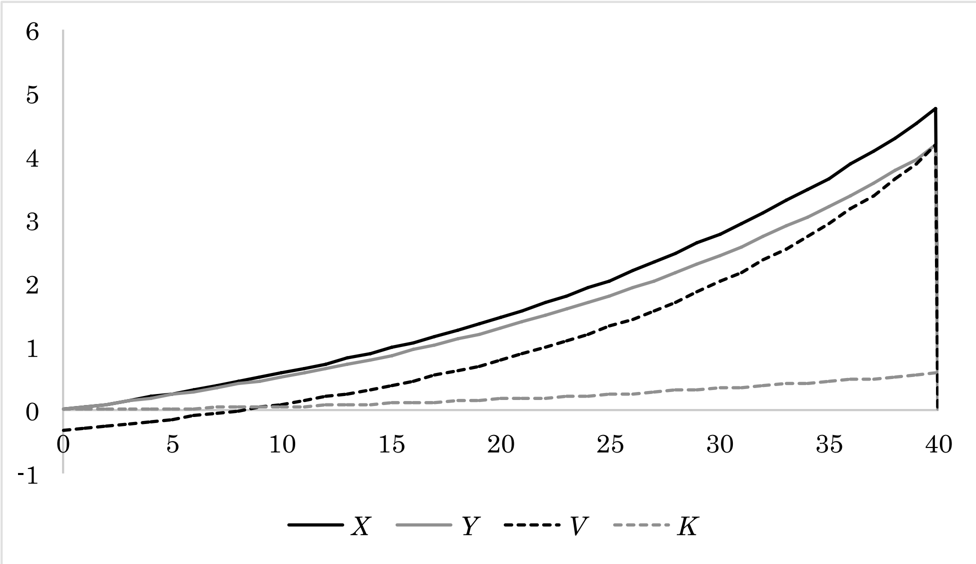

Figure 6.4: Approximated expected assets X, technical reserve Y, market reserve V and collective bonus potential K as a function of time.

6.4.11.2 Two-Policy Portfolio

We consider a portfolio consisting of two policies, Policy 1 and 2, which are identical to the one in Section 6.4.11.1. However, only Policy 1 is issued at time 0; Policy 2 is not issued until time 20. Hence, from time 0 to time 20, there is one policy in the portfolio; from time 20 to time 40, there are two policies in the portfolio; and from time 40 to time 60, there is, again, one policy in the portfolio. If not careful, this overlap in time easily causes an unfair redistribution between the two policies.

We let α_i, $k_i^{(\cdot)}$ and Y_i denote, respectively, the adjustment term, the upscaling factor and the market expected technical reserve for policy $i = 1, 2$. We work with the convention that all quantities (except the upscaling factor) are zero for Policy 2 until time 20 and zero for Policy 1 after time 40. By X, V and Y, we denote the total assets, the total market expected market reserve and the total market expected technical reserve of the portfolio.

We make the same market assumptions as in Section 6.4.11.1, and because of the longer time period, we now have

$$\varepsilon(t) = \#\{i = 1, \ldots, 60 : i \leq t\} \ .$$

The guarantee injection g is calculated on portfolio level and reads

$$g(t) = (L(t-) + \pi_g(t) - X(t-))^+ \,,$$

where the guaranteed liabilities L are given by

$$L(t) = \max\{V(t), Y(t)\} \ .$$

The total bonus to the policies in the portfolio is determined as a fraction θ_1 of the excess collective bonus potential K just before the bonus allocation over a threshold $\bar{K}$ if this fraction exceeds the positive part of the total natural risk bonus α for the policies of the portfolio, *i.e.*,

$$d\left(t\right) = \max\left\{\alpha\left(t\right), \theta_1 \left(K\left(t-\right) - \bar{K}\left(t-\right)\right)^+\right\} \ , \quad t \in \{1, 2, \ldots, 60\} \ ,$$

where α, K and $\bar{K}$ are given by

$$\begin{aligned} \alpha\left(t\right) &= \left(\alpha_1\left(t, r^*, k_1^{(\varepsilon(t-))}\right)\right)^+ + \left(\alpha_2\left(t, r^*, k_2^{(\varepsilon(t-))}\right)\right)^+ , \\ K\left(t\right) &= \left(X\left(t\right) - L\left(t\right)\right)^+ , \\ \bar{K}\left(t\right) &= \theta_2 L\left(t\right) \ . \end{aligned}$$

The threshold $\bar{K}$ can again be seen as a preferred minimum collective bonus potential for the portfolio. We keep $\theta_1 = 0.2$ and $\theta_2 = 0.1$. The bonus is divided between the policies of the portfolio in the following way: First, each policy receives its natural risk bonus given by the adjustment terms α_1 and α_2. We, thereby, use the collective bonus potential as a financial buffer only. Second, the remaining bonus (if any) is distributed as a technical interest rate margin, *i.e.*, proportional to the market expected technical reserves Y_1 and Y_2. In formulas, the bonus to policy $i = 1, 2$ is given by

$$d_i\left(t\right) = \left(\alpha_i\left(t, r^*, k_i^{(\varepsilon(t))}\right)\right)^+ + \left(d\left(t\right) - \alpha\left(t\right)\right)\frac{Y_i\left(t\right)}{Y\left(t\right)} \ , \quad t \in \{1, 2, \ldots, 60\} \ .$$

6.4.11.3 Constant Guarantee Fee Fraction

First, we stick to a guarantee fee π_g that is a constant fraction θ_3 of the positive part of the returns on the assets, *i.e.*,

$$\pi_g\left(t\right) = \theta_3\left(R_X\left(t\right) X\left(t-1\right)\right)^+ \ .$$

In addition to the yearly guarantee fee, the equity holders of the insurance company receive the remaining collective bonus potential at the expiration of Policy 2 as part of the final guarantee fee. We determine the fraction θ_3 according to the fairness criterion in Equation (6.8) applied on portfolio level. We take the 5000 sample paths simulated in Section 6.4.11.1, and for each sample path, we project X, Y_1 and Y_2 for different values of θ_3, using three difference equations almost identical to the ones in Equation (6.12). More specifically, we look for a θ_3, such that we get zero when approximating the time 0 market value from Equation (6.11) on portfolio level. We arrive at the guarantee fee fraction $\theta_3 = 0.35$. We notice that the fraction is higher than in the one-policy case, possibly to cover increased risk associated with an extra policy and a longer time horizon. Furthermore, we calculate the

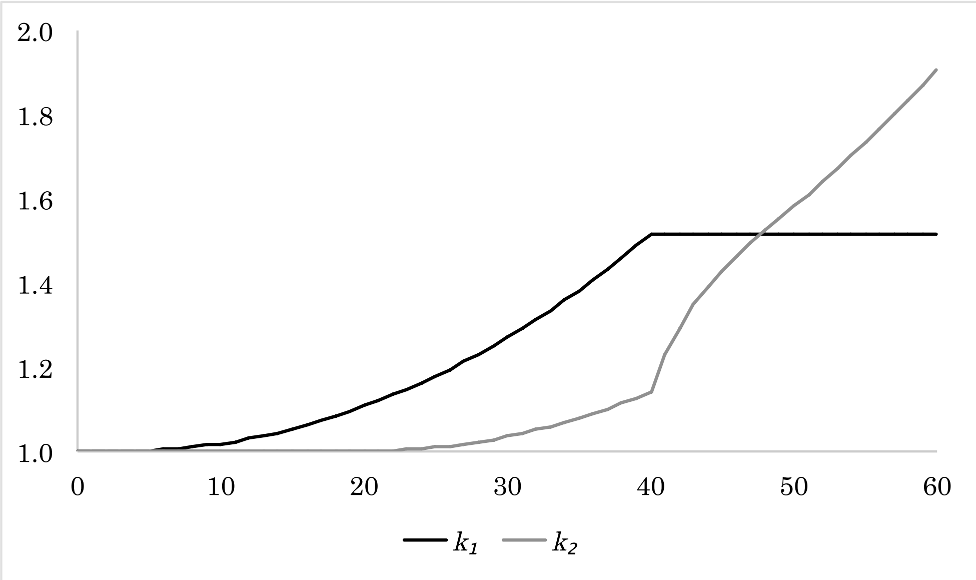

Figure 6.5: Approximated expected upscaling factors k_1 and k_2 as function of time.

average evolution of the upscaling factor, the average guarantee injection and guarantee fee levels, and the average evolution of the assets, reserves and collective bonus potential. We discover that the guarantee fee is only fair on portfolio level. Approximating the time 0 market value from Equation (6.11) for each of the policies individually, we get

$$W^1(0) = -0.061 \quad , \quad W^2(0) = 0.068 \ .$$

Hence, a significant amount of the systematic surplus is being redistributed from Policy 1 to Policy 2. We recall that W^1 and W^2 are the market values of the guaranteed and non-guaranteed payments; and not the usual market values that only include guaranteed payments. To illustrate the redistribution, we plot the average evolution of the upscaling factors in Figure 6.5. From the figure, it appears that Policy 2's final upscaling factor is much larger than Policy 1's. Furthermore, comparing with Figure 6.2, we see that Policy 1's final upscaling factor is significantly smaller than in the one-policy case, so it is not just a matter of both policies benefiting from being part of the two-policy portfolio and Policy 2 benefiting more from it than Policy 1.

6.4.11.4 Period-Dependent Guarantee Fee Fraction

To overcome the unfairness introduced by the constant guarantee fee fraction, we allow there to be a different fraction θ_3 determining the guarantee fee for each of the time periods $[0, 20]$, $(20, 40]$ and $(40, 60]$, *i.e.*,

$$\pi_g(t) = \theta_3^{i(t)} \left(R_X(t)\, X(t-1)\right)^+ \ ,$$

where

$$i(t) = \begin{cases} 1 & t \leq 20 \ , \\ 2 & 20 < t \leq 40 \ , \\ 3 & t > 40 \ . \end{cases}$$

In addition to the yearly guarantee fee, the equity holders of the insurance company still receive the remaining collective bonus potential at the expiration of Policy 2 as part of the final guarantee fee. First, we fix $\theta_3^1 = 0.31$, since this is the fair guarantee fee fraction from the one-policy portfolio. Second, we take the 5000 sample paths simulated in Section 6.4.11.1 and search (in the same way as before) for a value of θ_3^2 for which $W^1(0) = 0$, meaning that the guarantee fee determined by the pair $\left(\theta_3^1, \theta_3^2\right)$ is fair for Policy 1. We find the fair guarantee fee fraction $\theta_3^2 = 0.29$. Third, we search for a value of θ_3^3 for which $W^2(0) = 0$, meaning that the guarantee fee determined by the triplet $\left(\theta_3^1, \theta_3^2, \theta_3^3\right)$ is fair for Policy 2. We find the fair guarantee fee fraction $\theta_3^3 = 0.61$. We notice that the guarantee fee fraction is much higher in the last time period than in the first two time periods and that the guarantee fee fraction in the second time period is slightly smaller than in the first time period. This is best explained by the fact that Policy 2 inherits collective bonus potential from Policy 1. Policy 1 is compensated for this transfer via the lower guarantee fee fraction in the second time period, and Policy 2 pays for the transfer in terms of the high guarantee fee fraction in the last time period. This is reflected in Figure 6.6. Again, the final guarantee fee is much higher than the other guarantee fees, since it includes the remaining collective bonus potential at the expiration of Policy 2. In a multi-generation portfolio, this is transferred to the other policies, leading to a higher guarantee fee throughout the period.

In Figures 6.6–6.8, we plot the average evolution of the upscaling factors, the average guarantee injection and guarantee fee, and the average evolution of the assets, reserves and collective bonus potential. With the fair guarantee fee, we see from Figure 6.7 that the two policies' final upscaling factors are essentially equal; also to the final upscaling factor from Section 6.4.11.1. From Figure 6.8, we see how the different parts of the balance in Figure 6.1 evolve in expectation for the two-policy portfolio.

6.5 Unit-Linked Insurance

In this section, we consider unit-linked insurance. We touch upon different aspects of unit-linked insurance, and we present our two-account model for a general unit-linked insurance policy. Again, we include simple and illustrative survival model examples. We end the section with a numerical example that is a unit-linked version of the numerical example in the previous section.

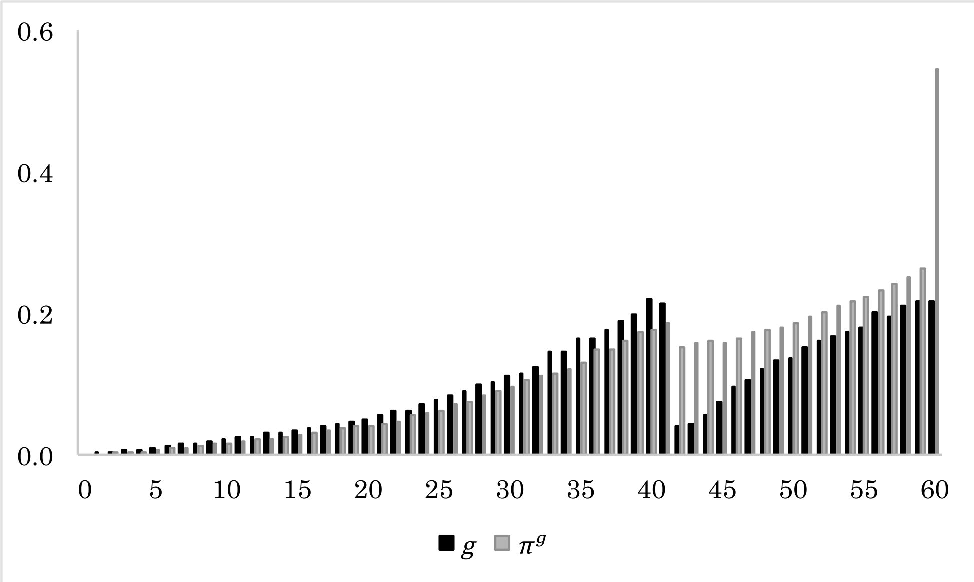

Figure 6.6: Approximated expected guarantee injection g and guarantee fee π_g as a function of time.

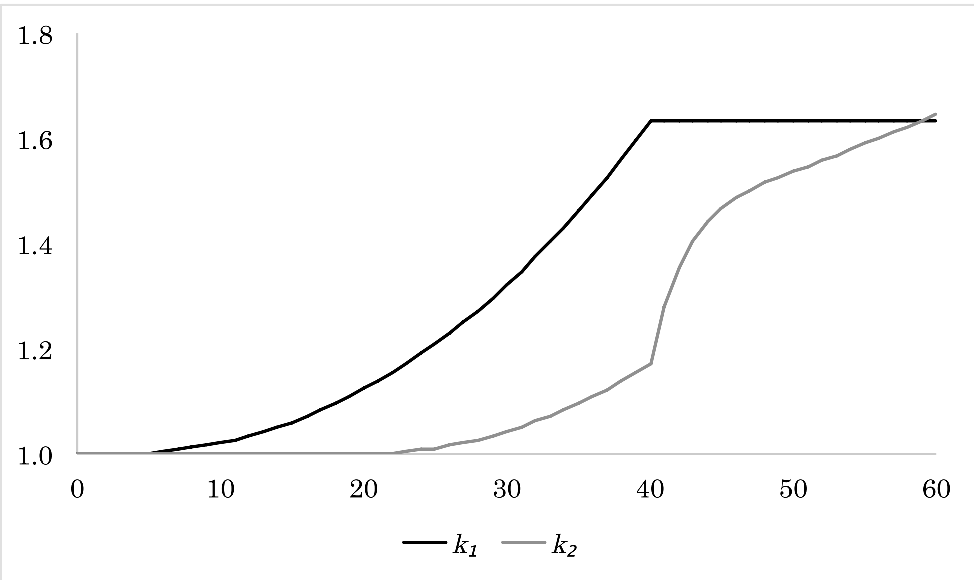

Figure 6.7: Approximated expected upscaling factors k_1 and k_2 as a function of time.

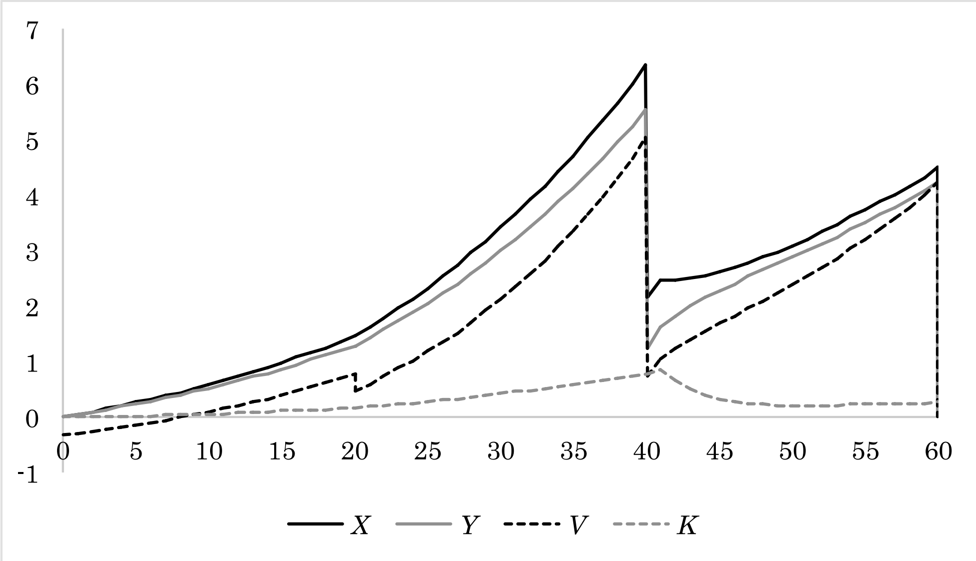

Figure 6.8: Approximated expected assets X, technical reserve Y, market reserve V and collective bonus potential K as a function of time.

