

Aspects of Valuation and Optimization in Life Insurance

Lars Frederik Brandt Henriksen

PhD Thesis

Supervisor: Mogens Steffensen, University of Copenhagen
Co-supervisor: Jepser Lund Pedersen, University of Copenhagen
Submitted: September 30, 2014.

DEPARTMENT OF MATHEMATICAL SCIENCES
FACULTY OF SCIENCE
UNIVERSITY OF COPENHAGEN

Author: Lars Frederik Brandt Henriksen
Lyngbyvej 32D, 4. TV
2100 København Ø
Denmark
lars.brandt.henriksen@gmail.com

Assessment Committee: Professor Stefan Weber,
University of Hannover,
Hannover, Germany

Professor Griselda Deelstra,
Université libre de Bruxelles and Vrije Universiteit Brussel,
Brussels, Belgium

Professor Jostein Paulsen,
University of Copenhagen,
Denmark (Chairman)

ISBN: 978-87-7078-965-3

Contents

Preface	iii
Summary	v
Sammenfatning	vii
List of papers	ix
1 Introduction	2
2 Local risk-minimization with longevity bonds	10
3 Markov chain modeling of policyholder behavior in life insurance and pension	38
4 Stress scenario generation for solvency and risk management	62
5 Affine processes and Markov chains: Interest Rate-Dependent Transition Rates in Life Insurance	88
6 Optimal surplus distribution problems for regulated funds with assets and liabilities	112
Bibliography	159

Preface

This thesis has been prepared in fulfillment of the requirements for the PhD degree at the Department of Mathematical Sciences, Faculty of Science, University of Copenhagen, Denmark. The work has been carried out under the supervision of Professor Mogens Steffensen and Jesper Lund Pedersen, University of Copenhagen, in the period from October 1, 2011 to September 30, 2014.

The main body of the thesis consists of an introduction to the material in the thesis, and five chapters on different but related topics. The five chapters are written as individual academic papers, and are thus self-contained, and can be read independently.

Acknowledgments

At first, I would like to thank Mogens Steffensen for setting up and supporting the entire project and many discussions and advises during the process. Second, I would like to thank my co-supervisor Jesper Lund Pedersen and my co-authors, Thomas Møller, Jeppe Woetmann, Mogens Steffensen, Christian Svensson, Marcus Christiansen, Kristian Schomacker, Benjamin Avanzi, and Bernard Wong, for the collaboration and many good times.

A special thanks to Benjamin Avanzi and Bernard Wong for facilitating, hosting and supervising me through my half year visit at University of New South Wales and to Benjamin Avanzi for inviting me to University of Montreal for two shorter stays.

Finally, heartfelt thanks to my beloved fiancée Randi for support and love.

*Lars Frederik Brandt Henriksen
Østerbro, September 2014*

Summary

This thesis consists, in excess to an introductory chapter, of five papers within the broad area of life insurance mathematics. The last of the papers can also be considered as within the area of non-life insurance mathematics. There is no unifying topic, but many recurrent subtopics.

In the first paper, we consider a market where the interest rate and mortality intensity are modelled as affine processes. The market consists of zero coupon bonds and longevity bonds and we do not assume that the price processes are martingales under the real measure. In this setup we find an optimal hedging strategy for a portfolio of general life insurance contracts by using the criterion of *local* risk-minimization.

The second paper studies expected policyholder behavior in a multistate Markov chain model with deterministic intensities. Valuation techniques in the cases where policyholder behavior is modelled to occur independently or dependently of insurance risk, respectively, are discussed and studied numerically. The impact (quantitatively and computationally) of different simplifying assumptions are investigated for representative insurance contracts.

In the third paper, we derive worst-case scenarios and reserves in a life insurance model in the case where the interest rate and the various transition intensities are mutually dependent. The calculations are based on deterministic optimal control theory. The results of a single insurance contract are extended to *inhomogeneous* portfolios of insurance contracts and various numerical studies are presented. These studies qualify the standard formula of Solvency II.

In the fourth paper, we study the class of affine processes. We obtain transform results which can be used for valuation of life insurance contracts modelled within general, hierarchical, multistate Markov chains. The affine setup makes the calculations computationally tractable because we only need to solve systems of ordinary differential equations and not partial differential equations. The setup allows for mutual dependence between interest rate and transition intensities which makes it possible to e.g. model interest rate dependent surrender rates.

Finally, the fifth paper obtains optimal surplus distribution strategies in a model with infinite time horizon where assets and liabilities are modelled by correlated, geometric Brownian motions. The controls considered are, that we either increase liabilities or decrease assets. The increase of liabilities could be used in the modelling of non-for-profit mutual funds or pension funds. On the other hand, the decrease in assets could be used for modelling of for-profit companies. We impose a simple solvency constraint and prove optimality of barrier strategies, where the barrier is defined in terms of the funding ratio. We also study barrier strategies within a model with a more advanced solvency constraint, where the allowance of controlling either liabilities or assets, when the funding ratio is between a lower and upper barrier, depends on which of the two barriers that have been crossed last. We also consider barrier strategies under the assumption that ruin must be prevented which is done by either decreasing the liabilities or by capital injections. All the results are illustrated numerically.

Sammenfatning

Denne afhandling består, i tillæg til et introduktionsafsnit, af fem artikler indenfor the brede felt af livsforsikringsmatematik. Den sidste artikel kan også anses som værende indenfor skadesforsikringsmatematik. De fem artikler har ikke noget overordnet emne, men der er mange underemner, der går igen i artiklerne.

I den første artikel betragter vi et marked, hvor renten og dødsintensiteten er modelleret som affine processer. Markedet består af nulkuponobligationer og longevity obligationer, og vi antager, at prisprocesserne ikke er martingaler under det sande mål. I dette setup finder vi den optimale afdækningsstrategi for en portefølje bestående af generelle livsforsikringskontrakter ved kriteriet lokal risikominimering.

Den anden artikel undersøger forventet policetageradfærd i en flertilstands-Markovkæde model med deterministiske intensiteter. Beregningsteknikker, for de tilfælde hvor policetagerens adfærd er modelleret som afhængig eller uafhængig af forsikringsrisikoen, diskuteres og studeres numerisk. Effekten (kvantitativt og beregningsmæssigt) af forskellige simplificerende antagelser undersøges for repræsentative forsikringskontrakter.

I den tredje artikel udleder vi worst-case scenarier og reserver i en livsforsikringsmodel i det tilfælde, hvor renten og de forskellige overgangssandsynligheder er indbyrdes afhængige. Beregningerne er baseret på deterministisk optimal kontrolteori. Resultaterne for en enkelt forsikringskontrakt udvides til inhomogene porteføljer af forsikringskontrakter, og diverse numeriske studier præsenteres. Disse studier kvalificerer Standardformelen fra Solvens II.

I den fjerde artikel betragter vi klassen af affine processer. Vi finder transformationsresultater, som kan bruges til beregning af livsforsikringskontrakter modelleret indenfor generelle, hierarkiske Markovkæder med flere tilstande. Det affine setup gør udregningerne beregningsmæssigt medgørlige, idet vi kun skal løse systemer af ordinære differentiaalligninger og ikke partielle differentiaalligninger. Setuppet tillader indbyrdes afhængighed mellem renten og overgangssandsynlighederne, som gør det muligt f.eks. at modellere renteafhængige genkøbsrater.

Endeligt finder vi i den femte artikel optimale strategier for udlodning af overskud i en model med uendelig tidshorisont, hvor aktiver og passiver er modelleret ved hjælp af korrelerede, geometriske Brownske bevægelser. Kontrollerne, der betragtes, er, at vi enten øger passiverne eller mindsker aktiverne. Forøgelsen af passiverne kan bruges i modelleringen af non-profit mutual funds eller pensionsfonde. På den anden side kan mindskelsen af aktiverne bruges i modelleringen af for-profit selskaber. Vi indfører en simpel solvensbegrænsning og viser optimalitet af barrierestrategier, hvor barrieren er defineret i forhold til funding-ratioen. Vi undersøger også barriererstrategier i en model med mere avancerede solvensbegrænsninger, hvor tilladelsen til at kontrollere enten passiver eller aktiver, når funding-ratioen er mellem en øvre og en nedre barriere, afhænger af hvilken af de to barrierer, der blev krydset sidst. Vi undersøger også barriererstrategier under den antagelse, at ruin skal forhindres. Dette opnås ved enten en mindskelse af passiverne eller ved en kapitalinjektion. Alle resultaterne er illustreret numerisk.

List of papers

This thesis is based on five papers:

- Lars Frederik Brandt Henriksen, Thomas Møller (2014): Local risk-minimization with longevity bonds
Applied Stochastic Models in Business and Industry
Online publication available at:
<http://onlinelibrary.wiley.com/doi/10.1002/asmb.2028/abstract>
- Lars Frederik Brandt Henriksen, Jeppe Woetmann Nielsen, Mogens Steffensen, Christian Svensson (2014): Markov chain modeling of policyholder behavior in life insurance and pension
European Actuarial Journal 4(1), pages 1-29, 2014,
<http://dx.doi.org/10.1007/s13385-014-0091-2>
- Marcus Christiansen, Lars Frederik Brandt Henriksen, Kristian Juul Schomacker, Mogens Steffensen (2014): Stress scenario generation for solvency and risk management
Accepted for publication in *Scandinavian Actuarial Journal*.
- Lars Frederik Brandt Henriksen (2014): Affine processes and Markov chains: Interest Rate-Dependent Transition Rates in Life Insurance
- Benjamin Avanzi, Lars Frederik Brandt Henriksen, Bernard Wong (2014): Optimal surplus distribution problems for regulated funds with assets and liabilities

Chapter 1

Introduction

This chapter gives an overview of the contributions of the thesis and highlights some of the connections between the five separate papers, which circles around many of the same themes but consider these from quite different perspectives. None of the papers are continuations of each other nor are they using results of each other as foundation. Several of the chapters address how to react to parts of the Solvency II legislation. These reactions include how to reduce the solvency capital requirement (SCR) and how to make the more advanced compulsory calculations more tractable seen from a numerical point of view.

1.1 Local risk-minimization

In Chapter 2 we consider the problem of hedging an insurance payment stream for a portfolio, where the lives of the policyholders are mutually independent and identically distributed conditional on the mortality intensity. The affine framework is chosen to make the calculations tractable such that the insurance reserves can be calculated by solving ordinary differential equations (ODEs) instead of partial differential equations (PDEs). The work is partly motivated by the increasing interest for mortality derivatives, see Blake et al. (2008), Blake et al. (2010) and Blake (2011), and the increasing improvements in life expectancy seen over the last decade, which exposes life insurance companies and pension funds to a non-diversifiable risk. This work is inspired by Dahl et al. (2008) but here we don't require the discounted price processes of the traded assets to be martingales under the real measure. Moreover, we focus on longevity bonds instead of survivor swaps. The quantitative impact studies relating to the Solvency II legislation have shown, that especially longevity is a major risk which requires a lot of capital. By hedging the longevity risk, the companies can reduce their solvency capital requirement, and the chapter analyses how this can be done in an optimal way.

The interest rate is modeled by an affine process, the financial market consists of zero-coupon bonds and a longevity bond, and the number of survivors is modeled via a double-stochastic process, where the mortality intensity is driven by a time-inhomogeneous CIR-model. The longevity bond is a mortality linked derivative paying a coupon, which is proportional to the actual number of survivors in a portfolio. That is, the payment process of the longevity bond is given by

$$dA^{\text{LB}}(x, t) = (n - N(x, t))dt, \quad 0 \leq t \leq T,$$

where n is the number of insured lives and $N(x, t)$ is the numbers of deaths in the portfolio at time t . The payment stream, we are hedging, is quite general and covers the most common life insurance products. Because the market is incomplete, we cannot hedge the insurance payment stream perfectly. Instead we use the quadratic criterion “local risk-minimization” to find the

best way of hedging the risk. Local risk-minimization was introduced by Schweizer (1988) as a generalization of risk-minimization to the case where the price processes of the traded assets are not martingales under the real measure, which often is the case in practice. The chapter extends Dahl et al. (2008) to the case where the price processes of the traded assets are not martingales under the measure used for determining the optimal strategies.

The criterion local risk-minimization, which is equivalent to risk-minimization in the martingale case, is that one minimizes the conditional expected quadratic deviation between the costs today and the costs at termination time in a local manner. That is, one tries to make the strategy as close to selffinancing as possible (in a quadratic sense). To obtain the risk-minimizing strategy we find the value of the reserve under a class of equivalent martingale measures, the real measure and a specific martingale measure called the minimal martingale measure. Here, one of the challenges is that we need to prove that the minimal martingale measure exists (is a well-defined probability measure) for the market, which in general is not the case when the price processes of the assets have jumps. The value function of the reserve under different measures helps us construct the local risk-minimizing strategy in an explicit way and finally checking that various technical conditions are satisfied. The overall form of the optimal strategy resembles the optimal strategy found in Dahl et al. (2008) in some ways but various terms are calculated under different measures. The size of the impact on the strategy of these differences depends of course on the differences between the measures and the other parameters of the model. Also note that local risk-minimization requires calculations of the reserve under the minimal martingale measure where interest rate and mortality intensity are not necessarily affine processes. That is, we are forced to solve PDEs. In summary, it is possible to find the local risk-minimizing strategy under the correct measure but it comes with a fairly high price in terms of technical challenges and computational workload.

1.2 Policyholder options and affine processes

Chapter 3 deals with calculation of life insurance reserves where we take expected policyholder behaviour and options into account in a computationally tractable way. The policyholder behavior we think of here are the surrender option and the free policy options, which means that for a disability product, we want to calculate reserves and cash flows in a Markov chain model like the one illustrated in Figure 1.1. There is a wide range of options for modelling of policyholder

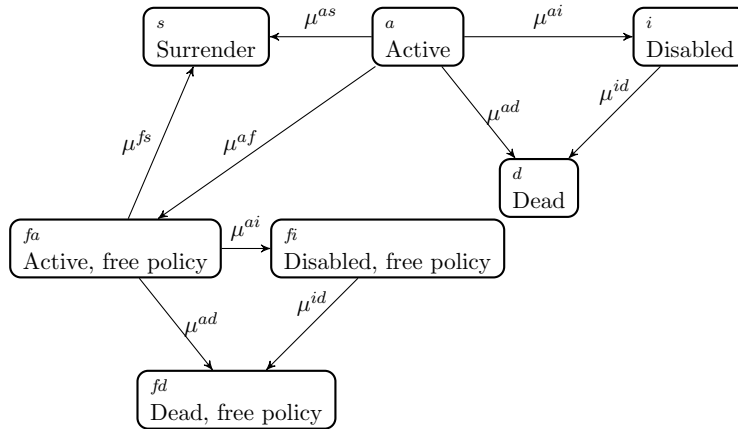


Figure 1.1: State space for a disability model with lapse.

behavior. One extreme option is to assume that the policyholder exercises options based on an economically optimal strategy meaning that the reserves should be valued like American type options. However, it is highly questionable whether the average policyholder has the information and knowledge to exercise these options when beneficial. Moreover, many contracts are linked to a job making these behavioral options more difficult to exercise. Another extreme approach is to assume that the intervention options are exercised completely incidentally, which allows us to model the transitions of the Markov chain by deterministic intensities. In between these two extremes are models where the intervention intensities depend on interest rates or other internal (to the model) variables. These models also include the cases, where the intervention intensities are split into rational parts and irrational parts. In Chapter 3 we take the extreme approach to have an intensity based model with deterministic intensities, whereas Chapter 5 deals with the approach in between, where intensities for example can be linked to the interest rate.

Figure 1.1 illustrates the case with dependence between insurance risk and behavior risk in the way that you e.g. cannot exercise the free policy option in state “Disabled”. This dependence between insurance risk and behavior essentially arises from the product design. One of the main points of this chapter is to study and relate formulas and numerical results in the two cases of dependence and independence, respectively. The chapter shows formalistic and computational advantages and disadvantages under different simplifying assumptions in the case where dependence is part of the contract design. The numerical part studies the effects (is it beneficial or not to take behavioral options into account for insurance companies) for some standard contracts. It also quantifies the impact of using various simplifying assumptions. An example of this quantification can be seen in Figure 1.2 where one also observes that the effect of using simplifying assumptions depends on the type of the contract.

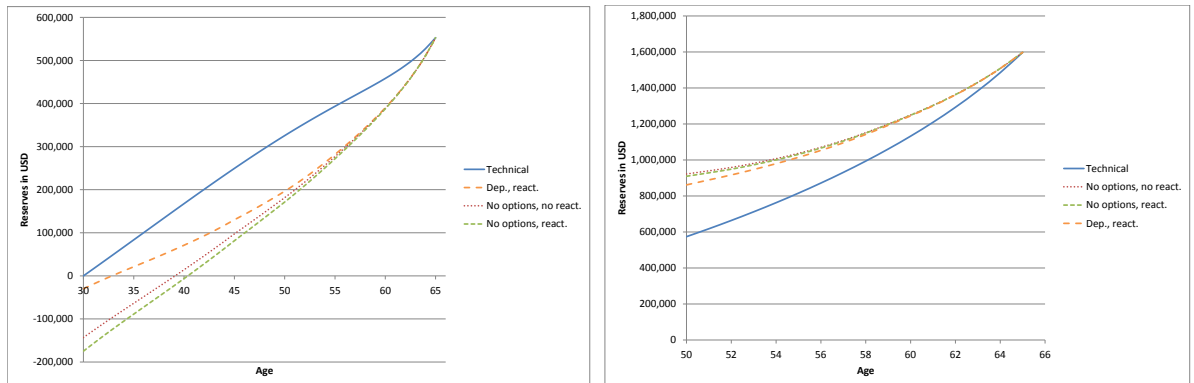


Figure 1.2: The technical reserve, the “dependent” reserve (including reactivation), and the two market reserves without behavioral options. Left plot is for a “new policy” and the right plot is for an “old policy”.

In Chapter 5 we present a theoretical framework for valuation of life insurance contracts within a Markov chain model by solving ODEs instead of PDEs in models where the interest rate and the transition rates are modeled as stochastic processes. This is e.g. motivated by the Solvency II legislation that demands that life insurance companies take into account future policyholder behavior such as the likelihood of lapse during the remaining period of an insurance contract. Moreover, the work of Kuo et al. (2003) and De Giovanni (2010) have been an inspiration for interest-dependent modelling of intervention intensities. More precisely, we model the interest rate and the transition rates as affine processes, where we allow for dependence between the

interest rate and the transition rates. Affine processes are stochastic processes with the property that the conditional characteristic functions are exponentially affine. To be able to calculate reserves in a stochastic setup is useful in itself as shown in Chapter 2. The theory in this chapter extends some of the results used in Chapter 2, which can be used if one for example wants to make local risk-minimization within more general Markov chain models than the life-death model discussed in Chapter 2. The reason why it is useful to allow for dependence between the processes, is outlined above. The class of affine processes is a popular choice for many types of stochastic modelling due to mathematical and computational tractability. The results are not only useful in life insurance but also areas like credit risk.

We obtain results for general hierarchical Markov chain models which is an generalization of the results in Duffie et al. (2000) to an arbitrary number of transitions within hierarchical Markov chain models. An example of such a model is given in Figure 1.1. Moreover, we give two different representations of transforms of affine processes, which are both useful but in different models. One of the representations has the form

$$\mathbb{E} \left[e^{u_0(v) + u_1(v)X(v) - \int_s^v \rho_0(\tau) + \rho_1(\tau)X(\tau)d\tau} \sum_{i=0}^n g_i(v)X^i(v) \middle| \mathcal{F}(s) \right] = e^{\beta(s,v)X(s)} \sum_{i=0}^n C_i(s,v)X^i(s), \quad (1.2.1)$$

where u_j , ρ_j , $j = 1, 2$, and g_i , $i = 0, \dots, n$, are deterministic functions, X is an affine process, and the functions β and C_i are solutions to a system of ODEs. We denote this representation the dense representation since it requires calculation of a minimum number of ODEs. To calculate reserves in state “Active” for the example illustrated in Figure 1.1, one needs to calculate terms of the form (1.2.1) for i up to three and hereafter integrate the terms over all possible jump times. It turns out that the class of affine processes is very useful and computational tractable for many Markov chain models used in life insurance but that the approach also has its limitations. These limitations arise when there are possibility of making many “jumps” within the Markov chain before reaching an absorbing state, since this requires calculation of a huge number of ODEs. This challenge is described in more details in one of the last parts of Chapter 5. However, the dimension of the underlying affine process does not give rise to computational problems, since the number of ODEs is sublinear in the dimension. Another limitation is that the affine setup is only useful in a limited number of non-hierarchical models.

1.3 Worst-case calculations

In Chapter 4 we calculate worst-case scenarios for probability weighted life insurance reserves. The chapter extends Christiansen and Steffensen (2013) in the following way: Like us, they search for optimal deterministic scenarios and obtain simple formulas for these but in a quite restricted class of models. The class is defined endogenously by requiring that certain argmax operations over transition intensities are constant with respect to the transition probabilities they generate, see Christiansen and Steffensen (2013, Proposition 4.1 and 4.2). The work in this paper is very much inspired from the structure of problems and solutions in that article, but we succeed in also finding the worst-case scenario without imposing their restrictive assumptions. Like in Chapter 2, we address the question of how to deal with the Solvency II legislation but in a quite different manner. In Chapter 2 we do this by considering an optimal hedge, and here we do so by suggesting an alternative to the standard formula or at least a way of qualifying the standard formula of Solvency II. The framework described in the chapter allows for the case where the interest rate and the various transition intensities are mutually dependent, which means that the worst-case scenarios are typically *not* “corner solutions”. As opposed to Chapter 5, the dependence is not

modeled directly but indirectly by the shape of the sets of scenarios over which we are maximising the reserves. Some types of dependencies one can think of in the context of life insurance is e.g. surrender intensities and interest rates that are high at the same time, that mortality intensities of a policyholder as active and disabled, respectively, are low at the same time, and that mortality intensities of the policyholders in a portfolio are low at the same time.

The first part of the chapter deals with describing the worst-case probability weighted reserve (and scenario) for a single policyholder by a system of ODEs. By worst-case scenario we mean, that for unknown interest and transition rates (ϕ, μ) , we want to find deterministic interest and transition rates $(\tilde{\phi}, \tilde{\mu})$ such that for the liabilities L , it holds that

$$P\left(L(t, \tilde{\phi}, \tilde{\mu}) \geq L(t, \phi, \mu)\right) \geq 1 - \alpha,$$

where $\alpha \in [0, 1)$. That is, we want to find a deterministic calculation basis such that the liabilities calculated with this basis with a certain probability are larger than the liabilities calculated with the real (stochastic) basis. This can be obtained by choosing $(\tilde{\phi}, \tilde{\mu}) = \operatorname{argmax}_{(\phi, \mu) \in M} L(t, \phi, \mu)$ for a set M such that $P((\phi, \mu) \in M) \geq 1 - \alpha$. The object of study in this paper is, given a set M , to calculate the $\operatorname{argmax}_{(\phi, \mu) \in M} L(t, \phi, \mu)$.

As shown in Christiansen and Steffensen (2013), we can use this to obtain an upper bound for the SCR given by

$$\sup_{(\phi, \mu) \in M} \{L(t, \phi, \mu)\} - L(t, \phi^{\text{BE}}, \mu^{\text{BE}}),$$

where ϕ^{BE} and μ^{BE} are best estimates for the interest rate and transition intensities, respectively. This study also includes a verification lemma, which has deterministic control theory as foundation, stating that the system of ODEs characterizes the worst-case reserve and results about existence of a worst-case scenario and reserve, which requires delicate functional analysis. One of the main focus points in this chapter is to obtain a tractable characterisation of the worst-case scenario and reserve by a system of ODEs, like it is the goal in Chapter 5. The ODEs have boundary conditions in different time points and two numerical methods (fixed point equation method and the shooting method) for obtaining numerical results are outlined.

Seen from a company and solvency perspective, this type of calculations are by far most interesting for a portfolio of life insurance contracts. The theory for a single policy can be used to cover special cases of homogeneous portfolios but does not in general cover inhomogeneous portfolios. The second half of Chapter 4 extends the theory to cover inhomogeneous portfolios in a computationally tractable way and investigate the results numerically. Figure 1.3 shows that there can be a big difference between calculating the worst-case intensities for a portfolio of policyholders compared to calculating the worst-case intensity for each policyholder individually. The solid lines are worst-case death intensities for the whole portfolio and the dotted lines are worst-case death intensities for the individual policyholders. The three representative policyholders have the same product but different ages. This difference is also illustrated in Table 1.1, where we have made a numerical study of the impact on SCR when using the standard formula, using the worst-case intensities for the portfolio and by using the worst-case intensities for the individual policyholders. The table should *not* be read in the way that we think the SCR of Solvency II is too low but it illustrates what results we obtain if we use the percentages of the mortality and longevity shocks of Solvency II as lower and upper bounds relative to a best-estimate mortality intensity for the sets we are maximizing the reserves with respect to. This leads to a transparent SCR where one exactly knows where the worst-case scenario comes from.

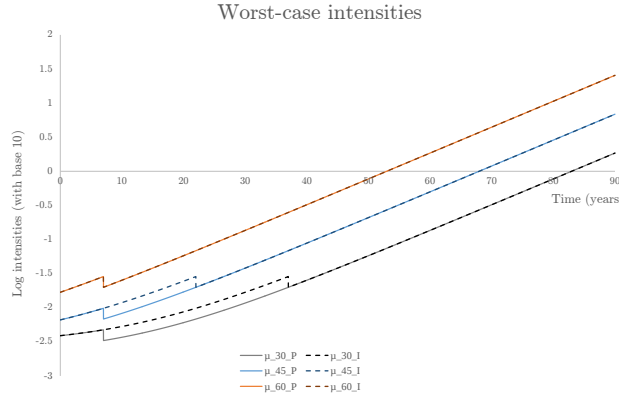


Figure 1.3: Worst-case intensities for the portfolio and for each individual policyholder.

Solvency II	Worst-case (PF)	Worst-case (separate)
1.59	1.83	2.19

Table 1.1: SCR calculated by different methods.

1.4 Optimal surplus distribution problems

In Chapter 6, we consider optimal surplus distribution strategies for pension funds and for-profit companies. The problems we are studying in this chapter belong to the class of stochastic control theory problems with infinite time horizon. The controls considered are that we either increase liabilities or decrease assets. The increase of liabilities could be used in the modeling of non-for-profit mutual funds or pension funds. On the other hand, the decrease in assets could be used for modelling of for-profit companies. The latter case is equivalent with paying out dividends. The work is an extension of Gerber and Shiu (2003) and the main contribution is to take different kinds of solvency constraints into account and to perform numerical explorations of these constraints.

We model assets and liabilities as a correlated two-dimensional geometric Brownian motion and allow for singular controls but we impose some solvency constraints. In the first part of the chapter we consider simple solvency constraints as the ones suggested in Paulsen (2003) which impose that the pension fund or company is not allowed to increase liabilities or pay out dividends if the funding ratio is below a level α_1 . It turns out that barrier strategies, where the barrier is for the funding ratio, are optimal and that the optimal barriers are the same for both the case where we increase liabilities and pay out dividends.