6.5.1 Product Specification and Two-Account Model

We consider a unit-linked insurance policy. The state-wise evolution of the policy is described in Section 6.3. The payments of the policy consist of a state-dependent payment stream $B^f + B^u - C$ where C is a fixed state-dependent premium stream ("C" for contribution), B^f is a fixed state-dependent benefit stream ("B" for benefits and superscript "f" for fixed) and B^u is a state-dependent benefit stream that is linked to the financial market (superscript "u" for unit-linked). More precisely, the assets of the policy are invested in a fund, and the benefit stream B^u depends on the value of the assets. The policy includes a guaranteed minimum retirement savings amount at the retirement date R, based on a guarantee account with a guaranteed interest rate r^* (for example, $r^* = 0$). We do not take costs into account.

We denote by X the assets of the policy and by Y the guarantee account. Again, the accounts X and Y are the backbone of our two-account model. The policy is issued before or at time 0, and the two accounts amount to $Y(0-) = x_0$ and $Y(0-) = y_0$ just before time 0. The policy terminates at time $T \geq R$. Thereafter, there are no payments.

The assets X are invested in a fund with stochastic return R_X. The guarantee account Y accumulates according to the guaranteed interest rate. In good times, the return rate on the assets exceeds the technical interest rate, and then, the assets outgrow the guarantee account. In that case, the guarantee account is upgraded (increased) according to the terms of the contract. For example, it may be stipulated in the contract that the guarantee account is always to make up at least 80% of the assets. Regardless of the developments in the financial market, the guarantee account is never to be downgraded (lowered), and at retirement, the maximum value of the assets and the guarantee account is paid out to the policyholder. In bad times where the guarantee account exceeds the assets at retirement, the equity holders of the insurance company step in with a capital injection taken from the company's equity. We speak of the possible capital injection

$$g = (Y(R-) - X(R-))^+ \tag{6.13}$$

as guarantee injection, and its role is to raise the assets at retirement in case of unfavorable developments in the financial market. The policyholder pays for the company's risk taking by having a guarantee fee deducted from the assets and paid to the equity holders of the insurance company. We assume that the insurance company's equity is always sufficient to cover the guarantee injection and that the guarantee injection and guarantee fees are settled via the equity. All of the above does not happen continuously, but at pre-specified deterministic time points $0 < t_1 < \ldots < t_n = R$ (for example, once a year), where the two accounts X and Y are updated. We let

$$\varepsilon(t) = \#\{i = 1, \ldots, n : t_i \leq t\}$$

count the number of updates prior to time t. The updates consist of upgrades, u, of the guarantee account (if the assets exceed the guarantee account in a pre-described way) and the deduction of the guarantee fee, π_g, in return for the possible guarantee injection g at retirement. At retirement R, the assets are updated with the guarantee injection, g, if the guarantee account exceeds the assets. We let $\varepsilon_R(t) = 1_{\{t \geq R\}}$ mark this final update. We stress the similarity with participating life insurance, although, there, the guarantee injection is settled at each update and not only at retirement. Compare for example Equation (6.13) to Equation (6.6). Again, we assume that the stochastic return on the assets, R_X, does not jump at time points with an account update. Furthermore, to ensure predictability, we assume that $u(t)$ and $\pi_g(t)$ are known at time $t-$ for all t with $\mathrm{d}\varepsilon(t) = 1$, *i.e.*, for all time points with an account update.

We assume that the market-linked benefit stream B^u is linear in X, *i.e.*,

$$\mathrm{d}B^u(t) = X(t-)\,\mathrm{d}B^p(t) \ ,$$

where B^p denotes a fixed state-dependent benefit stream (superscript "p" for profile). We write $X(t-)$ instead of just $X(t)$ to ensure that the asset process X is well-defined (see definition below). To ensure that, in expectation, the assets are paid out to the policyholder, we assume that $\Delta B^f(T) = \Delta C(T) = 0$ and $\mathbb{E}[\Delta B^p(T)] = 1$. Here, Δ denotes the jump part of the processes.

The fixed benefit stream B^f includes, for example, disability or death payments, whereas the market-linked benefit stream B^u includes, for example, deposit protection, surrender payments, a variable pure endowment or a variable life annuity. We note that the setup does not cover benefits that are non-linear in X. Thus, it does not cover a deposit protection of the form

$$\max\{X, D\},$$

where D is a fixed death sum. Furthermore, we exclude the policyholder behavior option "free policy", as it introduces a duration-dependent free policy conversion factor. Formally, the payment streams of the policy are given as

$$\begin{aligned}
\mathrm{d}C &= \sum_{j\in\mathcal{J}} I_j\,\mathrm{d}c_j\ ,\\
\mathrm{d}B^i &= \sum_{j\in\mathcal{J}} I_j\,\mathrm{d}b^i_j + \sum_{j,k\in\mathcal{J}:j\neq k} b^i_{jk}\,\mathrm{d}N_{jk}\ ,\quad i=f,p\ ,
\end{aligned}$$

where c_j, b^f_j and b^p_j are deterministic, state-wise payment streams and b^f_{jk} and b^p_{jk} are deterministic lump sum payments upon jumps. Examples of deterministic lump sum payments upon jumps include surrender payments and insurance coverage, such as a death sum, a disability sum or a sum upon critical illness.

Example 6.5 (Survival model)**.** We consider a unit-linked version of the simple participating life insurance policy in Examples 6.1–6.4. This example provides the basis for numerical illustrations later on. The state of the policy is described by the classical survival model with two states, 0 (alive) and 1 (dead). The policy expires at the retirement date, *i.e.*, $T=R$. The payments of the policy consist of a constant continuous premium payment π while alive, a term insurance sum b^{ad} upon death before expiration T and a pure endowment sum upon survival until expiration T. The size of the endowment sum is equal to the value of the assets at expiration divided by the (market) probability of surviving to expiration. There are no payments in the death state. For simplicity, we write $I=I_1$, $N=N_{01}$, $\mu^m=\mu^m_{01}$ and $p^m=p^m_{00}$, and we have

$$p^m(s,t) = e^{-\int_s^t \mu^m(v)\,\mathrm{d}v}\ ,\quad s\le t\ .$$

The payment streams of the policy read

$$\begin{aligned}
\mathrm{d}C(t) &= \pi I\,\mathrm{d}t\ ,\quad t\le T\ ,\\
\mathrm{d}B^f(t) &= b^{ad}\,\mathrm{d}N(t)\ ,\quad t\le T\ ,\\
\mathrm{d}B^p(t) &= \frac{I(t)}{p^m(0,T)}\,\mathrm{d}\varepsilon_T(t)\ ,\quad t\le T\ ,
\end{aligned}$$

where ε_T is the Dirac measure in T (for details, see Page 146). We divide by the probability $p^m(0,T)$ in $\mathrm{d}B^p(t)$ to ensure that the assets in expectation are paid out to the policyholder, *i.e.*, to account for inheritance from those who die before time T. We note that

$$\int_0^T \mathrm{d}B^p(t) = \frac{I(T)}{p^m(0,T)}\ .$$

6.5.2 Cash Flows

For projection on portfolio level, it is useful to consider market cash flows of the policy. Again, we use the term market cash flows for the expectation of the stochastic payment streams taken under the market basis. By ς, β^f and β^p, we denote the time 0 market cash flows for the premium stream C and the benefit streams B^f and B^p, *i.e.*,

$$\begin{aligned}
\varsigma(t) &= \mathbb{E}^m\left[\int_0^t \mathrm{d}C(s)\right] = \sum_{j\in\mathcal{J}} \int_0^t p_{0j}^m(0,s)\,\mathrm{d}c_j(s) \ , \\
\beta^i(t) &= \mathbb{E}^m\left[\int_0^t \mathrm{d}B^i(s)\right] \\
&= \sum_{j\in\mathcal{J}} \int_0^t p_{0j}^m(0,s)\left\{\mathrm{d}b_j^i(s) + \sum_{k\in\mathcal{J}:k\neq j} \mu_{jk}^m(s)\,b_{jk}^i(s)\,\mathrm{d}s\right\} \ , \quad i=f,p \ ,
\end{aligned}$$

where the expectation $\mathbb{E}^m$ is taken under the market basis and where p_{0j}^m is the market probability of transition from state 0 to j, which is determined by the transition intensities from the market basis.

Example 6.6 (Survival model continued)**.** For the simple policy in Example 6.5, the time 0 market premium and benefit cash flows read

$$\begin{aligned}
\varsigma(t) &= \int_0^t p^m(0,s)\,\pi\,\mathrm{d}s = \pi\int_0^t e^{-\int_0^s \mu^m(v)\,\mathrm{d}v}\,\mathrm{d}s \ , \ t\le T \ , \\
\beta^f(t) &= \int_0^t p^m(0,s)\,b^{ad}\mu^m(s)\,\mathrm{d}s = b^{ad}\int_0^t e^{-\int_0^s \mu^m(v)\,\mathrm{d}v}\mu^m(s)\,\mathrm{d}s \ , \ t\le T \ , \\
\beta^p(t) &= \int_0^t \frac{p^m(0,s)}{p^m(0,T)}\,\mathrm{d}\varepsilon_T(s) = I_{\{t\ge T\}} \ , \ t\le T \ .
\end{aligned}$$

6.5.3 Two-Account Projection

On portfolio level, the assets X and the guarantee account Y evolve according to the stochastic differential equations

$$\begin{aligned}
\mathrm{d}X(t) &= X(t-)\,\mathrm{d}R_X(t) - X(t-)\,\mathrm{d}\beta^p(t) - \mathrm{d}\beta^f(t) + \mathrm{d}\varsigma(t) \\
&\quad - \pi_g(t)\,\mathrm{d}\varepsilon(t) + g\,\mathrm{d}\varepsilon_R(t) \\
&\quad - 1_{\{t=R\}}\left(Y(R-) - X(R-)\right)^+ \mathrm{d}\beta^p(t) \ , \\
X(0-) &= x_0 \ , \\
\mathrm{d}Y(t) &= Y(t)\,r^*(t)\,\mathrm{d}t - X(t-)\,\mathrm{d}\beta^p(t) - \mathrm{d}\beta^f(t) + \mathrm{d}\varsigma(t) \\
&\quad + u(t)\,\mathrm{d}\varepsilon(t) \ , \ t\le R \ , \\
Y(0-) &= y_0 \ , \\
Y(t) &= 0 \ , \ t > R \ .
\end{aligned} \tag{6.14}$$

We recall that R_X is the stochastic return on the assets, g is the guarantee injection at retirement, π_g is the guarantee fee deducted from the assets and paid to the equity holders of the insurance company, u is the upgrade of the guarantee account, ε counts the number of guarantee fee payments and guarantee upgrades (typically annual) and ε_R marks the exercise of the guarantee at the retirement date. The last term in the equation for X ensures that the guarantee injection at time R is included in a possible lump sum payment at time R. The three quantities r^*, u and π_g are specified in the contract. They are non-negative, and they are determined in such a way that the contract is financially fair, *i.e.*, such that the equivalence principle is satisfied for the total payments under the market basis

$$x_0 = \mathbb{E}^Q \left[\int_0^T e^{-\int_0^s r(v)\,\mathrm{d}v} \,\mathrm{d}\left(B^u + B^f - C\right)(s) \right] . \tag{6.15}$$

Here, Q is the pricing measure and r is the stochastic short interest rate. Again, the stochastic element R_X enters via a sample path for the asset returns. In practice, one will often work with a discretized version of the stochastic differential equations in Equation (6.14). For an example, see Section 6.5.5.

We assume that the total payment stream of the policy $B^f + B^u - C$ is constructed in such a way that the assets X never become negative. This is, for example, satisfied if the expected premiums are continually enough to cover the expected fixed benefits. Generally, the natural premium for unit-linked insurance payments makes up only a small part of the total insurance and savings premium. Therefore, we do not see this last assumption as a critical limitation.

It is easily seen that the projection in unit-linked insurance is equivalent to the projection in participating life insurance. The contractual difference lies in the specification of how non-guaranteed payments arise (written in the contract *versus* decided fairly by the company along the way). Formally, the bonus updates d in Equation (6.4) are replaced by guarantee upgrades u in Equation (6.14), and the guarantee upgrades are determined by the assets and the guarantee account only, not by some collective reserves. Furthermore, the running guarantee g in Equation (6.4) is replaced by the final guarantee $g = (Y(R-) - X(R-))^+$ in Equation (6.14). Finally, both unit-linked accounts are based on market transition intensities, so the adjustment term α from Equation (6.4) vanishes. Apart from that, the driving stochastic differential equation are the same in participating life insurance and unit-linked insurance. In Section 6.5.5.1, we compare the two product types in a simple numerical example.

A real-life product example is the Danica Link from 2001 where $r^* = 0$, $\pi_g(t) = 1_{\{t<R\}}\kappa X(t-)$ and $u(t) = 1_{\{t<R\}}\left(\alpha\left(X(t-) - \pi_g(t)\right) - Y(t-)\right)^+$ for some constants $\kappa, \alpha \in (0,1)$ (see Steffensen and Waldstrøm (2009)). In other words, the guarantee account bears zero interest, but each year, it is upgraded,

so that it makes up at least a fraction α of the assets after the guarantee fee payment. At retirement, the difference between the guarantee account and the assets is added to the assets, if the guarantee account exceeds the assets. In return for this retirement guarantee, the policyholder pays a fraction κ of his assets to the equity holders of the insurance company each year.

6.5.4 Applications of Projections

Most importantly, the projections of X and Y can be used to calculate the total time 0 market cash flow CF and market value W for the contract, *i.e.*, to calculate

$$\begin{aligned} \mathrm{d}CF(t) &= \mathbb{E}^Q\left[\mathrm{d}B^u(t) + \mathrm{d}B^f(t) - \mathrm{d}C(t)\right] \\ &= \mathbb{E}^Q[X(t-)]\,\mathrm{d}\beta^p(t) + \mathrm{d}\beta^f(s) - \mathrm{d}\varsigma(s) \end{aligned}$$

and

$$\begin{aligned} W(0) &= \mathbb{E}^Q\left[\int_0^T e^{-\int_0^s r(v)\,\mathrm{d}v}\,\mathrm{d}\left(B^u + B^f - C\right)(s)\right] \\ &= \mathbb{E}^Q\left[\int_0^T e^{-\int_0^s r(v)\,\mathrm{d}v}\left(X(s-)\,\mathrm{d}\beta^p(s) + \mathrm{d}\beta^f(s) - \mathrm{d}\varsigma(s)\right)\right] . \end{aligned} \tag{6.16}$$

Again, r is the stochastic short interest rate. As in participating life insurance, the cash flow and market value distinguish themselves by including non-guaranteed payments as well as guaranteed payments. The Q-expectations are approximated as described in Section 6.4.10. If the short interest rate is deterministic, then Equation (6.16) simplifies to

$$W(0) = \int_0^T e^{-\int_0^s r(v)\,\mathrm{d}v}\,\mathrm{d}CF(s) . \tag{6.17}$$

The market value is useful for determining π_g, u and r^* according to the fairness criterion in Equation (6.15), which can be written as

$$W(0) = x_0 .$$

Again, the task of determining π_g, u and r^* is the classical trade-off between the aggressiveness of dividend allocation (expressed by u and r^*) and the option price (expressed by π_g) given the aggressiveness of the investment strategy (suppressed here).

The projections are also useful for determining the time 0 market value of the guarantee injection at time R, *i.e.*, to calculate

$$\mathbb{E}^Q\left[e^{-\int_0^R r(v)\,\mathrm{d}v}\left(Y(R-) - X(R-)\right)^+\right] .$$

6.5.5 Numerical Example

In this section, we go through a numerical example with a unit-linked version of the participating life insurance policy in Section 6.4.11. Working in a discrete projection setup, we show how to find a fair guarantee and guarantee fee strategy, and at the end of this section, we compare the unit-linked insurance policy with its participating life insurance counterpart. In unit-linked insurance, there is no interaction between policies as long as the insurance company has sufficient equity to meet its liabilities. In this paper, we do not model the insurance company's equity, but just assume that it is sufficient. Therefore, it is reasonable to consider just a single policy, but the example could easily be extended with more policies. The example is based on the 5000 scenarios generated via Monte Carlo simulation from Section 6.4.11. The number of simulated scenarios is, again, enough to ensure that our numerical results and graphs do not change between simulations.

The basics of the unit-linked policy are described in Examples 6.5–6.6. The policyholder is the 25-year-old female from Section 6.4.11. Her death is still governed by the market mortality intensity

$$\mu^m(t) = 0.8 \cdot \left(5 \cdot 10^{-4} + 5.3456 \cdot 10^{-5} \cdot e^{0.087498(25+t)}\right) .$$

The policy expires at time $T = R = 40$ when the policyholder is 65. For comparability, we fix the term insurance sum at $b^{ad} = 1$ and the premium at $\pi = 0.04614$ as in Section 6.4.11. Furthermore, we make the same market assumptions as in Section 6.4.11.1. The guarantee account is upgraded and the guarantee fee paid once a year. Hence, we have

$$\varepsilon(t) = \#\{i = 1, \dots, 40 : i \leq t\} .$$

We project the two accounts X and Y using steps of a size of one year by applying a discretized version of the stochastic differential equations for X and Y. For the discretization, we recall from Example 6.6 that β^p is a pure jump function and that ς and β^f are continuous functions. Hence, we get the stochastic difference equations

$$\begin{aligned}
X(t-) &= X(t-1)\left(1 + R_X(t)\right) - \int_{(t-1,t)} \left(\mathrm{d}\beta^f(s) - \mathrm{d}\varsigma(s)\right) , \\
X(t) &= X(t-) - X(t-)\,\Delta\beta^p(t) - \pi_g(t) \\
&\quad + 1_{\{t=R\}}\, g\left(1 - \Delta\beta^p(t)\right) , \qquad t = 1, \dots, 40 , \\
X(0) &= 0 , \qquad\qquad (6.18) \\
Y(t-) &= Y(t-1)\, e^{r^*} - \int_{(t-1,t)} \left(\mathrm{d}\beta^f(s) - \mathrm{d}\varsigma(s)\right) , \\
Y(t) &= Y(t-) - X(t-)\,\Delta\beta^u(t) + u(t) , \qquad t = 1, \dots, 40 , \\
Y(0) &= 0 .
\end{aligned}$$

We emphasize that the discretized projection in unit-linked insurance is practically the same as the discretized projection in participating life insurance, which eases the implementation. The guarantee account Y bears interest at the rate $r^* = 0$. The guarantee upgrade u is determined as a fraction θ_1 of the positive part of assets X less guarantee fee π_g and less guarantee account Y, all taken just before the guarantee upgrade, *i.e.*,

$$u(t) = \theta_1 \left((X(t-) - \pi_g(t)) - Y(t-) \right)^+ \ , \ t = 1, \ldots, 40 \ .$$

We fix $\theta_1 = 0.8$. At expiration, the guarantee $g = (Y(40-) - X(40-))^+$ is added to the assets to ensure that they match the guarantee account. The guarantee fee π_g is a fraction θ_2 of the positive part of the returns on the assets, *i.e.*,

$$\pi_g(t) = \theta_2 \left(R_X(t) X(t-1) \right)^+ \ .$$

We determine the fraction θ_2 according to the fairness criterion in Equation (6.15). Furthermore, using this guarantee fee, we consider

- the expected evolution of the assets $t \mapsto \mathbb{E}^Q[X(t)]$ and the guarantee account $t \mapsto \mathbb{E}^Q[Y(t)]$,
- the expected level for the guarantee upgrade $\mathbb{E}^Q[u(t)]$, $t = 1, \ldots, 40$.

We take the 5000 sample paths simulated in Section 6.4.11.1, and for each sample path, we project X and Y for different values of θ_2, using the difference equations in Equation (6.18). More specifically, we look for a θ_2, such that we get zero when approximating the time 0 market value $W(0)$ from Equation (6.17). We arrive at the fair guarantee fee fraction $\theta_2 = 0.1$. Furthermore, we calculate the average evolution of the assets and the guarantee account, the average guarantee fee and guarantee levels and the average guarantee upgrade.

In Figures 6.9–6.10, we plot the average evolution of the assets and the guarantee account and the average level of the guarantee upgrades. From Figure 6.9, we see that the final guarantee injection on average raises the assets by around 30%, even though the guarantee account does not bear interest. This is because the final guarantee kicks in for all sample paths where the assets finish below their (previous) maximum value. From Figure 6.10, we see how the yearly guarantee upgrades increase over time. This is explained by the fact that the guarantee upgrades are ultimately a fraction of the excess return on the assets. On average, the assets increase over time, and so does the average excess return and, hence, the guarantee upgrades.

6.5.5.1 Unit-Linked *versus* Participating Life

Finally, we compare the unit-linked insurance policy with its participating life insurance counterpart from Section 6.4.11. The comparison is straightforward,

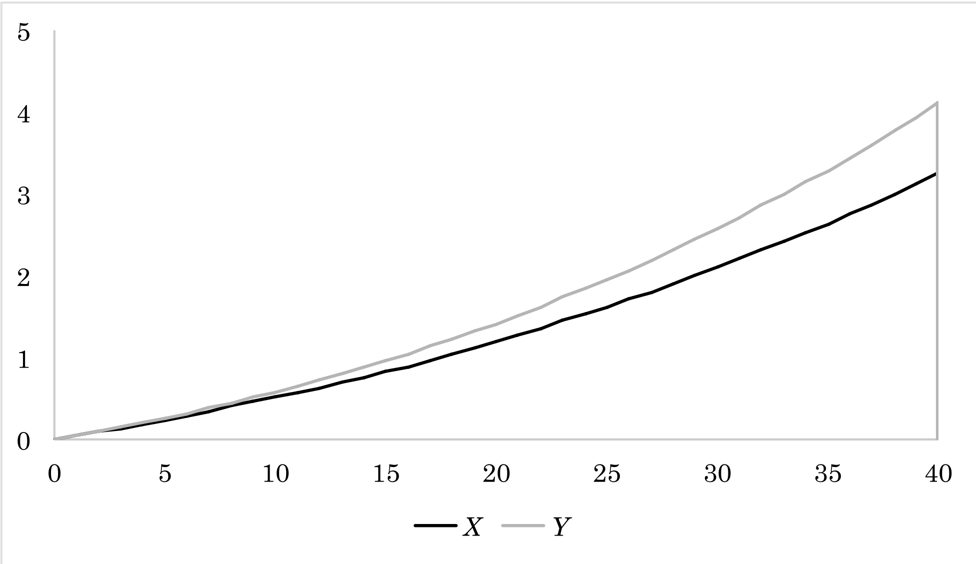

Figure 6.9: Approximated expected assets X and guarantee account Y as a function of time.

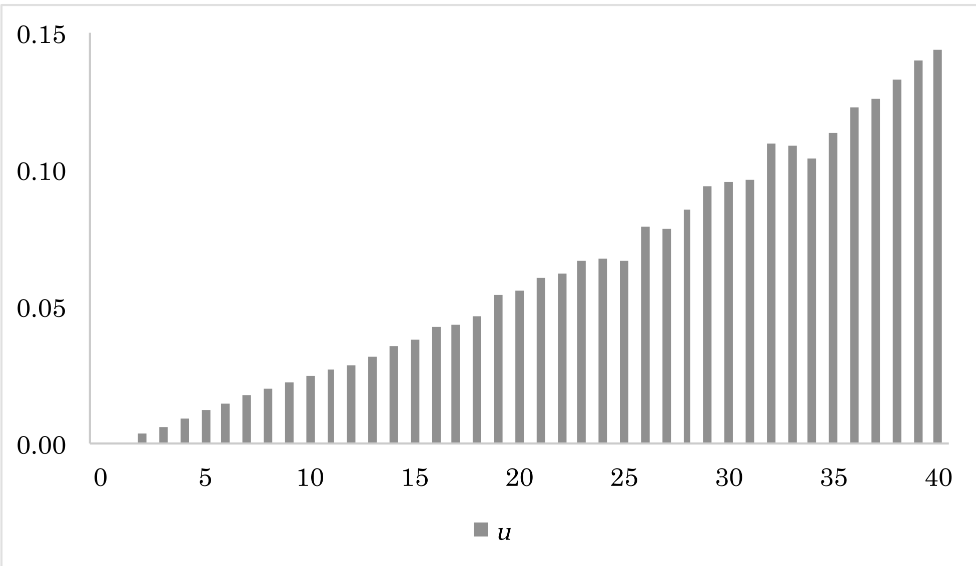

Figure 6.10: Approximated expected guarantee upgrade u as a function of time.

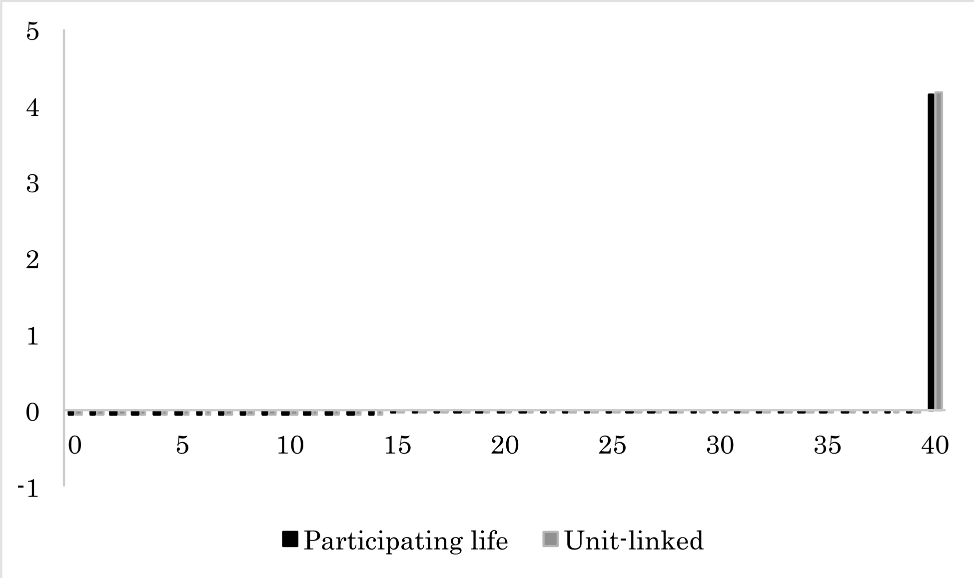

Figure 6.11: Approximated expected cash flows for the participating life insurance policy and the unit-linked insurance policy as a function of time.

since the two policies are modeled in the same framework. We plot the average cash flows for the two policies side-by-side in Figure 6.11. The two cash flows are practically the same. This was to be expected in order for the two contracts to be fair, since the premiums and death benefits are the same. To refine the picture, we plot the empirical distributions of the final payments for the two policies in Figure 6.12. We notice that the empirical distributions differ significantly. In particular, the unit-linked insurance policy has a bigger downside than the participating life insurance policy, and the average cash flow for the unit-linked insurance policy is held up by a few, very large, final payments and a heavier right-tale in general. This emphasizes the fact that, even though unit-linked insurance and participating life insurance are two sides of the same coin, the products may differ in riskiness. Our two-account model with event risk is a valuable tool in quantifying these differences.

6.6 Conclusions

We have introduced a two-account model with event risk, such as death and disability, for the purpose of modeling life insurance contracts taking into account both guaranteed and non-guaranteed payments in participating life insurance as well as in unit-linked insurance. We have formalized how the bonus schemes “consolidation” and “additional benefits” work and interact in participating life insurance, and we have formalized how guarantees can be implemented in unit-linked insurance. We have addressed similarities and

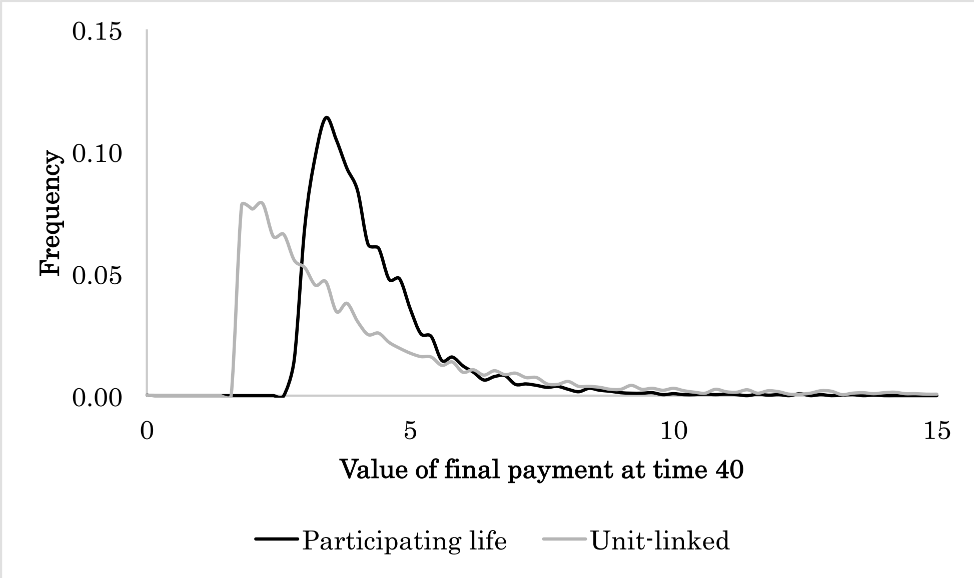

Figure 6.12: Empirical distribution of final payments for the participating life insurance policy and the unit-linked insurance policy.

differences between participating life insurance and unit-linked insurance, and for both product types, we have provided numerical examples to demonstrate the possible applications of our two-account model. Our numerical examples highlight the risk of unfair redistribution across policies in a seemingly homogeneous participation life insurance portfolio. Furthermore, the examples illustrate the potential difference in riskiness between a participating life insurance product and a unit-linked insurance product that are identical in expectation, but by (product) nature are different in guarantee structure. Our model is based on economic scenarios, which makes it flexible with respect to the change of financial input. We have illustrated the use of our model by conducting scenario analysis based on Monte Carlo simulation, but the model applies to scenarios in general and to worst-case and best-estimate scenarios in particular. Our work distinguishes itself from the previous literature by the inclusion of event risk and by the common framework for the valuation of guaranteed and non-guaranteed payments, in participating life and unit-linked insurance. Furthermore, the two-account structure makes it easy to illustrate general concepts, such as the interaction between realized return and bonus allocation (in participating life insurance) or the interaction between realized returns and the final guarantee (in unit-linked insurance). Finally, our paper provides a unique formalization of the most common bonus schemes in the Danish life insurance and pensions industry.

Chapter 7

Scenario-based Life Insurance Prognoses in a Multi-State Markov Model

Ninna Reitzel Jensen (2015)

Abstract: Traditional life insurance and pension prognoses from the policyholder's perspective do not illustrate financial riskiness or the effect of financial guarantees. We address this issue by introducing stochastic scenarios. Our model applies to participating life insurance as well as unit-linked insurance, and it is formulated in a general multi-state Markov model. In addition to illustrating financial riskiness, our model allows for tailor-made best-estimate prognoses in any financial market. We illustrate the use of our model by conducting scenario analysis based on Monte Carlo simulation, but the model applies to scenarios in general and to worst-case and best-estimate scenarios in particular. Our paper offers moderate mathematical complexity and a common framework for the valuation of life insurance payments across product types, and it fills the existing gap in the literature with respect to prognoses from the policyholder's perspective.

Keywords: Prognoses, bonus, economic scenarios, participating life insurance, unit-linked insurance, stochastic differential equations.

7.1 Introduction

In today's world of highly complex life insurance and pension products, the topic of life insurance and pension prognoses from the policyholder's perspective has never been more relevant. In particular, the rise of unit-linked insurance products increases the demand for prognoses that better illustrate financial riskiness. Traditional life insurance and pension prognoses are based on best-estimate expectations about future returns. Thereby, the aspect of financial riskiness is neglected, and more risky products tend to appear more attractive. Furthermore, the impact of financial guarantees is not apparent.

We address these issues by introducing stochastic scenarios. Our model applies to participating life insurance as well as unit-linked insurance, and it is formulated in a general multi-state Markov model. In addition to illustrating financial riskiness, our model allows for tailor-made best-estimate prognoses in any financial market without increased mathematical complexity. In our modeling, we condition on the policyholder starting and staying in a certain state of life, typically "alive and active". We have chosen this fixed path approach to provide policyholders with the best possible economic forecast given that they continue their course of life. A similar approach is suggested, but not pursued, in Section 5.7 of Norberg (2001).

In a recent paper, Jensen and Schomacker (2015) [Chapter 6 of this thesis], we introduced stochastic scenarios in participating life and unit-linked insurance to utilize pricing, hedging, market valuation, and solvency assessments of guaranteed and non-guaranteed payments, and for examining bonus allocation strategies. There, each scenario consists of two sample paths: one for the short interest rate, and one for the return of the fund that the policyholder and/or the insurance company invest in. In this paper, we introduce stochastic scenarios in participating life and unit-linked insurance to utilize tailor-made bonus, benefit, and retirement savings prognoses that illustrate financial riskiness. Here, each scenario consists of, either, a sample path for the bonus allocation (in participating life insurance), or, a sample path for the short interest rate and a sample path for the return of the fund that the policyholder invests in (in unit-linked insurance).

The scenarios may be worst-case scenarios, scenarios generated via Monte Carlo simulation or best-estimate scenarios. For a given scenario, the policyholder's account is projected into the future. For scenarios generated via Monte Carlo simulation, one obtains a valid savings, benefit, or bonus prognosis by averaging over sufficiently many projections (as is common practice with Monte Carlo simulation). For worst-case or best-estimate scenarios, a single projection is enough to obtain the corresponding worst-case or best-estimate prognosis. For Monte Carlo simulation, we refer to Glasserman (2004). For the generation of worst-case scenarios, we refer to Christiansen et al. (2014).

Prognoses from the policyholder's perspective are widely used in the life insurance and pension industry, but the topic is hardly covered in the literature. In Norberg (2001), the author treats prognoses in participating life insurance, but the model is only tractable for a very simple financial environment and does not apply to unit-linked insurance. To the knowledge of the author, this present paper is the first to address risk-based prognoses from the policyholder's prospective in participating life and unit-linked insurance in a general financial market. By risk-based prognoses, we mean prognoses that illustrate financial riskiness by going beyond best-estimate expectations about future returns. [Post-submission comment: For an expansion of this paragraph, see Appendix A.]