In the second and third part of the chapter we consider a model with more advanced solvency constraints and a model where ruin must be prevented. The idea of advanced solvency constraints was first suggested in Avanzi and Wong (2012) and is illustrated in Figure 1.4. The interpretation of the advanced solvency constraint in the pension fund case is the following: We assume that the fund is not fully funded if its funding ratio is below α_1 (although it is not severe enough to declare bankruptcy). Hence it is not realistic to assume that a fund that has just recovered and just reached α_1 could immediately start distributing some of its excess profits. In other words, $\alpha_2 > \alpha_1$ describes a situation where downcrossing α_1 would trigger some alarm and put the fund in an emergency state under which no distribution is allowed (and perhaps, the fund is closely monitored by the regulator). This state would revert back to normal when the process upcrosses level $\alpha_2 > \alpha_1$ again. For the case of a for-profit company, $\alpha_1 = \alpha_2$ may lead to erratic periods of

dividend payments if the barrier is equal to α_1 , which is unrealistic in practice; see also Avanzi and Wong (2012) for a discussion of this. In this setup, we also study barrier strategies.

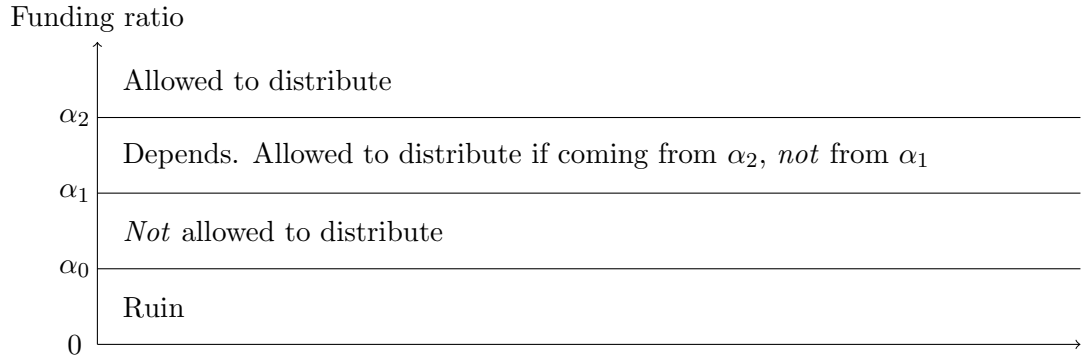


Figure 1.4: Graphical illustration of model with advanced solvency constraints.

In the part where ruin must be prevented, we evaluate strategies where lowering of liabilities are forced at the ruin level for pension funds and where capital injections are forced at the ruin level for for-profit companies. Since lowering of liabilities and injection of capital comes with a price (modelled by a multiplicative factor), it is only optimal to do so when absolutely necessary, i.e. at the ruin level. We also investigate the more realistic assumption, that the price of capital injection is dependent on the funding ratio by the rationale that the lower the funding ratio the greater the cost of raising capital. This will make it more attractive to inject before the ruin level. The calculations come with lots of numerical illustrations elucidating whether solvency constraints are good or bad and whether forced injections are optimal relative to declaring ruin.

Chapter 2

Local risk-minimization with longevity bonds

Abstract: This paper studies the criterion of local risk-minimization for life insurance contracts in a financial market which includes longevity bonds. The longevity bond is a bond specifying payments which are linked to the current number of survivors in a given portfolio of insured lives. The number of survivors is modeled via a double-stochastic process, where the mortality intensity is driven by a time-inhomogeneous CIR-model. In addition to the longevity bond, the financial market is assumed to consist of a traditional bond and a savings account. We define the price process of the longevity bond by introducing a pricing measure. The paper extends previous work in the literature to the case where the traded assets are not martingales under the measure used for determining the optimal strategies. We compare our results under the real measure with the former results of globally risk-minimizing strategies, obtained using an equivalent martingale measure.

Keywords: Longevity bond, minimal martingale measure, local risk-minimization.

2.1 Introduction

The focus on biometric risks in life insurance is increasing vigorously these years. This is in particular motivated by the forthcoming Solvency II legislation. Based on the results of Fifth Quantitative Impact Study for Solvency II (QIS V), see EIOPA (2011), one realises that the longevity stress is the second largest risk of the life module. Hedging this risk by means of longevity derivatives could potentially reduce the solvency capital requirement.

In this paper we discuss the problem of finding a locally risk-minimizing hedging strategy for life insurance contracts in a financial market containing longevity bonds. The study of mortality derivatives has increased in popularity during the last years. The motivation is the increasing average lifetime in many countries, which exposes the life insurance companies to a major non-diversifiable risk. This risk comes along with the financial risk. One challenge for the life insurance companies in the coming years will be to control or minimize the combined risk. A possible way to do this is to hedge the risk by trading mortality derivatives. This has to a limited extent been done alongside the introduction of mortality derivatives in the financial markets the last few years. The potential market for mortality derivatives is huge, since the global longevity risk exposure in the private pension sector is estimated to about \$25 trillions, see Blake and Biffis (2012). The general interest in mortality derivatives is increasing rapidly these years, see e.g. Blake et al. (2008), and some people also advocate for the introduction of government-issued longevity bonds,

see Blake et al. (2010). A current account of the market for longevity derivatives can be found in Blake (2011).

Here, we give a short review of the literature on risk-minimization and its applications in life insurance. Risk-minimization was introduced by Föllmer and Sondermann (1986), and the criterion has been applied in several papers subsequently. Dahl et al. (2008) carries out risk-minimization for insurance payment streams for different markets in the case where the interest rate process and the mortality intensity process are continuous. Schweizer (1988) introduces the theory of local risk-minimization and Schweizer (2008) gives a streamlined theoretical description of local risk-minimization for multidimensional payment processes.

The criterion local risk-minimization has recently been used by various authors in insurance applications. Barbarin (2007) finds locally risk-minimizing strategies for insurance contracts with surrender options in a setting without jumps. As examples of the theory used in a setting with jumps, we mention Riesner (2006) and Vandaele and Vanmaele (2008), where locally risk-minimizing strategies are determined for an insurance contract in a market with a risky asset containing jumps and a deterministic interest rate.

We apply the criterion of local risk-minimization within the model of Dahl et al. (2008) in the case where the discounted price processes of the traded assets are not martingales under the objective measure. Local risk-minimization may be viewed as a more natural way of minimizing the hedging-error than risk-minimization under some equivalent martingale measure, since the price processes of the traded assets evolve with respect to the objective measure.

The paper is organized as follows: In Section 2.2 we introduce a financial market with zero coupon bonds, a model for the mortality intensity and a portfolio of insured lives. Both the interest rate and the mortality intensity are modeled as in Dahl et al. (2008). We present the combined model for the financial market and the portfolio of insured lives and study a class of equivalent martingale measures. In Section 2.3 and 2.4 we examine mortality derivatives in the form of longevity bonds and a general insurance contract and find their stochastic representations. The theory of local risk-minimization is reviewed in Section 2.5. In Section 2.6, we prove that the minimal martingale measure exists (is a well-defined probability measure) for the market, and we find an expression for this measure. In Section 2.7, we finally derive locally risk-minimizing strategies for the market under the objective measure and compare the result with some results obtained using the criterion global risk-minimization. The appendix presents proofs of some technical results.

2.2 The model

This section contains a brief review of the model and some main results of Dahl et al. (2008). Let T be a fixed finite time horizon and $(\Omega, \mathcal{F}, P)$ a probability space with filtration $\mathbb{F} = (\mathcal{F}(t))_{0 \leq t \leq T}$ satisfying the usual conditions of right-continuity and completeness. All random variables and stochastic processes are defined on $(\Omega, \mathcal{F}, P)$. We also define the subfiltrations $\mathbb{G} = (\mathcal{G}(t))_{0 \leq t \leq T}$, $\mathbb{I} = (\mathcal{I}(t))_{0 \leq t \leq T}$ and $\mathbb{H} = (\mathcal{H}(t))_{0 \leq t \leq T}$ generated by two standard Brownian motions W^r (determines the interest rate), W^μ (determines the mortality intensity) and a counting process N (counting the deaths in the insurance portfolio), respectively. These processes are defined later. We define the filtration $\mathbb{F}$ by $\mathcal{F}(t) = (\mathcal{G}(t) \vee \mathcal{I}(t) \vee \mathcal{H}(t))$.

2.2.1 Interest rate model

The short rate is determined by a standard Vasiček model, i.e. the short rate dynamics under P are

$$dr(t) = (\gamma^r - \delta^r r(t))dt + \sigma^r dW^r(t),$$

with $r(0) = r_0 > 0$ fixed. Here, γ^r , δ^r and σ^r are constants, and W^r is a standard Brownian motion.

2.2.2 The financial market

The basic financial market consists of two traded assets: A savings account and a zero-coupon bond with maturity T . Later on, the market will be extended with a longevity bond.

The price process of the savings account is given by

$$dB(t) = r(t)B(t)dt, \quad B(0) = 1,$$

and the price process of the zero coupon bond expiring at time T is given by

$$P(t, T) = E^{Q^r} \left[e^{-\int_t^T r(u)du} \middle| \mathcal{F}(t) \right].$$

Here, the unique equivalent martingale measure Q^r for the market is defined by

$$\frac{dQ^r}{dP} = \Lambda^r(T), \tag{2.2.1}$$

where

$$d\Lambda^r(t) = \Lambda^r(t)h^r(r(t), t)dW^r(t), \quad \Lambda^r(0) = 1. \tag{2.2.2}$$

We assume that $E^P(\Lambda^r(T)) = 1$. In addition, we assume that

$$h^r(r(t), t) = - \left(\frac{\tilde{c}}{\sigma^r} + \frac{cr(t)}{\sigma^r} \right),$$

where c and $\tilde{c}$ are constants. It follows from Girsanov's Theorem that the dynamics of the short rate under Q^r are given by

$$dr(t) = (\gamma^{r,Q} - \delta^{r,Q}r(t))dt + \sigma^r dW^{r,Q}(t),$$

where $W^{r,Q}$ is a standard Brownian Motion under Q^r , $\gamma^{r,Q} = \gamma^r - \tilde{c}$ and $\delta^{r,Q} = \delta^r + c$. Under Q^r the interest rate process is affine, and it is well-known that the price process of a zero coupon bond is given by

$$P(t, T) = e^{\alpha(t,T) - \beta(t,T)r(t)},$$

where

$$\begin{aligned} \beta(t, T) &= \frac{1}{\delta^{r,Q}} \left(1 - e^{-\delta^{r,Q}(T-t)} \right), \\ \alpha(t, T) &= \frac{(\beta(t, T) - T + t)(\gamma^{r,Q}\delta^{r,Q} - \frac{1}{2}(\sigma^r)^2)}{(\delta^{r,Q})^2} - \frac{(\sigma^r)^2(\beta(t, T))^2}{4\delta^{r,Q}}. \end{aligned}$$

Under Q^r , the dynamics of the discounted price process of the zero coupon bond are given by

$$dP^*(t, T) = -\sigma^r \beta(t, T)P^*(t, T)dW^{r,Q}(t). \tag{2.2.3}$$

2.2.3 The mortality intensity

The mortality intensity is modeled by a 1-dimensional stochastic process. As in Dahl et al. (2008), we start by defining an underlying process ζ , which determines the relative change in the mortality intensity. Let $\zeta(x, t)$ be a Markov process with dynamics

$$d\zeta(x, t) = (\gamma(x, t) - \delta(x, t)\zeta(x, t))dt + \sqrt{\zeta(x, t)}\sigma(x, t)dW^\mu(t), \quad (2.2.4)$$

$\zeta(x, 0) = 1$. Note that ζ is a process depending only on the variable t , not on x . The parameter x just indicates that we consider a cohort, where each individual has the same age x at time 0. In contrast to Dahl et al. (2008), we work with a 1-dimensional ζ -process. The functions γ , δ and σ are deterministic functions only depending on t , and W^μ is a 1-dimensional standard Brownian motion.

We let the mortality intensity $\mu(x, t)$ be given by $\mu(x, t) = \mu^\circ(x, t)\zeta(x, t)$, where $\mu^\circ(x, t)$ is a positive, deterministic function. By modeling the mortality intensity as the product $\mu^\circ\zeta$, we can view μ° as the time-0 mortality intensity and ζ as a process describing the changes of the mortality intensity relative to the time-0 mortality intensity.

By Itô's formula and by using (2.2.4) we get

$$\begin{aligned} d\mu(x, t) &= \mu^\circ(x, t)d\zeta(x, t) + (d\mu^\circ(x, t))\zeta(x, t) \\ &= \Theta(x, t, \mu(x, t))dt + \sigma^\mu(t, \mu(x, t))dW^\mu(t), \end{aligned} \quad (2.2.5)$$

where

$$\begin{aligned} \Theta(x, t, \mu(x, t)) &= \mu^\circ(x, t)\gamma(x, t) + \left(\left(\frac{d}{dt}\mu^\circ(x, t) \right) \mu^\circ(x, t)^{-1} - \delta(x, t) \right) \mu(x, t), \\ \sigma^\mu(t, \mu(x, t)) &= \sqrt{\mu(x, t)}\sqrt{\mu^\circ(x, t)}\sigma(x, t). \end{aligned}$$

To ensure that the mortality intensity is strictly positive, we assume that

$$2\gamma(x, t) \geq (\sigma(x, t))^2.$$

2.2.4 The lifetimes in the portfolio

We consider a portfolio consisting of n insured lives. They are all of same age x at time 0. We assume that the lives are mutually independent and identically distributed conditional on the mortality intensity. The remaining lifetimes are given by the positive random variables $T_1, \dots, T_n$. The probability of surviving until time t , given the mortality intensity up to time t , is given by

$$P(T_1 > t | \mathcal{I}(t)) = e^{-\int_0^t \mu(x, s)ds}.$$

The number of deaths in the portfolio is given by the counting process

$$N(x, t) = \sum_{i=1}^n 1_{(T_i \leq t)}.$$

The stochastic intensity process $\lambda(x, \cdot)$ related to $N(x, \cdot)$ is informally given by

$$\lambda(x, t)dt = E^P(dN(x, t) | \mathcal{I}(t-) \vee \mathcal{H}(t-)) = (n - N(x, t-))\mu(x, t)dt.$$

That is, λ is the mortality intensity times the number of survivors. The martingale corresponding to the counting process $N(x, t)$ is given by $dM(x, t) = dN(x, t) - \lambda(x, t)dt$.

2.2.5 Change of measure in the combined model

We consider a combined model comprising the financial market, the mortality intensity and the portfolio of insured lives. We assume that the financial market is independent of the mortality intensity and the portfolio of lives, i.e. $\mathcal{G}(t) \perp (\mathcal{H}(t), \mathcal{I}(t))$. The assumption of independence implies that processes which are martingales with respect to a subfiltration are also martingales with respect to $\mathbb{F}$.

We consider a class of equivalent martingale measures for the combined financial and insurance market. We let a martingale measure Q be given by

$$\frac{dQ}{dP} = \Lambda(T), \quad (2.2.6)$$

where

$$d\Lambda(t) = \Lambda(t-)(h^r(r(t), t)dW^r(t) + g(t)dM(x, t)), \quad \Lambda(0) = 1. \quad (2.2.7)$$

We assume that $E^P(\Lambda(T)) = 1$. The change of measure given by (2.2.6) is a generalization of the change of measure given by (2.2.1). We maintain the assumptions described in Section 2.2.2 for the first term of the likelihood process Λ . Moreover, we assume that the deterministic function g depends on t only, is differentiable, and that $g > -1$.

This measure change is slightly less general than the one studied in Dahl et al. (2008), which also includes a change of measure for W^μ . Here, we need to restrict ourselves to the measure change given by (2.2.6) to ensure that the minimal martingale measure found below in Section 2.6 is a well-defined probability measure.

The dynamics of the likelihood process $\Lambda(t)$ consist of two parts. The first part is related to the interest rate and the second part to the counting process N . By Girsanov's Theorem, the dynamics of ζ and μ are unchanged under Q .

Under the measure Q , the stochastic intensity process is given by

$$\lambda^Q(x, t) = (n - N(x, t-))\mu^Q(x, t),$$

where

$$\mu^Q(x, t) = (1 + g(t))\mu(x, t).$$

The associated Q -martingale is given by

$$dM^Q(x, t) = dN(x, t) - \lambda^Q(x, t)dt. \quad (2.2.8)$$

We end this section by writing the dynamics of the process $\mu^Q(x, \cdot)$ in an affine form. A straightforward calculation shows that the dynamics of μ^Q can be written as

$$d\mu^Q(x, t) = (\gamma^{\mu, Q}(x, t) - \delta^{\mu, Q}(x, t)\mu^Q(x, t))dt + \sqrt{\mu^Q(x, t)}\sigma^{\mu, Q}(x, t)dW^\mu(t),$$

where

$$\begin{aligned} \gamma^{\mu, Q}(x, t) &= (1 + g(t))\mu^\circ(x, t)\gamma(x, t), \\ \delta^{\mu, Q}(x, t) &= \delta(x, t) - \frac{\frac{d}{dt}g(t)}{1 + g(t)} - \frac{\frac{d}{dt}\mu^\circ(x, t)}{\mu^\circ(x, t)}, \\ \sigma^{\mu, Q}(x, t) &= \sqrt{(1 + g(t))\mu^\circ(x, t)}\sigma(x, t). \end{aligned}$$

2.2.6 Survival probability under Q

We provide a characterization of the survival probabilities under this equivalent martingale measure. This characterization is used frequently when studying longevity bonds and insurance contracts, which contain risk related to survival probabilities. We define the conditional expected survival probabilities under the measure Q by

$$\mathcal{S}^Q(x, t, T) = \mathbb{E}^Q \left[e^{-\int_t^T \mu^Q(x, s) ds} \middle| \mathcal{F}(t) \right].$$

The associated Q -martingale is given by

$$\mathcal{S}^{Q, M}(x, t, T) = \mathbb{E}^Q \left[e^{-\int_0^T \mu^Q(x, s) ds} \middle| \mathcal{F}(t) \right] = e^{-\int_0^t \mu^Q(x, s) ds} \mathcal{S}^Q(x, t, T).$$

The martingale property follows from the rule of iterated expectations.

By the results in Dahl et al. (2008), we get that the survival probability can be written in the following form:

$$\mathbb{E}^Q \left[e^{-\int_t^T \mu^Q(x, s) ds} \middle| \mathcal{F}(t) \right] = e^{\alpha^\mu(t, T) - \beta^\mu(t, T) \mu^Q(x, t)}, \quad (2.2.9)$$

where β^μ and α^μ satisfy the following differential equations

$$\begin{aligned} \frac{\partial}{\partial t} \beta^\mu(t, T) &= \delta^{\mu, Q}(x, t) \beta^\mu(t, T) + \frac{1}{2} (\sigma^{\mu, Q}(x, t))^2 (\beta^\mu(t, T))^2 - 1, & \beta^\mu(T, T) &= 0, \\ \frac{\partial}{\partial t} \alpha^\mu(t, T) &= \gamma^{\mu, Q}(x, t) \beta^\mu(t, T), & \alpha^\mu(T, T) &= 0. \end{aligned}$$

2.3 Extended financial market

We extend the financial market introduced in Section 2.2.2 with a longevity bond, which is a so-called mortality derivative, see also Dahl et al. (2008), who consider a market with survivor swaps. The longevity bond pays a coupon, which is proportional to the actual number of survivors in a given portfolio. Longevity bonds are financial instruments that can be used to hedge the longevity risk. The survivor swap pays a coupon which is the difference between the actual number of survivors in a portfolio and a fixed number of survivors. There are some resemblances between these two mortality derivatives but also important differences. One difference is that the survivor swap requires a smaller initial investment than the longevity bond. Another difference is that the payment process of the longevity bond is non-negative, whereas the payment process of the survivor swap can be both negative and positive.

2.3.1 Longevity bonds

In the following sections we consider a fixed, arbitrary equivalent martingale measure Q from the class of measures given by (2.2.6). Under this fixed measure we preserve the independence between the basic financial market and the mortality intensity. This follows since the likelihood process does not contain mixed terms, but only terms related to either the financial market or to the mortality intensity. To start with we find the price of a longevity bond on the insurance portfolio.

The payment process of the longevity bond is given by

$$dA^{\text{LB}}(x, t) = (n - N(x, t))dt, \quad 0 \leq t \leq T,$$

and $A^{\text{LB}}(x, 0) = 0$. We denote by $Z^{*,Q}(x, t)$ the conditional expected value under Q at time t of the discounted value of payments from A^{LB} , i.e. we let

$$Z^{*,Q}(x, t) = \mathbb{E}^Q \left[\int_0^T e^{-\int_0^\tau r(s)ds} dA^{\text{LB}}(x, \tau) \middle| \mathcal{F}(t) \right] = A^{*,\text{LB}}(x, t) + \tilde{Z}^{*,Q}(x, t),$$

where $A^{*,\text{LB}}$ is the discounted payment process of the longevity bond and

$$\begin{aligned} \tilde{Z}^{*,Q}(x, t) &= e^{-\int_0^t r(u)du} \mathbb{E}^Q \left[\int_t^T e^{-\int_t^\tau r(s)ds} dA^{\text{LB}}(x, \tau) \middle| \mathcal{F}(t) \right] \\ &= (n - N(x, t)) \int_t^T P^*(t, \tau) \mathcal{S}^Q(x, t, \tau) d\tau. \end{aligned}$$

We assume that an asset with the discounted price process $Z^{*,Q}$ can be traded dynamically in the financial market. It follows by the rule of iterated expectations that the discounted price process $Z^{*,Q}$ is indeed a Q -martingale. Thus, the measure Q is also an equivalent martingale measure for the extended financial market $(B^*, P^*(\cdot, T), Z^{*,Q})$ consisting of a savings account, a zero coupon bond and a longevity bond.

We assume that $\tilde{Z}^{*,Q}$ is continuously differentiable with respect to t and twice continuously differentiable with respect to r and μ . By use of Itô's formula, it is possible to find a stochastic representation of $Z^{*,Q}(x, t)$. The stochastic representation is given by

$$dZ^{*,Q}(x, t) = \eta^{Z,Q}(t) dW^{r,Q}(t) + \nu^{Z,Q}(t) dM^Q(t) + \rho^{Z,Q}(t) dW^{\mu,Q}(t), \quad (2.3.1)$$

where

$$\begin{aligned} \eta^{Z,Q}(t) &= -\sigma^r (n - N(x, t-)) \int_t^T \beta(t, \tau) P^*(t, \tau) \mathcal{S}^Q(x, t, \tau) d\tau, \\ \nu^{Z,Q}(t) &= -\int_t^T P^*(t, \tau) \mathcal{S}^Q(x, t, \tau) d\tau, \\ \rho^{Z,Q}(t) &= -\sigma \sqrt{\mu^o(x, t) \mu(x, t)} (n - N(x, t-)) \\ &\quad \times (1 + g(t)) \int_t^T \beta^\mu(t, \tau) P^*(t, \tau) \mathcal{S}^Q(x, t, \tau) d\tau. \end{aligned}$$

For details and similar calculations for survivor swaps, see Dahl et al. (2008). Note that $\eta^{Z,Q}$, $\nu^{Z,Q}$ and $\rho^{Z,Q}$ are functions of t , N , r and μ .

Remark 2.3.1. *We have chosen to extend the market with a longevity bond on the insurance portfolio. We are aware of the fact, that it is more common to trade longevity bonds on a population rather than the insurance portfolio. We have chosen to stick to the longevity bond on the insurance portfolio, since the approach and techniques are the same as for the longevity bond on a population and because the setup with a 1-dimensional mortality intensity keeps the calculations and results more simple.*

2.4 Insurance contracts

In this section, we introduce an insurance contract which specifies premiums paid by the policyholders and benefits paid by the insurance company. The payment process is constructed in such a way that premiums are negative and benefits are positive, and we study the payment process for an entire portfolio of insured lives. The payment process allows for a single premium $\pi^s(0)$

paid at time 0, continuously paid premiums π^c , lump sum payments $a^r(\bar{T})$ upon retirement and a^d upon death, and life annuity payments a^p from retirement until death. The payment process is identical to the one considered in Dahl et al. (2008) and is given by

$$\begin{aligned} dA(t) = & -n\pi^s(0)d1_{\{t \geq 0\}} - \pi^c(t)(n - N(x, t))1_{\{0 \leq t \leq \bar{T}\}}dt + a^d(t)dN(x, t) \\ & + (n - N(x, \bar{T}))a^r(\bar{T})d1_{\{t \geq \bar{T}\}} + a^p(t)(n - N(x, t))1_{\{\bar{T} \leq t \leq T\}}dt. \end{aligned} \quad (2.4.1)$$

Here, $\bar{T}$ is the time of retirement, n is the number of policyholders in the portfolio at time 0 and $N(x, t)$ is the number of deaths until time t . Analogous to the rest of the paper, we denote by A^* the discounted version of the payment process.

2.4.1 Market reserves

The so-called intrinsic value process associated with the payment process A is given by

$$\begin{aligned} V^{*,Q}(t) &= \mathbb{E}^Q \left[\int_{0-}^T dA^*(\tau) \middle| \mathcal{F}(t) \right] \\ &= \int_{0-}^t e^{-\int_0^\tau r(s)ds} dA(\tau) + \mathbb{E}^Q \left[\int_t^T e^{-\int_0^\tau r(s)ds} dA(\tau) \middle| \mathcal{F}(t) \right] \\ &= A^*(t) + \tilde{V}^{*,Q}(t), \end{aligned} \quad (2.4.2)$$

where the process

$$\tilde{V}^{*,Q}(t) = \mathbb{E}^Q \left[\int_t^T e^{-\int_0^\tau r(s)ds} dA(\tau) \middle| \mathcal{F}(t) \right]$$

is called the discounted market reserve. The discounted market reserve is the number of survivors times the discounted market reserve for a single policy holder, which we denote by $\tilde{V}_p^{*,Q}(t)$, i.e.

$$\tilde{V}^{*,Q}(t) = (n - N(x, t))\tilde{V}_p^{*,Q}(t), \quad (2.4.3)$$

where

$$\begin{aligned} \tilde{V}_p^{*,Q}(t) = & \int_t^T P^*(t, \tau) \mathcal{S}^Q(x, t, \tau) \left(-\pi^c(\tau)1_{\{0 \leq \tau \leq \bar{T}\}}d\tau + a^d(\tau)f^{\mu,Q}(x, t, \tau)d\tau \right. \\ & \left. + a^p(\tau)1_{\{\bar{T} \leq \tau \leq T\}}d\tau \right) + P^*(t, \bar{T})\mathcal{S}^Q(x, t, \bar{T})a^r(\bar{T})1_{\{t < \bar{T}\}}. \end{aligned} \quad (2.4.4)$$

For a proof of (2.4.3) and (2.4.4), we refer to Dahl and Møller (2006).

In (2.4.4), $f^{\mu,Q}(x, t, \tau)$ is called the forward mortality intensity and is given by

$$f^{\mu,Q}(x, t, \tau) = -\frac{\partial}{\partial \tau} \log \mathcal{S}^Q(x, t, \tau).$$

For further details about the forward mortality intensity, see e.g. Dahl et al. (2008).