We model unit-linked insurance policies using two interacting accounts described by stochastic differential equations. One account measures the assets, and the other account is a technical account. For each scenario, the stochastic differential equations simplify to deterministic differential equations that can be solved numerically. A numerical solution can, for example, be obtained by applying a simple numerical discretization. Thereby, our model is simple to implement.

Participating life insurance differs from unit-linked insurance by having collective funds. In particular, the amount of bonus allocated to a policy depends on the evolution of the whole portfolio. In this paper, we take sample paths for the bonus allocation as stochastic input. The sample paths can be generated using the approach from Jensen and Schomacker (2015). Given the bonus allocation, participating life insurance policies can be modeled using just one account, namely the technical reserve of the policy. Apart from the stochastic input and the number of accounts, we model participating life and unit-linked insurance in the same framework. By doing so, we are able to compare the two. In their nature, unit-linked and participating life insurance seem different, but, in fact, they are not. The main difference lies in the specification of how non-guaranteed payments arise, stated in the contract from the beginning (unit-linked insurance) or determined fairly by the company along the way (participating life insurance).

In Section 7.2, we discuss scenario-based projection in general. Our main focus is on projection level and which measure to project under (physical or pricing measure). In Section 7.3, we formalize a common model for the state-wise evolution of the policies under consideration. In Section 7.4, we consider participating life insurance. We present a one-account model for a general participating life insurance policy. We condition on the policy staying in the same state. We end the section with a numerical example building on a simple policy in the classical survival model. The example illustrates how scenario-based calculations can be used for prognoses that illustrate financial riskiness. In Section 7.5, we consider unit-linked insurance. We touch upon different aspects of unit-linked insurance, and we present a two-account model for a general unit-linked insurance policy. We end the section with a numerical example that is a unit-linked version of the numerical example in the previous section. Again, the example illustrates how scenario-based calculations can be used for prognoses that illustrate financial riskiness. We compare the unit-linked insurance policy to its participating life insurance counterpart, making good use of our common modeling framework.

7.2 Projection in General

We assume that the stochastic scenarios arise from a financial model equipped with a physical measure P and a risk-neutral pricing measure Q.

For pricing, hedging, market valuation, and solvency assessments of guaranteed and non-guaranteed payments and for examining bonus allocation strategies, it is the expected evolution of the policy—both financially and across states—that is relevant. Hence, the evolution of the policy is considered on an average "portfolio level". The projections are carried out under the pricing measure since the focus is on pricing and valuation. This was the topic of Jensen and Schomacker (2015).

For retirement savings, benefit, and bonus prognoses, it is the expected financial evolution of the policy that is relevant. The policyholder needs to know what to expect in a certain state, not the expectation across states of life. Hence, the evolution of the policy is considered on an individual "policy level". However, in participating life insurance, the amount of bonus allocated to a policy depends on the financial evolution *and* the expected state-wise evolution of the policy. Hence, for the purpose of prognoses in participating life insurance, the assets and the reserves must, first, be projected on portfolio level to produce a sample path for the bonus allocation. Second, the sample path for the bonus allocation, typically expressed via a bonus basis, is used to project the reserve on an individual path-wise policy level. Unit-linked insurance does not require the same two-stage approach. In either case, the projections are carried out under the physical measure since the focus is on the actual bonus, retirement savings, and benefits. Projection on policy level is the topic of this paper.

7.3 Valuation Bases and Insurance Model

For a discussion of the different calculation bases in participating life insurance, see Jensen and Schomacker (2015). Below, we mark elements of the technical basis by superscript "$*$" and elements of the market basis by superscript "m".

In participating life insurance as well as in unit-linked insurance, we consider a policy whose state-wise evolution is governed by a continuous-time Markov process Z with a finite state space $\mathcal{J}$, starting in 0. For $k, j \in \mathcal{J}, j \neq k$, we define the counting process N_{jk} and the indicator process I_k by

$$\begin{aligned} N_{jk}(t) &= \#\{s \leq t : Z(s-) = j, Z(s) = k\} \ , \\ I_k(t) &= 1_{\{Z(t)=k\}} \ . \end{aligned}$$

With this definition, $N_{jk}(t)$ counts the number of jumps from state j to state k until time t, and $I_k(t)$ indicates sojourn in state k at time t. Under the technical basis, we model the evolution of Z by the transition intensities $t \mapsto \mu^*_{jk}(t)$, $j, k \in \mathcal{J}, j \neq k$, and under the market basis, we model the evolution of Z by the transition intensities $t \mapsto \mu^m_{jk}(t)$, $j, k \in \mathcal{J}, j \neq k$. The corresponding technical and market transition probabilities from state j to state k over the time-interval $[t, s]$ are denoted by $p^*_{jk}(t, s)$ and $p^m_{jk}(t, s)$, and with $\circ = *, m$ indicating the basis, we have

$$\mu_{jk}^{\circ}(t) = \lim_{h \downarrow 0} \frac{p_{jk}^{\circ}(t, t+h)}{h} .$$

We assume that the process Z governing the state of the policy is independent of the financial market, and under both P and Q, the evolution of Z is described by the transition intensities from the market basis.

7.4 Participating Life Insurance

In participating life insurance, the conservative technical basis gives rise to a systematic surplus that is to be paid back to the policyholders in terms of bonus. For a short survey on bonus schemes, see Møller and Steffensen (2007).

7.4.1 Bonus scheme

We consider a bonus scheme consisting of two steps: first, consolidation, and then—when the policy is consolidated on a sufficiently low technical interest rate (if ever)—additional benefits. By consolidation, we mean that the technical interest rate is lowered without changing the guaranteed payments. Consolidation is primarily used for policies with a technical interest rate that is "too high" compared to the market interest rate. Consolidation does not benefit the policyholder in terms of more favorable payments immediately after bonus payments, but it helps to ensure that the liabilities of the policy can be met. By additional benefits, we mean that bonus is used to increase parts of the guaranteed benefits proportionally, whereas the remaining benefits, the premiums, and the technical interest rate are maintained. Additional benefits is primarily used for policies with a low technical interest rate compared to the market interest rate. For a detailed description of the two bonus schemes, see Jensen and Schomacker (2015). Consolidation (in Danish "styrkelse") is much used in the Danish market, but it can easily be skipped below, heading straight for the bonus scheme additional benefits.

7.4.2 The Policy

We consider a participating life insurance policy with guaranteed payments based on a technical basis. The state-wise evolution of the policy is described in Section 7.3. The payments of the policy consist of a state-dependent guaranteed payment stream

$$B^u + B^f - C ,$$

where C is the premium stream ("C" for contributions), B^u is the benefit stream for the benefits that are increased ("B" for benefits, and superscript

"u" for upscaled), and B^f is the benefit stream for the benefits that are kept fixed (superscript "f" for fixed). They are given by

$$\begin{aligned} \mathrm{d}C &= \sum_{j \in \mathcal{J}} I_j \, \mathrm{d}c_j \ , \\ \mathrm{d}B^i &= \sum_{j \in \mathcal{J}} I_j \, \mathrm{d}b^i_j + \sum_{j,k \in \mathcal{J}: k \neq j} b^i_{jk} \, \mathrm{d}N_{jk} \ , \quad i = f, u \ , \end{aligned}$$

where c_j, b^f_j, and b^u_j are deterministic, state-wise payment streams, and b^f_{jk} and b^u_{jk} are deterministic lump sum payments upon jumps. We, hereby, exclude policyholder behavior options such as surrender and free policy since they imply non-deterministic payments, but for the purpose of prognoses, one typically assumes that the policyholder continues his course of life. Hence, for practical purposes, our assumption is unproblematic. The policy terminates at time T. Thereafter, there are no payments. For a simple example in the classical survival model, see Jensen and Schomacker (2015).

We denote by Y the technical reserve for the policy given that the policy starts and stays in state 0 (presumably "alive and active"). The policy is issued before or at time 0, and the account amounts to $Y(0-) = y_0$ just before time 0. The technical reserve Y accumulates according to the technical basis. In good times, the return rate on the insurance company's investments exceeds the technical interest rate. Parts of the excess return are allocated to the policy in terms of bonus which adds to the technical reserve, but parts are saved for times when the return rate is less favourable. Bonus is allocated at pre-specified, deterministic time points $0 < t_1 < \ldots < t_n = T$. We let

$$\varepsilon(t) = \#\{i = 1, \ldots, n : t_i \leq t\}$$

count the number of bonus allocations prior to time t. For all t with $\mathrm{d}\varepsilon(t) = 1$, i.e. for all time points with a bonus allocation, we assume that the bonus allocation $d_0(t)$ is known at time $t-$. This is to ensure predictability and, thereby, stochastic integrability. We write d_0 instead of just d to emphasize that we are dealing with the bonus allocation in state 0.

In Jensen and Schomacker (2015), it is described how to determine a fair bonus allocation strategy. In this paper, we take sample paths for the bonus allocation as stochastic input. We assume that the sample paths reflect the physical measure P.

7.4.3 Bonus Mechanisms

Bonus allocated to the policy is, first, used to lower the technical interest rate until it hits a pre-described level r^*. Typically, this level coincides with the technical interest rate for new policies. Thereafter, bonus is used to increase the benefits B^u. We let $r^{*(n)}$ denote the technical interest rate after the n-th bonus accrual and $k^{(n)}$ denote the upscaling of the benefits B^u after the n-th bonus accrual. After the n-th bonus accrual, the guaranteed benefit stream for the policy is given by

$$B^{(n)} = k^{(n)} B^u + B^f \ .$$

We note that $r^{*(n)}$ and $k^{(n)}$ depend on the stochastic bonus allocation and are therefore stochastic. However, for each bonus allocation scenario, we have a procedure for calculating them which is presented in Section 7.4.6. The upscaling factor starts at one, i.e. $k^{(0)} = 1$, and we have $k^{(n)} = 1$ for all n with $r^{*(n)} > r^*$, and if $k^{(n)} > 1$, then necessarily $r^{*(n)} = r^*$. This is because we do not increase the guaranteed benefits until the technical interest rate has been lowered to r^*.

For all t with $\mathrm{d}\varepsilon(t) = 1$, we assume that the technical interest rate $r^{*(\varepsilon(t))}$ and upscaling factor $k^{(\varepsilon(t))}$ are calculated at time $t-$. Again, this is to ensure predictability. Furthermore, additional benefits are in effect from time $t-$, such that benefits paid out at time t include the upscaling $k^{(\varepsilon(t))}$. The latter ensures that a policyholder with a final lump sum payment actually benefits from the last bonus update.

7.4.4 Technical Reserves and Risk Premiums

We denote by $V_j^{f,*}(\cdot, \rho)$ and $V_j^{u,*}(\cdot, \rho)$ the state-wise technical reserves for the payment streams $B^f - C$ and B^u given that the policy is in state j and that the technical interest rate is ρ. We have

$$\begin{aligned} V_j^{f,*}(t, \rho) &= \mathbb{E}^* \left[\int_t^T e^{-\rho(s-t)} \, \mathrm{d}\left(B^f - C\right)(s) \middle| Z(t) = j \right] \\ &= \int_t^T e^{-\rho(s-t)} \sum_{l \in \mathcal{J}} p_{jl}^*(t, s) \left\{ \mathrm{d}b_l^f(s) + \sum_{k \in \mathcal{J}: k \neq l} \mu_{lk}^*(s)\, b_{lk}^f(s)\, \mathrm{d}s - \mathrm{d}c_l(s) \right\}, \\ V_j^{u,*}(t, \rho) &= \mathbb{E}^* \left[\int_t^T e^{-\rho(s-t)} \, \mathrm{d}B^u(s) \middle| Z(t) = j \right] \\ &= \int_t^T e^{-\rho(s-t)} \sum_{l \in \mathcal{J}} p_{jl}^*(t, s) \left\{ \mathrm{d}b_l^u(s) + \sum_{k \in \mathcal{J}: k \neq l} \mu_{lk}^*(s)\, b_{lk}^u(s)\, \mathrm{d}s \right\}, \end{aligned} \tag{7.1}$$

where $\mathbb{E}^*$ denotes technical expectation and p_{jl}^* is the technical probability of transition from state j to l. Both are determined by the transition intensities from the technical basis. The state-wise technical reserves can be calculated numerically by use of Thiele's differential equations; see Hoem (1969).

We denote by $V_j^*(\cdot, \rho, k)$ the state-wise technical reserve for the (partly upscaled by k) payment stream $B^f + kB^u - C$ given that the policy is in state j and that the technical interest rate is ρ, i.e.

$$V_j^*(t, \rho, k) = k V_j^{u,*}(t, \rho) + V_j^{f,*}(t, \rho) \ , \quad j \in \mathcal{J} \ . \tag{7.2}$$

Finally, with a slight abuse of notation, we denote by $V_j^*(\cdot,\rho)$ the state-wise technical reserve for the initial payment stream $B^f + B^u - C$ given that the policy is in state j and that the technical interest rate is ρ, i.e.

$$V_j^*(t,\rho) = V_j^{u,*}(t,\rho) + V_j^{f,*}(t,\rho) = V_j^*(t,\rho,1) \ , \quad j \in \mathcal{J} \ . \tag{7.3}$$

In exchange for the policyholder's insurance coverage, a technical risk premium is continuously deducted from the technical reserve. The risk premium covers all risk associated with future payments upon leaving state 0. We denote the technical risk premium stream in state 0 by s_0^*, and it is given by

$$\begin{aligned} \mathrm{d}s_0^*(t) = \mathbb{E}^*\Big[\sum_{j\in\mathcal{J}:j\neq 0} \Big(& k^{(\varepsilon(t))} b_{0j}^u(t) + b_{0j}^f(t) + V_j^*\left(t, r^{*(\varepsilon(t))}, k^{(\varepsilon(t))}\right) \\ & - V_0^*\left(t, r^{*(\varepsilon(t))}, k^{(\varepsilon(t))}\right)\Big) \mathrm{d}N_{0j}(t) \,\Big|\, Z(t) = 0, r^{*(\varepsilon(t))}, k^{(\varepsilon(t))} \Big] \\ = \sum_{j\in\mathcal{J}:j\neq 0} & \mu_{0j}^*(t) \left(k^{(\varepsilon(t))} b_{0j}^u(t) + b_{0j}^f(t) + V_j^*\left(t, r^{*(\varepsilon(t))}, k^{(\varepsilon(t))}\right)\right) \mathrm{d}t \\ & - Y(t) \sum_{j\in\mathcal{J}:j\neq 0} \mu_{0j}^*(t)\,\mathrm{d}t, \end{aligned}$$

where V_j^*, $j \in \mathcal{J}$, are the state-wise technical reserves defined in Equation (7.2).

7.4.5 Account Projection

This subsection constitutes the largest conceptual departure from Jensen and Schomacker (2015) with respect to participating life insurance. There are two main differences: Firstly, instead of considering the expected evolution of the policy across states, we condition on the policy starting and staying in state 0. In Jensen and Schomacker (2015), we projected the policyholder's accounts by adding the expected premiums and subtracting the expected benefits. In that way, we obtained the expected evolution of the policyholder's accounts across states which is relevant for the purpose of e.g. market valuation. Below, we add and subtract the actual premiums and benefits in state 0. In addition, we subtract the technical risk premium associated with jumps out of the state. As a result, we obtain the evolution in state 0 which is relevant for the purpose of prognoses from the policyholder's perspective. Secondly, the stochastic input is different. In Jensen and Schomacker (2015), the stochastic element consisted of the short interest rate and the return on the assets of the policy. Below, the stochastic element consists of a sample path for the bonus allocation in state 0, derived via account projection on portfolio level, see Jensen and Schomacker (2015). In exchange for having the bonus allocation as input, we only need to project the technical reserve and not the assets of the policy as in Jensen and Schomacker (2015).

Conditional on the policy starting and staying in state 0, the technical reserve Y of the policy evolves according to the stochastic differential equation

$$\begin{aligned} \mathrm{d}Y(t) &= Y(t)\, r^{*(\varepsilon(t))}(t)\,\mathrm{d}t + \mathrm{d}c_0(t) - k^{(\varepsilon(t))}\mathrm{d}b_0^u(t) - \mathrm{d}b_0^f(t) \\ &\qquad - \mathrm{d}s_0^*(t) + d_0(t)\,\mathrm{d}\varepsilon(t) \ , \\ Y(0-) &= y_0 \ . \end{aligned} \tag{7.4}$$

Here, d_0 is the stochastic bonus allocation in state 0, and ε counts the number of bonus allocations. Furthermore, c_0 is the premium stream in state 0, b_0^u and b_0^f are the initial benefit streams in state 0, and s_0^* is technical risk premium stream in state 0. The stochastic element d_0 enters via a sample path for the bonus allocation. This way of formalizing the technical reserve, including bonus, is a generalization of Møller and Steffensen (2007).

7.4.6 Procedure for Determining the Technical Interest Rate and the Upscaling Factor

The procedure for determining the technical interest rate and upscaling factor that we present below might, at first glance, seem identical to the procedure in Jensen and Schomacker (2015). However, there is an important conceptual difference. In Jensen and Schomacker (2015), the procedure is based on the distribution of the policy across states. Hence, the resulting technical interest rate and upscaling factor are averaged quantities that do not reflect the evolution of an actual policy. Moreover, one has to decide which basis to use for the distribution of the policy; the realistic market basis or the computationally simpler technical basis. For a brief discussion, see Jensen and Schomacker (2015). Below, the procedure is based on the policy being in state 0, and the resulting technical interest rate and upscaling factor reflect the possible evolution of a policy in state 0. Also, there is no decision to be made on valuation basis.

We fix a time point t with $\mathrm{d}\varepsilon(t) = 1$ such that there is a bonus allocation at time t. First, we assume that $r^{*(\varepsilon(t-))} > r^*$, so that the policy is still in the consolidation phase of the bonus scheme. Then, necessarily, $k^{(\varepsilon(t-))} = 1$ (since we consolidate first), and the technical interest rate $r^{*(\varepsilon(t))}$ is determined as the solution to the equation

$$Y(t-) + d_0(t) = V_0^*\left(t-, r^{*(\varepsilon(t))}\right) \ ,$$

where V_0^* is the state-wise technical reserve defined in Equation (7.3). Hence, $r^{*(\varepsilon(t))}$ is the technical interest rate that complies with the equivalence principle on policy level. If the solution $r^{*(\varepsilon(t))}$ is strictly smaller than r^*, then $r^{*(\varepsilon(t))}$ is set to r^*, and the remaining bonus

$$Y(t-) + d_0(t) - V_0^*(t-, r^*)$$

is used to raise the upscaling factor $k^{(\varepsilon(t))}$ as below. Otherwise, we set $k^{(\varepsilon(t))} = 1$. We emphasize that consolidation increases the technical reserve without changing the guaranteed payments. Thereby, the liabilities of the insurance company are unaffected by the bonus allocation. An increased technical reserve combined with unchanged liabilities corresponds to a more well-founded policy. Therefore, we use the term consolidation.

Now, assume that $r^{*(\varepsilon(t-))} = r^*$. Then, the policy is in the additional benefits phase of the bonus scheme, and we set $r^{*(\varepsilon(t))} = r^*$. The upscaling factor $k^{(\varepsilon(t))}$ is determined as the solution to the equation

$$d_0(t) = \left(k^{(\varepsilon(t))} - k^{(\varepsilon(t-))}\right) V_0^{u,*}(t-, r^*) \ ,$$

i.e.

$$k^{(\varepsilon(t))} = k^{(\varepsilon(t-))} + \frac{d_0(t)}{V_0^{u,*}(t-, r^*)} \ .$$

Here, $V_0^{u,*}$ is the technical reserve in state 0 for the benefit stream B^u given that the interest rate is r^*. The reserve is given in Equation (7.1). Hence, $k^{(\varepsilon(t))}$ is the upscaling factor that satisfies the equivalence principle on policy level.

There is no reason to consider the case $r^{*(\varepsilon(t-))} < r^*$. For details, see Jensen and Schomacker (2015).

7.4.7 Benefit and Bonus Prognoses from the Policyholder's Perspective

The projections of Y on policy level are useful for bonus and benefit prognoses with confidence intervals, conditional on the policy starting and staying in state 0. The projections reflect the financial evolution of the policy given that the policyholder continue his course of life which is exactly what most prognoses from the policyholder's perspective focus on. For benefit prognoses, it is relevant to calculate the expected upscaling factor at the retirement date R,

$$\mathbb{E}^P_{R,0}\left[k^{(\varepsilon(R))}\right] \ ,$$

and the expected benefit stream

$$t \mapsto \mathbb{E}^P_{t,0}\left[k^{(\varepsilon(t))}\right]\left(\mathrm{d}b_0^u(t) + \sum_{j\in\mathcal{J}:j\neq 0} b_{0j}^u \,\mathrm{d}N_{0j}(t)\right) + \mathrm{d}b_0^f(t) + \sum_{j\in\mathcal{J}:j\neq 0} b_{0j}^f \,\mathrm{d}N_{0j}(t) \,,$$

given that the policy starts and stays in state 0. Here, $\mathbb{E}^P_{t,0}$ denotes P-expectation given sojourn in state 0 up to time t, i.e.

$$\mathbb{E}^P_{t,0}[\ \cdot\ ] = \mathbb{E}^P[\ \cdot\ |\forall s \leq t : Z(s) = 0] \ ,$$

where Z is the stochastic process governing the state of the policy.

If the projections are based on bonus allocation scenarios generated via Monte Carlo simulation, then for each t, the expectation $\mathbb{E}^P_{t,0}\left[k^{(\varepsilon(t))}\right]$ is approximated by averaging over a sufficient number of P-projections up to time t. If, instead, the projections are of the worst-case or best-estimate type (and, hence, singular), then $\mathbb{E}^P_{t,0}\left[k^{(\varepsilon(t))}\right]$ is just the single projected value. To illustrate the financial riskiness of the policy, one can calculate quantiles for $k^{(\varepsilon(t))}$, based on Monte Carlo-based projections.

For bonus prognoses, it is relevant to calculate the expected bonus stream

$$t \mapsto \left(\mathbb{E}^P_{t,0}\left[k^{(\varepsilon(t))}\right] - 1\right)\left(\mathrm{d}b^u_0 + \sum_{j\in\mathcal{J}:j\neq 0} b^u_{0j}\,\mathrm{d}N_{0j}(t)\right) .$$

The expected bonus stream is approximated in the same way as the expected benefit stream.

7.4.8 Numerical Example

The example below is based on 5000 bonus allocation scenarios generated via Monte Carlo simulation. We have made sure that the number of simulated scenarios is sufficiently high for our numerical results and graphs not to change between simulations, but we do not go into details about the robustness of the simulations.

We consider a participating life insurance policy which is identical to the example policy in Jensen and Schomacker (2015). The state of the policy is described by the classical survival model with two states, 0 (alive) and 1 (dead). For simplicity, we write $\mu^* = \mu^*_{01}$. The payments of the policy consist of a constant continuous premium payment π while alive, a term insurance sum b^{ad} upon death before expiration T, and a pure endowment sum b^a upon survival until expiration T. Under the bonus scheme "additional benefits", bonus is used to increase the endowment sum. There are no payments in the death state.

The policyholder is a female aged 25 at time 0 when the policy is issued. We fix $r^* = 0.02$, and we assume that $r^{*(0)} = r^*$ which is natural for a newly-issued policy. Thereby, we only consider the bonus scheme "additional benefits". The death of the policyholder is governed by the technical mortality intensity

$$\mu^*(t) = 5\cdot 10^{-4} + 5.3456\cdot 10^{-5}\cdot e^{0.087498(25+t)} .$$

For the last three decades, this has served as a standard mortality intensity for adult women in Denmark. It is part of the so-called G82 technical basis that was set forth as a Danish industry standard in 1982. The market mortality intensity is given by

$$\mu^m(t) = 0.8\mu^*(t) . \tag{7.5}$$

The policy expires at time $T = 40$ when the policyholder is 65. In the notation from the previous sections, we have

$$\begin{aligned}
&\mathrm{d}b_0^f(t) = \mathrm{d}b_1^f(t) = \mathrm{d}c_1(t) = \mathrm{d}b_1^u(t) = \mathrm{d}b_{01}^u(t) = 0\ ,\\
&\mathrm{d}c_0(t) = \pi\,\mathrm{d}t\ ,\\
&\mathrm{d}b_0^u(t) = b^a 1_{\{t=40\}}\ ,\\
&\mathrm{d}b_0^f(t) = 0\ .
\end{aligned}$$

Since there are no payments in the dead state, we get $V_1^*(t,\cdot,\cdot) = 0$. Hence, the technical risk premium stream is

$$\mathrm{d}s_0^*(t) = \mu^*(t)\left(b^{ad} - Y(t)\right)\,\mathrm{d}t\ .$$

We fix the term insurance sum at $b^{ad} = 1$ and the pure endowment sum at $b^a = 3$ as in Jensen and Schomacker (2015). The equivalence premium is determined via the equivalence relation

$$V_0^*(0, r^*) = 0\ .$$

For this simple policy, we have

$$V_0^*(0, r^*) = b^a e^{-\int_0^T (r^* + \mu^*(v))\,\mathrm{d}v} + \int_0^T e^{-\int_0^s (r^*+\mu^*(v))\,\mathrm{d}v}\left(b^{ad}\mu^*(s) - \pi\right)\mathrm{d}s\ ,$$

so we get the premium

$$\pi = \frac{b^a e^{-\int_0^T (r^*+\mu^*(v))\,\mathrm{d}v} + b^{ad}\int_0^T e^{-\int_0^s (r^*+\mu^*(v))\,\mathrm{d}v}\mu^*(s)\,\mathrm{d}s}{\int_0^T e^{-\int_0^s (r^*+\mu^*(v))\,\mathrm{d}v}\,\mathrm{d}s}\ .$$

Using numerical methods, we obtain $\pi = 0.04614$.

The bonus d_0 is allocated once a year and is given by

$$d_0(t) = \left((\mu^m(t) - \mu^*(t))\left(b^{ad} - Y(t-1)\right)\right)^+ + \delta(t)\,Y(t-1)\ .$$

Here, the first term is risk bonus. The second term is interest rate bonus, and the excess interest rate δ enters via sample paths. Using standard Monte Carlo methods, we simulate 5000 sample paths for the excess interest rate δ using the approach from Section 4.11.1 in Jensen and Schomacker (2015). For details, we refer to Jensen and Schomacker (2015), but to keep this paper self-contained, we provide this recapitulation:

1. We assume a deterministic short interest rate $r = 0.04$. Assets corresponding to the policy are invested in a simple Black–Scholes stock with drift 0.07 and volatility 0.2 under the physical measure P. The excess

return gives rise to a collective bonus potential K which is defined as the maximum of zero and assets less guaranteed liabilities, i.e.

$$K(t) = (X(t) - L(t))^{+} \ .$$

Here, the guaranteed liabilities L is the maximum of the market reserve V and the technical reserve V^* for the guaranteed payments, i.e.

$$L(t) = \max\{V(t), V^*(t)\} \ .$$

For details, see Jensen and Schomacker (2015).

2. The assets and reserves X, V^*, and V are simulated using standard Monte Carlo methods (on portfolio level). For each sample path, the interest rate bonus d is determined as a fraction θ_1 of the excess collective bonus potential K over a threshold $\bar{K}$, i.e.

$$d(t) = \theta_1 \left(K(t-) - \bar{K}(t-)\right)^{+} \ .$$

The threshold $\bar{K}$ can be seen as a preferred minimum collective bonus potential and is given by

$$\bar{K}(t) = \theta_2 L(t) \ .$$

3. The interest rate bonus d is converted to an excess interest rate δ via the formula

$$\delta(t) = \frac{d(t)}{V^*(t-1)} \ .$$

We fix $\theta_1 = 0.2$ and $\theta_2 = 0.1$. In exchange for the right to bonus, the policyholder pays a guarantee fee which is subtracted from the assets. The guarantee fee is a fraction θ_3 of the positive part of the returns on the assets, and it is determined such that the policy is financially fair. By financially fair, we mean that the equivalence principle is satisfied for the total payments under the pricing measure, i.e.

$$\mathbb{E}^Q \left[\int_0^T e^{-rs} \, \mathrm{d}\left(B^{(\varepsilon(s))} - C\right)(s) \right] = x_0 \ .$$

The fair guarantee fee fraction amounts to $\theta_3 = 0.31$. [Post-submission comment: For a revised recapitulation, see Appendix A.]

For each sample path of δ, we project Y, using a discretized version of the stochastic differential equation in Equation (7.4). In Figure 7.1, we plot the average and guaranteed evolution of the policyholder's account. We also plot quantiles for the evolution of the account. The expected endowment sum at retirement, $k^{(40)}b^a$, amounts to 6.2 which means that 52% of the endowment

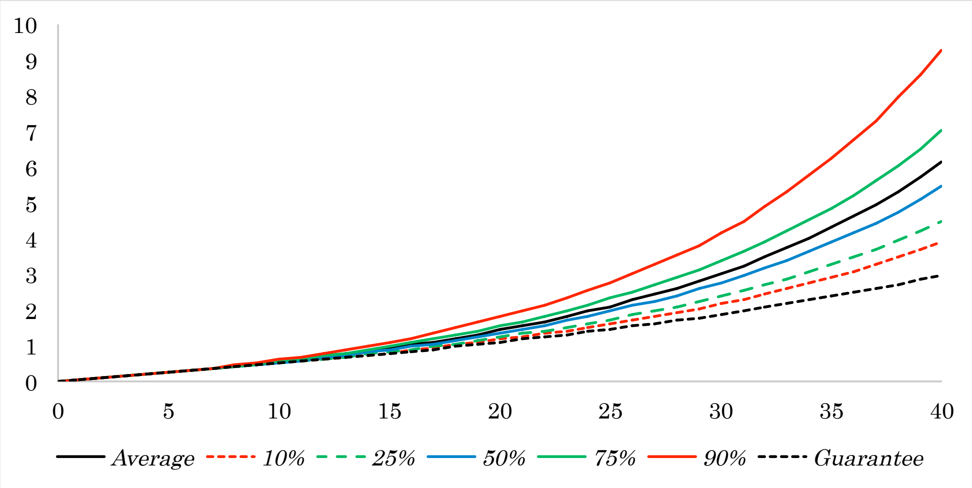

Figure 7.1: Average, quantiles and guaranteed evolution of the policyholder's account Y as a function of time.

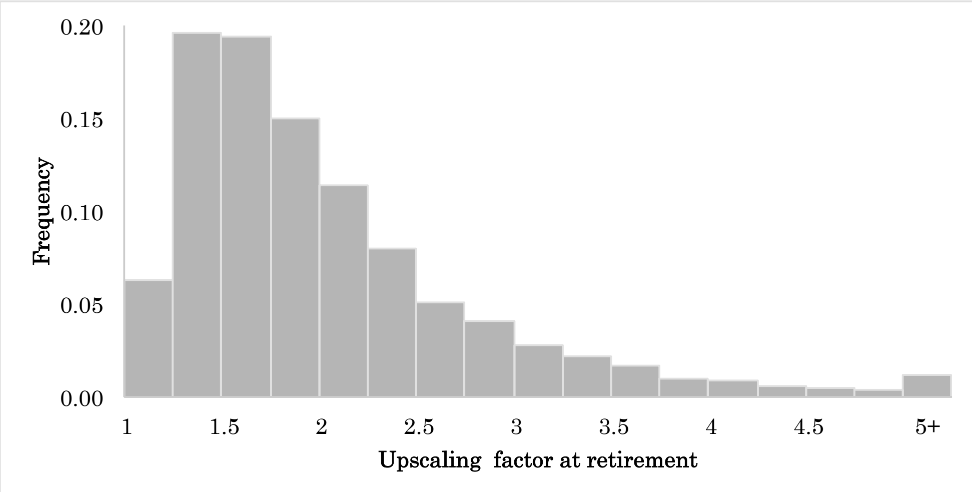

Figure 7.2: Empirical distribution of the final upscaling factor $k^{(40)}$.

sum comes from bonus. This is primarily due to the low technical interest rate. With regards to prognoses, the average provides the policyholder with a best-estimate of the endowment sum at retirement, given that the policyholder is still alive at retirement. In addition, the policyholder gets a best-estimate of the evolution of the technical reserve which is useful for surrender where the policyholder typically receives the technical reserve. The quantiles add to the picture by illustrating the financial riskiness of the policy. The message is clear; even a participating life insurance policy inflicts financial risk.

In Figure 7.2, we plot the empirical distribution of the upscaling factor at retirement. The upscaling factor can never fall below 1, but for 63% of the sample paths, the endowment sum at retirement is less than the average of 6.2. This is valuable information for the policyholder, and in traditional prognoses, the riskiness would not be apparent. Our model is a unique tool in providing this information.

7.5 Unit-Linked Insurance

In unit-linked insurance, (parts of) the benefits are directly linked to the financial market. To ensure a certain living standard after retirement, unit-linked insurance policies often come with an embedded financial guarantee. The guarantee can take on many forms. We focus on a guarantee consisting of a guaranteed minimum retirement savings amount at the retirement date.

7.5.1 Two-Account Model

We consider a unit-linked insurance policy. The state-wise evolution of the policy is described in Section 7.3. The policy includes a guaranteed minimum retirement savings amount at the retirement date R, based on a guarantee account with a guaranteed interest rate r^* (for example, $r^* = 0$). We do not take costs into account. We denote by X the assets of the policy and by Y the guarantee account, given that the policy starts and stays in state 0 (presumably "alive and active"). The policy is issued before or at time 0, and the two accounts amount to $X(0-) = x_0$ and $Y(0-) = y_0$ just before time 0. The policy terminates at time $T \geq R$. Thereafter, there are no payments.

The assets X are invested in a fund with stochastic return R_X. The guarantee account Y accumulates according to the guaranteed interest rate. In good times, the return rate on the assets exceeds the technical interest rate, and then, the assets outgrow the guarantee account. In that case, the guarantee account is upgraded (increased) according to the terms of the contract. Regardless of the developments in the financial market, the guarantee account is never to be downgraded (lowered), and at retirement, the maximum value of the assets and the guarantee account is paid out to the policyholder. In bad times when the guarantee account exceeds the assets at retirement, the equity holders of the insurance company step in with a capital injection taken from the company's equity. We speak of the possible capital injection

$$g = (Y(R-) - X(R-))^+$$

as guarantee injection, and its role is to raise the assets at retirement in case of unfavorable developments in the financial market. The policyholder pays for the company's risk taking by having a guarantee fee deducted from the assets and paid to the equity holders of the insurance company.