We assume that $\tilde{V}_p^{*,Q} \in C^{1,2,2}$, i.e. $\tilde{V}_p^{*,Q}$ is continuously differentiable with respect to t and twice continuously differentiable with respect to r and μ . In the following proposition, we state a stochastic representation of the insurance contract.

Proposition 2.4.1. *The intrinsic value process $V^{*,Q}(t)$ admits the representation*

$$dV^{*,Q}(t) = \eta^{V,Q}(t)dW^{r,Q}(t) + \nu^{V,Q}(t)dM^Q(t) + \rho^{V,Q}(t)dW^{\mu,Q}(t),$$

where

$$\begin{aligned} \eta^{V,Q}(t) &= -(n - N(x, t-)) \sigma^r \left(\int_t^T P^*(t, \tau) \beta(t, \tau) \mathcal{S}^Q(x, t, \tau) \right. \\ &\quad \times \left(-\pi^c(\tau) 1_{\{0 \leq \tau \leq \bar{T}\}} d\tau + a^d(\tau) f^{\mu,Q}(x, t, \tau) d\tau + a^p(\tau) 1_{\{\bar{T} \leq \tau \leq T\}} d\tau \right) \\ &\quad \left. + P^*(t, \bar{T}) \beta(t, \bar{T}) \mathcal{S}^Q(x, t, \bar{T}) a^r(\bar{T}) 1_{\{t < \bar{T}\}} \right), \\ \nu^{V,Q}(t) &= B(t)^{-1} a^d(t) - \tilde{V}_p^{*,Q}(t), \\ \rho^{V,Q}(t) &= -(n - N(x, t-)) \sigma \sqrt{\mu^o(x, t) \mu(x, t) (1 + g(t))} \\ &\quad \times \left(\int_t^T P^*(t, \tau) \beta^\mu(t, \tau) \mathcal{S}^Q(x, t, \tau) \left(-\pi^c(\tau) 1_{\{0 \leq \tau \leq \bar{T}\}} \right. \right. \\ &\quad \left. \left. + a^p(\tau) 1_{\{\bar{T} \leq \tau \leq T\}} + a^d(\tau) \left(f^{\mu,Q}(x, t, \tau) - \frac{\frac{\partial}{\partial \tau} \beta^\mu(t, \tau)}{\beta^\mu(t, \tau)} \right) \right) d\tau \right. \\ &\quad \left. + P^*(t, \bar{T}) \beta^\mu(t, \bar{T}) \mathcal{S}^Q(x, t, \bar{T}) a^r(\bar{T}) 1_{\{t < \bar{T}\}} \right). \end{aligned}$$

We leave out the proof but note that it is based on techniques similar to the ones used in the proof of (Dahl et al., 2008, Lemma 3.2).

2.5 Local risk-minimization

Since the market $(B^*, P^*(\cdot, T), Z^{*,Q})$ is incomplete, one cannot in general hedge the insurance payment stream perfectly. That is, the risk cannot be totally eliminated, and one needs to choose between different approaches in order to minimize the risk. A special case deviating from this general observation would be insurance contracts consisting of annuity payments only. In this case we would be able to hedge the payment stream perfectly, since we are trading longevity bonds on the insurance portfolio.

Föllmer and Sondermann (1986) introduced the measurement of risk by a quadratic criterion and various quadratic hedging methods have been studied hereafter. In the case where the discounted price processes of the traded assets are local martingales, the criterion of risk-minimization is often applied, but in the case where the discounted price processes are *not* local martingales, risk-minimization cannot in general be applied. For details and discussions of this result see (Schweizer, 2001, Proposition 3.1), which states that a contingent claim in general does not admit a risk-minimizing strategy if the discounted price processes of the assets are not local martingales.

The discounted price processes in the present market are in general not local martingales under the objective measure P , but one can make a change of measure from the objective measure to an equivalent martingale measure and apply the criterion of risk-minimization under this measure. Doing so, we have not accomplished risk-minimization (under the objective measure, as desired), but only found strategies which, depending on the change of measure, can be more or less close to a strategy minimizing the conditional second moments of the increments of the cost process.

This approximation can in some cases be the best one can hope to obtain. Another approach is to apply the more general criterion of local risk-minimization.

The aim of local risk-minimization is essentially to minimize the conditional second moments of the increments of the cost process in the case where the price processes of the traded assets are not local martingales under the measure P . The cost process measures any additional investments needed to finance the strategy and pay the insurance obligations. The difference between local risk-minimization and risk-minimization (under an equivalent martingale measure different from P) is that local risk-minimization is formulated in terms of the objective measure. In some cases, the so-called minimal martingale measure can be useful, which we describe below in Section 2.5.2.

2.5.1 A short review of local risk-minimization

This section is based on Schweizer (2008) and contains a short review of the criterion of local risk-minimization; for more details, see Schweizer (2008).

Let $X = (X(t))_{t \in [0, T]}$, where $T > 0$ is finite, be a d -dimensional semimartingale defined on a filtered probability space. The process X has a Doob-Meyer decomposition given by

$$X = X_0 + M^X + A^X = X_0 + M^X + \int d\langle M^X \rangle \lambda^X, \quad (2.5.1)$$

where M^X is a d -dimensional local martingale, A^X is a d -dimensional predictable process of bounded variation and λ^X is a d -dimensional predictable process. We note that $\langle M^X \rangle = \langle M^X, M^X \rangle$, where $\langle M^X \rangle_t \in \mathbb{R}^d \times \mathbb{R}^d$, is the predictable variation process, and that

$$A_i^X(t) = \sum_{j=1}^d \lambda_j^X(t) d\langle M_i^X, M_j^X \rangle(t),$$

where $\langle M_i^X, M_j^X \rangle$ is the predictable covariation process of M_i^X and M_j^X . One can think of X as the discounted price processes of the d risky assets.

We assume that X fulfills the so-called structure condition, which is a technical condition from Schweizer (2008). The condition is that M^X is locally P -square-integrable starting in 0 and that A^X has the form $A^X = \int d\langle M^X \rangle \lambda^X$ for an $\mathbb{R}^d$ -valued predictable process $\lambda^X \in L_{\text{loc}}^2(M^X)$ so that the mean-variance tradeoff process $K := \int \lambda^X dA^X = \int (\lambda^X)^{\text{tr}} d\langle M^X \rangle \lambda^X$ satisfies $K(T) < \infty$ P -a.s. We use the notation $(\cdot)^{\text{tr}}$ for the transpose of a vector.

In the following, we consider an L^2 portfolio strategy $\varphi = (\vartheta, \varsigma)$ such that the value process of the portfolio φ , defined by

$$V(\varphi, t) = \vartheta(t) \cdot X(t) + \varsigma(t),$$

is right-continuous and square-integrable. The process ϑ is d -dimensional and ϑ_i describes the number of units of asset i . The process ς is 1-dimensional and describes the number of units of the risk free asset. A strategy belongs to L^2 if ϑ is predictable, ς is a real-valued adapted process, and

$$\mathbb{E} \left[\int_0^T \vartheta^{\text{tr}}(s) d\langle M^X \rangle(s) \vartheta(s) + \left(\int_0^T |\vartheta^{\text{tr}}(s) dA^X(s)| \right)^2 \right] < \infty. \quad (2.5.2)$$

Consider a payment process $A^* = (A^*(t))_{t \in [0, T]}$, where $A^*(t)$ are the discounted payments related to the insurance contract during $[0, t]$. The cost process of an L^2 -strategy corresponding to the payment process A^* is defined by

$$C(\varphi, t) = A^*(t) + V(\varphi, t) - \int_0^t \vartheta(s) dX(s), \quad t \in [0, T].$$

One can interpret the cost process as the discounted amount one needs to provide in order to pay the insurance obligations and to fund the portfolio strategy φ . In the following, we consider 0-achieving strategies, i.e. strategies where $V(\varphi, T) = 0$ P -a.s. The risk process is given by

$$R(\varphi, t) = \mathbb{E}^P \left[(C(\varphi, T) - C(\varphi, t))^2 \middle| \mathcal{F}(t) \right]. \quad (2.5.3)$$

For a given strategy, the risk process measures the conditional expected quadratic deviation between the costs today and the costs at termination time T . We want to minimize the risk process in order to hedge the payment process as closely as possible. If the payment stream is hedgeable, the cost process is constant, and the risk process is zero-valued.

Before we can define the criterion of local risk-minimization, we need to introduce the concept of a small perturbation. To do this, we define a partition of $[0, T]$ as a set τ , where $\tau = \{\tau_0, \tau_1, \dots, \tau_n\}$, such that $\tau_0 = 0$, $\tau_n = T$ and $\tau_i < \tau_{i+1}$. A sequence of partitions $(\tau^n)_{n \in \mathbb{N}}$ is tending to the identity, if the maximum distance between the points τ_i converges to 0, i.e. if $\lim_{n \rightarrow \infty} \max(\tau_{i+1} - \tau_i | \tau_i, \tau_{i+1} \in \tau^n) \rightarrow 0$. Moreover, we define $\tilde{\sigma}$ by the equation $d\langle M^X \rangle = \tilde{\sigma} dD$, where D is a strictly increasing, predictable 1-dimensional process with $D(0) = 0$, and with $\langle M_i^X, M_j^X \rangle \ll D$ for all i, j . That is, $\langle M_i^X, M_j^X \rangle$ has a density with respect to D .

Definition 2.5.1. A pair $\Delta = (\delta, \varepsilon)$ consisting of an $\mathbb{R}^d$ -valued predictable process δ and an adapted real-valued process ε is called a small perturbation if $\langle \int \delta dM^X \rangle$ and $|\delta^{tr} \tilde{\sigma} \lambda|$ are bounded (uniformly in t, ω), the process $V(\Delta) = \delta \cdot X + \varepsilon$ is square-integrable, and $V(\Delta, T) = 0$ P -a.s. For each subinterval $(s, t]$ of $[0, T]$, we then define the small perturbation

$$\Delta_{|(s, t]} = \begin{cases} (\delta 1_{(s, t]}, \varepsilon 1_{[s, t)}) & \text{if } t < T, \\ (\delta 1_{(s, t]}, \varepsilon 1_{[s, t]}) & \text{if } t = T. \end{cases} \quad (2.5.4)$$

A small perturbation is used for the purpose of making a small modification of a strategy. For an L^2 -strategy φ , a small perturbation Δ and a partition τ , we define

$$r^\tau[\varphi, \Delta; A^*] = \sum_{t_i, t_{i+1} \in \tau} \frac{R(\varphi + \Delta_{|(t_i, t_{i+1}]}, t_i) - R(\varphi, t_i)}{\mathbb{E}[D(t_{i+1}) - D(t_i) | \mathcal{F}(t_i)]} 1_{(t_i, t_{i+1}]}. \quad (2.5.5)$$

We are now ready to define local risk-minimization.

Definition 2.5.2. For a fixed payment stream A^* , an L^2 -strategy φ is called locally risk-minimizing for A^* if for every small perturbation Δ and every increasing sequence $(\tau^n)_{n \in \mathbb{N}}$ of partitions tending to the identity, we have

$$\liminf_{n \rightarrow \infty} r^{\tau^n}[\varphi, \Delta; A^*] \geq 0 \quad P_D\text{-a.e.}, \quad (2.5.6)$$

where P_D is given by $P_D(E) = \int_0^T 1_E(\omega, s) dD(\omega, s)$.

The intuition behind Definition 2.5.2 is that the stochastic process $r^{\tau^n}[\varphi, \Delta; A^*]$ measures the total change of risk. The total risk (in terms of the risk process) is given by the sum of risks for all subintervals until time T when perturbing the strategy φ locally by Δ . The total change of risk when perturbing φ needs to be positive for φ to be locally risk-minimizing. The denominator of (2.5.5), $E[D(t_{i+1}) - D(t_i)|\mathcal{F}(t_i)]$, specifies a time scale for the measurement of the risk.

Fortunately, we do not need to work with this rather cumbersome definition, because we have a theorem stating how local risk-minimization is related to the cost process, and a theorem stating how local risk-minimization is related to the so-called Föllmer-Schweizer decomposition.

The following main result of local risk-minimization is from (Schweizer, 2008, Theorem 1.6). Recall that two square-integrable martingales are called strongly orthogonal if their product is a martingale.

Theorem 2.5.3. *Suppose the $\mathbb{R}^d$ -valued semimartingale X with decomposition (2.5.1) satisfies the structure condition and let A^* be a payment stream. If A^X is continuous, the following are equivalent for an L^2 -strategy φ :*

1. φ is locally risk-minimizing for A^* . (2.5.7)

2. φ is 0-achieving and mean self-financing, and the cost process $C(\varphi)$ is strongly orthogonal to M^X . (2.5.8)

In particular, the concept “locally risk-minimizing” does not depend on the choice of D .

The next result, Theorem 2.5.5, creates a connection between local risk-minimization and the so-called Föllmer-Schweizer decomposition, see (Schweizer, 2008, Proposition 5.2). It is a simple application of Theorem 2.5.3 and we use the result to obtain the locally risk-minimizing strategy for our market. First, we define the decomposition.

Definition 2.5.4. (Föllmer-Schweizer decomposition)

Let H be an $\mathcal{F}(T)$ -measurable random variable with $E^P[H^2] < \infty$. Then H admits a Föllmer-Schweizer decomposition, if H can be written as

$$H = H(0) + \int_0^T \vartheta^H(s) dX(s) + L^H(T) \quad P - a.s.,$$

where $H(0)$ is $\mathcal{F}(0)$ -measurable with $E^P[H^2(0)] < \infty$, ϑ^H is predictable and fulfilling (2.5.2), and L^H is a square integrable martingale null at time 0 and strongly orthogonal to M^X .

Theorem 2.5.5. *Suppose the $\mathbb{R}^d$ -valued semimartingale X satisfies the structure condition and A^X is continuous. Then a payment stream A^* admits a locally risk-minimizing L^2 -strategy φ if and only if $A^*(T)$ admits a Föllmer-Schweizer decomposition. In that case, $\varphi = (\vartheta, \varsigma)$ is given by*

$$\vartheta = \vartheta^{A_T^*}, \quad \varsigma = V^{A_T^*} - (\vartheta^{A_T^*})^{tr} X, \quad (2.5.9)$$

with

$$V^{A_T^*}(t) = A_T^{*,(0)} + \int_0^t \vartheta^{A_T^*}(s) dX(s) + L^{A^*}(t) - A^*(t), \quad t \in [0, T], \quad (2.5.10)$$

and then

$$C^{A^*}(\varphi, t) = A_T^{*,(0)} + L^{A^*}(t), \quad t \in [0, T]. \quad (2.5.11)$$

Proof: See Schweizer (2008).

We analyze the so-called minimal martingale measure as a starting point for the study of local risk-minimization within our model. The minimal martingale measure is an important tool when one wants to determine locally risk-minimizing strategies in the case where the underlying assets are continuous. In this case, the Föllmer-Schweizer decomposition under the original measure P can be obtained from a Galtchouk-Kunita-Watanabe decomposition (under some integrability conditions) under the minimal martingale measure, which leads to the locally risk-minimizing strategy. If the price processes are not continuous, this is in general not the case (see for example Vandaele and Vanmaele (2008) for an example of this phenomenon). The price processes are not continuous in our setting, so we cannot follow this path. However, the minimal martingale measure still plays an important role as we show in Section 2.7.

2.5.2 The minimal martingale measure

As noted above, the Föllmer-Schweizer decomposition can be obtained from the Galtchouk-Kunita-Watanabe decomposition under the minimal martingale measure if the discounted price processes are continuous. This makes the task of finding the locally risk-minimizing strategy similar to the task of finding a risk-minimizing strategy. We cannot benefit from this result, because the continuity condition is violated by the jumps of the counting process N , which generate jumps in the price process of the longevity bond. However, it is still relevant to determine the minimal martingale measure since the measure appears when applying the criterion of local risk-minimization.

The concept of a minimal martingale measure was formally defined in Föllmer and Schweizer (1990). Consider a d -dimensional semimartingale X with decomposition given by (2.5.1). A martingale measure $\hat{P}$ is said to be the minimal martingale measure, if $\hat{P}$ is equivalent to the physical measure P , $\hat{P} = P$ on $\mathcal{F}(0)$, and every square-integrable P -martingale strongly orthogonal to M^X remains a martingale under $\hat{P}$.

The minimal martingale measure was introduced for the first time in Schweizer (1988), and a formal definition and an existence result (for $d = 1$ and X continuous) were given in Föllmer and Schweizer (1990). For the more general case, see e.g. Föllmer and Schweizer (2010), the minimal martingale measure (if it exists) is defined by the likelihood process given by

$$\begin{aligned} \hat{\Lambda}(t) &= \mathcal{E} \left(- \int_0^t \lambda^X(\tau) dM^X(\tau) \right) \\ &= e^{-\int_0^t \lambda^X(\tau) dM^X(\tau) - \frac{1}{2} \int_0^t (\lambda^X)^{\text{tr}}(\tau) d[M^X](\tau) \lambda^X(\tau)} \\ &\quad \cdot \prod_{s \leq t} \left(\left(1 - (\lambda^X)^{\text{tr}}(s) \Delta M^X(s) \right) e^{(\lambda^X)^{\text{tr}}(s) \Delta M^X(s) + \frac{1}{2} ((\lambda^X)^{\text{tr}}(s) \Delta M^X(s))^2} \right). \end{aligned} \tag{2.5.12}$$

Here, $\mathcal{E}$ is the stochastic exponential and $[\cdot]$ is the quadratic variation process. We have that the minimal martingale measure exists if the three following conditions are fulfilled: $\hat{\Lambda}$ is strictly positive, which is the case if $(\lambda^X)^{\text{tr}} M^X < 1$, $\hat{\Lambda}$ is a true P -martingale, and $\hat{\Lambda}$ is square-integrable.

Under these conditions the minimal martingale measure is given by $\frac{d\hat{P}}{dP} = \hat{\Lambda}(T)$.

2.6 The minimal martingale measure for the market

As stated in the previous section, we cannot be sure that the minimal martingale measure exists. In this section, we find the minimal martingale measure for the financial market and prove that the minimal martingale measure exists under certain additional integrability conditions. We start by writing the dynamics of the zero coupon bond and the longevity bond under the original measure P .

It follows from Girsanov's Theorem and equation (2.2.3) that the dynamics under P of the discounted zero coupon bond are

$$dP^*(t, T) = h^r(r(t), t)P^*(t, T)\beta(t, T)\sigma^r dt - P^*(t, T)\beta(t, T)\sigma^r dW^r(t), \quad (2.6.1)$$

and from Girsanov's Theorem and equation (2.3.1), that the dynamics under P of the discounted longevity bond are

$$\begin{aligned} dZ^{*,Q}(x, t) &= -\nu^{Z,Q}(t)\lambda(x, t)g(t)dt - \eta^{Z,Q}(t)h^r(r(t), t)dt \\ &\quad + \eta^{Z,Q}(t)dW^r(t) + \nu^{Z,Q}(t)dM(x, t) + \rho^{Z,Q}(t)dW^\mu(t). \end{aligned} \quad (2.6.2)$$

We henceforth use the notation $X = (P^*, Z^*)$ and let A_1^X denote the drift part of the zero coupon bond and let M_1^X denote the martingale part. Similarly, we denote by A_2^X the drift of the longevity bond and by M_2^X the martingale part.

Hence, the dynamics of the drift parts are

$$dA_1^X(t) = h^r(r(t), t)P^*(t, T)\beta(t, T)\sigma^r dt, \quad (2.6.3)$$

$$dA_2^X(t) = -\nu^{Z,Q}(t)\lambda(x, t)g(t)dt - \eta^{Z,Q}(t)h^r(r(t), t)dt, \quad (2.6.4)$$

and the predictable variation and covariation processes are

$$\begin{aligned} d\langle M_1^X \rangle(t) &= (P^*(t, T)\beta(t, T)\sigma^r)^2 dt, \\ d\langle M_2^X \rangle(t) &= (\eta^{Z,Q}(t))^2 dt + (\nu^{Z,Q}(t))^2 \lambda(x, t)dt + (\rho^{Z,Q}(t))^2 dt, \\ d\langle M_1^X, M_2^X \rangle(t) &= -P^*(t, T)\beta(t, T)\sigma^r \eta^{Z,Q}(t)dt. \end{aligned}$$

Define processes $\lambda_1^X(t)$ and $\lambda_2^X(t)$ as the solution to the vector equation

$$dA^X(t) = d\langle M^X \rangle(t)\lambda^X(t).$$

The solution to the equation may be written informally as

$$\lambda_1^X(t) = -\frac{d\langle M_1^X, M_2^X \rangle(t)dA_2^X(t) - d\langle M_2^X \rangle(t)dA_1^X(t)}{d\langle M_1^X \rangle(t)d\langle M_2^X \rangle(t) - (d\langle M_1^X, M_2^X \rangle(t))^2}, \quad (2.6.5)$$

$$\lambda_2^X(t) = \frac{d\langle M_1^X \rangle(t)dA_2^X(t) - d\langle M_1^X, M_2^X \rangle(t)dA_1^X(t)}{d\langle M_1^X \rangle(t)d\langle M_2^X \rangle(t) - (d\langle M_1^X, M_2^X \rangle(t))^2}. \quad (2.6.6)$$

Proposition 2.6.1. *Assume that $\hat{\Lambda}$ is a square-integrable martingale. Then the minimal martingale measure for the present market denoted by $\hat{P}$ exists. The minimal martingale measure is given by*

$$\frac{d\hat{P}}{dP} = d\hat{\Lambda}(T),$$

where the likelihood process $\hat{\Lambda}$ is given by

$$\begin{aligned} d\hat{\Lambda}(t) &= \hat{\Lambda}(t-)\left(h^r(r(t),t)dW^r(t) \right. \\ &\quad \left. - \lambda_2^X(t)\left(\nu^{Z,Q}(t)dM(x,t) + \rho^{Z,Q}(t)dW^\mu(t)\right)\right), \\ \hat{\Lambda}(0) &= 1, \end{aligned} \tag{2.6.7}$$

and where

$$\lambda_2^X(t) = \frac{-\nu^{Z,Q}(t)\lambda(x,t)g(t)}{(\nu^{Z,Q}(t))^2\lambda(x,t) + (\rho^{Z,Q}(t))^2}. \tag{2.6.8}$$

Proof: It follows from Section 2.5.2 that the minimal martingale measure (if it exists) is given by the measure change, where the likelihood process is

$$\begin{aligned} d\hat{\Lambda}(t) &= \hat{\Lambda}(t-)\left(-\left(\lambda_1^X(t)\left(-P^*(t,T)\beta(t,T)\sigma^r\right) + \lambda_2^X(t)\eta^{Z,Q}(t)\right)dW^r(t) \right. \\ &\quad \left. - \lambda_2^X(t)\left(\nu^{Z,Q}(t)dM(x,t) + \rho^{Z,Q}(t)dW^\mu(t)\right)\right). \end{aligned} \tag{2.6.9}$$

To guarantee that the minimal martingale measure is in fact a martingale measure, we must have that the Girsanov kernel corresponding to the Brownian motion W^r for the short rate process is equal to h^r . This enables us to rewrite the likelihood process (2.6.9) in the simpler way given by (2.6.7), where

$$h^r = \lambda_1^X P^*(\cdot, T)\beta(\cdot, T)\sigma^r - \lambda_2^X \eta^{Z,Q}. \tag{2.6.10}$$

The condition that $\hat{\Lambda}$ is strictly positive is equal to the condition

$$\lambda_2^X(t)\nu^{Z,Q}(t) < 1. \tag{2.6.11}$$

We want to check that condition (2.6.11) holds true. By use of (2.6.6) and by doing some simplifications we get that

$$\lambda_2^X(t)\nu^{Z,Q}(t) = \frac{-(\nu^{Z,Q}(t))^2\lambda(x,t)g(t)}{(\nu^{Z,Q}(t))^2\lambda(x,t) + (\rho^{Z,Q}(t))^2}. \tag{2.6.12}$$

The term $(\nu^{Z,Q}(t))^2\lambda(x,t)$ is greater than or equal to 0 and less than or equal to the denominator of the right hand side of (2.6.12). That is, the condition (2.6.11) is fulfilled for $g > -1$. To obtain the expression for λ_2^X given by (2.6.8), we divide (2.6.12) by $\nu^{Z,Q}(t)$. □

We introduce the notation M^c and M^d for the continuous and the purely discontinuous martingale parts of $-\int \lambda^X dM^X$, respectively. By (Protter and Shimbo, 2008, Theorem 9), $\hat{\Lambda}$ is a martingale, if $-\int \lambda^X dM^X$ is a locally square-integrable martingale such that $\Delta\left(-\int \lambda^X dM^X\right) > -1$ and

$$\mathbb{E}^P \left[e^{\frac{1}{2}\langle M^c \rangle(T) + \langle M^d \rangle(T)} \right] < \infty.$$

Under the measure $\hat{P}$,

$$dW^{r,Q}(t) = dW^r(t) - h^r(r(t),t)dt \quad \text{and} \quad d\hat{W}^\mu(t) = dW^\mu(t) + \lambda_2^X(t)\nu^{Z,Q}(t)dt$$

are standard Brownian motions, and N is a pure jump process, where N has intensity $\hat{\lambda}(x, t) = \lambda(x, t) (1 - \lambda_2^X(t) \nu^{Z, Q}(t))$ and jumps of size 1.

We denote by $\hat{M}^Q$ the martingale under $\hat{P}$ associated with N . We note that we do not maintain the independence of the stochastic sources under the minimal martingale measure and that the minimal martingale measure does not belong to the class of measures given by (2.2.6).

The intuition of the result in Proposition 2.6.1 is that the drift of the price process is divided between the processes M and W^μ based on their variation processes. The requirement that the minimal martingale measure is a well-defined probability measure implies that the measure change (2.2.6) differs from the measure change in Dahl et al. (2008). If we had chosen the same measure change as in Dahl et al. (2008), the right hand side of (2.6.12) would have been

$$\frac{- (\nu^{Z, Q}(t))^2 \lambda(x, t) g(t) - \nu^{Z, Q}(t) \rho^{Z, Q}(t) h^\mu(t, \zeta(t))}{(\nu^{Z, Q}(t))^2 \lambda(x, t) + (\rho^{Z, Q}(t))^2} \quad (2.6.13)$$

for a given class of functions h^μ . In general we cannot find assumptions for the functions g and h^μ which ensure that (2.6.13) is smaller than 1.

From the above comments we see, that it is not trivial to have a market with a well defined minimal martingale measure. However, it will not impose any problems to the minimal martingale measure, if the market consists of a survivor swap instead of a longevity bond. The minimal martingale measure is still a probability measure for $g > -1$, since we can repeat the proof of Proposition 2.6.1, where we exchange the terms relating to the stochastic representation of the longevity bond by some similar terms for the survivor swap, which can be found in Dahl et al. (2008).