The two accounts X and Y are updated at pre-specified deterministic time points $0 < t_1 < \ldots < t_n = T$ (for example, once a year). We let

$$\varepsilon(t) = \#\{i = 1, \ldots, n : t_i \leq t\}$$

count the number of updates prior to time t. The updates consist of upgrades, u, of the guarantee account (if the assets exceed the guarantee account in a pre-described way) and deductions of the guarantee fee, π_g, in return for the possible guarantee injection g at retirement. At retirement R, the assets are updated with the guarantee injection, g, if the guarantee account exceeds the assets. We let $\varepsilon_R(t) = 1_{\{t \geq R\}}$ mark this final update. After retirement, the guarantee account falls away, and the assets evolve without any underlying guarantee until termination T.

We assume that the stochastic return on the assets, R_X, does not jump at time points with an account update. Furthermore, to ensure predictability, we assume that $u(t)$ and $\pi_g(t)$ are known at time $t-$ for all time points t with an account update. We take sample paths for the return on the assets as stochastic input. We assume that the sample paths reflect the physical measure P.

7.5.2 Product Specification

The payments of the policy consist of a state-dependent payment stream

$$B^f + B^u - C \ ,$$

where C is a fixed state-dependent premium stream ("C" for contribution), B^f is a fixed state-dependent benefit stream ("B" for benefits and superscript "f" for fixed), and B^u is a state-dependent benefit stream that is linked to the financial market (superscript "u" for unit-linked). More precisely, B^u is linear in the assets X, i.e.

$$\mathrm{d}B^u(t) = X(t-)\,\mathrm{d}B^p(t) \ ,$$

where B^p denotes a fixed state-dependent benefit stream (superscript "p" for profile). We write $X(t-)$ instead of just $X(t)$ to ensure that the asset process X is well-defined (see definition below). We assume that the total payment stream of the policy, $B^f + B^u - C$, is constructed in such a way that the assets X never become negative. This is, for example, satisfied if the premium stream in state 0 is continually enough to cover the risk premium stream (defined below) for the fixed benefit stream.

The fixed benefit stream B^f includes insurance payments such as disability or death payments whereas the market-linked benefit stream B^u includes, for example, deposit protection, surrender payments, a variable pure endowment or a variable life annuity. Formally, the payment streams of the policy are given as

$$\mathrm{d}C = \sum_{j\in\mathcal{J}} I_j \,\mathrm{d}c_j \ ,$$
$$\mathrm{d}B^i = \sum_{j\in\mathcal{J}} I_j \,\mathrm{d}b^i_j + \sum_{j,k\in\mathcal{J}:j\neq k} b^i_{jk}\,\mathrm{d}N_{jk} \ , \quad i = f, p \ ,$$

where c_j, b^f_j and b^p_j are deterministic, state-wise payment streams and b^f_{jk} and b^p_{jk} are deterministic lump sum payments upon jumps. For a simple example in the classical survival model, see Jensen and Schomacker (2015). To ensure that the assets are paid out to the policyholder, we assume that $\Delta b^f_0(T) = \Delta c_0(T) = 0$ and $\Delta b^p_0(T) = 1$. Here, Δ denotes the jump part of the processes.

7.5.3 Risk Premiums

We let $V^{f,m}_j(t)$ denote the state-wise market reserve at time t for the benefit stream B^f given that the policy is in state j at time t. We have

$$\begin{aligned} V^{f,m}_j(t) &= \mathbb{E}^m\left[\int_t^T e^{-\int_t^s r(v)\,\mathrm{d}v}\,\mathrm{d}B^f(s)\,\middle|\, Z(t) = j, r(t)\right] \\ &= \int_t^T e^{-\int_t^s r_t(v)\,\mathrm{d}v} \sum_{l\in\mathcal{J}} p^m_{jl}(t,s)\left\{\mathrm{d}b^f_l(s) + \sum_{k\in\mathcal{J}:k\neq l} \mu^m_{lk}(s)\, b^f_{lk}(s)\,\mathrm{d}s\right\}, \end{aligned}$$

where $\mathbb{E}^m$ denotes market expectation, r is the stochastic short interest rate, and r_t is the yield curve seen from time t.

By s^f_0, we denote the market risk premium stream in state 0 associated with the benefit stream B^f. The risk premium covers all risk associated with future payments of B^f upon leaving state 0. We have

$$\begin{aligned} \mathrm{d}s^f_0(t) &= \mathbb{E}^m\left[\sum_{j\in\mathcal{J}:j\neq 0}\left(b^f_{0j}(t) + V^{f,m}_j(t) - V^{f,m}_0(t)\right)\mathrm{d}N_{0j}(t)\,\middle|\, Z(t) = 0\right] \\ &= \sum_{j\in\mathcal{J}:j\neq 0} \mu^m_{0j}(t)\left(b^f_{0j}(t) + V^{f,m}_j(t) - V^{f,m}_0(t)\right)\mathrm{d}s \ . \end{aligned}$$

By s^u_0, we denote the market risk premium stream in state 0 associated with the benefit stream B^u. The risk premium s^u_0 covers unit-linked lump sum payments upon jumps out of state 0. Letting $P_j(t)$ indicate whether the policyholder keeps his assets upon transition to state j at time t, we get

$$\begin{aligned} \mathrm{d}s^u_0(t) &= X(t-)\,\mathbb{E}^m\left[\sum_{j\in\mathcal{J}:j\neq 0}\left(b^p_{0j}(t) + P_j(t) - 1\right)\mathrm{d}N_{0j}(t)\,\middle|\, Z(t) = 0\right] \\ &= X(t)\sum_{j\in\mathcal{J}:j\neq 0} \mu^m_{0j}(t)\left(b^p_{0j}(t) + P_j(t) - 1\right)\mathrm{d}s \ . \end{aligned}$$

We emphasize that for states k where the policy holder does not keep his assets, corresponding to $P_k(t) = 0$, the policyholder actually receives $X(t)\,\mu_{0k}^{m}(t)$ in exchange for giving up his assets upon transition to state k at time t.

7.5.4 Two-Account Projection

This subsection constitutes the largest conceptual departure from Jensen and Schomacker (2015) with respect to unit-linked insurance. There is one main difference: Instead of considering the expected evolution of the policy across states, we condition on the policy starting and staying in state 0. As for participating life insurance, we do this by adding and subtracting the actual premiums and benefits in state 0 instead of adding and subtracting the expected premiums and benefits. In addition, we subtract the market risk premium associated with jumps out of the state. As opposed to participation life insurance, the stochastic input in unit-linked insurance is the same in this paper and Jensen and Schomacker (2015). In both cases, the stochastic element consists of a sample path for the short interest rate, r, and the return, R_X, of the fund that the policyholder invests in.

Conditional on the policy starting and staying in state 0, the assets X and the guarantee account Y of the policy evolve according to the stochastic differential equations

$$
\begin{aligned}
\mathrm{d}X(t) &= X(t-)\,\mathrm{d}R_X(t) + \mathrm{d}c_0(t) - \mathrm{d}b_0^f(t) - X(t-)\,\mathrm{d}b_0^p(t) \\
&\quad - \mathrm{d}s_0^f(t) - \mathrm{d}s_0^u(t) \\
&\quad - \pi_g(t)\,\mathrm{d}\varepsilon(t) + \left(Y(R-) - X(R-)\right)^+ \mathrm{d}\varepsilon_R(t)\ , \\
X(0-) &= x_0\ , \\
\mathrm{d}Y(t) &= Y(t)\,r^*(t)\,\mathrm{d}t + \mathrm{d}c_0(t) - \mathrm{d}b_0^f(t) - X(t-)\,\mathrm{d}b_0^p(t) \\
&\quad - \mathrm{d}s_0^f(t) - \mathrm{d}s_0^u(t) + u(t)\,\mathrm{d}\varepsilon(t)\ , \qquad t \le R\ , \\
Y(0-) &= y_0\ , \\
Y(t) &= 0\ ,\ t > R\ .
\end{aligned}
\tag{7.6}
$$

We recall that π_g is premium for the included retirement guarantee, u is the upgrade of the guarantee account, ε counts the number of guarantee premium payments and guarantee upgrades (typically annual), and ε_R marks the exercise of the guarantee at the retirement date. This way of formalizing the assets and the guarantee account is a generalization of Steffensen and Waldstrøm (2009) to a set-up with insurance risk. The stochastic element R_X enters via a sample path for the asset returns. The short interest rate, r, enters via the market risk premium.

The guarantee account is set to zero after time R because it falls away after retirement, and the assets evolve without any underlying guarantee until termination T. The three quantities r^*, u, and π_g are specified in the contract. They are non-negative, and they are determined in such a way that the contract is financially fair, i.e. such that the equivalence principle is satisfied for the total payments under the pricing measure:

$$x_0 = \mathbb{E}^Q \left[\int_0^T e^{-\int_0^s r(v)\,\mathrm{d}v} \,\mathrm{d}\left(B^u + B^f - C \right)(s) \right] . \tag{7.7}$$

7.5.5 Savings and Benefit Prognoses from the Policyholder's Perspective

The projections of X and Y on policy level are useful for retirement savings and benefit prognoses with confidence intervals, conditional on the policy starting and staying in the state 0. Again the projections reflect the financial evolution of the policy given that the policyholder continue his course of life which is exactly what most prognoses from the policyholder's perspective focus on. For retirement savings prognoses, it is relevant to calculate the expected assets at the retirement date R,

$$\mathbb{E}^P_{R,0}\left[X(R)\right] ,$$

given that the policy starts and stays in state 0. Here, $\mathbb{E}^P_{t,0}$ still denotes P-expectation given sojourn in state 0 up to time t. For benefit prognoses, it is relevant to calculate the expected benefit stream

$$t \mapsto \mathbb{E}^P_{t,0}\left[X(t)\right] \left(\mathrm{d}b^p_0(t) + \sum_{j \in \mathcal{J}: j \neq 0} b^p_{0j}\,\mathrm{d}N_{0j}(t) \right) + \mathrm{d}b^f_0(t) + \sum_{j \in \mathcal{J}: j \neq 0} b^f_{0j}\,\mathrm{d}N_{0j}(t) ,$$

If the projections are based on scenarios generated via Monte Carlo simulation, then for each t, the expectation $\mathbb{E}^P_{t,0}\left[X(t)\right]$ is approximated by averaging over a sufficient number of P-projections up to time t. If, instead, the projections are of the worst-case or best-estimate type (and, hence, singular), then $\mathbb{E}^P_{t,0}\left[X(t)\right]$ equals the single projected value. To illustrate the riskiness of the policy, one can calculate quantiles for $X(t)$, based on Monte Carlo-based projections.

7.5.6 Numerical Example

The example below is based on 5000 scenarios generated via Monte Carlo simulation. The number of simulated scenarios is, again, enough to ensure that our numerical results and graphs do not change between simulations.

We consider a unit-linked insurance policy which is identical to the example policy in Jensen and Schomacker (2015). The policy is a unit-linked version of the participating life insurance policy in Section 7.4.8. The state of the policy is described by the classical survival model with two states, 0 (alive) and 1 (dead). For simplicity, we write $\mu^m = \mu^m_{01}$ and $p^m = p^m_{00}$. The policy expires at the retirement date, i.e. $T = R$. The payments of the policy consist of a constant continuous premium payment π while alive, a term insurance sum b^{ad} upon death before expiration T, and a pure endowment sum upon survival until expiration T. The size of the endowment sum is equal to the value of the assets at expiration. There are no payments in the death state.

The policyholder is the 25-year-old female from Section 7.4.8. Her death is still governed by the market mortality in Equation (7.5). The policy expires at time $T = R = 40$ when the policyholder is 65. In the notation from the previous sections, we have

$$\begin{aligned} \mathrm{d}b_0^f(t) &= \mathrm{d}b_1^f(t) = \mathrm{d}c_1(t) = \mathrm{d}b_1^p(t) = \mathrm{d}b_{01}^p(t) = 0 \ , \\ \mathrm{d}c_0(t) &= \pi\,\mathrm{d}s \ , \\ \mathrm{d}b_0^p(t) &= 1_{\{t=40\}} \ , \\ \mathrm{d}b_{01}^f(t) &= b^{ad} \ . \end{aligned}$$

Since there are no fixed benefit payments in either state, we get $V_0^{f,m}(t) = V_1^{f,m}(t) = 0$. Furthermore, the policyholder does not keep her assets upon death, so $P_1(t) = 0$. As a consequence, we have the following risk premium streams

$$\begin{aligned} \mathrm{d}s_0^f(t) &= \mu^m(t)\,b^{ad}\mathrm{d}s \ , \\ \mathrm{d}s_0^u(t) &= -\mu^m(t)\,X(t-)\,\mathrm{d}s \ . \end{aligned}$$

For comparability, we fix the term insurance sum at $b^{ad} = 1$ and the premium at $\pi = 0.04614$ as in Section 7.4.8. We assume a deterministic market interest rate $r = 0.04$, and the assets of the policy are invested in a fund with log-normal returns that are paid out once a year, i.e.

$$\mathrm{d}R_X(t) = \frac{S(t) - S(t-1)}{S(t-1)} \ , \quad t = 1, \ldots, 40 \ ,$$

where S is a geometric Brownian motion. We basically consider a simple Black–Scholes financial market. We assume that the fund size S has drift 0.07 and volatility 0.2. The financial market is identical to the market in the participating life insurance example in Section 7.4.8.

The guarantee account Y bears interest at the rate $r^* = 0$. The guarantee account is upgraded and the guarantee fee paid once a year. The guarantee upgrade u is determined as a fraction θ_1 of the positive part of the assets X less the guarantee account Y, all taken just before the guarantee upgrade, i.e.

$$u(t) = \theta_1 \left(X(t-) - Y(t-)\right)^+ \ , \ t = 1, \ldots, 40 \ .$$

We fix $\theta_1 = 0.8$. In words, the guarantee upgrade makes up 80% of the excess return on the assets. At expiration, the guarantee $g = (Y(40-) - X(40-))^+$ is added to the assets to ensure that they match the guarantee account. The guarantee fee π_g is a fraction θ_2 of the positive part of the returns on the assets, i.e.

$$\pi_g(t) = \theta_2 \left(R_X(t)\,X(t-1)\right)^+ \ .$$

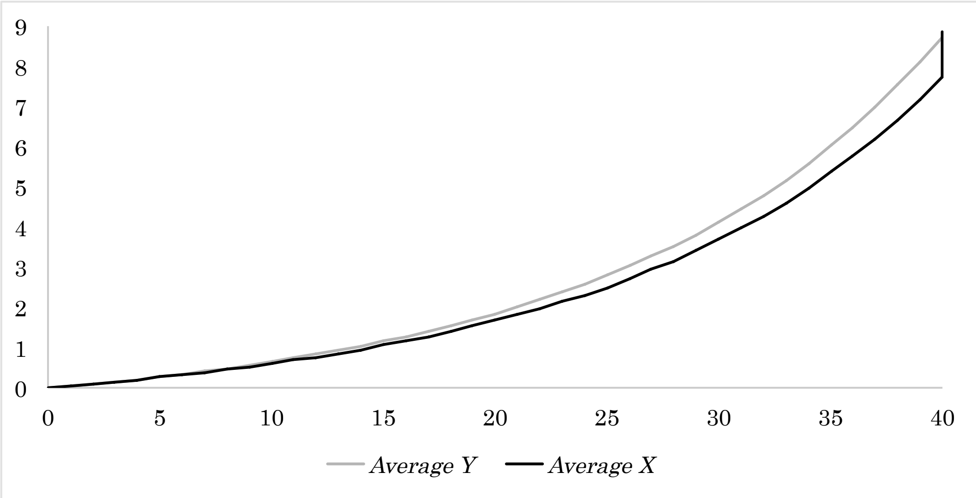

Figure 7.3: Average evolution of the assets X and guarantee account Y as a function of time.

In Jensen and Schomacker (2015), the authors consider an identical policy and market, and the fraction θ_2 is determined according to the fairness criterion in Equation (7.7). The fair guarantee fee fraction amounts to $\theta_2 = 0.1$ which we apply in the following. Using standard Monte Carlo methods, we simulate 5000 sample paths for the asset returns R_X, and for each sample path, we project X and Y using discretized version of the stochastic differential equations in Equation (7.6).

In Figure 7.3, we plot the average evolution of the assets and the guarantee account. The average endowment sum at retirement, $X\,(40-) + g$, amounts to 8.9. The final guarantee injection on average raises the assets by 14%, even though the guarantee account does not bear interest. This is because the final guarantee kicks in for all sample paths where the assets finish below their (previous) maximum value. With regards to prognoses, the average provides the policyholder with a best-estimate of the endowment sum at retirement, given that the policyholder is still alive at retirement. In addition, the policyholder gets a best-estimate of the evolution of the assets which is useful for surrender where the policyholder typically receives the assets, regardless of the value of the guarantee account.

In Figure 7.4, we plot the empirical distribution of the endowment sum at retirement. With regards to prognoses, the figure illustrates the significant financial riskiness of the policy. The policy has a big potential upside, but also a big downside risk. The distribution is skew, and for 73% of the sample paths, the endowment sum at retirement is less than the average of 8.9. This is important to know for the policyholder. In Figure 7.5, we plot quantiles and the worst outcome for the evolution of the assets. The message is the

Figure 7.4: Empirical distribution of endowment sum at retirement, $X(40-) + g$.