2.7 Local risk-minimization for the market

In this section, we find the locally risk-minimizing strategy for the market and compare it with the strategy found by risk-minimization under the minimal martingale measure. Moreover, we compare the strategy with the one found by following the procedure of Dahl et al. (2008).

The approach for finding the locally risk-minimizing strategy is to find the Föllmer-Schweizer decomposition of $A^*(T)$ and prove that the conditions in Theorem 2.5.5 are fulfilled. To achieve this, we impose certain restrictions for the decomposition, see (2.7.6)-(2.7.10), that lead to a set of equations that must be satisfied.

For this purpose we consider a general measure $\check{P}$ given by

$$\frac{d\check{P}}{dP} = \check{\Lambda}(T), \quad (2.7.1)$$

where the likelihood process is given by

$$d\check{\Lambda}(t) = \check{\Lambda}(t-) (\varrho^M(t) dM(x, t) + \varrho^{W^\mu}(t) dW^\mu(t) + \varrho^{W^r}(t) dW^r(t)), \quad \check{\Lambda}(0) = 1.$$

We have suppressed that the Girsanov kernels can depend on the underlying stochastic processes $r(t)$, $\mu(x, t)$ and $N(x, t-)$. We note that the minimal martingale measure belongs to the class of measures given by (2.7.1).

Under the measure $\check{P}$, $d\check{W}^r(t) = dW^r(t) - \varrho^{W^r}(t)dt$ is a standard Brownian motion, $d\check{W}^\mu(t) = dW^\mu(t) - \varrho^{W^\mu}(t)dt$ is a standard Brownian motion and N is a counting process with intensity $\lambda(x, t) (1 + \varrho^M(t))$ and jumps of size 1.

In the following, we denote by $\check{M}$ the martingale corresponding to the jump process N under the measure $\check{P}$. We assume that ϱ^M , ϱ^{W^μ} and ϱ^{W^r} are chosen such that the discounted price process

of the zero coupon bond and the discounted price process of the longevity bond are martingales under $\check{P}$. Under the original measure P , the dynamics of these discounted price processes are given by (2.6.1) and (2.6.2), respectively. The dynamics under the measure $\check{P}$ are given by

$$\begin{aligned} dP^*(t, T) &= h^r(r(t), t)P^*(t, T)\beta(t, T)\sigma^r dt - P^*(t, T)\beta(t, T)\sigma^r d\check{W}^r(t) \\ &\quad - P^*(t, T)\beta(t, T)\sigma^r \varrho^{W^r}(t)dt \end{aligned} \quad (2.7.2)$$

and

$$\begin{aligned} dZ^{*,Q}(x, t) &= -\nu^{Z,Q}(t)\lambda(x, t)g(t)dt - \eta^{Z,Q}(t)h^r(r(t), t)dt + \eta^{Z,Q}(t)d\check{W}^r(t) \\ &\quad + \nu^{Z,Q}(t)d\check{M}(t) + \rho^{Z,Q}(t)d\check{W}^\mu(t) + \nu^{Z,Q}(t)\lambda(x, t)\varrho^M(t)dt \\ &\quad + \rho^{Z,Q}(t)\varrho^{W^\mu}(t)dt + \eta^{Z,Q}(t)\varrho^{W^r}(t)dt. \end{aligned} \quad (2.7.3)$$

That is, $\check{P}$ is a martingale measure if and only if the drift of the two discounted risky assets equals 0. That is, we need the following equations to be fulfilled:

$$0 = -P^*(t, T)\beta(t, T)\sigma^r \varrho^{W^r}(t) + h^r(r(t), t)P^*(t, T)\beta(t, T)\sigma^r,$$

and

$$\begin{aligned} 0 &= \nu^{Z,Q}(t)\lambda(x, t)\varrho^M(t) + \rho^{Z,Q}(t)\varrho^{W^\mu}(t) + \eta^{Z,Q}(t)\varrho^{W^r}(t) \\ &\quad - \nu^{Z,Q}(t)\lambda(x, t)g(t) - \eta^{Z,Q}(t)h^r(r(t), t). \end{aligned}$$

For the purpose of explicitly finding a Föllmer-Schweizer decomposition of $A^*(T)$, we need a stochastic representation of the expected discounted payments of the insurance contract under $\check{P}$. That is, we need a stochastic representation of

$$V^{*,\check{P}}(t) = E^{\check{P}} \left[\int_{0-}^T dA^*(\tau) \middle| \mathcal{F}(t) \right], \quad (2.7.4)$$

which is given in Lemma 2.7.1. To obtain this, we assume that $E^{\check{P}} \left[\int_t^T e^{-\int_0^\tau r(s)ds} dA(\tau) \middle| \mathcal{F}(t) \right]$ is continuously differentiable with respect to t and twice continuously differentiable with respect to r and μ .

Lemma 2.7.1. *The $\check{P}$ -martingale $V^{*,\check{P}}(t)$ admits the representation*

$$V^{*,\check{P}}(t) = V^{*,\check{P}}(0) + \int_0^t \check{\eta}^V(\tau)d\check{W}^r(\tau) + \int_0^t \check{\rho}^V(\tau)d\check{W}^\mu(\tau) + \int_0^t \check{\nu}^V(\tau)d\check{M}(\tau),$$

where

$$\begin{aligned} \check{\eta}^V(t) &= e^{-\int_0^t r(s)ds} G_r(t, Y(t-))\sigma^r, \\ \check{\rho}^V(t) &= e^{-\int_0^t r(s)ds} G_\mu(t, Y(t-))\sigma^\mu(t, \mu(x, t)), \\ \check{\nu}^V(t) &= e^{-\int_0^t r(s)ds} \left(a^d(t) + G(t, r(t), N(x, t-) + 1, \mu(x, t)) \right. \\ &\quad \left. - G(t, r(t), N(x, t-), \mu(x, t)) \right), \end{aligned}$$

and

$$G(t, r(t), N(x, t), \mu(x, t)) = G(t, Y(t)) \equiv E^{\check{P}} \left[\int_t^T e^{-\int_t^\tau r(s)ds} dA(\tau) \middle| \mathcal{F}(t) \right].$$

The function G satisfies the following partial differential equation:

$$\begin{aligned}
0 = & G_t + G_{x_r} ((\gamma^r - \delta^r x_r) + \varrho^{W^r}(t) \sigma^r) \\
& + G_{x_\mu} \left(\Theta(x, t, x_\mu) + \varrho^{W^\mu}(t) \sqrt{x_\mu} \sqrt{\mu^\circ(x, t)} \sigma(x, t) \right) \\
& + \frac{1}{2} G_{x_r x_r} (\sigma^r)^2 + \frac{1}{2} G_{x_\mu x_\mu} x_\mu \mu^\circ(x, t) \sigma(x, t)^2 \\
& + \left(a^d(t) + \Delta G \right) \lambda(x, t) (1 + \varrho^M(t)) \\
& - \pi^c(t) (n - x_N) 1_{\{0 \leq t \leq \bar{T}\}} + a^p(t) (n - x_N) 1_{\{\bar{T} \leq t \leq T\}} - x_r G,
\end{aligned} \tag{2.7.5}$$

where $\Delta G = G(t, x_r, x_N + 1, x_\mu) - G(t, x_r, x_N, x_\mu)$.

The boundary condition is $G(T, x_r, x_N, x_\mu) = (n - x_N) a^r(\bar{T}) 1_{\{T = \bar{T}\}}$.

Moreover, we have an extra boundary condition if $\bar{T} < T$. The condition is:

$$G(\bar{T}^-, x_r, x_N, x_\mu) = G(\bar{T}, x_r, x_N, x_\mu) + (n - x_N) a^r(\bar{T}).$$

Note that we suppressed the arguments (t, x_r, x_N, x_μ) of G in Lemma 2.7.1. The PDE for G would be hard to solve numerically. This is because the solution depends on the entire grid and because G depends on four variables, which requires calculations in many grid points.

We use the notation G_t , G_{x_r} and G_{x_μ} for the derivatives of G with respect to the first, second, and fourth argument, respectively. For the double derivative of G with respect to the first argument we use the notation $G_{x_r x_r}$. The proof of the proposition is postponed to Appendix 2.A.

To find the risk-minimizing strategy, we need to find the Föllmer-Schweizer decomposition of the discounted payment stream at the terminal time T . In order to do so, we introduce the auxiliary process $\check{V}(\varphi, \cdot)$. Following the approach of Colwell and Elliott (1993) and Vandaele and Vanmaele (2008), it is possible to put up conditions for $\check{V}(\varphi, \cdot)$, which can be used to find the locally risk-minimizing strategy. The conditions are as follows:

- $\check{V}(\varphi, T) = \int_{0-}^T dA^*(\tau) = A^*(T).$ (2.7.6)

- $\check{V}(\varphi, t) = \check{V}(\varphi, 0) + \int_0^t \phi_1(s) dP^*(s, T) + \int_0^t \phi_2(s) dZ^{*,Q}(x, s) + \Gamma(t).$ (2.7.7)

- Γ is a martingale under P and strongly orthogonal to the martingale parts of the discounted risky assets. (2.7.8)

- The price processes of the discounted risky assets are martingales under $\check{P}$. (2.7.9)

- $\check{V}(\varphi, \cdot)$ is a martingale under $\check{P}$. (2.7.10)

These conditions help to specify the Föllmer-Schweizer decomposition, which leads to the locally risk-minimizing strategy in Theorem 2.7.2.

Condition (2.7.6) follows because $\check{V}(\varphi, T)$ should be equal to the quantity we want to find a Föllmer-Schweizer decomposition for. The conditions (2.7.7) and (2.7.8) are related to the requirements of the Föllmer-Schweizer decomposition. Condition (2.7.9) says that the measure defined by (2.7.1) is an equivalent martingale measure. Finally, we use condition (2.7.10) to obtain a unique measure $\check{P}$.

Theorem 2.7.2. Let ξ^* denote the number of zero coupon bonds, let ϖ^* denote the number of longevity bonds and let η^* be the amount deposited in the savings account. The locally risk-minimizing strategy φ^* is given by

$$\begin{aligned}\varphi^*(t) &= (\xi^*(t), \varpi^*(t), \eta^*(t)) \\ &= (\phi_1(t), \phi_2(t), \check{V}(\varphi, t) - A^*(t) - \phi_1(t)P^*(t, T) - \phi_2(t)Z^{*,Q}(t)),\end{aligned}\quad (2.7.11)$$

where ϕ_1 is given by

$$\begin{aligned}\phi_1(t) &= \frac{\check{\eta}^V(t)}{-P^*(t, T)\beta(t, T)\sigma^r} \\ &\quad + \frac{\eta^{Z, Q}(t) \left(\nu^{Z, Q}(t) \check{\nu}^V(t) \lambda(x, t) + \rho^{Z, Q}(t) \check{\rho}^V(t) \right)}{P^*(t, T)\beta(t, T)\sigma^r \left((\nu^{Z, Q}(t))^2 \lambda(x, t) + (\rho^{Z, Q}(t))^2 \right)},\end{aligned}\quad (2.7.12)$$

ϕ_2 is given by

$$\phi_2(t) = \frac{\nu^{Z, Q}(t) \check{\nu}^V(t) \lambda(x, t) + \rho^{Z, Q}(t) \check{\rho}^V(t)}{(\nu^{Z, Q}(t))^2 \lambda(x, t) + (\rho^{Z, Q}(t))^2} \quad (2.7.13)$$

and $\check{P} = \hat{P}$. The associated risk process is given by

$$\begin{aligned}R(\varphi^*, t) &= \int_t^T E^P \left[(\check{\rho}^V(\tau) - \phi_2(\tau) \rho^{Z, Q}(\tau))^2 \middle| \mathcal{F}(t) \right] d\tau \\ &\quad + \int_t^T E^P \left[(\check{\nu}^V(\tau) - \phi_2(\tau) \nu^{Z, Q}(\tau))^2 \lambda(x, \tau) \middle| \mathcal{F}(t) \right] d\tau.\end{aligned}$$

The proof of the theorem is postponed to Appendix 2.B.

We end this section by comparing the strategy found using risk-minimization under the minimal martingale measure and the strategy found using local risk-minimization. In order to do so, we first state the risk-minimizing strategy under the minimal martingale measure and give a description of how to obtain the result. The risk-minimizing strategy under the minimal martingale measure is given by

$$\begin{aligned}\bar{\varphi}(t) &= (\bar{\xi}^*(t), \bar{\varpi}^*(t), \bar{\eta}^*(t)) \\ &= (\bar{\phi}_1(t), \bar{\phi}_2(t), V^{*, \check{P}}(t) - A^*(t) - \bar{\phi}_1(t)P^*(t, T) - \bar{\phi}_2(t)Z^{*,Q}(t)),\end{aligned}\quad (2.7.14)$$

where

$$\begin{aligned}\bar{\phi}_1(t) &= \frac{\check{\eta}^V(t)}{-P^*(t, T)\beta(t, T)\sigma^r} \\ &\quad + \frac{\eta^{Z, Q}(t) \left(\nu^{Z, Q}(t) \check{\nu}^V(t) \hat{\lambda}(x, t) + \rho^{Z, Q}(t) \check{\rho}^V(t) \right)}{\left((\nu^{Z, Q}(t))^2 \hat{\lambda}(x, t) + (\rho^{Z, Q}(t))^2 \right) P^*(t, T)\beta(t, T)\sigma^r}\end{aligned}\quad (2.7.15)$$

and

$$\bar{\phi}_2(t) = \frac{\nu^{Z, Q}(t) \check{\nu}^V(t) \hat{\lambda}(x, t) + \rho^{Z, Q}(t) \check{\rho}^V(t)}{(\nu^{Z, Q}(t))^2 \hat{\lambda}(x, t) + (\rho^{Z, Q}(t))^2}.\quad (2.7.16)$$

We can get the result in the same manner as the results in Dahl et al. (2008), or we can calculate the coefficients given in (2.7.15) and (2.7.16) by use of the covariation processes. With informal notation, the coefficients are given by

$$\begin{pmatrix} \bar{\phi}_1(t) \\ \bar{\phi}_2(t) \end{pmatrix} = \begin{pmatrix} \frac{d\langle Z^{*,Q} \rangle(t) d\langle V^{*,\hat{P}}, P^* \rangle(t) - d\langle Z^{*,Q}, P^* \rangle(t) d\langle V^{*,\hat{P}}, Z^{*,Q} \rangle(t)}{\frac{d\langle Z^{*,Q} \rangle(t) d\langle P^* \rangle(t) - (d\langle Z^{*,Q}, P^* \rangle(t))^2}{d\langle P^* \rangle(t) d\langle V^{*,\hat{P}}, Z^{*,Q} \rangle(t) - d\langle Z^{*,Q}, P^* \rangle(t) d\langle V^{*,\hat{P}}, P^* \rangle(t)}} \\ \frac{d\langle Z^{*,Q} \rangle(t) d\langle P^* \rangle(t) - (d\langle Z^{*,Q}, P^* \rangle(t))^2}{d\langle Z^{*,Q} \rangle(t) d\langle P^* \rangle(t) - (d\langle Z^{*,Q}, P^* \rangle(t))^2} \end{pmatrix}, \quad (2.7.17)$$

where

$$V^{*,\hat{P}}(t) = E^{\hat{P}} \left[\int_{0-}^T dA^*(\tau) \middle| \mathcal{F}(t) \right].$$

Thus, the locally risk-minimizing strategy is given by (2.7.11), and the risk-minimizing strategy under the minimal martingale measure is given by (2.7.14). Since $\hat{P} = \hat{P}$, it follows that the two strategies are closely related. In fact, the only difference is that the compensator of the counting process N , which is the jump intensity, since the jump sizes are 1, under $\hat{P}$ is exchanged with the compensator of N under the original measure P . That is, $\hat{\lambda}(x, t)$, which appears in the case of risk-minimization under the minimal martingale measure, is exchanged by $\lambda(x, t)$ in the case of local risk-minimization. This difference is due to the fact that the predictable variation process of the N -term under P does not equal the predictable variation process under $\hat{P}$. However, the predictable variation processes of the Brownian motion terms are the same under P and $\hat{P}$.

2.7.1 Comparison with the results of Dahl et al. (2008) (Dahl et al. (2008))

We conclude by comparing the results of Theorem 2.7.2 with the results of (Dahl et al., 2008, Proposition 5.1). There are some difficulties in doing so, because the setups in the two papers are a little different. Firstly, there are two portfolios of lives in Dahl et al. (2008) and they are driven by a two-dimensional standard Brownian motion. In addition, Dahl et al. (2008) uses survivor swaps and not longevity bonds to hedge the mortality risk. Finally, in Dahl et al. (2008) the Brownian motion W^μ is affected by the change of measure from P to Q .

Taking these differences into account, we can adapt the result from Dahl et al. (2008) to the setup in the present paper and get that the risk-minimizing strategy under an equivalent martingale measure Q is given by Theorem 2.7.3. The result can be found by a calculation similar to (2.7.17), but done under the measure Q .

Theorem 2.7.3. *Let $\tilde{\xi}^*$ denote the number of zero coupon bonds, let $\tilde{\omega}^*$ denote the number of longevity bonds, and let $\tilde{\eta}^*$ be the deposit in the savings account. The risk-minimizing strategy under the measure Q is given by*

$$\begin{aligned} \tilde{\varphi}^*(t) &= (\tilde{\xi}^*(t), \tilde{\omega}^*(t), \tilde{\eta}^*(t)) \\ &= \left(\tilde{\phi}_1(t), \tilde{\phi}_2(t), \tilde{V}(\tilde{\varphi}, t) - A^*(t) - \tilde{\phi}_1(t)P^*(t, T) - \tilde{\phi}_2(t)Z^{*,Q}(t) \right), \end{aligned} \quad (2.7.18)$$

where

$$\begin{aligned} \tilde{\phi}_1(t) &= \frac{\eta^V(t)}{-P^*(t, T)\beta(t, T)\sigma^r} \\ &\quad + \frac{\eta^{Z,Q}(t) (\nu^{Z,Q}(t)\nu^V(t)\lambda^Q(x, t) + \rho^{Z,Q}(t)\rho^V(t))}{\left((\nu^{Z,Q}(t))^2 \lambda^Q(x, t) + (\rho^{Z,Q}(t))^2 \right) P^*(t, T)\beta(t, T)\sigma^r}, \end{aligned} \quad (2.7.19)$$

and where

$$\tilde{\phi}_2(t) = \frac{\nu^{Z,Q}(t)\nu^V(t)\lambda^Q(x,t) + \rho^{Z,Q}(t)\rho^V(t)}{(\nu^{Z,Q}(t))^2\lambda^Q(x,t) + (\rho^{Z,Q}(t))^2}. \quad (2.7.20)$$

There are two differences between the locally risk-minimizing strategy found in Theorem 2.7.2 and the risk-minimizing strategy found under the measure Q given in Theorem 2.7.3. The differences exist because certain quantities in the expressions of the optimal number of risky assets are calculated under different probability measures for the locally risk-minimizing strategy and the risk-minimizing strategy under Q , respectively. First, we note that the compensator of the jump process is calculated under P and Q , respectively. In addition, we see that the integrands of the stochastic representation of the intrinsic value process are calculated under the minimal martingale measure and under the measure Q , respectively. The size of the impact on the strategy of these differences depends of course on the difference of the two measures and the other parameters of the model.

Appendix

2.A Proof of Lemma 2.7.1

We utilize that

$$V^{*,\check{P}}(t) = A^*(t) + e^{-\int_0^t r(s)ds} \mathbf{E}^{\check{P}} \left[\int_t^T e^{-\int_t^\tau r(s)ds} dA(\tau) \middle| \mathcal{F}(t) \right], \quad (2.A.1)$$

and that $V^{*,\check{P}}$ is a martingale. Thus, we do not focus on the drift of $V^{*,\check{P}}$. We denote by $\theta_i, i = 1, 2, 3, 4$, some processes concerning the drift. We do not need to specify the exact form of these processes, since the drift of a martingale is zero. Let $Y = (r, N, \mu)^{\text{tr}}$. By Itô's Lemma we get

$$\begin{aligned} dV^{*,\check{P}}(t) &= dA^*(t) - r(t)e^{-\int_0^t r(s)ds} G(t, Y(t))dt \\ &\quad + e^{-\int_0^t r(s)ds} dG(t, Y(t)), \end{aligned} \quad (2.A.2)$$

and

$$\begin{aligned} dG(t, Y(t)) &= \theta_1(t)dt + G_r(t, Y(t-))dr(t) + G_\mu(t, Y(t-))d\mu(x, t) \\ &\quad - (n - N(x, \bar{T}))a^r(\bar{T})d1_{\{t \geq \bar{T}\}} + G(t, Y(t-) + \Delta Y(t)) - G(t, Y(t-)). \end{aligned}$$

Calculations yield that

$$\begin{aligned} dG(t, Y(t)) &= \theta_2(t)dt + G_r(t, Y(t-))\sigma^r d\check{W}^r(t) + G_\mu(t, Y(t-))\sigma^\mu(t, \mu(x, t-))d\check{W}^\mu(t) \\ &\quad + G(t, r(t-), N(x, t-) + \Delta N(x, t), \mu(t-)) - G(t, r(t-), N(x, t-), \mu(t-)) \\ &\quad - (n - N(x, \bar{T}))a^r(\bar{T})d1_{\{t \geq \bar{T}\}}. \end{aligned}$$

We use that the jump sizes of N are 1 and write the dynamics of G in terms of the compensated jump process.

$$\begin{aligned} dG(t, Y(t)) &= \theta_3(t)dt + G_r(t, Y(t-))\sigma^r d\check{W}^r(t) + G_\mu(t, Y(t-))\sigma^\mu(t, \mu(x, t))d\check{W}^\mu(t) \\ &\quad + (G(t, r(t-), N(x, t-) + 1, \mu(t-)) - G(t, r(t-), N(x, t-), \mu(t-)))d\check{M}(t) \\ &\quad - (n - N(x, \bar{T}))a^r(\bar{T})d1_{\{t \geq \bar{T}\}}. \end{aligned}$$

Since

$$dA^*(t) = \theta_4(t) + e^{-\int_0^t r(s)ds} a^d(t)d\check{M}(t) + e^{-\int_0^t r(s)ds} (n - N(x, \bar{T}))a^r(\bar{T})d1_{\{t \geq \bar{T}\}}$$

and $\theta_3(t) = -\theta_4(t)$, the stochastic representation given in the theorem follows.

To obtain the PDE for G we again utilize that $V^{*,\check{P}}$ is a martingale, thus it has drift equal to 0, and is given by (2.A.1). Moreover, we have that (using shorthand notation)

$$\begin{aligned} dG(t, Y(t)) &= G_t dt + G_r dr(t) + G_\mu d\mu(t) + \frac{1}{2} G_{r,r} d[r, r]_t + \frac{1}{2} G_{\mu,\mu} d[\mu, \mu]_t \\ &\quad + G(t, r(t-), N(x, t-)) + \Delta N(x, t), \mu(x, t-) - G(t, r(t-), N(x, t-), \mu(x, t-)) \\ &\quad - (n - N(x, \bar{T})) a^r(\bar{T}) d\mathbf{1}_{\{t \geq \bar{T}\}}. \end{aligned}$$

Note that some zero-valued quadratic covariation terms have been disregarded in the equation.

The dynamics of the interest rate under the measure $\check{P}$ are

$$dr(t) = (\gamma^r - \delta^r r(t))dt + \varrho^{W^r}(t) \sigma^r dt + \sigma^r d\check{W}^r(t),$$

whereas the dynamics of the mortality intensity under $\check{P}$ are

$$\begin{aligned} d\mu(x, t) &= \Theta(x, t, \mu(x, t)) dt + \varrho^{W^\mu}(t) \sigma^\mu(t, \mu(x, t)) dt \\ &\quad + \sigma^\mu(t, \mu(x, t)) d\check{W}^\mu(t). \end{aligned}$$

In addition, the compensator of the process N is given by $\lambda(x, t) (1 + \varrho^M(t))$.

We split the dynamics into a drift part and martingale part and use that the drift part is equal to 0. We multiply the drift of the right hand side of (2.A.2) by $e^{\int_0^t r(s) ds}$ and obtain the partial differential equation given by (2.7.5).

The first boundary condition follows, since, with probability 1, there is no lump sum payments at time T , unless $T = \bar{T}$, and because the lump sum payment at time T is given by $(n - N(x, \bar{T})) a^r(\bar{T})$ if $T = \bar{T}$. If $\bar{T} < T$ and $n - N(x, \bar{T}) > 0$, there is a jump in the G -process corresponding to the lump sum payment at time $\bar{T}$. This is reflected by the second boundary condition. This completes the proof. □

2.B Proof of Theorem 2.7.2

First, we use (2.7.10), (2.7.6), and (2.7.4) to obtain:

$$\check{V}(\varphi, t) = E^{\check{P}} [\check{V}(\varphi, T) | \mathcal{F}(t)] = E^{\check{P}} [A^*(T) | \mathcal{F}(t)] = V^{*,\check{P}}(t). \quad (2.B.1)$$

Next, we find the dynamics of the process Γ in (2.7.7) under the measure P . Since we know that Γ is a martingale under P , thus, has no drift, we get an equation, which has to be fulfilled by the decomposition of $\check{V}(\varphi, \cdot)$.

We use (2.7.7), (2.B.1) and the representation given in Lemma 2.7.1 to obtain the following representation of Γ under $\check{P}$:

$$\begin{aligned} \Gamma(t) &= \check{V}(\varphi, t) - \check{V}(\varphi, 0) - \int_0^t \phi_1(\tau) dP^*(\tau, T) - \int_0^t \phi_2(\tau) dZ^{*,Q}(x, \tau) \\ &= \int_0^t \check{\eta}^V(\tau) d\check{W}^r(\tau) + \int_0^t \check{\rho}^V(\tau) d\check{W}^\mu(\tau) + \int_0^t \check{\nu}^V(\tau) d\check{M}(\tau) \\ &\quad - \int_0^t \phi_1(\tau) dP^*(\tau, T) - \int_0^t \phi_2(\tau) dZ^{*,Q}(x, \tau). \end{aligned}$$

By writing the dynamics of the stochastic sources and the price processes of the discounted claims under the measure P , we get the following representation:

$$\begin{aligned}
\Gamma(t) &= \int_0^t \check{\eta}^V(\tau) dW^r(\tau) + \int_0^t \check{\rho}^V(\tau) dW^\mu(\tau) + \int_0^t \check{\nu}^V(\tau) dM(x, \tau) \\
&\quad - \int_0^t \check{\eta}^V(\tau) \varrho^{Wr}(\tau) d\tau - \int_0^t \check{\rho}^V(\tau) \varrho^{W^\mu}(\tau) d\tau - \int_0^t \check{\nu}^V(\tau) \lambda(x, \tau) \varrho^M(\tau) d\tau \\
&\quad + \int_0^t \phi_1(\tau) (P^*(\tau, T) \beta(\tau, T) \sigma^r dW^r(\tau) - h^r(r(\tau), \tau) P^*(\tau, T) \beta(\tau, T) \sigma^r d\tau) \\
&\quad - \int_0^t \phi_2(\tau) \left(-\nu^{Z,Q}(\tau) \lambda(x, \tau) g(\tau) d\tau - \eta^{Z,Q}(\tau) h^r(r(\tau), \tau) d\tau \right. \\
&\quad \left. + \eta^{Z,Q}(\tau) dW^r(\tau) + \nu^{Z,Q}(\tau) dM(x, \tau) + \rho^{Z,Q}(\tau) dW^\mu(\tau) \right).
\end{aligned}$$

Since Γ is a martingale under P we get the following equation:

$$\begin{aligned}
&\phi_2(t) (\nu^{Z,Q}(t) \lambda(x, t) g(t) + \eta^{Z,Q}(t) h^r(r(t), t)) - \phi_1(t) h^r(r(t), t) P^*(t, T) \beta(t, T) \sigma^r \\
&= \check{\eta}^V(t) \varrho^{Wr}(t) + \check{\rho}^V(t) \varrho^{W^\mu}(t) + \check{\nu}^V(t) \lambda(x, t) \varrho^M(t).
\end{aligned} \tag{2.B.2}$$

That is, we can write Γ in this more appealing way:

$$\begin{aligned}
\Gamma(t) &= \int_0^t \left(\check{\eta}^V(\tau) + \phi_1(\tau) P^*(\tau, T) \beta(\tau, T) \sigma^r - \phi_2(\tau) \eta^{Z,Q}(\tau) \right) dW^r(\tau) \\
&\quad + \int_0^t (\check{\rho}^V(\tau) - \phi_2(\tau) \rho^{Z,Q}(\tau)) dW^\mu(\tau) \\
&\quad + \int_0^t (\check{\nu}^V(\tau) - \phi_2(\tau) \nu^{Z,Q}(\tau)) dM(x, \tau).
\end{aligned} \tag{2.B.3}$$

We get the next equations from the condition that Γ is strongly orthogonal to the martingale parts of the discounted risky assets under P . We start by calculating the covariation processes between Γ and the zero coupon bond and the longevity bond, respectively.

$$\begin{aligned}
&d \langle \Gamma(\cdot), P^*(\cdot, T) \rangle(t) \\
&= -P^*(t, T) \beta(t, T) \sigma^r \left(\check{\eta}^V(t) + \phi_1(t) P^*(t, T) \beta(t, T) \sigma^r - \phi_2(t) \eta^{Z,Q}(t) \right) dt.
\end{aligned} \tag{2.B.4}$$

We denote by $Z_{\text{MG}}^{*,Q}$ the martingale part of the price process of the longevity bond under Q . The process has dynamics

$$dZ_{\text{MG}}^{*,Q}(t) = \eta^{Z,Q}(t) dW^r(t) + \nu^{Z,Q}(t) dM(x, t) + \rho^{Z,Q}(t) dW^\mu(t).$$

We find the covariation process between Γ and $Z^{*,Q}$ to be

$$\begin{aligned}
&d \langle \Gamma(\cdot), Z^{*,Q}(\cdot) \rangle(t) \\
&= d \left\langle \Gamma(\cdot), Z_{\text{MG}}^{*,Q}(\cdot) \right\rangle(t) \\
&= \eta^{Z,Q}(t) (\check{\eta}^V(t) + \phi_1(t) P^*(t, T) \beta(t, T) \sigma^r - \phi_2(t) \eta^{Z,Q}(t)) dt \\
&\quad + \nu^{Z,Q}(t) (\check{\nu}^V(t) - \phi_2(t) \nu^{Z,Q}(t)) \lambda(x, t) dt \\
&\quad + \rho^{Z,Q}(t) (\check{\rho}^V(t) - \phi_2(t) \rho^{Z,Q}(t)) dt.
\end{aligned} \tag{2.B.5}$$

By setting the two covariation processes (2.B.4) and (2.B.5) equal to 0, we obtain a system of equations, which we can solve for $(\phi_1(t), \phi_2(t))$. The solution is given by

$$\begin{aligned} \phi_1(t) = & \frac{\check{\eta}^V(t)}{-P^*(t, T)\beta(t, T)\sigma^r} \\ & + \frac{\eta^{Z, Q}(t) \left(\nu^{Z, Q}(t) \check{\nu}^V(t) \lambda(x, t) + \rho^{Z, Q}(t) \check{\rho}^V(t) \right)}{P^*(t, T)\beta(t, T)\sigma^r \left((\nu^{Z, Q}(t))^2 \lambda(x, t) + (\rho^{Z, Q}(t))^2 \right)} \end{aligned} \quad (2.B.6)$$

and

$$\phi_2(t) = \frac{\nu^{Z, Q}(t) (\check{\nu}^V(t)) \lambda(x, t) + \rho^{Z, Q}(t) \check{\rho}^V(t)}{(\nu^{Z, Q}(t))^2 \lambda(x, t) + (\rho^{Z, Q}(t))^2}. \quad (2.B.7)$$

Now we have found ϕ_1 and ϕ_2 for a given measure $\check{P}$, but we still need to determine the measure $\check{P}$.

We get the next equations from the fact that the discounted risky assets have to be martingales under $\check{P}$. The equation

$$h^r(r(t), t) P^*(t, T) \beta(t, T) \sigma^r = P^*(t, T) \beta(t, T) \sigma^r \varrho^{W^r}(t) \quad (2.B.8)$$

needs to hold in order to guarantee that the zero coupon bond is a martingale. Moreover, the equation

$$\begin{aligned} & \nu^{Z, Q}(t) \lambda(x, t) g(t) + \eta^{Z, Q}(t) h^r(r(t), t) \\ & = \nu^{Z, Q}(t) \lambda(x, t) \varrho^M(t) + \rho^{Z, Q}(t) \varrho^{W^\mu}(t) + \eta^{Z, Q}(t) \varrho^{W^r}(t) \end{aligned} \quad (2.B.9)$$

needs to hold for the longevity bond to be a martingale.

We insert the values of ϕ_1 and ϕ_2 in equation (2.B.2). Together with (2.B.8) and (2.B.9) this lead to a system of 3 equations. Now, the task is to find values for ϱ^{W^r} , ϱ^{W^μ} and ϱ^M fulfilling the equations. In order to do so, we insert the coefficients given by the likelihood process (2.6.7) of the minimal martingale measure and find that they constitute a claim-independent solution to the equation system. This approach is inspired by the results of Vandaele and Vanmaele (2008). In the following, it is worth noticing that we have a nice representation of $\lambda_2^X(t) \nu^{Z, Q}(t)$ given by (2.6.8).

First note, that (2.B.8) holds for $\varrho^{W^r} = h^r(r(t), t)$, which is the Girsanov kernel related to W^r of the minimal martingale measure. We now insert the coefficients from the minimal martingale measure into equation (2.B.9). This yields the equation:

$$\begin{aligned} & \nu^{Z, Q}(t) \lambda(x, t) g(t) + \eta^{Z, Q}(t) h^r(r(t), t) \\ & = -\lambda_2^X(t) \left(\nu^{Z, Q}(t) \lambda(x, t) \nu^{Z, Q}(t) + (\rho^{Z, Q}(t))^2 \right) + \eta^{Z, Q}(t) h^r(r(t), t). \end{aligned} \quad (2.B.10)$$

Using (2.6.8), we see that

$$\begin{aligned} & \nu^{Z, Q}(t) \lambda(x, t) g(t) + \eta^{Z, Q}(t) h^r(r(t), t) \\ & = \frac{\nu^{Z, Q}(t) \lambda(x, t) g(t) \left((\nu^{Z, Q}(t))^2 \lambda(x, t) + (\rho^{Z, Q}(t))^2 \right)}{(\nu^{Z, Q}(t))^2 \lambda(x, t) + (\rho^{Z, Q}(t))^2} + \eta^{Z, Q}(t) h^r(r(t), t). \end{aligned}$$

We reduce this and get

$$\nu^{Z,Q}(t)\lambda(x,t)g(t) + \eta^{Z,Q}(t)h^r(r(t),t) = \nu^{Z,Q}(t)\lambda(x,t)g(t) + \eta^{Z,Q}(t)h^r(r(t),t),$$

which shows that equation (2.B.9) holds.

Inserting ϕ_1 , ϕ_2 , and the coefficients from the minimal martingale measure into equation (2.B.2) yields

$$\begin{aligned} & \frac{(\nu^{Z,Q}(t) (\check{\nu}^V(t)) \lambda(x,t) + \rho^{Z,Q}(t) \check{\rho}^V(t)) (\nu^{Z,Q}(t) \lambda(x,t) g(t) + \eta^{Z,Q}(t) h^r(r(t),t))}{(\nu^{Z,Q}(t))^2 \lambda(x,t) + (\rho^{Z,Q}(t))^2} \\ & + h^r(r(t),t) \check{\eta}^V(t) \\ & + \frac{h^r(r(t),t) P^*(t,T) \beta(t,T) \sigma^r (\eta^{Z,Q}(t) (\nu^{Z,Q}(t) \check{\nu}^V(t) \lambda(x,t) + \rho^{Z,Q}(t) \check{\rho}^V(t)))}{-P^*(t,T) \beta(t,T) \sigma^r ((\nu^{Z,Q}(t))^2 \lambda(x,t) + (\rho^{Z,Q}(t))^2)} \\ & = \check{\eta}^V(t) h^r(r(t),t) + \frac{(\nu^{Z,Q}(t) \lambda(x,t) g(t)) (\check{\rho}^V(t) \rho^{Z,Q}(t) + \check{\nu}^V(t) \lambda(x,t) \nu^{Z,Q}(t))}{(\nu^{Z,Q}(t))^2 \lambda(x,t) dt + (\rho^{Z,Q}(t))^2}. \end{aligned}$$

Since several of these terms cancel out, we get the equation

$$\begin{aligned} & \frac{(\nu^{Z,Q}(t) (\check{\nu}^V(t)) \lambda(x,t) + \rho^{Z,Q}(t) \check{\rho}^V(t)) (\nu^{Z,Q}(t) \lambda(x,t) g(t))}{(\nu^{Z,Q}(t))^2 \lambda(x,t) + (\rho^{Z,Q}(t))^2} + h^r(r(t),t) \check{\eta}^V(t) \\ & = \check{\eta}^V(t) h^r(r(t),t) + \frac{\nu^{Z,Q}(t) \lambda(x,t) g(t) (\check{\rho}^V(t) \rho^{Z,Q}(t) + \check{\nu}^V(t) \lambda(x,t) \nu^{Z,Q}(t))}{(\nu^{Z,Q}(t))^2 \lambda(x,t) dt + (\rho^{Z,Q}(t))^2}. \end{aligned}$$

This shows, that equation (2.B.2) indeed holds.

Before we can apply Theorem 2.5.5, we also need to check, that the decomposition is in fact a Föllmer-Schweizer decomposition. Note that $\check{V}(\varphi, T) = A^*(T)$ is $\mathcal{F}(T)$ -measurable and that $A^*(T)$ has second order moment, since all payments, the time length and the number of policy holders are bounded. In addition, it follows that $\check{V}(\varphi, 0) = -n\pi^s(0) + n\check{V}_p^{*,Q}(t)$ is $\mathcal{F}(0)$ -measurable and clearly has second order moment under P . We see from (2.7.12) and (2.7.13), that ϕ_1 and ϕ_2 are clearly predictable. Earlier in the proof, we have shown, that Γ is a martingale and strongly orthogonal to M^X . By (2.B.3), we also have that Γ is null at time 0. In principle, we also need to show that

$$\begin{aligned} & \mathbb{E} \left[\int_0^T (\phi_1(s), \phi_2(s)) d\langle M^X \rangle(s) (\phi_1(s), \phi_2(s))^{\text{tr}} \right] \\ & + \mathbb{E} \left[\left(\int_0^T |(\phi_1(s), \phi_2(s)) dA^X(s)| \right)^2 \right] < \infty, \end{aligned} \tag{2.B.11}$$

and that

$$\Gamma \text{ is square-integrable.} \tag{2.B.12}$$

We have showed that the coefficients of the minimal martingale measure fulfill the necessary equations and can be used to find the Föllmer-Schweizer decomposition of the claim.

We finally check some extra conditions in order to ensure, that we can use Theorem 2.5.5. We write down these conditions with the notation from equation (2.5.1):

$$A^X \text{ is continuous,} \tag{2.B.13}$$

$$M^X \text{ is locally square integrable starting in 0,} \tag{2.B.14}$$

$$\lambda^X \text{ is predictable and in } L_{\text{loc}}^2(M^X). \tag{2.B.15}$$

Condition (2.B.13) is obviously fulfilled (the drift parts of the price processes of the discounted risky assets are continuous, since the Lebesgue measure for the time points of the jumps is zero). We show that M^X is locally square integrable by showing the more restrictive condition that M^X is square integrable. We do this by showing that $[M^X]$ is integrable, which we do by showing that each entry of the vector is integrable. The first entry is:

$$\begin{aligned} \mathbb{E}^P \left[\left[\int_0^T P^*(t, T) \beta(t, T) \sigma^r dW^r(t) \right] \right] &= \mathbb{E}^P \left[\int_0^T (P^*(t, T) \beta(t, T) \sigma^r)^2 d[W^r(t)] \right] \\ &= \mathbb{E}^P \left[\int_0^T (P^*(t, T) \beta(t, T) \sigma^r)^2 dt \right] \\ &= \int_0^T \mathbb{E}^P \left[(P^*(t, T))^2 \right] (\beta(t, T) \sigma^r)^2 dt < \infty, \end{aligned} \quad (2.B.16)$$

where we have used that a log-normal distribution has variance. The second entry is:

$$\begin{aligned} &\mathbb{E}^P \left[\left[\int_0^T \eta^{Z, Q}(t) dW^r(t) + \nu^{Z, Q}(t) dM(x, t) + \rho^{Z, Q}(t) dW^\mu(t) \right] \right] \\ &= \mathbb{E}^P \left[\int_0^T (\eta^{Z, Q}(t))^2 dt + (\nu^{Z, Q}(t))^2 dN(x, t) + (\rho^{Z, Q}(t))^2 dt \right] \\ &\leq \int_0^T (\sigma^r n)^2 \int_t^T \beta^2(t, \tau) \mathbb{E}^P \left[(P^*(t, \tau))^2 \right] d\tau dt \\ &\quad + \int_0^T \int_t^T \mathbb{E}^P \left[(P^*(t, \tau))^2 \lambda(x, \tau) \right] d\tau dt \\ &\quad + \int_0^T (\sigma n(1 + g(t)))^2 \mu^\circ(x, t) \int_t^T (\beta^\mu(t, \tau))^2 \mathbb{E}^P \left[\mu(x, t) (P^*(t, \tau))^2 \right] d\tau dt \\ &< \infty, \end{aligned} \quad (2.B.17)$$

where we in the first inequality have used that $\mathcal{S}^Q < 1$, since the mortality intensity is positive, and we have used Jensen's inequality to interchange the integral and the quadratic function. For the second inequality we have used that a log-normal distribution has variance and that a CIR-process has first moment. We also note that M^X is starting in 0. That is, we have that (2.B.14) holds.

It is clear from (2.6.8) and (2.6.10) that λ^X is predictable. Finally, we prove that $\lambda^X \in L_{loc}^2(M^X)$ by showing that

$$\mathbb{E}^P \left[\int_0^T (\lambda^X)^{\text{tr}}(s) d\langle M^X \rangle(s) \lambda^X(s) \right] = \mathbb{E}^P \left[\int_0^T (\lambda^X)^{\text{tr}}(s) dA^X(s) \right] < \infty, \quad (2.B.18)$$

where the equality follows by the definition of λ^X . In the above equation, λ_2^X is given by (2.6.8) and λ_1^X is given by equation (2.6.10). We can write λ_1^X as

$$\begin{aligned} \lambda_1^X &= \frac{h^r + \lambda_2^X \eta^{Z, Q}}{P^*(\cdot, T) \beta(\cdot, T) \sigma^r} \\ &= \frac{h^r}{P^*(\cdot, T) \beta(\cdot, T) \sigma^r} + \frac{-\nu^{Z, Q}(t) \lambda(x, t) g(t) \eta^{Z, Q}}{\left((\nu^{Z, Q}(t))^2 \lambda(x, t) + (\rho^{Z, Q}(t))^2 \right) P^*(\cdot, T) \beta(\cdot, T) \sigma^r}. \end{aligned} \quad (2.B.19)$$

Moreover, dA^X is given by (2.6.3) and (2.6.4). Thus, we get that $E^P \left[\int_0^T (\lambda^X)^{\text{tr}}(s) dA^X(s) \right]$ is equal to

$$\begin{aligned}
& E^P \left[\int_0^T (h^r(r(s), s))^2 + \frac{-\nu^{Z,Q}(s)\lambda(x, s)g(s)\eta^{Z,Q}(s)h^r(r(s), s)}{(\nu^{Z,Q}(s))^2\lambda(x, s) + (\rho^{Z,Q}(s))^2} ds \right] \\
& + E^P \left[\int_0^T \left(\frac{-\nu^{Z,Q}(s)\lambda(x, s)g(s) (-\nu^{Z,Q}(s)\lambda(x, s)g(s) - \eta^{Z,Q}(s)h^r(r(s), s))}{(\nu^{Z,Q}(s))^2\lambda(x, s) + (\rho^{Z,Q}(s))^2} \right) ds \right] \\
& = E^P \left[\int_0^T \left(\frac{\tilde{c}}{\sigma^r} + \frac{cr(t)}{\sigma^r} \right)^2 + \frac{(\nu^{Z,Q}(s)\lambda(x, s)g(s))^2}{(\nu^{Z,Q}(s))^2\lambda(x, s) + (\rho^{Z,Q}(s))^2} ds \right] \\
& \leq E^P \left[\int_0^T \left(\frac{\tilde{c}}{\sigma^r} + \frac{cr(t)}{\sigma^r} \right)^2 + \lambda(x, s) (g(s))^2 ds \right] \\
& = \int_0^T E^P \left[\left(\frac{\tilde{c}}{\sigma^r} + \frac{cr(t)}{\sigma^r} \right)^2 + \lambda(x, s) (g(s))^2 \right] ds < \infty,
\end{aligned}$$

where the last equality follows because the Vasiček process has variance and the CIR-process has first moment. By Theorem 2.5.5 we have that the first part of the theorem holds. The risk process is given by

$$\begin{aligned}
R(\varphi^*, t) &= E^P \left[(C(\varphi^*, T) - C(\varphi^*, t))^2 \middle| \mathcal{F}(t) \right] \\
&= E^P \left[\left(L^{A^*}(T) - L^{A^*}(t) \right)^2 \middle| \mathcal{F}(t) \right] \\
&= E^P \left[(\Gamma(T) - \Gamma(t))^2 \middle| \mathcal{F}(t) \right] \\
&= E^P \left[\left(\int_t^T \left(\check{\eta}^V(\tau) + \phi_1(\tau)P^*(\tau, T)\beta(\tau, T)\sigma^r - \phi_2(\tau)\eta^{Z,Q}(\tau) \right) dW^r(\tau) \right. \right. \\
&\quad \left. \left. + \int_t^T (\check{\rho}^V(\tau) - \phi_2(\tau)\rho^{Z,Q}(\tau)) dW^\mu(\tau) \right. \right. \\
&\quad \left. \left. + \int_t^T (\check{\nu}^V(\tau) - \phi_2(\tau)\nu^{Z,Q}(\tau)) dM(x, \tau) \right)^2 \middle| \mathcal{F}(t) \right] \\
&= \int_t^T E^P \left[\left(\check{\eta}^V(\tau) + \phi_1(\tau)P^*(\tau, T)\beta(\tau, T)\sigma^r - \phi_2(\tau)\eta^{Z,Q}(\tau) \right)^2 \middle| \mathcal{F}(t) \right] d\tau \\
&\quad + \int_t^T E^P \left[\left(\check{\rho}^V(\tau) - \phi_2(\tau)\rho^{Z,Q}(\tau) \right)^2 \middle| \mathcal{F}(t) \right] d\tau \\
&\quad + \int_t^T E^P \left[\left(\check{\nu}^V(\tau) - \phi_2(\tau)\nu^{Z,Q}(\tau) \right)^2 \lambda(x, \tau) \middle| \mathcal{F}(t) \right] d\tau \\
&= \int_t^T E^P \left[\left(\check{\rho}^V(\tau) - \phi_2(\tau)\rho^{Z,Q}(\tau) \right)^2 \middle| \mathcal{F}(t) \right] d\tau \\
&\quad + \int_t^T E^P \left[\left(\check{\nu}^V(\tau) - \phi_2(\tau)\nu^{Z,Q}(\tau) \right)^2 \lambda(x, \tau) \middle| \mathcal{F}(t) \right] d\tau.
\end{aligned}$$

The last equality follows since the covariation process between Γ and the zero coupon bond is 0, see equation (2.B.4). This completes the proof. Note however, that one in principle should verify that the conditions (2.B.11) and (2.B.12) are fulfilled.

□

Chapter 3

Markov chain modeling of policyholder behavior in life insurance and pension

Abstract: We calculate reserves regarding expected policyholder behavior. The behavior is modeled to occur incidentally similarly to insurance risk. The focus is on multi-state modelling of insurance risk and behavioral risk in terms of free policy risk and surrender risk. We discuss valuation techniques in the cases where behavior is modeled to occur independently or dependently of insurance risk, respectively. Ordinary differential equations make it easier to work with dependence between insurance risk and behavior risk. We analyze the effects of the underlying behavioral assumptions for two contracts. For a “new” contract with low technical interest rate relative to the market interest rate, we obtain the lowest reserve by counting in dependence. For an “old” contract with high technical interest rate relative to the market interest rate, the picture is more blurred, depending on assumptions on reactivation (recovery) and independence.

Keywords: Surrender option, free policy option, ordinary differential equation, recovery, dependence.

3.1 Introduction

We characterize reserves under finite-state Markov chain modeling of policyholder behavior and illustrate numerically the effects on values from modeling behavior in various ways. By behavior we think, in particular, of policyholder interventions like transcription to free policy and surrender. The reserves are characterized by ordinary differential equations having more or less explicit solutions, depending on the behavior model and the insurance risk model. These solutions are particularly tractable if one assumes independence between insurance risk and behavior, although such independence is often ruled out by contract design: Disability annuitants and life annuitants are typically not allowed to exercise such behavior options. In the illustrations we calculate values for standard contracts in order to analyze the effects of taking into account behavior in various ways. In particular, we study the consequences of assuming independence between insurance risk and behavior.

Dependence between insurance risk and behavior is a classical object of study since this is what notions like moral hazard and adverse selection are, formalistically, all about. The idea of for-

malising the dependence via a multi-state Markov chain may be generalized to cover also these aspects. However, our focus here is on the absence of the annuitant's behavioral options. Still, we are interested in the case of independence, since it offers some shortcuts with respect to calculations. These shortcuts may be attractive if they ease calculations a lot but only change the values a little. Actually, the rationale for this paper is to study and relate formulas and numerical results in the two cases of dependence and independence, respectively.

Current developments in insurance accounting and solvency rules take an explicit approach to behavior. For calculating reserves it is to an increasing extent required to take into account behavior. Behavior should be thought of as actions taken by the policyholder that influence either the risk in the processes that drive the payment streams of an insurance contract or the payments themselves. In this paper we pay special attention to the surrender option, i.e. the option to terminate the contract in exchange for a lump sum payment, and the free policy option, i.e. the option to stop paying the premium against a reduction of benefits. Among other options held by the policyholder may be the annuitization option in case the default coverage is a pension sum that can then, on basis of technical assumptions about interest and mortality, be converted to an annuity. Although this distinction is not necessarily a clean cut in practice, annuitization is an option that can be exercised upon retirement and can therefore be thought of as a European type option. The surrender and free policy options can, in general, be exercised at any point in time and can therefore be thought of as American type options. Another option that is sometimes mentioned explicitly is the option to raise the premiums. Typically, such an option is provided in connection with an occupational pension scheme, where e.g. premiums are calculated as a percentage of the salary.

There exists a range of approaches to modeling of behavior. One extreme position to take is to assume that the policyholder exercises his options based on an economically optimal strategy, i.e. in order to maximize the value of the payment stream of the contract. This approach is taken in Steffensen (2002) leading to characterization of values by so-called variational inequalities known in a financial context from American option pricing. Compared to a standard American option it is a delicate feature of the contract that both a free policy and a surrender option exist. Furthermore, if the free policy option is exercised, the contract continues under different terms and possibly still including a surrender option. This is all dealt with by Steffensen (2002). The same extreme American option approach is taken by e.g. Grosen and Jørgensen (2000), Bacinello (2003), and Siu (2005). A primitive approximation of the value obtained from this approach is to reserve, at any point in time, the larger of the value based on no exercise and the surrender value. This is what has been called a “now or never”-reserve since it corresponds to optimizing over two intervention strategies corresponding to exercising now or never. Due to its tractability, this is often seen as a first approach to take intervention options into account in practical accounting rules. Clearly, this approximation underestimates the true value since the optimal strategy may be to exercise somewhere between now and never.

Another extreme approach is to assume that the intervention options are exercised completely incidentally. Then intervention risk can be treated formalistically as a diversifiable insurance risk, although they are different concepts and the treatment considerably complicates, in general, the states of the world that have to be taken into account. This approach is taken e.g. by Buchardt and Møller (2013) and Buchardt et al. (2013).

A modern approach in accounting and solvency is to base reserves on expected policyholder behavior rather than rational policyholder behavior. This approach is taken e.g. in preliminary formulations of both Solvency II and IFRS. This draws attention towards the latter of the two extremes. However, the expected policyholder behavior is explicitly required to take into account also e.g. the economic environment and/or whether the option is beneficial. This appears to be

one step back towards the first extreme without really going that far. Such intermediary modeling is an interesting object of study with a lot of challenges concerning the statistical material available, economic intuition and mathematical tractability of the studied objects depending on the driving factors. Simple ideas are to let the intervention intensity depend on interest rates, as was done by De Giovanni (2010), or a relation between the intervention value and (some notion of) the market reserve. These ideas address the questions regarding the economic environment and whether the option is beneficial.

There exists a large amount of empirical literature discussing explanatory variables. These range from macro variables like interest rates, e.g. studied by Kuo et al. (2003) and Tsai et al. (2002), to micro variables like for instance policyholder age or policy duration. Some authors study macro and micro explanatory variables simultaneously, see e.g. Kim (2005) and Milhaud (2013). Note that working with macro variables creates dependence across policyholders in a portfolio which is crucial for solvency issues, see Loisel and Milhaud (2011). We refer to Eling and Kiesenbauer (2014) and references therein for a comprehensive literature overview.

In this paper we take the extreme approach to assume that the intervention options are exercised completely incidentally. This does not mean that we do not believe that working with interest rate or reserve dependent intervention intensities is interesting, important, challenging, or relevant. We are just focusing on something else. Also, we focus on something rather different from Buchardt and Møller (2013) who mainly concentrate on representation and calculation of cash flows in one of the special cases of our study, and Buchardt et al. (2013) who in a more theoretical framework deal with duration dependence in the risk model. Dependence between mortality and surrender risk was actually studied by Valdez (2001) but he worked in a model setup and with a focus quite different from ours. The relation between our approach and the idea of counting in external risk factors with respect to behavior is discussed in the conclusion raising several issues for future research.