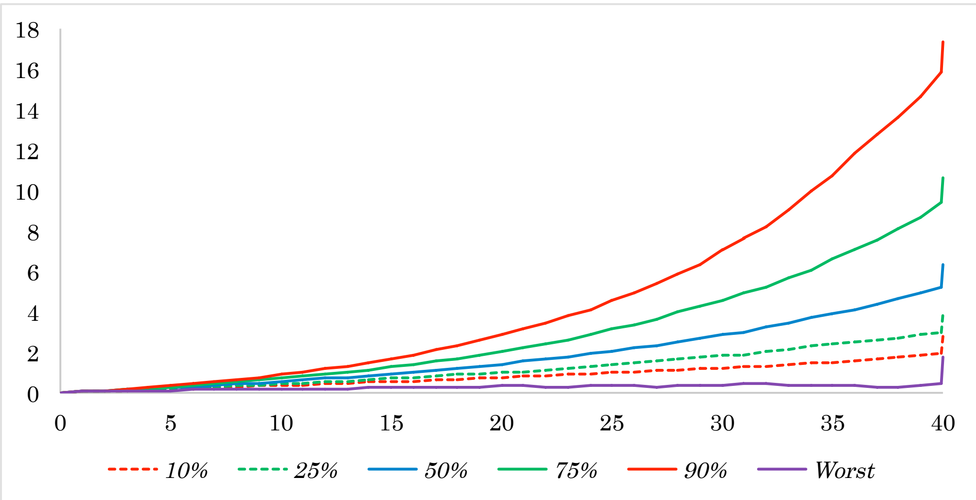

Figure 7.5: Quantiles and worst outcome for the assets X as a function of time.

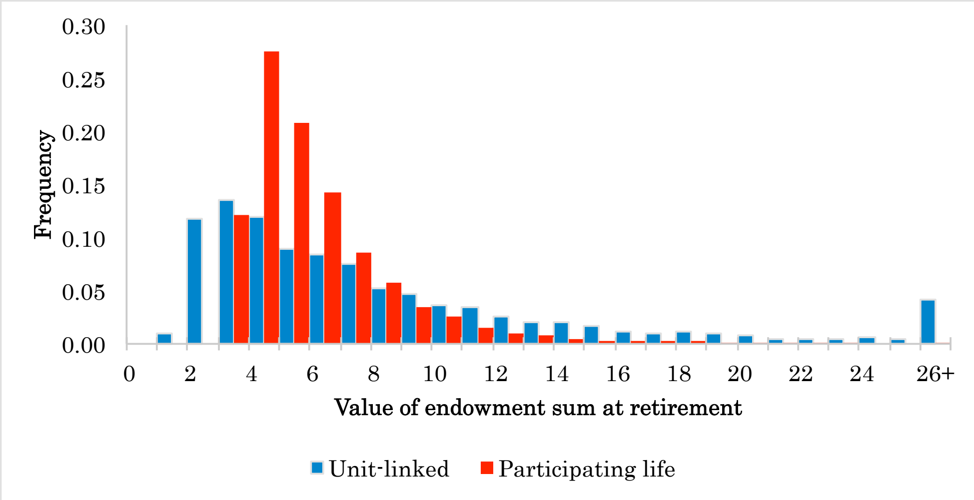

Figure 7.6: Empirical distribution of the endowment sum at retirement for the participating life insurance policy and the unit-linked insurance policy.

same as for Figure 7.4, and Figure 7.5 adds to the picture by illustrating the distributional evolution of the assets, in addition to the value at retirement.

From Figure 7.3 and 7.5, the impact of the final guarantee injection is very clear. This is valuable for the policyholder when assessing the usefulness of an embedded guarantee.

7.5.7 Unit-Linked *versus* Participating Life

To round off, we compare the unit-linked insurance policy with its participating life insurance counterpart from Section 7.4.8. The comparison is straightforward since the two policies are modeled in the same framework. The two policies have the same the premiums and death benefits, and both policies are designed to be financially fair. Hence, under the pricing measure, the expected endowment sum at retirement are the same for the two policies. In Figure 7.6, we plot the empirical distributions of the endowment sum at retirement for the two policies. We notice that the empirical distribution of the endowment sum at retirement differs significantly under the physical measure. In particular, the unit-linked insurance policy has a bigger downside than the participating life insurance policy. This is highly relevant information for the policyholder when choosing between a participating life and unit-linked insurance product. Our model is a valuable tool in providing this information.

7.6 Conclusion

In this paper, we have introduced stochastic scenarios in participating life and unit-linked insurance to utilize tailor-made bonus, benefit, and retirement savings prognoses that illustrate financial riskiness. Each scenario consists of, either, a sample path for the bonus allocation (in participating life insurance), or, a sample path for the short interest rate and a sample path for the return of the fund that the policyholder invests in (in unit-linked insurance). The paper is a self-contained continuation of Jensen and Schomacker (2015), and to the knowledge of the author, it is the first paper to address risk-based prognoses from the policyholder's perspective in participating life and unit-linked insurance in a general financial market. By risk-based prognoses, we mean prognoses that illustrate financial riskiness by going beyond best-estimate expectations about future returns. In a general multi-state Markov model, we model participating life and unit-linked insurance in the same framework which makes comparison easy. For both product types, we have provided numerical examples to demonstrate the possible applications of our model. We have illustrated the use of our model by conducting scenario analysis based on Monte Carlo simulation, but the model applies to scenarios in general and to worst-case and best-estimate scenarios in particular. Our model is a valuable tool in providing tailor-made best-estimate prognoses, illustrating financial riskiness, and visualizing the impact of financial guarantees. All, with moderate mathematical complexity, thanks to our scenario-based approach.

Bibliography

Bank, P. and D. Baum (2004). Hedging and Portfolio Optimization in Financial Markets with a Large Trader. *Mathematical Finance 14*(1), 1–18. doi: 10.1111/j.0960-1627.2004.00179.x.

Basak, S. and G. Chabakauri (2010). Dynamic Mean-Variance Asset Allocation. *Review of Financial Studies 23*(8), 2970–3016. doi: 10.1093/rfs/hhq028.

Bauer, D., D. Bergmann, and R. Kiesel (2010, 5). On the Risk-Neutral Valuation of Life Insurance Contracts with Numerical Methods in View. *ASTIN Bulletin 40*, 65–95. doi: 10.2143/AST.40.1.2049219.

Bauer, D., R. Kiesel, A. Kling, and J. Ruß (2006). Risk-neutral valuation of participating life insurance contracts. *Insurance: Mathematics and Economics 39*(2), 171–183. doi: 10.1016/j.insmatheco.2006.02.003.

Bellman, R. (1957). *Dynamic Programming.* Princeton University Press, Princeton, New Jersey.

Bertsimas, D. and A. W. Lo (1998). Optimal control of execution costs. *Journal of Financial Markets 1*(1), 1–50. doi: 10.1016/S1386-4181(97)00012-8.

Björk, T. (2009). *Arbitrage Theory in Continuous Time* (3rd ed.). Oxford University Press.

Björk, T. and A. Murgoci (2010). A General Theory of Markovian Time Inconsistent Stochastic Control Problems. *Working paper, September 17, 2010.* `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1694759`.

Björk, T., A. Murgoci, and X. Y. Zhou (2014). Mean–Variance Portfolio Optimization with State-Dependent Risk Aversion. *Mathematical Finance 24*(1), 1–24. doi: 10.1111/j.1467-9965.2011.00515.x.

Bohnert, A. (2015). The Impact of Guarantees on the Performance of Pension Saving Schemes: Insights from the Literature. *Risks 3*(4), 515–542. doi: 10.3390/risks3040515.

Bohnert, A. and N. Gatzert (2012). Analyzing surplus appropriation schemes in participating life insurance from the insurer's and the policyholder's perspective. *Insurance: Mathematics and Economics 50*(1), 64–78. doi: 10.1016/j.insmatheco.2011.09.004.

Bruhn, K. and M. Steffensen (2013). Optimal smooth consumption and annuity design. *Journal of Banking & Finance 37*(8), 2693–2701. doi: 10.1016/j.jbankfin.2013.03.015.

Bullard, J. and J. Feigenbaum (2007). A leisurely reading of the life-cycle consumption data. *Journal of Monetary Economics 54*(8), 2305–2320. doi: 10.1016/j.jmoneco.2007.06.025.

Chacko, G. and L. M. Viceira (2005). Dynamic Consumption and Portfolio Choice with Stochastic Volatility in Incomplete Markets. *The Review of Financial Studies 18*(4), 1369–1402. doi: 10.1093/rfs/hhi035.

Chellathurai, T. and T. Draviam (2007). Dynamic portfolio selection with fixed and/or proportional transaction costs using non-singular stochastic optimal control theory. *Journal of Economic Dynamics and Control 31*(7), 2168–2195. doi: 10.1016/j.jedc.2006.06.006.

Chen, Z. and L. Epstein (2002). Ambiguity, Risk, and Asset Returns in Continuous Time. *Econometrica 70*(4), 1403–1443. doi: 10.1111/1468-0262.00337.

Christiansen, M. C., L. F. B. Henriksen, K. J. Schomacker, and M. Steffensen (2014). Stress scenario generation for solvency and risk management. *Scandinavian Actuarial Journal*. doi: 10.1080/03461238.2014.971860.

Czichowsky, C. (2013). Time-consistent mean-variance portfolio selection in discrete and continuous time. *Finance and Stochastics 17*(2), 227–271. doi: 10.1007/s00780-012-0189-9.

Davis, M. H. A. and A. R. Norman (1990). Portfolio Selection with Transaction Costs. *Mathematics of Operations Research 15*(4), 676–713. doi: 10.1287/moor.15.4.676.

Duffie, D. and L. G. Epstein (1992a). Asset Pricing with Stochastic Differential Utility. *Review of Financial Studies 5*(3), 411–436. doi: 10.1093/rfs/5.3.411.

Duffie, D. and L. G. Epstein (1992b). Stochastic Differential Utility. *Econometrica 60*(2), 353–394. doi: 10.2307/2951600.

Ekeland, I. and T. A. Pirvu (2008). Investment and consumption without commitment. *Mathematics and Financial Economics 2*(1), 57–86. doi: 10.1007/s11579-008-0014-6.

Epstein, L. G. and S. E. Zin (1989). Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica 57*(4), 937–969. doi: 10.2307/1913778.

Epstein, L. G. and S. E. Zin (1991). Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis. *Journal of Political Economy 99*(2), 263–286. doi: 10.1086/261750.

Feigenbaum, J. (2008). Can mortality risk explain the consumption hump? *Journal of Macroeconomics 30*(3), 844–872. doi: 10.1016/j.jmacro.2007.08.010.

Gatzert, N. and A. Kling (2007). Analysis of Participating Life Insurance Contracts: A Unification Approach. *Journal of Risk and Insurance 74*(3), 547–570. doi: 10.1111/j.1539-6975.2007.00224.x.

Glasserman, P. (2004). *Monte Carlo Methods in Financial Engineering.* Springer-Verlag.

Goldman, S. M. (1980). Consistent Plans. *The Review of Economic Studies 47*(3), 533–537. doi: 10.2307/2297304.

Gourinchas, P.-O. and J. A. Parker (2002). Consumption over the Life Cycle. *Econometrica 70*(1), 47–89. doi: 10.1111/1468-0262.00269.

Graf, S., A. Kling, and J. Ruß (2011). Risk analysis and valuation of life insurance contracts: Combining actuarial and financial approaches. *Insurance: Mathematics and Economics 49*(1), 115–125. doi: 10.1016/j.insmatheco.2011.02.008.

Grosen, A. and P. L. Jørgensen (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. doi: 10.1016/S0167-6687(99)00041-4.

Guillén, M., P. L. Jørgensen, and J. P. Nielsen (2006). Return smoothing mechanisms in life and pension insurance: Path-dependent contingent claims. *Insurance: Mathematics and Economics 38*(2), 229–252. doi: 10.1016/j.insmatheco.2005.06.014.

Guillén, M., A. K. Konicz, J. P. Nielsen, and A. M. Pérez-Marín (2013a). Do not pay for a danish interest guarantee. the law of the triple blow. *Annals of Actuarial Science 7*, 192–209. doi: 10.1017/S1748499512000176.

Guillén, M., J. P. Nielsen, A. M. Pérez-Marín, and K. S. Petersen (2013b). Performance measurement of pension strategies: a case study of danish life-cycle products. *Scandinavian Actuarial Journal 2013*(1), 49–68. doi: 10.1080/03461238.2010.546138.

Gârleanu, N. and L. H. Pedersen (2013). Dynamic Trading with Predictable Returns and Transaction Costs. *The Journal of Finance 68*(6), 2309–2340. doi: 10.1111/jofi.12080.

Gârleanu, N. and L. H. Pedersen (2014). Dynamic Portfolio Choice with Frictions. *Working paper, March, 2014*. `http://docs.lhpedersen.com/DynamicPortfolioChoiceWithFrictions.pdf`.

Hall, R. E. (1988). Intertemporal Substitution in Consumption. *Journal of Political Economy 96*(2), 339–357. doi: 10.1086/261539.

Hansen, M. and K. R. Miltersen (2002). Minimum Rate of Return Guarantees: The Danish Case. *Scandinavian Actuarial Journal 2002*(4), 280–318. doi: 10.1080/03461230110106282.

Hoem, J. M. (1969). Markov Chain Models in Life Insurance. *Blätter der DGVFM 9*(2), 91–107. doi: 10.1007/BF02810082.

Insurance Regulation Committee of the International Actuarial Association (2013). Stress testing and scenario analysis. *Committee paper*. `http://www.actuaries.org/CTTEES_SOLV/Documents/StressTestingPaper.pdf`.

Jensen, B., P. L. Jørgensen, and A. Grosen (2001). A Finite Difference Approach to the Valuation of Path Dependent Life Insurance Liabilities. *The GENEVA Papers on Risk and Insurance Theory 26*(1), 57–84. doi: 10.1023/A:1011264408187.

Jensen, N. R. and K. J. Schomacker (2015). A Two-Account Life Insurance Model for Scenario-Based Valuation Including Event Risk. *Risks 3*(2), 183–218. doi: 10.3390/risks3020183.

Jensen, N. R. and M. Steffensen (2015). Personal finance and life insurance under separation of risk aversion and elasticity of substitution. *Insurance: Mathematics and Economics 62*, 28–41. doi: 10.1016/j.insmatheco.2015.02.006.

Jørgensen, P. L. and P. Linnemann (2012). A comparison of three different pension savings products with special emphasis on the payout phase. *Annals of Actuarial Science 6*, 137–152. doi: 10.1017/S1748499511000352.

Kihlstrom, R. (2009). Risk aversion and the elasticity of substitution in general dynamic portfolio theory: Consistent planning by forward looking, expected utility maximizing investors. *Journal of Mathematical Economics 45*(9–10), 634–663. doi: 10.1016/j.jmateco.2008.08.008.

Kling, A., A. Richter, and J. Ruß (2007). The Impact of Surplus Distribution on the Risk Exposure of With Profit Life Insurance Policies Including Interest Rate Guarantees. *Journal of Risk and Insurance 74*(3), 571–589. doi: 10.1111/j.1539-6975.2007.00225.x.

Korn, R. and H. Kraft (2002). A Stochastic Control Approach to Portfolio Problems with Stochastic Interest Rates. *SIAM Journal on Control and Optimization 40*(4), 1250–1269. doi: 10.1137/S0363012900377791.

Kraft, H., F. Seifried, and M. Steffensen (2013). Consumption-portfolio optimization with recursive utility in incomplete markets. *Finance and Stochastics 17*, 161–176. doi: 10.1007/s00780-012-0184-1.

Kraft, H. and F. T. Seifried (2010). Foundations of Continuous-Time Recursive Utility: Differentiability and Normalization of Certainty Equivalents. *Mathematics and Financial Economics 3*, 115–138. doi: 10.1007/s11579-010-0030-1.

Kraft, H. and F. T. Seifried (2014). Stochastic differential utility as the continuous-time limit of recursive utility. *Journal of Economic Theory 151*, 528–550. doi: 10.1016/j.jet.2013.12.007.

Kraft, H. and M. Steffensen (2008). Optimal Consumption and Insurance: a Continuous-Time Markov Chain Approach. *ASTIN Bulletin 38*(1), 231–257. doi: 10.2143/AST.38.1.2030412.

Kreps, D. M. and E. L. Porteus (1978). Temporal Resolution of Uncertainty and Dynamic Choice Theory. *Econometrica 46*(1), 185–200. doi: 10.2307/1913656.

Kreps, D. M. and E. L. Porteus (1979). Dynamic Choice Theory and Dynamic Programming. *Econometrica 47*(1), 91–100. doi: 10.2307/1912348.

Kronborg, M. T. and M. Steffensen (2015). Inconsistent Investment and Consumption Problems. *Applied Mathematics & Optimization 71*(3), 473–515. doi: 10.1007/s00245-014-9267-z.

Laibson, D. (1997). Golden Eggs and Hyperbolic Discounting. *Quarterly Journal of Economics 112*(2), 443–478. doi: 10.1162/003355397555253.

Liu, C. and H. Zheng (2016). Asymptotic analysis for target asset portfolio allocation with small transaction costs. *Insurance: Mathematics and Economics 66*, 59–68. doi: 10.1016/j.insmatheco.2015.10.014.

Liu, J. (2007). Portfolio Selection in Stochastic Environments. *Review of Financial Studies 20*(1), 1–39. doi: 10.1093/rfs/hhl001.

Longstaff, F. A. (2001). Optimal Portfolio Choice and the Valuation of Illiquid Securities. *The Review of Financial Studies 14*(2), 407–431. doi: 10.1093/rfs/14.2.407.