We are interested in discussing the dependence between insurance risk and behavior that essentially arises from the product design. We do this in a finite state Markov chain framework. That allows us to characterize conditional expected values by ordinary differential equations and to specify their solutions. Their structures make it clear in what sense one can choose between a complicated differential equation based approach and a solution based on simplifying assumptions. Or said in a different way: Keep it simple or keep it right! To make the good choice here, it is of course relevant to qualify this one-liner. How simple is simple? And if simple means wrong, then how wrong is wrong? These questions are discussed from a theoretical point of view throughout the first part of the paper in Sections 3.2 - 3.5 and addressed numerically in Section 3.6 after which Section 3.7 concludes. In a thorough analysis where several aspects are taken into account, including varying over the value of the intervention options, the answer is not surprising: It depends! This paper illuminates on what it depends. A conclusion is that it really does matter for the entry values whether one takes the “simple” or the “right” approach. This makes a case for our advanced methods. All numerical results are obtained by Actulus® Calculation Platform that has been developed to deal with such questions of importance to the life insurance and pensions sector by numerically solving the relevant systems of ordinary differential equations.

3.2 Risk and Behavior Models

In this section we present the idea of considering a combined model for risk and behavior as being decomposed into two separate models for risk and behavior, respectively, that are or are

not probabilistically dependent of each other. We think of a risk state model Z^{risk} and a behavior state model $Z_{behavior}$ and consider the state model $(Z^{risk}, Z_{behavior})$.

Given (the whole process history of) $Z_{behavior}$, Z^{risk} is assumed to be a finite-state Markov chain taking values in $\mathbf{Z}^{risk}$. Thus, conditional on $Z_{behavior}$, there exist transition intensities $\mu^{jk}(t)$ for $j, k \in \mathbf{Z}^{risk}$ and $t \geq 0$, such that for all $k \in \mathbf{Z}^{risk}$, $\int_0^t \mu^{Z^{risk}(s)k}(s) ds$ is (conditional on $Z_{behavior}$) a compensator for the counting process counting the number of jumps into risk state k . The transition intensities may be independent of $Z_{behavior}$ and in that case Z^{risk} is, even unconditionally, a finite-state Markov chain. A canonical multi-state example of a risk model is the disability model illustrated in Figure 3.1. We have labeled the states {active, disabled, dead} by the letters $\{a, i, d\}$. This risk model is a key example below and in the numerical illustrations in particular.

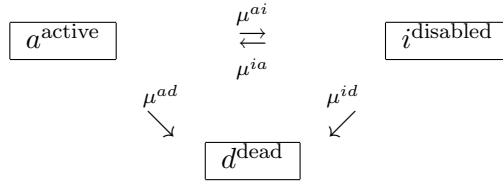


Figure 3.1: Disability risk model.

Symmetrically to the paragraph above we can now introduce the Markov chain $Z_{behavior}$ given (the whole process history of) Z^{risk} . We denote by $\nu_{jk}(t)$ for $j, k \in \mathbf{Z}_{behavior}$ and $t \geq 0$ the transition intensities and illustrate the canonical example of a behavior model in Figure 3.2. We have labeled the states {premium payment, free policy, surrender} by the letters $\{p, f, s\}$. This behavior model is a key example below and in the numerical illustrations in particular.

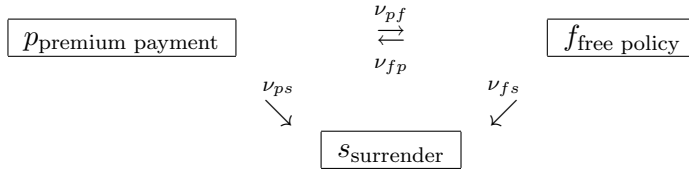


Figure 3.2: Behavior model.

As can be seen above, we refer consequently to specifications and states in the risk model by superscripts and to specifications and states in the behavior model by subscripts. When Z^{risk} given $Z_{behavior}$ and $Z_{behavior}$ given Z^{risk} are Markov models, the combined model $Z = (Z^{risk}, Z_{behavior})$ is a Markov model. Thus, there exist risk transition intensities $\mu_l^{jk}(t)$ for $j, k \in \mathbf{Z}^{risk}$, $l \in \mathbf{Z}_{behavior}$ and $t \geq 0$, such that for all $k \in \mathbf{Z}^{risk}$, $\int_0^t \mu_{Z_{behavior}(s)}^{Z^{risk}(s)k}(s) ds$ is a compensator for the counting process counting the number of jumps into risk state k . Similarly, there exist risk transition intensities $\nu_{jk}^l(t)$ for $j, k \in \mathbf{Z}_{behavior}$, $l \in \mathbf{Z}^{risk}$ and $t \geq 0$, such that for all $k \in \mathbf{Z}_{behavior}$, $\int_0^t \nu_{Z_{behavior}(s)k}^{Z^{risk}(s)}(s) ds$ is a compensator for the counting process counting the number of jumps into behavior state k .

We introduce the notation $p_{lm}^{jk}(t, s)$ for the transition probability that the risk process moves from j to k and the behavior process moves from l to m over (t, s) . In the case where the two sub-models for Z^{risk} and $Z_{behavior}$ are independent, i.e. the transition intensities μ do not depend on $Z_{behavior}$ and the transition intensities ν do not depend on Z^{risk} , we can simplify this probability into a product of probabilities with respect to each sub-model, i.e.

$$p_{lm}^{jk}(t, s) = p^{jk}(t, s) p_{lm}(t, s).$$

In case of independence we here specify the transition probabilities in the two models exemplified above. If μ^{ai} and μ^{ia} are both positive, we have no closed-form expressions for the probabilities $(p^{aa}, p^{ii}, p^{ai}, p^{ia}, p^{ad}, p^{id})$. However, in the case of no reactivation, i.e. $\mu^{ia} = 0$, we do:

$$\begin{aligned} p^{aa}(t, s) &= e^{-\int_t^s (\mu^{ai}(\tau) + \mu^{ad}(\tau)) d\tau}, & p^{ii}(t, s) &= e^{-\int_t^s \mu^{id}(\tau) d\tau}, \\ p^{ai}(t, s) &= \int_t^s p^{aa}(t, \tau) \mu^{ai}(\tau) p^{ii}(\tau, s) d\tau, & p^{ia}(t, s) &= 0. \end{aligned}$$

The probabilities p^{ad} and p^{id} are calculated residually by conditional probabilities summing to 1. Correspondingly, if ν_{pf} and ν_{fp} are both strictly positive, we have no closed-form expressions for the probabilities $(p_{pp}, p_{ff}, p_{pf}, p_{fp}, p_{ps}, p_{fs})$. However, in the case of no premium resumption, i.e. $\nu_{fp} = 0$, we do:

$$\begin{aligned} p_{pp}(t, s) &= e^{-\int_t^s (\nu_{pf}(\tau) + \nu_{ps}(\tau)) d\tau}, & p_{ff}(t, s) &= e^{-\int_t^s \nu_{fs}(\tau) d\tau}, \\ p_{pf}(t, s) &= \int_t^s p_{pp}(t, \tau) \nu_{pf}(\tau) p_{ff}(\tau, s) d\tau, & p_{fp}(t, s) &= 0. \end{aligned}$$

The probabilities p_{ps} and p_{fs} are calculated residually by conditional probabilities summing to 1.

A specific model for behavior is a model for the demand from policyholders. By a probabilistic model for demand we formalise a tendency that policyholders hold certain types of contracts. There are many motivations for thinking of the two processes Z^{risk} and $Z_{behavior}$ as being dependent. Two classical features in risk trading represent each one direction of influence between the two sub-models. On one hand, adverse selection means that policyholders with certain risks tend to demand certain contracts. We can reflect this in our model by letting the transition intensities in the behavior model be more or less explicitly dependent on the risk process. On the other hand, moral hazard means that policyholders with certain behavior/demand tend to cause certain levels of risks. We can reflect this in our model by letting the transition intensities in the risk model depend, more or less explicitly, on the behavior model. Thus, causal effects between the models have different directions and each direction has a given economic interpretation. At the end of the day, though, we observe a combined process where it may be difficult or even impossible to detect the direction of causal effects from the experienced dependence.

The canonical behavior model illustrated in Figure 3.2 above also forms a model for demand of certain types of payment profiles. There may be effects of adverse selection, i.e. policyholders in different risk states tend to exercise their free policy and surrender options differently. Moreover, there may be effects of moral hazard, i.e. mortality and disability rates differ for policyholders in the premium payment and free policy states, respectively. Below we pay full attention to a simple effect in the policy design where the dependence between the risk and behavior models is part of the contract. It is common practice that e.g. only policyholders in the risk state “active” are allowed to transcribe into a free policy or surrender. A standard contractual formulation is that such exercise options fall away when the contract goes from a premium paying contract to a benefit receiving contract, either by transition of state or by transition of time. We assume throughout that the risk of policyholders resuming their premium payment after having been transcribed to free policy is zero, i.e. $\nu_{fp} = 0$. This is often a harmless assumption since a contract, when premiums are resumed, is typically handled as a new contract and should therefore not be taken into account. With such a dependence coming exclusively from the behavioral options in the contract, we have illustrated the two-dimensional model in Figure 3.3.

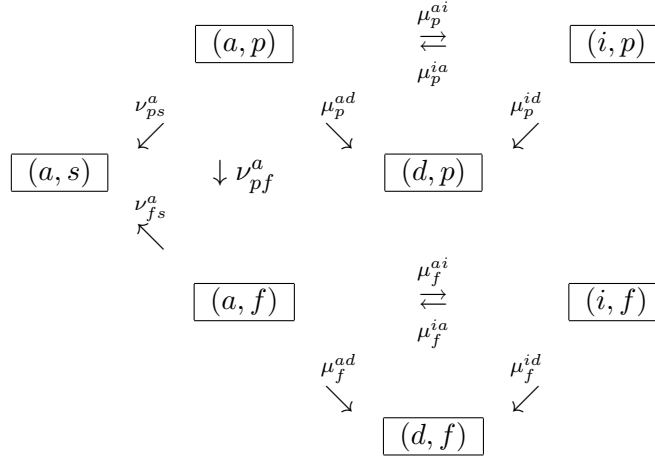


Figure 3.3: Combined model.

3.3 Values and Cash Flows in Risk and Behavior Models

In this section we describe the contractual payments and present formulas for calculation of their conditional expected present values. We assume a general risk model in combination with the canonical behavior model illustrated in Figure 3.2 with the premium resumption rate set to zero.

We take as starting point a contract that, in the first place, specifies its payments in the risk model, conditional on the behavior model being in the premium payment state p . We assume that the contract pays net benefits to the policyholder at rate b^j as long as the policyholder is in state j and a lump sum net benefit b^{jk} upon a transition from j to k . By net benefit we mean (positive) benefits paid to the policyholder minus (positive) premiums paid by the policyholder. The contract terminates at time n . We can now formalize the expected payment rate at time s given that the policyholder is in risk state k at time s as

$$c^k(s) = b^k(s) + \sum_{l:l \neq k} \mu_p^{kl}(s) b^{kl}(s),$$

where the sum here and in the rest of the paper is over states in $\mathbf{Z}^{risk}$.

We assume that the contract specifies that upon surrender from risk state k at time t , all future payments are canceled and a surrender sum $G^k(t)$ is paid out in return. Moreover, we assume that the contract specifies that if the policy is transcribed into a free policy while the policyholder is in risk state h at time t , the future payments are changed in the following way. The negative elements of b^j and b^{jk} , i.e. premiums, are set to zero whereas positive elements of b^j and b^{jk} , denoted by b^{j+} and b^{jk+} , are multiplied by a so-called free policy factor which depends, exclusively, on t and h and we denote by $f^h(t)$. First we introduce the expected payment rate of positive payments (before multiplication by f^h) as

$$c^{k+}(s) = b^{k+}(s) + \sum_{l:l \neq k} \mu_f^{kl}(s) b^{kl+}(s).$$

We can also write the expected payment rate at time s given that the policyholder is in risk state k at time s and jumped into the behavior state free policy at time t while being in risk state h as

$$f^h(t) c^{k+}(s) = f^h(t) \left(b^{k+}(s) + \sum_{l:l \neq k} \mu_f^{kl}(s) b^{kl+}(s) \right).$$

The surrender sum and the effects on the benefits in case the free policy option is exercised are parts of the contract terms. The contract specifies what happens in case of surrender or free policy. It should be clear that, if the policyholder decides to stop the premium payments prematurely, the benefits have to be recalculated accordingly in order to prevent speculation. The free policy factor typically reduces the benefits proportionally. A reduction occurs since a part of the benefits are related to the future premiums that fall away upon the exercise. A proportional reduction is motivated by a wish to treat all benefits equally. One can imagine a series of alternative recalculations of payments upon transition into the free policy state. E.g. the policyholder may want his risk coverages (like term insurance and disability annuities) to either fall away or to be fully kept upon transcription, and then the saving coverages (like deferred life annuities) are changed residually. We elaborate briefly on such alternatives in Section 3.5 below. However, we develop the valuation formulas under the assumption that all future benefits are changed proportionally. That even goes for the surrender payment in the following sense. If the policy was transcribed into a free policy while the policyholder was in risk state h at time t , and the policy is surrendered at time $u > t$ while the policyholder is in risk state j , the policy pays out a surrender sum $f^h(t) G^{j+}(u)$. All payment coefficients, including f and G , are assumed to be bounded and continuous. All formulas can easily be generalized to coefficients with countably many discontinuities by using that the reserves are still continuous, although not differentiable, in these points.

It is important to note the following. Since the process Z is Markovian the intensity of making a jump at time t depends on the position of Z only. However, this does not mean that the payment rate at time t only depends on Z . Since the expected payment rate at time s , $f^h(t) c^{k+}(s)$, depends on t and h through the free policy factor, we have introduced a specific duration dependence in the payment process which is not present in the probability model.

For the rest of the paper we consider a deterministic interest rate, r , but skip the time dependence in the formulas. In the numerical calculations in Section 3.6 we use a forward rate observed in the market at a fixed valuation date.

3.3.1 Given the free policy state

In this subsection we present a differential equation characterizing the reserve defined as the expected present value of future payments at time $t > \tau$ given that the policy jumped to the free policy state while the policyholder was in risk state h at time τ . We also specify its solution. Here and throughout we refer to Steffensen (2000) for all differential equations and their solutions. Denoting by $V_f^j(t)_{\tau h}$ the reserve if the policyholder is in risk state j at time t and became a free policy while being in risk state h at time τ , we can characterize this reserve by the differential equation (here and in the rest of the paper we skip the specification of the trivial boundary condition, $V_f^j(n)_{\tau h} = 0$)

$$\begin{aligned} \frac{d}{dt} V_f^j(t)_{\tau h} &= r V_f^j(t)_{\tau h} - f^h(\tau) b^{j+}(t) - \sum_{k:k \neq j} \mu_f^{jk}(t) \left(f^h(\tau) b^{jk+}(t) + V_f^k(t)_{\tau h} - V_f^j(t)_{\tau h} \right) \\ &\quad - \nu_{fs}^j(t) \left(f^h(\tau) G^{j+}(t) - V_f^j(t)_{\tau h} \right). \end{aligned}$$

The solution can be written as

$$V_f^j(t)_{\tau h} = f^h(\tau) \int_t^n e^{-\int_t^s r} \sum_k p_{ff}^{jk}(t, s) \left(c^{k+}(s) + \nu_{fs}^k(s) G^{k+}(s) \right) ds,$$

where $e^{-\int_t^s r}$ is shorthand notation for $e^{-\int_t^s r(u)du}$ and $p_{ff}^{jk}(t, s)$ is the probability that the policyholder moves from j to k in the risk model while staying in state f in the behavior model. This interpretation of $p_{ff}^{jk}(t, s)$ relies on the assumption that $\nu_{fp} = 0$, such that $p_{ff}^{jk}(t, s) = p_{\overline{ff}}^{jk}(t, s)$, where the notation $\overline{ff}$ denotes that the transition from f to f is made by staying uninterruptedly in state f . The solution is not in closed form, since the transition probabilities, in general, do not exist in closed form. From the integral solution we can see that the reserve consists of payments during sojourn in the free policy state (the $c^{k+}(s)$ terms) and payments paid upon surrender (the $G^{k+}(s)$ terms).

In the special case where the behavior and the risk models are independent, we have the simple form,

$$p_{ff}^{jk}(t, s) = p_{ff}(t, s) p^{jk}(t, s),$$

such that

$$V_f^j(t)_{\tau h} = f^h(\tau) \int_t^{\tau} e^{-\int_t^s r} p_{ff}(t, s) \sum_k p^{jk}(t, s) \left(c^{k+}(s) + \nu_{fs}(s) G^{k+}(s) \right) ds.$$

This formula is particularly convenient since it can be built around the “original” expected cash flow rates $\sum_k p^{jk}(t, s) c^{k+}(s)$ and the rates $\sum_k p^{jk}(t, s) \nu_{fs}(s) G^{k+}(s)$. Thus, when making use of the integral solution, one may, for computational convenience, be inclined to assume probabilistic independence. This is not correct, but the numerical errors may or may not be negligible. We elaborate on this in Section 3.6. On the other hand, when working with the differential equations, using the correct model does not introduce any additional complexity. Hence, there is no excuse for not performing the right calculations.

3.3.2 Given the premium payment state

In this subsection we present a differential equation characterizing the reserve given that the policy is in the premium payment state and in risk state j at time t . We also specify its solution. Denoting this reserve by $V^j(t)$, tacitly skipping the subscript p on all reserves below, we can characterize this reserve by the differential equation,

$$\begin{aligned} \frac{d}{dt} V^j(t) &= r V^j(t) - b^j(t) - \sum_{k:k \neq j} \mu_p^{jk}(t) \left(b^{jk}(t) + V^k(t) - V^j(t) \right) \\ &\quad - \nu_{pf}^j(t) \left(V_f^j(t)_{tj} - V^j(t) \right) - \nu_{ps}^j(t) \left(G^j(t) - V^j(t) \right). \end{aligned} \quad (3.3.1)$$

Note that the reserve $V_f^j(t)_{tj}$ is obtained by solving the differential equation for $V_f^j(t)_{\tau j}$ for fixed τ and subsequently replacing τ by t . The solution can be written as

$$\begin{aligned} V^j(t) &= \int_t^{\tau} e^{-\int_t^s r} \sum_k p_{pp}^{jk}(t, s) \left(c^k(s) + G^k(s) \nu_{ps}^k(s) \right) ds \\ &\quad + \int_t^{\tau} e^{-\int_t^s r} \sum_k W^{jk}(t, s) \left(c^{k+}(s) + \nu_{fs}^k(s) G^{k+}(s) \right) ds, \end{aligned} \quad (3.3.2)$$

where $p_{pp}^{jk}(t, s)$ is the probability that the policyholder moves from j to k in the risk model while staying in state p in the behavior model. This interpretation of $p_{pp}^{jk}(t, s)$ relies on the assumption that $\nu_{fp} = 0$, such that $p_{pp}^{jk}(t, s) = p_{\overline{pp}}^{jk}(t, s)$. Furthermore,

$$W^{jk}(t, s) = \int_t^s \sum_h p_{pp}^{jh}(t, \tau) \nu_{pf}^h(\tau) p_{ff}^{hk}(\tau, s) f^h(\tau) d\tau.$$

In the integral solution we can see that the reserve consists of payments during sojourn in the premium payment state (the $c^k(s)$ terms), payments due upon surrender from the premium payment state (the $G^k(s)$ terms), payments during sojourn in the free policy state (the $c^{k+}(s)$ terms), and payments upon surrender from the free policy state (the $G^{k+}(s)$ terms). The ratio $W^{jk}(t, s) / p_{pf}^{jk}(t, s)$ is the expected free policy factor given that the policyholder jumps in the risk model from j to k and in the behavior model from p to f over (t, s) .

In the special case where the behavior and the risk models are independent, we have the simple form (recall the simple forms for p_{pp} , p_{ff} , and p_{pf} from Section 3.2)

$$\begin{aligned} p_{pp}^{jk}(t, s) &= p_{pp}(t, s) p^{jk}(t, s), \\ p_{ff}^{jk}(t, s) &= p_{ff}(t, s) p^{jk}(t, s), \\ p_{pf}^{jk}(t, s) &= p_{pf}(t, s) p^{jk}(t, s), \end{aligned}$$

such that

$$\begin{aligned} V^j(t) &= \int_t^n e^{-\int_t^s r} p_{pp}(t, s) \sum_k p^{jk}(t, s) \left(c^k(s) + \nu_{ps}(s) G^k(s) \right) ds \\ &\quad + \int_t^n e^{-\int_t^s r} \sum_k W^{jk}(t, s) \left(c^{k+}(s) + \nu_{fs}(s) G^{k+}(s) \right) ds, \\ W^{jk}(t, s) &= \int_t^s p_{pp}(t, \tau) \nu_{pf}(\tau) p_{ff}(\tau, s) \sum_h p^{jh}(t, \tau) p^{hk}(\tau, s) f^h(\tau) d\tau. \end{aligned}$$

This formula appears convenient since the first line can be built around the “original” expected cash flow rates $\sum_k p^{jk}(t, s) c^k(s)$ and the rates $\sum_k p^{jk}(t, s) \nu_{ps}(s) G^k(s)$. Thus, when making use of the integral solution, one may, for computational convenience, again be inclined to assume probabilistic independence. However, this shortcut is not as appealing as it seems. In spite of the probabilistic independence, the second line is still an involved quantity. In order to really benefit from “original” expected cash flow rates, we further need to assume that $f^h(\tau)$ does not depend on h , in case we denote f^h by f . Then

$$W^{jk}(t, s) = p^{jk}(t, s) W(t, s),$$

with

$$W(t, s) = \int_t^s p_{pp}(t, \tau) \nu_{pf}(\tau) p_{ff}(\tau, s) f(\tau) d\tau, \quad (3.3.3)$$

such that the second line becomes

$$\int_t^n e^{-\int_t^s r} W(t, s) \sum_k p^{jk}(t, s) \left(c^{k+}(s) + \nu_{fs}(s) G^{k+}(s) \right) ds.$$

Finally, we have now reached the most elegant, but incorrect, expression since all elements are built around “original” cash flow rates $\sum_k p^{jk}(t, s) c^k(s)$ and $\sum_k p^{jk}(t, s) c^{k+}(s)$ and the rates $\sum_k p^{jk}(t, s) \nu_{ps}(s) G^k(s)$ and $\sum_k p^{jk}(t, s) \nu_{fs}(s) G^{k+}(s)$.

3.4 Important special cases

In this section we specialize the main results from Section 3.3 to two particularly important special cases. We consider the canonical disability model illustrated in Figure 3.1 and the survival model, respectively. In both cases we present the relevant differential equations and their more or less explicit solutions depending on the assumptions about the underlying model or contract. Particular attention is paid to the various simplifying assumptions that one can make in order to ease calculations. We concentrate on the valuation of policies that are in the behavior state “premium payment” since this is where the main calculation challenges arise.

3.4.1 The disability model

First we consider the disability model. We assume that all payments are zero in the state “dead” and label the different states according to Figure 3.1. Then the reserve corresponding to the policyholder being premium paying and active at time t can be characterized by a special case of (3.3.1) given by

$$\begin{aligned} \frac{d}{dt}V^a(t) &= rV^a(t) - b^a(t) \\ &\quad - \mu^{ai}(t)(b^{ai}(t) + V^i(t) - V^a(t)) - \mu^{ad}(t)(b^{ad}(t) - V^a(t)) \\ &\quad - \nu_{pf}^a(t)(V_f^a(t) - V^a(t)) - \nu_{ps}^a(t)(G^a(t) - V^a(t)). \end{aligned}$$

Here, the second line contains the risk premia related to the state transitions in the risk model whereas the third line contains risk premia related to the state transitions in the behavior model.

The general solution represented in (3.3.2) becomes

$$\begin{aligned} V^a(t) &= \int_t^n e^{-\int_t^s r} \left(\begin{aligned} &p_{pp}^{aa}(t, s)(c^a(s) + \nu_{ps}^a(s)G^a(s)) \\ &+ p_{pp}^{ai}(t, s)(c^i(s) + \nu_{ps}^i(s)G^i(s)) \end{aligned} \right) ds \\ &\quad + \int_t^n e^{-\int_t^s r} \left(\begin{aligned} &W^{aa}(t, s)(c^{a+}(s) + \nu_{fs}^a(s)G^{a+}(s)) \\ &+ W^{ai}(t, s)(c^{i+}(s) + \nu_{fs}^i(s)G^{i+}(s)) \end{aligned} \right) ds, \end{aligned}$$

where

$$\begin{aligned} W^{aa}(t, s) &= \int_t^s \left(\begin{aligned} &p_{pp}^{aa}(t, \tau)\nu_{pf}^a(\tau)p_{ff}^{aa}(\tau, s)f^a(\tau) \\ &+ p_{pp}^{ai}(t, \tau)\nu_{pf}^i(\tau)p_{ff}^{ia}(\tau, s)f^i(\tau) \end{aligned} \right) d\tau, \\ W^{ai}(t, s) &= \int_t^s \left(\begin{aligned} &p_{pp}^{aa}(t, \tau)\nu_{pf}^a(\tau)p_{ff}^{ai}(\tau, s)f^a(\tau) \\ &+ p_{pp}^{ai}(t, \tau)\nu_{pf}^i(\tau)p_{ff}^{ii}(\tau, s)f^i(\tau) \end{aligned} \right) d\tau. \end{aligned}$$

We now make the realistic assumption that the contract specifies that behavioral events only take place as long as the policyholder is active. This means that $\nu_{pf}^i(t) = \nu_{ps}^i(t) = 0$ such that W^{aa} and W^{ai} reduce to

$$\begin{aligned} W^{aa}(t, s) &= \int_t^s p_{pp}^{aa}(t, \tau)\nu_{pf}^a(\tau)p_{ff}^{aa}(\tau, s)f^a(\tau) d\tau, \\ W^{ai}(t, s) &= \int_t^s p_{pp}^{aa}(t, \tau)\nu_{pf}^a(\tau)p_{ff}^{ai}(\tau, s)f^a(\tau) d\tau. \end{aligned}$$

These formulas are, of course, not explicit due to the allowance for a positive reactivation rate. This makes the probabilities impossible to calculate explicitly. However, if we do not allow for

reactivation, we get simpler expressions for the probabilities, e.g.

$$\begin{aligned} p_{pp}^{aa}(t, s) &= \bar{p}_{pp}^{aa}(t, s) = e^{-\int_t^s \mu_p^{ai} + \mu_p^{ad} + \nu_{pf}^a + \nu_{ps}^a}, \\ p_{ff}^{aa}(t, s) &= \bar{p}_{ff}^{aa}(t, s) = e^{-\int_t^s \mu_f^{ai} + \mu_f^{ad} + \nu_{fs}^a}. \end{aligned}$$

This has simplifying consequences for the calculation of $p_{ff}^{ai}(t, s)$, $W^{aa}(t, s)$, and $W^{ai}(t, s)$.

Instead of assuming that the contract allows for behavioral events from the “active” state only, we now make the “opposite” assumption and say that the risk and behavior models are independent. We know from the previous section that this may help make the calculations much simpler. We get the following expression for the reserve,

$$\begin{aligned} V^a(t) &= \int_t^n e^{-\int_t^s r} p_{pp}(t, s) \left(\begin{array}{l} p^{aa}(t, s) (c^a(s) + \nu_{ps}(s) G^a(s)) \\ + p^{ai}(t, s) (c^i(s) + \nu_{ps}(s) G^i(s)) \end{array} \right) ds \\ &\quad + \int_t^n e^{-\int_t^s r} \left(\begin{array}{l} W^{aa}(t, s) (c^{a+}(s) + \nu_{fs}(s) G^{a+}(s)) \\ + W^{ai}(t, s) (c^{i+}(s) + \nu_{fs}(s) G^{i+}(s)) \end{array} \right) ds, \end{aligned}$$

where

$$\begin{aligned} W^{aa}(t, s) &= \int_t^s p_{pp}(t, \tau) \nu_{pf}(\tau) p_{ff}(\tau, s) \left(\begin{array}{l} p^{aa}(t, \tau) p^{aa}(\tau, s) f^a(\tau) \\ + p^{ai}(t, \tau) p^{ia}(\tau, s) f^i(\tau) \end{array} \right) d\tau, \\ W^{ai}(t, s) &= \int_t^s p_{pp}(t, \tau) \nu_{pf}(\tau) p_{ff}(\tau, s) \left(\begin{array}{l} p^{aa}(t, \tau) p^{ai}(\tau, s) f^a(\tau) \\ + p^{ai}(t, \tau) p^{ii}(\tau, s) f^i(\tau) \end{array} \right) d\tau. \end{aligned}$$

There are still several difficulties with this representation. Even though some elements relate to conditional expected cash flows from the original contract, we note that we cannot calculate the transition probabilities explicitly as long as we allow for reactivation. Furthermore, the expected cash flows from the free policy state are still quite complicated and do not relate to the original cash flows in an easy manner. Referring to the results in the previous section, we propose the additional assumption that $f^i = f^a$. Then

$$\begin{aligned} W^{aa}(t, s) &= p^{aa}(t, s) W(t, s), \\ W^{ai}(t, s) &= p^{ai}(t, s) W(t, s), \end{aligned}$$

with W defined as in (3.3.3), such that

$$\begin{aligned} V^a(t) &= \int_t^n e^{-\int_t^s r} p_{pp}(t, s) \left(\begin{array}{l} p^{aa}(t, s) (c^a(s) + \nu_{ps}(s) G^a(s)) \\ + p^{ai}(t, s) (c^i(s) + \nu_{ps}(s) G^i(s)) \end{array} \right) ds \\ &\quad + \int_t^n e^{-\int_t^s r} W(t, s) \left(\begin{array}{l} p^{aa}(t, s) (c^{a+}(s) + \nu_{fs}(s) G^{a+}(s)) \\ + p^{ai}(t, s) (c^{i+}(s) + \nu_{fs}(s) G^{i+}(s)) \end{array} \right) ds. \end{aligned}$$

Finally, the ingredients relate to the original cash flows. In these cash flows appear the transition probabilities. How accessible they are, depends on whether or not we allow for reactivation. If we do not, the probabilities are explicit and we have reached the “simplest” representation of our reserve.

3.4.2 The Survival model

Now we specialize to the survival model by skipping the disability state in Subsection 3.4.1. We skip the superscript a in the reserve V^a since all quantities are conditional on the policyholder being alive. This is a special case that deserves special attention. Namely, if there is no payments

in the death state, the simplifying and, in general, harmful assumptions are in this particular case harmless. This is so because all behavioral options and payments fall away upon death. The general solution represented in (3.3.2) becomes

$$\begin{aligned} V(t) &= \int_t^n e^{-\int_t^s r} p_{pp}^{aa}(c^a(s) + \nu_{ps}(s) G^a(s)) ds \\ &\quad + \int_t^n e^{-\int_t^s r} W^{aa}(t, s) (c^{a+}(s) + \nu_{fs}(s) G^a(s)) ds, \end{aligned}$$

where

$$W^{aa}(t, s) = \int_t^s p_{pp}^{aa}(t, \tau) \nu_{pf}(\tau) p_{ff}^{aa}(\tau, s) f^a(\tau) d\tau.$$

If we assume independence between the models we essentially assume that the mortality rate is not affected by the state of the behavior model and the transition intensities in the behavior model are not affected by the state of the risk model. The latter assumption is harmless since we have assumed that there are no payments in the death state. This means that it does not affect the value to allow for transcription into a free policy or surrendering among dead policyholders. If instead there were payments in the death state it would, of course, make a difference for these payments whether we allow for behavioral intervention or not. Under the assumption of independence we get the simplifications,

$$\begin{aligned} V(t) &= \int_t^n e^{-\int_t^s r} p_{pp}(t, s) p^{aa}(t, s) (c^a(s) + \nu_{ps}(s) G^a(s)) ds \\ &\quad + \int_t^n e^{-\int_t^s r} W^{aa}(t, s) (c^{a+}(s) + \nu_{fs}(s) G^{a+}(s)) ds, \end{aligned}$$

where

$$\begin{aligned} W^{aa}(t, s) &= \int_t^s p_{pp}(t, \tau) \nu_{pf}(\tau) p_{ff}(\tau, s) p^{aa}(t, \tau) p^{aa}(\tau, s) f^a(\tau) d\tau \\ &= p^{aa}(t, s) \int_t^s p_{pp}(t, \tau) \nu_{pf}(\tau) p_{ff}(\tau, s) f^a(\tau) d\tau. \end{aligned}$$

Then we have reached an expression based on the original cash flows. In the survival model, the final simplification, $f^h = f$, is not necessary. However, we stress that this is true only because we have no payments in the death state.

3.5 The free policy factor

In the calculations above we have assumed that all future benefits are multiplied by the same factor $f^j(t)$ upon transcription into free policy while being in risk state j at time t . We have not discussed what this f^j should be, though, and we have not discussed the situation where different factors apply to different future benefits. If f^j applies to all future benefits, a natural idea is to let f^j be determined by

$$f^j(t) = \frac{V^{j*}(t)}{V^{j+*}(t)},$$

where “*” denotes valuation of the contractual payments corresponding to a technical basis consisting of (r^*, μ^*) , possibly different from our valuation basis, (r, μ) , and “+” means that only

positive payments (benefits) are taken into account. To see why this idea is natural, consider the free policy sum at risk $V_f^j(t)_{tj} - V^j(t)$. Since $V_f^j(t)_{tj} = f^j(t) V^{j+}(t)$, we have that

$$V_f^j(t)_{tj} - V^j(t) = f^j(t) V^{j+}(t) - V^j(t).$$

Now, a particular version of our valuation basis would of course be the technical basis. In that case, as $f^j(t) = \frac{V^{j*}(t)}{V^{j+*}(t)}$, we get

$$V_f^{j*}(t)_{tj} - V^{j*}(t) = f^j(t) V^{j+*}(t) - V^{j*}(t) = 0.$$

In that sense we can say that under the technical basis the policyholder pays himself fully for the free policy risk. An important consequence of this approach is that one can disregard the free policy option for technical valuation purposes, e.g. for setting an equivalence premium. We can require from the free policy factor, that the free policy sum at risk is zero under the technical basis. Under this constraint, we can consider situations where different reduction factors apply to different benefits.

If a group of benefits (“keep”) is fully kept upon transcription while another group of benefits is deleted (“delete”), a residual group is reduced by the factor $f^j(t)$. It is found by solving the equation

$$0 = V^{j+(keep)*}(t) + f^j(t) \left(V^{j+*}(t) - V^{j+(keep)*}(t) - V^{j+(delete)*}(t) \right) - V^{j*}(t). \quad (3.5.1)$$

Hence

$$f^j(t) = \frac{V^{j*}(t) - V^{j+(keep)*}(t)}{V^{j+*}(t) - V^{j+(keep)*}(t) - V^{j+(delete)*}(t)}. \quad (3.5.2)$$

If there is a prioritized order in which the benefits are to be kept and the rest deleted, one could start filling the group “keep” as long as $V^{j*}(t) \geq V^{j+(keep)*}(t)$. The first benefit that violates this inequality should be reduced by the f^j in (3.5.2) and the residual benefits should be deleted. Two special cases of this are when either the group “keep” or the group “delete” is empty. If the group “keep” is empty and the group “delete” is not, then

$$f^j(t) = \frac{V^{j*}(t)}{V^{j+*}(t) - V^{j+(delete)*}(t)}$$

for the residual benefits. Note that this may lead to $f > 1$. If the group “delete” is empty and the group “keep” is not, then

$$f^j(t) = \frac{V^{j*}(t) - V^{j+(keep)*}(t)}{V^{j+*}(t) - V^{j+(keep)*}(t)},$$

for the residual benefits. Note that this leads to $f \leq 1$. However, in principle we may now end up with $f < 0$.

In all these cases the free policy sum at risk is

$$V_f^j(t)_{tj} - V^j(t) = V^{j+(keep)}(t) + f^j(t) \left(V^{j+}(t) - V^{j+(keep)}(t) - V^{j+(delete)}(t) \right) - V^j(t). \quad (3.5.3)$$

Plugging f^j defined in (3.5.2) into (3.5.3) under the technical basis gives the desired result that the technical free policy sum at risk is equal to zero, cf. (3.5.1).

3.6 Numerical results and discussion

We consider here the situation in Section 3.4.1, i.e. where the risk chain is a disability model with states “active” (a), “disabled” (i) and “dead” (d) (cf. Figure 3.1) and the behavioral chain consists of the states “premium payment” (p), “free policy” (f) and “surrender” (s) (cf. Figure 3.2). For the various setups (modeling of dependent/independent chains, using the same/different free policy factors in risk states and including/disregarding reactivation from disability) mentioned in Section 3.4.1, regarding the disability model, we compute the reserve conditional on the policyholder being in the risk state “active” and the behavioral state “premium payment”. The purpose of this is to quantify the implications of using the various alternatives to the correct model, which is the dependent model where reactivation is included. Here “dependent” corresponds to including policyholder options only from the risk state “active”, cf. Section 3.6.1 below. All numerical calculations are performed by numerically solving the system of ordinary differential equations (3.3.1). In case of no reactivation the system has a hierarchical structure and the equations corresponding to the different states can be solved one at a time. In case of a positive reactivation rate, the equations of the system are solved simultaneously. The results are obtained by Actulus® Calculation Platform.

3.6.1 Model parameters

The two computational bases in play are the technical (sometimes referred to as “first order”) basis and the market (sometimes referred to as “third order”) basis. In other words, we omit including a separate so-called “second order” basis (which is sometimes used to model bonus distribution schemes). As usual, the payments resulting from exercise of surrender and free policy options are defined such that the corresponding sums at risk under the technical basis are zero (cf. Section 3.6.2). Hence we disregard these options under the technical basis, see also the discussion in Section 3.5.

We start off by defining the transition intensities for the risk chain under the bases, see Table 3.1. Throughout this section these are independent of the behavioral chain implying e.g. that the mortality of a premium paying policyholder is the same as the mortality of a holder of a free policy at the same age.

From	To	Technical basis	Market basis
active	dead	$\mu^{*ad}(age) = 0.0005 + 10^{5.728-10+0.038(age)}$	$\mu^{ad} = \mu^{*ad}$
active	disabled	$\mu^{*ai}(age) = 0.0006 + 10^{4.71609-10+0.06(age)}$	$\mu^{ai} = \mu^{*ai}$
disabled	dead	$\mu^{*id}(age) = \mu^{*ad}(age)$	$\mu^{id} = \mu^{*id}$
disabled	active	$\mu^{*ia}(age) \equiv 0$	$\mu^{ia}(age) = e^{-0.06(age)}$ or 0

Table 3.1: Transition intensities, risk chain.

We remark that for the technical basis we use the standard intensities for a female occurring in the Danish G82 risk table. We let the mortality and disability intensities of the two bases be the same whereas the interest rates, on the other hand, are different (cf. Section 3.6.2 below). Furthermore, we consider both the case when the market reactivation intensity, μ^{ia} , is non-zero and when it is zero (as the latter assumption generally simplifies the semi-closed formulae, cf. Section 3.4.1).

Finally, we introduce the transition intensities in the behavioral model, see Table 3.2. We consider two situations.

1. The behavioral intensities are dependent of the policyholder's current state in the risk chain in the sense that they are zero unless the policyholder is in the risk state "active". This corresponds to disallowing policyholder interventions in the risk states "disabled" and "dead".
2. The behavioral intensities are independent of the risk chain. This corresponds to allowing for policyholder interventions in the risk states "disabled" and "dead". Since there are no payments in the state "dead", the latter consequence is irrelevant.

From	To	Independent model	Dependent model
premium payment	free policy	$\nu_{pf}(age) = e^{-0.07(age)}$	$\tilde{\nu}_{pf} = 1_{\{\text{active}\}} \cdot \nu_{pf}$
premium payment	surrender	$\nu_{ps}(age) = \nu_{pf}(age)$	$\tilde{\nu}_{ps} = 1_{\{\text{active}\}} \cdot \nu_{ps}$
free policy	surrender	$\nu_{fs}(age) = \nu_{pf}(age)$	$\tilde{\nu}_{fs} = 1_{\{\text{active}\}} \cdot \nu_{fs}$

Table 3.2: Transition intensities, behavioral chain.

Hence, as is the general assumption throughout this paper, we do not model the option of "re-entering" the premium paying state from the free policy state, and furthermore the state "surrender" is absorbing. There does not seem to exist a standard parametrization for the behavioral transition intensities. The idea behind their forms used here is simply that the inclination to exercise policyholder options decreases with the age of the policyholder. For a fixed age of contract initiation, which is the case in this paper, this is equivalent to accounting for a loyalty effect.

3.6.2 Contracts

3.6.2.1 The surrender and free policy options

The surrender option gives the policyholder the choice of abandoning the contract in exchange for a lump sum payment, $G^j(t)$, where j is the risk state in which the policyholder resides at time t . In what follows, we let $G^j(t)$ be the value of the contract under the technical basis, i.e. $G^j(t) := V^{j*}(t)$. In reality the sum received is sometimes reduced by a constant factor but we disregard that in what follows. Note that this choice of G^j makes the sum at risk upon surrender equal to zero under the technical basis. The free policy option gives the policyholder the choice of stopping the premium payment. The contract is then kept but the benefits are scaled by a certain factor, $f^j(t)$.

Before we turn to specifying the factors f^j , we note that the size of the technical reserve in relation to the market reserve is what determines if the surrender option increases or decreases the market value of the contract. In a situation where the technical reserve is higher than the market reserve, it is to be considered profitable for the policyholder to surrender and conversely when it is lower. In order to account for both situations, we consider two different contracts as specified below.

In some cases the policyholder options are only allowed from the risk state "active". It is then standard to define $f^a(t) = \frac{V^{a*}(t)}{V^{a*+}(t)}$, i.e. the quotient between the technical reserve and the technical benefit reserve (premium removed, all benefits kept). When allowing policyholder options also from the "disabled" state, the most natural choice appears to be $f^i(t) = \frac{V^{i*}(t)}{V^{i*+}(t)} \equiv 1$. However, as is mentioned in Section 3.4.1 above, setting $f^i := f^a$ yields even simpler closed-form solutions, and we therefore consider both these variants of f^i . We refer to the situations as using "Separate f " and "Same f ", respectively.

3.6.2.2 Common

We outline the common features of the two contracts considered.

- Contract expiry at age 65
- Premium payment of intensity 20,000 USD p.a.
- Disability annuity of intensity 100,000 USD p.a.
- Term insurance at 400,000 USD
- Pure endowment at expiry (corresponding to a life annuity) determined at initiation time of the contract such that it gives the contract a technical value (i.e. V^{a*}) of zero at the time of initiation

Furthermore, the market interest rate, r , is the forward rate equivalent to the yield curve as published by the Danish FSA at 2013-04-08 (cf. appendix 3.A).

3.6.2.3 New contract

We consider here the situation where a relatively young policyholder has just signed the contract, and where the technical interest rate (r^*) is low relative to the current market interest rate. This makes the technical reserve higher than the market reserve. More precisely we have the following additional parameters.

- Contract initiation at age 30
- Age 30 at the time of calculation ($t = 0$)
- $r^* = 1\%$ p.a. (continuously compounded)
- Pure endowment at expiry of 552,796 USD (corresponding to the reserve of a life annuity, at expiry, of 38.070 USD p.a. computed under the market value basis)

3.6.2.4 Old contract

We consider here the situation where an older policyholder signed the contract 20 years earlier, and where the technical interest rate (r^*) is high relative to the current market interest rate. This makes the technical reserve lower than the market reserve. The rationale for this is that when the contract was signed, r^* was indeed low compared to the contemporary market interest rate. More precisely we have the following additional parameters.

- Contract initiation at age 30
- Age 50 at the time of calculation ($t = 0$)
- $r^* = 5\%$ p.a. (continuously compounded)
- Pure endowment at expiry of 1,597,593 USD (corresponding to the reserve of a life annuity, at expiry, of 110,023 USD p.a. computed under the market value basis)

3.6.2.5 Remarks

We comment on the significant difference in the pure endowments of the two contracts. The endowment sum computed at initiation of the contract should be viewed as what was then guaranteed. If the insurance company can obtain a higher interest than the technical interest rate, this sum is typically increased via surplus bonus as time goes by. Assuming, in the context of the new contract, that the insurance company realizes an interest rate of 5% p.a. and uses all surplus contributions to increase, immediately and continuously, the guaranteed pure endowment sum, it will be exactly the same as that computed for the old contract when the policyholder reaches the age of 50.

3.6.3 Numerical results

We now display and discuss the numerical results obtained in the two contractual contexts.

We first mention that model variations occur in two “dimensions”. On the one hand regarding whether or not we include reactivation from disability and on the other hand whether or not we model the risk and behavioral chains independently. Finally, in the case when the chains are modeled independently, we consider the two situations when the free policy factors are the same, $f^i(t) := f^a(t)$, from the states “active” and “disabled”, and when they are not (they are defined in Subsection 3.6.2.1 above).

Figure 3.4 illustrates the various computational setups that we consider. For each box we have computed the corresponding reserve for both contracts. The numerical results allow us to find variations in the reserve by alternating the model “one step at a time”. In Figure 3.4 the arrows illustrate a situation where we start off by using a model without reactivation and independent chains and then additionally consider dependence and reactivation, one at a time. Hence, we have a way of decomposing the change between two models. In the case below the parts being a consequence of introducing dependence and reactivation, respectively. We emphasize that the top left box, corresponding to the correct model, is always considered as the benchmark result for comparisons.

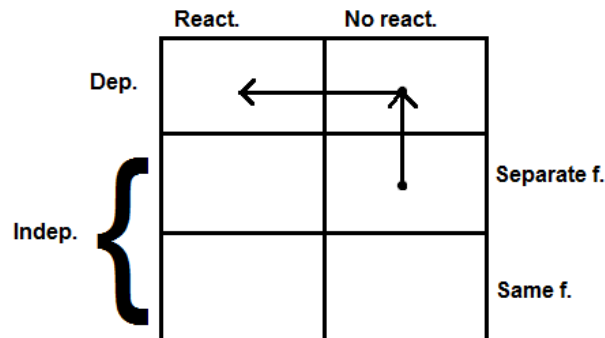


Figure 3.4: The included model variations and a possible path between two of them.

Some inequalities between reserves computed under different setups can be obtained in general by theoretical considerations. For example, it is fairly obvious that using $f^i(t) := f^a(t) \leq 1$ rather than $f^i(t) \equiv 1$ in the independent models yields, fixing the remaining parameters, a smaller market reserve. We show, however, by means of examples that other inequalities are indeed dependent on the concrete setup (e.g. the reactivation intensity).

3.6.3.1 New contract

Figure 3.5 below displays four reserves as functions of time and conditional on being in the state “active, premium payment”: The technical reserve (“Technical”), the reserve in the model including reactivation from disability and the aforementioned dependence between the risk and behavioral Markov chains (“Dep., react.”), the reserve disregarding behavioral options and reactivation (“No options, no react.”), and the reserve disregarding behavioral options only (“No options, react.”). The market reserves are smaller than the technical reserve since the policy is newly issued on the “safe side”. Disregarding reactivation (here illustrated under - but not only true for - no behavioral options) increases the reserve since less premiums and more benefits will be paid in the future. Disregarding behavioral options (here illustrated including reactivation) substantially decreases the reserve, setting a warning sign to the authorities that these options have a significant value and must be taken into account. The reason why the behavioral options increase the reserve is that they take the (larger) technical reserve as starting point for all fairness considerations and they therefore draw the reserve up towards the technical reserve. Larger behavioral intensities would draw the reserve including behavioral options up towards to the technical reserve.

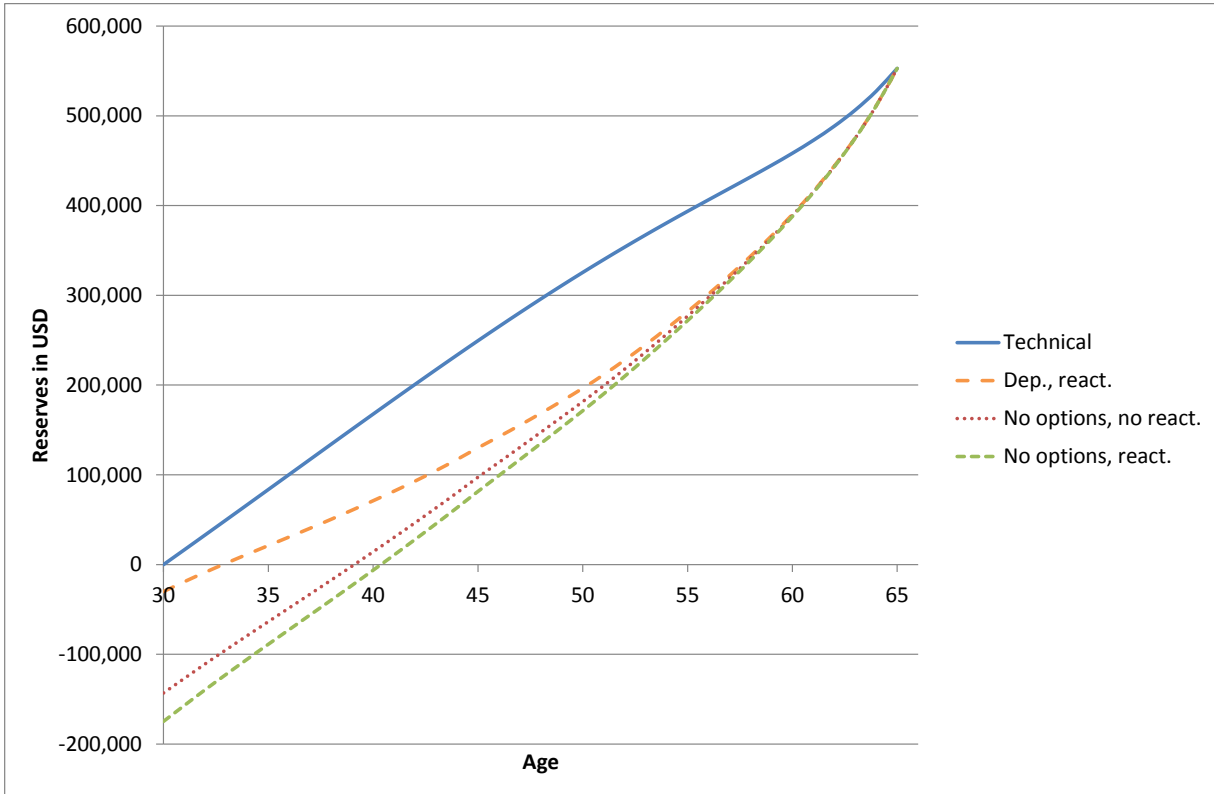


Figure 3.5: The technical reserve, the “dependent” reserve (including reactivation), and the two market reserves without behavioral options.

We now present all computed reserves evaluated at a few time points in Table 3.3. These are the technical reserve and eight different market reserves. The market reserve based on dependence and reactivation ($V^{\text{Dep., react.}}$) is spoken of as the correct reserve, whereas the other seven market reserves are calculated under simplifying assumptions. We also list the free policy factor in Table 3.3. Furthermore, we plot the difference between the five market reserves including behavioral options and the correct market reserve in Figure 3.6. More precisely, for a given reserve V^{Model} ,

where Model is e.g. “Dep., no react.”, we have plotted the function

$$t \mapsto V^{\text{Model}}(t) - V^{\text{Dep., react.}}(t).$$

The two market reserves without behavioral options are illustrated in Figure 3.5. It is seen how the reduction factor tends from 0 to 1 as a larger and larger proportion of the agreed premiums are paid. Here, we stress that “same f ” versus “separate f ” refers to whether the tabled reduction factor or 1 is used in case of transcription to free policy from the disability state.

Age	30	35	40	45	50	55	60	65
Technical	0	83,621	167,653	249,401	325,518	393,614	458,275	552,796
Indep., no react., separate f	-19,493	34,300	85,716	143,915	206,457	287,067	391,278	552,796
Dep., no react.	-21,625	31,852	83,269	141,833	205,131	286,489	391,171	552,796
Indep., no react., same f	-23,212	30,728	82,425	141,259	204,800	286,340	391,145	552,796
Indep., react., separate f	-26,450	25,026	74,769	133,055	198,205	282,578	390,216	552,796
Indep., react., same f	-28,341	22,876	72,484	130,987	196,794	281,913	390,088	552,796
Dep., react.	-30,014	21,198	71,130	130,124	196,436	281,852	390,094	552,796
No options, no react.	-143,050	-63,576	14,272	97,326	181,550	276,817	388,905	552,796
No options, react.	-174,674	-88,755	-6,249	81,507	171,357	271,825	387,800	552,796
f	0.000	0.153	0.300	0.440	0.573	0.702	0.838	1.000

Table 3.3: Computed reserves in descending order (in USD).