Magill, M. J. P. and G. M. Constantinides (1976). Portfolio selection with transactions costs. *Journal of Economic Theory 13*(2), 245–263. doi: 10.1016/0022-0531(76)90018-1.

Mehra, R. and E. C. Prescott (1985). The equity premium: A puzzle. *Journal of Monetary Economics 15*(2), 145–161. doi: 10.1016/0304-3932(85)90061-3.

Merton, R. C. (1971). Optimum Consumption and Portfolio Rules in a Continuous-Time Model. *Journal of Economic Theory 3*(4), 373–413. doi: 10.1016/0022-0531(71)90038-X.

Miltersen, K. R. and S.-A. Persson (2003). Guaranteed Investment Contracts: Distributed and Undistributed Excess Return. *Scandinavian Actuarial Journal 2003*(4), 257–279. doi: 10.1080/03461230110106417.

Moazeni, S., T. F. Coleman, and Y. Li (2010). Optimal Portfolio Execution Strategies and Sensitivity to Price Impact Parameters. *SIAM Journal on Optimization 20*(3), 1620–1654. doi: 10.1137/080715901.

Musiela, M. and T. Zariphopoulou (2010). Stochastic Partial Differential Equations and Portfolio Choice. In C. Chiarella and A. Novikov (Eds.), *Contemporary Quantitative Finance*, pp. 195–216. Springer-Verlag Berlin Heidelberg. doi: 10.1007/978-3-642-03479-4_11.

Møller, T. and M. Steffensen (2007). *Market-Valuation Methods in Life and Pension Insurance.* Cambridge University Press.

Norberg, R. (1991). Reserves in Life and Pension Insurance. *Scandinavian Actuarial Journal 1991*(1), 3–24. doi: 10.1080/03461238.1991.10557357.

Norberg, R. (1999). A theory of bonus in life insurance. *Finance and Stochastics 3*(4). doi: 10.1007/s007800050067.

Norberg, R. (2001). On Bonus and Bonus Prognoses in Life Insurance. *Scandinavian Actuarial Journal 2001*(2), 126–147. doi: 10.1080/03461230152592773.

Oksendal, B. and A. Sulem (2002). Optimal Consumption and Portfolio with Both Fixed and Proportional Transaction Costs. *SIAM Journal on Control and Optimization 40*(6), 1765–1790. doi: 10.1137/S0363012900376013.

Pliska, S. R. and K. Suzuki (2004). Optimal tracking for asset allocation with fixed and proportional transaction costs. *Quantitative Finance 4*(2), 233–243. doi: 10.1080/14697680400000027.

Pollak, R. A. (1968). Consistent Planning. *Review of Economic Studies 35*(2), 201–208. doi: 10.2307/2296548.

Protter, P. E. (2005). *Stochastic Integration and Differential Equations* (2nd ed.), Volume 21 of *Stochastic Modelling and Applied Probability.* Springer-Verlag.

Richard, S. F. (1975). Optimal consumption, portfolio and life insurance rules for an uncertain lived individual in a continuous time model. *Journal of Financial Economics 2*, 187–203. doi: 10.1016/0304-405X(75)90004-5.

Schroder, M. and C. Skiadas (1999). Optimal Consumption and Portfolio Selection with Stochastic Differential Utility. *Journal of Economic Theory 89*(1), 68–126. doi: 10.1006/jeth.1999.2558.

Schroder, M. and C. Skiadas (2005). Lifetime consumption-portfolio choice under trading constraints, recursive preferences, and nontradeable income. *Stochastic Processes and their Applications 115*(1), 1–30. doi: 10.1016/j.spa.2004.08.001.

Silvestrov, D. and A. Martin-Löf (Eds.) (2014). *Modern Problems in Insurance Mathematics*. EAA Series. Springer International Publishing.

Skiadas, C. (1998). Recursive utility and preferences for information. *Economic Theory 12*(2), 293–312. doi: 10.1007/s001990050222.

Skiadas, C. (2013). Smooth Ambiguity Aversion toward Small Risks and Continuous-Time Recursive Utility. *Journal of Political Economy 121*(4), 775–792. doi: 10.1086/671179.

Solvency II Directive (2009). DIRECTIVE 2009/138/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 25 November 2009 on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II). `http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32009L0138&from=EN`.

Soner, H. M. and M. Vukelja (2013). Utility maximization in an illiquid market. *Stochastics 85*(4), 692–706. doi: 10.1080/17442508.2013.795568.

Steffensen, M. (2006). Surplus-Linked Life Insurance. *Scandinavian Actuarial Journal 2006*(1), 1–22. doi: 10.1080/03461230500199533.

Steffensen, M. and S. Waldstrøm (2009). A two-account model of pension saving contracts. *Scandinavian Actuarial Journal 2009*(3), 169–186. doi: 10.1080/03461230802090406.

Stroock, D. (1983). Some applications of stochastic calculus to partial differential equations. In *Ecole d'Eté de Probabilités de Saint-Flour XI—1981*, Volume 976 of *Lecture Notes in Mathematics*, pp. 267–382. Berlin: Springer-Verlag.

Strotz, R. H. (1955). Myopia and Inconsistency in Dynamic Utility Maximization. *Review of Economic Studies 23*(3), 165–180. doi: 10.2307/2295722.

Wachther, J. A. (2002). Portfolio and Consumption Decisions under Mean-Reverting Returns: An Exact Solution for Complete Markets. *Journal of Financial and Quantitative Analysis 37*(1), 63–91. doi: 10.2307/3594995.

Yaari, M. E. (1965). Uncertain Lifetime, Life Insurance, and the Theory of the Consumer. *The Review of Economic Studies 32*(2), 137–150. doi: 10.2307/2296058.

Zaglauer, K. and D. Bauer (2008). Risk-neutral valuation of participating life insurance contracts in a stochastic interest rate environment. *Insurance: Mathematics and Economics 43*(1), 29–40. doi: 10.1016/j.insmatheco.2007.09.003.

Zhou, X. Y. and D. Li (2000). Continuous-Time Mean-Variance Portfolio Selection: A Stochastic LQ Framework. *Applied Mathematics and Optimization 41*(1), 19–33. doi: 10.1007/s002450010003.

Appendix A

Post-Submission Changes to the Thesis

This appendix has not been assessed by the Assessment Committee.

A.1 Post-Submission Changes to Chapter 4

A.1.1 Proof of Theorem 4.1

The follwing proof of Theorem 4.1 in Chapter 4 is almost identical to the proof of Theorem 3.1 in Kraft et al. (2013).

Let $(c, \pi, d) \in \mathcal{A}(x_0)$ be an arbitrary admissible control. Applying Itô's formula, we obtain

$$J\left(t, X_t^{c,\pi,d}\right) - U\left(X_T^{c,\pi,d}\right) = -\int_t^T L_s^{c,\pi,d}[J]\left(s, X_s^{c,\pi,d}\right) \mathrm{d}s - M^T + M_t \ ,$$

where M is a martingale thanks to the assumption that $\int_0^\cdot J_x\left(t, X_t^{c,\pi,d}\right)\pi_t\sigma \,\mathrm{d}W_t$ is a martingale, and where

$$\begin{aligned} L_s^{c,\pi,d}[J](t,x) = {} & J_t(t,x) + (rx + \pi_s\lambda - c_s - \hat{\mu}(s)\, d_s + w(s))\, J_x(t,x) \\ & + \frac{1}{2}\pi_s^2\sigma^2 J_{xx}(t,x) \ . \end{aligned}$$

Taking conditional expectation given $X_t^{c,\pi,d}$ and subtracting

$$V_t^{c,\pi,d} = \mathbb{E}_t\left[\int_t^T f\left(s, c_s, d_s + X_s^{c,\pi,d}, V_s^{c,\pi,d}\right) \mathrm{d}s + u\left(X_T^{c,\pi,d}\right)\right] \ ,$$

we get

$$\begin{aligned} & J\left(t, X_t^{c,\pi,d}\right) - V_t^{c,\pi,d} \\ & = -\mathbb{E}_t\left[\int_t^T \left\{L_s^{c,\pi,d}[J]\left(s, X_s^{c,\pi,d}\right) + f\left(s, c_s, d_s + X_s^{c,\pi,d}, V_s^{c,\pi,d}\right)\right\} \mathrm{d}s\right] \ . \end{aligned}$$

The Hamilton-Jacobi-Bellman equation in Equation (4.9) implies that

$$\begin{aligned}
&L_s^{c,\pi,d}[J]\left(s, X_s^{c,\pi,d}\right) + f\left(s, c_s, d_s + X_s^{c,\pi,d}, V_s^{c,\pi,d}\right) \\
&= L_s^{c,\pi,d}[J]\left(s, X_s^{c,\pi,d}\right) + f\left(s, c_s, d_s + X_s^{c,\pi,d}, J\left(t, X_s^{c,\pi,d}\right)\right) \\
&\quad + f\left(s, c_s, d_s + X_s^{c,\pi,d}, V_s^{c,\pi,d}\right) - f\left(s, c_s, d_s + X_s^{c,\pi,d}, J\left(t, X_s^{c,\pi,d}\right)\right) \\
&\geq f\left(s, c_s, d_s + X_s^{c,\pi,d}, V_s^{c,\pi,d}\right) - f\left(s, c_s, d_s + X_s^{c,\pi,d}, J\left(t, X_s^{c,\pi,d}\right)\right) ,
\end{aligned}$$

and, hence, the regularity condition in Equation (4.8) ensures that

$$\begin{aligned}
&L_s^{c,\pi,d}[J]\left(s, X_s^{c,\pi,d}\right) + f\left(s, c_s, d_s + X_s^{c,\pi,d}, V_s^{c,\pi,d}\right) \\
&\leq k\left(V_s^{c,\pi,d} - J\left(s, X_s^{c,\pi,d}\right)\right) \quad \text{on } \left\{V_s^{c,\pi,d} \geq J\left(s, X_s^{c,\pi,d}\right)\right\} .
\end{aligned}$$

Altogether, the process $\{Y_t\}_{t\in[0,T]} = \left\{J\left(t, X_t^{c,\pi,d}\right) - V_t^{c,\pi,d}\right\}_{t\in[0,T]}$ can be written as

$$Y_t = \mathbb{E}_t\left[\int_t^T H_s \,\mathrm{d}s\right] \text{ with } H_t \geq kY_t \text{ on } \{Y_t \leq 0\} .$$

Now, applying Theorem A.2 in Kraft et al. (2013), we get that $J\left(t, X_t^{c,\pi,d}\right) - V_t^{c,\pi,d} \geq 0$ for all $t \in [0,T]$. In particular, we have $J(0,x_0) = J\left(0, X_0^{c,\pi,d}\right) \geq V_0^{c,\pi,d}$. Since $(c,\pi,d) \in \mathcal{A}(x_0)$ is arbitrary, we obtain

$$J(0,x_0) \geq \max_{(c,\pi,d)\in\mathcal{A}(x_0)} V_0^{c,\pi,d} .$$

Conversely, under the assumptions of Theorem 4.1, there exists a control (c^*,π^*,d^*) such that

$$\begin{aligned}
&L_s^{c^*,\pi^*,d^*}[J]\left(s, X_s^{c^*,\pi^*,d^*}\right) + f\left(s, c_s^*, d_s + X_s^{c^*,\pi^*,dv}, V_s^{c^*,\pi^*,d^*}\right) \\
&= f\left(s, c_s^*, d_s + X_s^{c^*,\pi^*,d^*}, V_s^{cv,\pi^*,d^*}\right) - f\left(s, c_s^*, d_s^* + X_s^{c^*,\pi^*,d^*}, J\left(t, X_s^{c^*,\pi^*,d^*}\right)\right) .
\end{aligned}$$

Hence, the previous argument applies to both processes

$$\left\{J\left(t, X_t^{c^*,\pi^*,d^*}\right) - V_t^{c^*,\pi^*,d^*}\right\}_{t\in[0,T]} \text{ and } \left\{V_t^{c^*,\pi^*,d^*} - J\left(t, X_t^{c^*,\pi^*,d^*}\right)\right\}_{t\in[0,T]} .$$

As a consequence, we obtain $J(0,x_0) = V_0^{c^*,\pi^*,d^*}$. All in all, J is the value function of the problem in Equation (4.7), and (c^*,π^*,d^*) is the optimal control.

A.2 Post-Submission Changes to Chapter 7

A.2.1 Expansion of Paragraph in the Introduction

The 10-line paragraph starting with

> "Prognoses from the policyholder's perspective are widely used..."

from the introduction to Chapter 7 is expanded as follows:

Prognoses from the policyholder's perspective are widely used in the life insurance and pension industry, but the topic is hardly covered in the existing literature. In Norberg (2001), the author treats prognoses in participating life insurance, but the model is only tractable for a very simple financial environment and does not apply to unit-linked insurance. To the knowledge of the author, this present paper is one of the first papers to address risk-based prognoses from the policyholder's prospective in participating life and unit-linked insurance in a general financial market and with general insurance risk. By risk-based prognoses, we mean prognoses that illustrate financial riskiness by going beyond best-estimate expectations about future returns. Some life insurance and pension companies illustrate financial riskiness by supplying the policyholder with a prognosis based on a couple of deterministic scenarios, e.g. a low, medium, and high interest rate scenario (in participating life insurance) or a low, medium, and high stock return scenario (in unit-linked insurance). A related strand of literature focuses on the analysis of policyholder payoff distributions under the physical measure. Bohnert (2015) provides a thorough overview of the literature on the performance of pension savings schemes from the policyholder's perspective. For the case of Denmark, Bohnert (2015) mentions the papers by Guillén et al. (2013a), Guillén et al. (2013b), and Jørgensen and Linnemann (2012). Guillén et al. (2013a) investigate the performance of participating life insurance schemes for policies containing a guaranteed minimum rate of return whereas Guillén et al. (2013b) investigate the performance of life-cycle products. Jørgensen and Linnemann (2012) compare three different pension savings products: a traditional participating life insurance scheme, a market-based unit-linked insurance scheme, and a formula based smoothed investment-linked annuity scheme. Compared to the performance literature, this present paper focuses more on the calculation engine behind prognoses and less on the outcome of prognoses in terms of performance measurement. The paper adds to the existing literature and to the industry practice by formally describing how the policyholder's account evolve and by formalizing how the bonus schemes "consolidation" and "additional benefits" work and interact for a single participating life insurance policy and how guarantees are reflected in the payments of a single unit-linked insurance policy. Furthermore, this paper includes insurance risk which is typically not considered in the performance literature.

A.2.2 Revision of How to Derive Bonus Sample Paths

The revised recapitulation of how to derive bonus sample paths in Subsection 7.4.8 of Chapter 7 replaces the paragraph starting with

"1. We assume a deterministic short interest rate $r = 0.04$."

and ending with

"The fair guarantee fee fraction amounts to $\theta_3 = 0.31$."

The revision is as follows:

1. In Jensen and Schomacker (2015), we project on portfolio level by adding and subtracting expected premiums and benefits. Hence, all the quantities in this recapitulation constitute market expectations across states of the policy ("portfolio averages") rather than state-wise quantities. We denote by $\tilde{X}$ the assets of the policy, including its share of the collective bonus potential, and by $\tilde{V}$ and $\tilde{V}^*$ the market and technical reserve of the policy (projected on portfolio level).

2. We assume a deterministic short interest rate $r = 0.04$. The assets of the policy are invested in a simple Black-Scholes stock with drift 0.07 and volatility 0.2 under the physical measure P. The excess return gives rise to a collective bonus potential K which is defined as the maximum of zero and assets less guaranteed liabilities, i.e.

$$K(t) = \left(\tilde{X}(t) - L(t)\right)^+ \ .$$

   Here, the guaranteed liabilities L is the maximum of the market reserve $\tilde{V}$ and the technical reserve $\tilde{V}^*$ for the guaranteed payments, i.e.

$$L(t) = \max\left\{\tilde{V}(t), \tilde{V}^*(t)\right\} \ .$$

   For details, see Jensen and Schomacker (2015).

3. The assets and reserves $\tilde{X}$, $\tilde{V}^*$, and $\tilde{V}$ are simulated using standard Monte Carlo methods. For each sample path, the interest rate bonus d is determined as a fraction θ_1 of the excess collective bonus potential K over a threshold $\bar{K}$, i.e.

$$d(t) = \theta_1 \left(K(t-) - \bar{K}(t-)\right)^+ \ .$$

   The threshold $\bar{K}$ can be seen as a preferred minimum collective bonus potential and is given by

$$\bar{K}(t) = \theta_2 L(t) \ .$$

4. We fix $\theta_1 = 0.2$ and $\theta_2 = 0.1$. In exchange for the right to bonus, the policyholder pays a guarantee fee which is subtracted from the assets. The guarantee fee is a fraction θ_3 of the positive part of the returns on the assets, and it is determined such that the policy is financially fair. By financially fair, we mean that the equivalence principle is satisfied for the total payments under the pricing measure, i.e.

$$\mathbb{E}^Q\left[\int_0^T e^{-rs}\,\mathrm{d}\left(B^{(\varepsilon(s))} - C\right)(s)\right] = x_0 \ .$$

   The fair guarantee fee fraction amounts to $\theta_3 = 0.31$.

5. The interest rate bonus d is converted to an excess interest rate δ via the formula

$$\delta\,(t) = \frac{d\,(t)}{\tilde{V}^*\,(t-1)} \ .$$