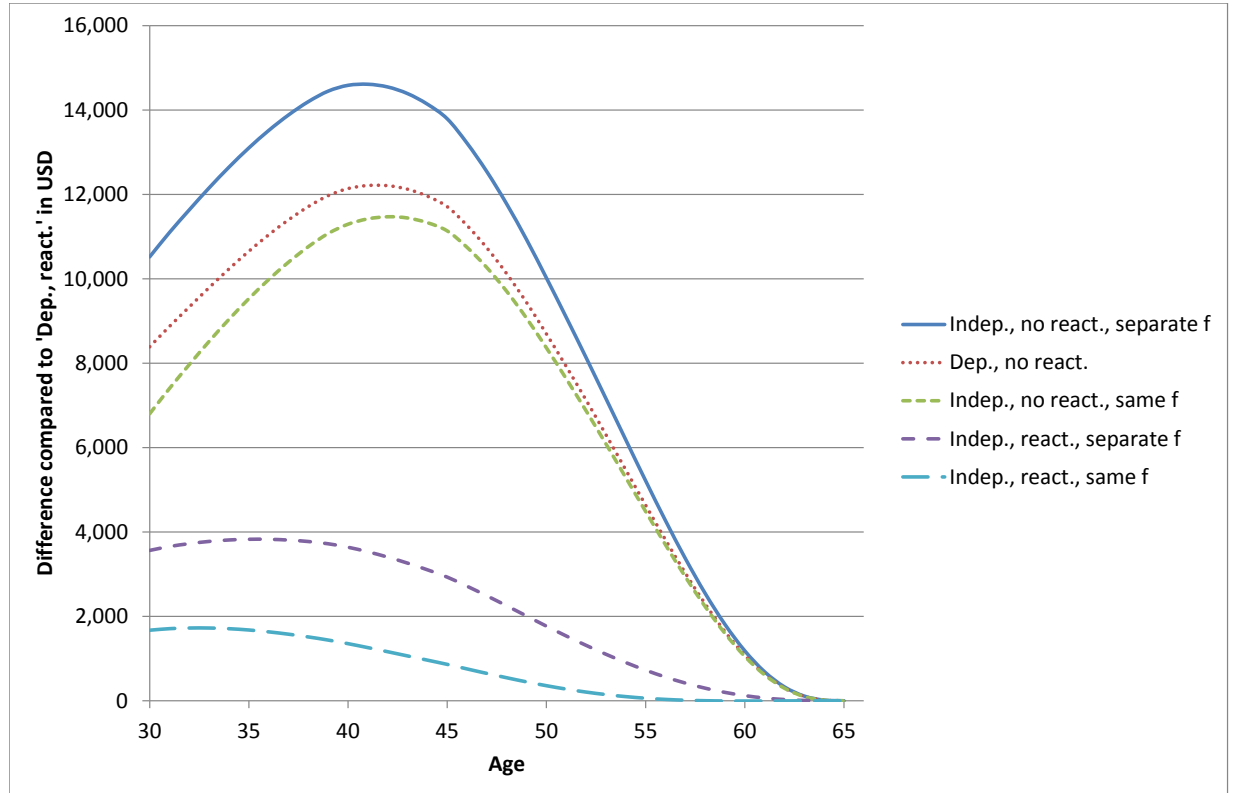


Figure 3.6: Reserve differences compared to “Dep., react.”.

We draw some conclusions from Figures 3.5 and 3.6 and Table 3.3.

1. Given that behavioral options are taken into account, we see that an insurance company would in fact benefit from using the correct model since this gives the smallest reserve. By incorporating reactivation, the reserves become smaller. For the intervention options it holds that working with dependence gives that fewer policyholders (compared to the independent case) exercise options that are technically fair and therefore the attraction up towards the technical reserve is smaller.
2. One cannot conclude from the above that using the model yielding the simplest closed form expression for the reserve, “Indep., no react., same f ”, is “at least on the safe side” of the correct reserve. Namely, from the numbers above it is clear that letting the reactivation intensity approach zero, the reserve in “Dep., react” converges to that in “Dep., no react.”, which is larger than the one in “Indep., no react., same f ”. Hence the order relation between the two reserves is in fact intensity dependent.
3. It is by no means a surprise that the largest market reserve is “Indep., no react., separate f ”. We omit a formal argument but note that when assuming independent chains we have a surrender and free policy option in the disability state. When using separate f , i.e. $f^i \equiv 1$, exercising the free policy option from the disability state gives the policyholder $V^{i+}(t)$, in other words the value of his own contract with premium payment streams removed from all states. The surrender option, when exercised from the disability state, gives the policyholder an amount equal to the value of his contract computed under an interest rate lower than the market interest rate and furthermore disregarding reactivation and free policy (namely, the technical reserve, $V^{i*}(t)$). Hence the disability state is as expensive as possible and further omitting reactivation from this state additionally increases the reserve (the policyholder can never resume the premium payment instead of receiving the disability annuity).
4. Note that the difference between the reserves “Indep., no react., separate f ” and “Dep., no react.” is caused solely by the surrender option. To see this, note that transcribing to free policy from the disability state gives the corresponding benefit reserve, $V^{i+}(t) = V^i(t)$ (no reactivation), scaled by the free policy factor $f^i(t) \equiv 1$.
5. The numerical results allow for a study of the isolated effect of counting in reactivation. This is done by comparing the reserve for “Dep., no react.” with the reserve for “Dep., react.” (or any other set of reserves where everything but the “react./no react.” dimension is fixed). We see that reactivation reduces the reserve considerably at low ages where the reserves are low, but even after 20 years where the reserves are around 200,000 USD the reduction is in the order of 5%.

3.6.3.2 Old contract

Figure 3.7 below displays four reserves for the old contract similarly to Figure 3.5 for the new contract. Now the market reserves are larger than the technical reserve since the technical basis is not on the same “safe side” today as it was upon initiation 20 years ago. Disregarding reactivation (here illustrated under - but not only true for - no behavioral options) increases the reserve as for the new contract since this qualitative effect does not rely on a particular relation between the bases. Disregarding behavioral options (here illustrated including reactivation) increases the reserve. The qualitative effect is opposite that of Figure 3.5 because the technical reserve is now smaller. Counting in behavioral options that take the (smaller) technical reserve into account for all fairness considerations draws the reserve towards the technical reserve. Larger behavioral intensities would draw the reserve including behavioral options down towards the technical reserve. The order of magnitude of the differences are comparable with the order of

magnitude of the differences in Figure 3.5, when focusing on age 50 and onwards. In Figure 3.5 the effects are nominally large but relatively (to the reserves themselves) small in the first 20 years, where both the probability of being reactivated and exercising behavioral options are large. In the last 15 years the probability of ever being reactivated (given that you are active) is small and the behavioral options are relatively unlikely to be exercised. The order of magnitude of the effect from behavioral options can easily change if we work with other behavioral intensities than the ones suggested here.

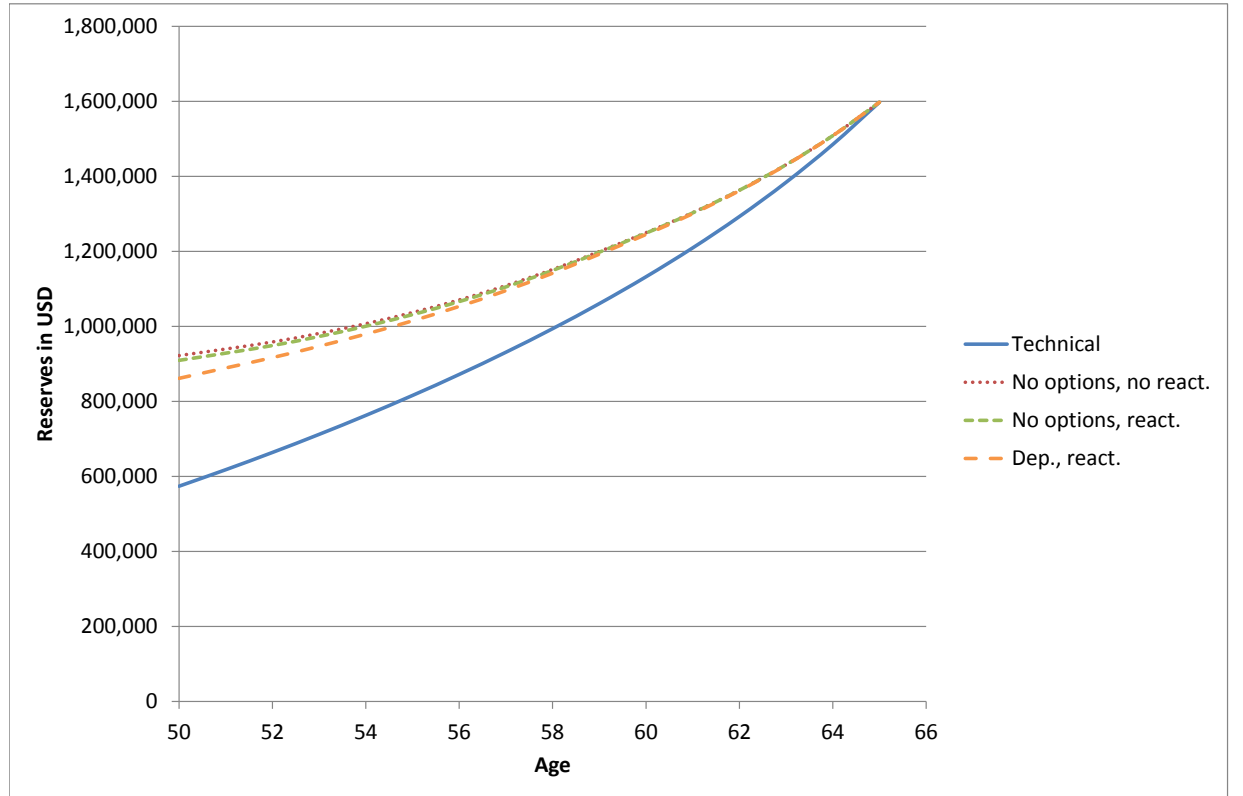


Figure 3.7: The technical reserve, the “dependent” reserve (including reactivation), and the two market reserves without behavioral options.

In Figures 3.7 and 3.8 and Table 3.4 we present the numerical results for the old contract in the same form as the numerical results for the new contract. We draw some conclusions from these results.

1. Given that behavioral options are taken into account, we see that the different assumptions move the reserve in different directions, and no general conclusion can be drawn. Working with reactivation gives, of course, a smaller reserve but among the reserves with reactivation, actually the correct model with dependence is the larger of the three reserves, “Dep., react.”, “Indep., react., separate f ”, and “Indep., react., same f ”. This sets a warning sign to the authorities that these options must be taken correctly into account. The effect is opposite compared to the new contract. Working with independence means that more policyholders exercise (technically fair) intervention options drawing down the reserve towards the technical reserve.
2. As mentioned in the comments to Figure 3.7 above, the effect from calculating with any simplifying assumption is smaller for the old contract than it was for the new contract. This

Age	50	55	60	65
Technical	573,984	815,950	1,132,248	1,597,593
No options, no react.	922,259	1,036,998	1,249,800	1,597,593
No options, react.	909,487	1,031,425	1,248,647	1,597,593
Dep., no react.	872,815	1,020,138	1,246,235	1,597,593
Indep., no react., separate f	870,710	1,019,318	1,246,089	1,597,593
Indep., no react., same f	869,112	1,018,675	1,245,971	1,597,593
Dep., react.	861,537	1,014,869	1,245,103	1,597,593
Indep., react., separate f	860,342	1,014,298	1,244,981	1,597,593
Indep., react., same f	858,947	1,013,702	1,244,867	1,597,593
f	0.754	0.854	0.933	1.000

Table 3.4: The technical reserve and the market reserves in descending order (in USD).

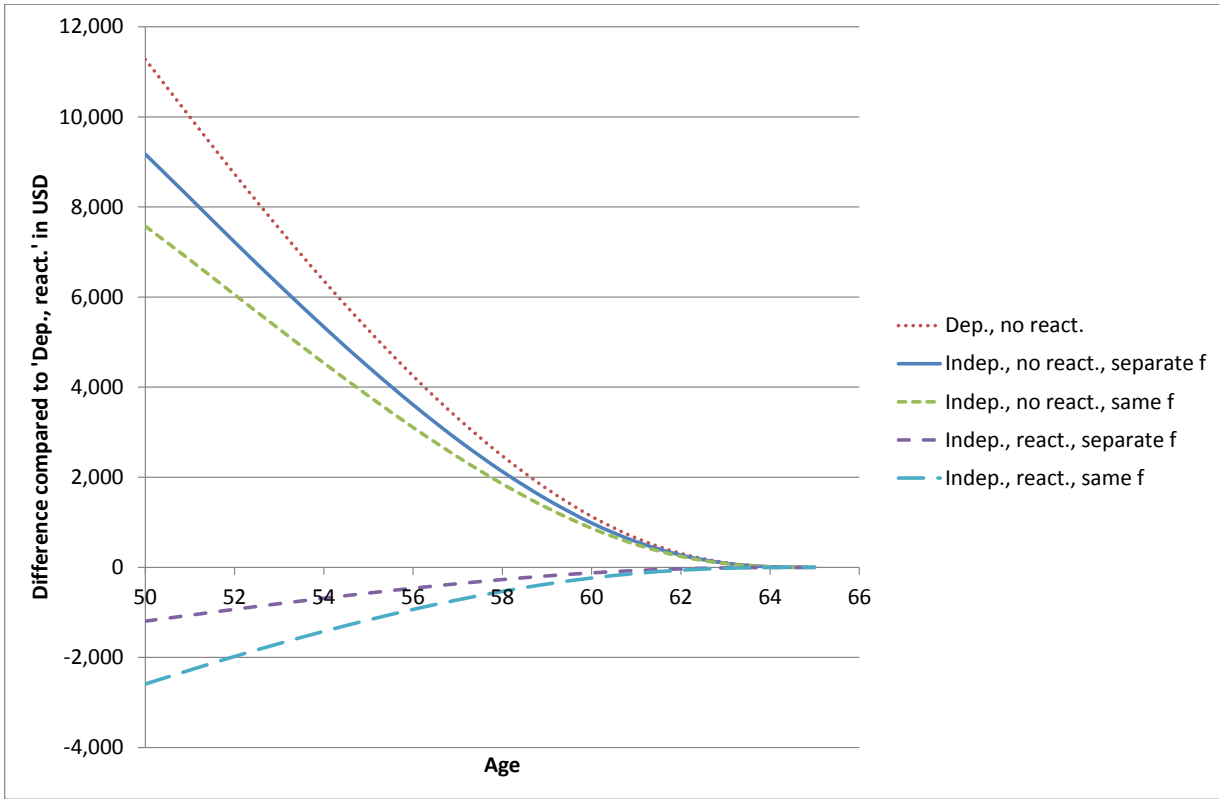


Figure 3.8: Reserve differences compared to “Dep., react.”.

is because any impact from behavioral options is reduced towards the end of the contract where the behavior intensities are small.

3. The reserve obtained when using the correct model “Dep., react.” is below those in the models yielding the simplest semi-closed form solutions, “Indep., no react., separate f ” and “Indep., no react., same f ”. Note, however, that when making the reactivation intensity small the reserve in “Dep., react.” tends to that in “Dep., no react.” which is larger than both. Hence, the relations between reserves are intensity dependent.
4. For precisely the same reason as in Section 3.6.3.1, we note that the difference between the reserves “Indep., no react., separate f ” and “Dep., no react.” is caused solely by the

surrender option.

5. The isolated effect of counting in reactivation can again be studied by comparing the reserve for “Dep., no react.” with the reserve for “Dep., react.” or similar. Reactivation reduces the reserves by an order of magnitude of 1%-2%.

3.7 Conclusion

We have characterized reserves by differential equations and explicit formulas and illustrated them numerically under different assumptions about underlying policyholder behavior. We have shown the formalistic and computational advantages and disadvantages under different simplifying assumptions in the case where dependence is part of the contract design. We have numerically illustrated the effects of these assumptions. The conclusion is that both the sign and the magnitude of these effects are blurred and depend on the relation between the assumptions of the technical basis and the market basis. Here, we conclude by a couple of remarks about what we have *not* done, also as a hint for future research.

- We have used Markov chains to model dependence between insurance risk and behavior stipulated in the contract. We have also mentioned, but not studied, cases where the dependence arises from the fundamental notions of adverse selection and moral hazard. It could be interesting to use the ideas of the paper as an approach to these notions. That would require specific multi-state models catching the effects observed.
- The numerical results are obtained by solving systems of ordinary differential equations, allowing for cycles in the Markov chain, e.g. in connection with reactivation. The effects from reactivation are pointed out in the tables but further studies on the sensitivity with respect to reactivation rate could be of interest.
- Since we take the approach of completely incidental exercise of the options we wish to add a comment on the applicability of our framework beyond this case. In case of external risk factors influencing policyholder behavior, the manuscript can be seen as a starting point in one of two different directions. One direction is to assume a Markovian structure of the external factors and then insist on exploiting this structure in order to represent all values by solutions to differential equations. The ordinary differential equations in this manuscript then make up special cases of e.g. the general partial differential equations that are necessary in order to deal with diffusive economic risk. As such they can serve as supporting the understanding and as a test point, formalistically and numerically, for more general cases. Another direction is to consider the values obtained here as corresponding to a given realization of an external (possibly non-Markovian) risk factor. Allowing e.g. behavioral intensities to depend on the realization, we can calculate the generalized values by simulation and calculation of empirical means. This method would delicately combine the ordinary differential equation approach to micro risk and the Monte-Carlo approach to macro risk. We leave both directions as studies for future research.

Appendix

3.A The Danish FSA yield curve used in the market basis

In Figure 3.9, we plot the (discretely compounded) yield curve, R , published on 2013-04-08 by the Danish FSA, from which we extracted the equivalent continuous forward rate used in the market basis. We also plot the technical interest rates used in the two contracts considered in Section 3.6.2, here discretely compounded ($R^* = e^{r^*} - 1$).

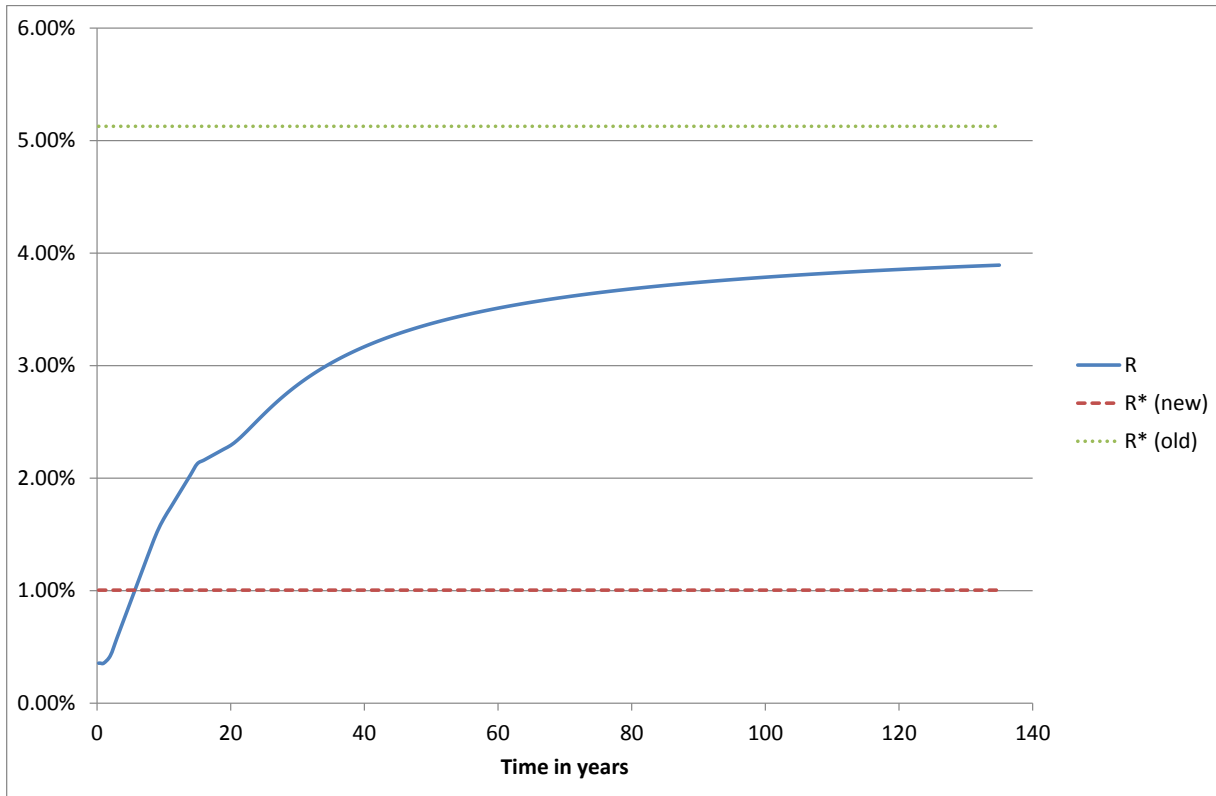


Figure 3.9: The interest rates used in the bases of the examples.

Chapter 4

Stress scenario generation for solvency and risk management

Abstract: We derive worst-case scenarios in a life insurance model in the case where the interest rate and the various transition intensities are mutually dependent. Examples of this dependence are that a) surrender intensities and interest rates are high at the same time, b) mortality intensities of a policyholder as active and disabled, respectively, are low at the same time, and c) mortality intensities of the policyholders in a portfolio are low at the same time. The set from which the worst-case scenario is taken reflects the dependence structure and allows us to relate the worst-case scenario-based reserve, qualitatively, to a Value-at-Risk-based calculation of solvency capital requirements. This brings out perspectives for our results in relation to qualifying the standard formula of Solvency II or using a scenario-based approach in internal models. Our results are powerful for various applications and the techniques are non-standard in control theory, exactly because our worst-case scenario is deterministic and not adapted to the stochastic development of the portfolio. The formalistic results are exemplified in a series of numerical studies.

Keywords: Life insurance, worst-case scenario, deterministic control, Solvency II, multistate Markov chain.

4.1 Introduction

From a specific set of scenarios we find the scenario that leads to the worst-case, i.e. largest, value of future obligations. The idea of such worst-case scenarios has a wide range of applications in life insurance pricing, management, and regulation. They include settlement of premiums and surrender values as well as calculation of risk margins and solvency capital requirements (SCR). In particular we draw the attention to the scenario-based standard formula in Solvency II, see Steffen (2008) for an overview of the Solvency II project. So far it has been difficult to say something qualitatively about the relation between a standard scenario and the original Value-at-Risk-based SCR. We generate scenarios such that SCRs based on these scenarios are proven sufficient under the Value-at-Risk measure within a given model. As input to our calculation, we take a set of interest and transition rates such that the probability of realizing interest and transition rates within that set is bounded. For such a set we calculate the scenario that maximizes the reserve. We do not generally address the difficult but interesting question of detecting such a set although we exemplify possible sets in some examples.

Our results are applicable to an inhomogeneous portfolio of contracts. Then the worst-case sce-

nario generated is worst-case for a whole portfolio with different policyholder ages and contracts. This makes the approach particularly useful in the discussion about portfolio SCRs. The results can qualify this discussion in two dimensions. First, for a regulator who wants to develop a stress scenario-based SCR we provide a scenario corresponding to bounds on shortfall probabilities. This qualified standard formula should be derived for a stylized market-realistic portfolio. Second, an insurance company can replace the standard SCR formula by a so-called (partial) internal model-based calculation and still exploit advantages working with scenarios. Our results show how such “internal stress scenarios” can be derived.

In relation to recent academic literature on worst-case scenario generation, we emphasize that our stress scenarios are deterministic in the sense that e.g. the portfolio worst-case mortality intensity at a future time point is calculated today and will not depend on the survivors of the portfolio at that future time point. This means that we are, briefly speaking, finding optimal deterministic processes maximizing an expectation in a stochastic environment. This is a non-standard exercise that contains methodological and computational challenges. The upside is that, once they are overcome, the resulting scenario can be better understood, communicated, implemented and extrapolated for usage in other (similar) portfolios. The idea to look for deterministic worst-case intensities is in sharp contrast to e.g. Li and Szimayer (2011, 2014) who, in a different framework, also study worst-case intensities. They, however, use more standard stochastic control techniques to derive intensities that are adapted to the development of the contract under study. This development amounts, in their cases, to the development of asset prices but, in a more general setting, it could be the number of, or even the names of, survivors in a portfolio. Such adapted scenarios may be useful for other applications but not necessarily for solvency issues. Thus, the conceptual innovation in relation to Li and Szimayer (2011, 2014) is that our worst case is deterministic and its derivation therefore draws on other techniques.

The present work extends the results of Christiansen and Steffensen (2013) in the following way: Like us, they search for optimal deterministic scenarios and obtain simple formulas for these but in a quite restricted class of models. The class is defined endogenously by requiring that certain argmax operations over transition intensities are constant with respect to the transition probabilities they generate, see Christiansen and Steffensen (2013, Proposition 4.1 and 4.2). The work in this paper is very much inspired from the structure of problems and solutions in that article, but we succeed in finding the worst-case scenario also outside their restrictive assumptions. This allows for studying much more realistic and important cases like a hierarchical disability model or a model for a portfolio of heterogeneous contracts hit by the same worst case. Thus, we develop a powerful tool for various applications while sticking to the natural but challenging idea of Christiansen and Steffensen (2013) that the worst-case should be deterministic. Thus, the innovation in relation to Christiansen and Steffensen (2013) is that we find, by means different than theirs, the worst case for interesting products and portfolios that they rule out.

In order to show how the studies in this paper can be applied to solvency calculations we choose to give the reader, already here in the introduction, a glance of the formalistic argument. Details can be found in Christiansen and Steffensen (2013). For (real, unknown) interest and transition rates (ϕ, μ) , we want to find deterministic interest and transition rates $(\tilde{\phi}, \tilde{\mu})$ such that for the liabilities L , it holds that

$$P\left(L(t, \tilde{\phi}, \tilde{\mu}) \geq L(t, \phi, \mu)\right) \geq 1 - \alpha, \quad (4.1.1)$$

where $\alpha \in [0, 1)$. That is, we want to find a deterministic calculation basis such that the liabilities calculated with this basis with a certain probability are larger than the liabilities calculated with the real (stochastic) basis. This can be obtained by choosing $(\tilde{\phi}, \tilde{\mu}) = \operatorname{argmax}_{(\phi, \mu) \in M} L(t, \phi, \mu)$

for a set M such that $P((\phi, \mu) \in M) \geq 1 - \alpha$. We do not pay any attention to how the set M is formed, except for in a few numerical examples. The object of study in this paper is, given a set M , to calculate the argmax, $(\tilde{\phi}, \tilde{\mu})$.

As shown in Christiansen and Steffensen (2013), we can use this to obtain an upper bound for the SCR given by

$$\sup_{(\phi, \mu) \in M} \{L(t, \phi, \mu)\} - L(t, \phi^{\text{BE}}, \mu^{\text{BE}}),$$

where ϕ^{BE} and μ^{BE} are best estimates for the interest rate and transition intensities, respectively.

The quality of an SCR based on a standard stress scenario compared to a Value-at-Risk-based calculation has been intensively discussed in the literature. Doff (2008) analyses, critically, the Solvency II proposal and the shortcomings of the standard stress model, which is also discussed by Devineau and Loisel (2009). Specific attention has been given to longevity risk and Olivieri and Pitacco (2008) study the pitfalls of approaching longevity risk by means of stress scenarios. A comprehensive numerical study that compares the stress scenarios to the Value-at-Risk calculation can be found in Börger (2010). To the authors' knowledge, all comparative studies are quantitative in the sense that the various principles for SCR calculations are numerically related to each other. If the stress-based SCR is significantly smaller than the Value-at-Risk based SCR, the whole idea of inducing financial stability from the standard formula may be criticised for being an optical illusion. We distinguish ourselves from this quantitative discussion by searching for a stress, such that the SCR derived from this stress scenario is at least as large as the SCR that can be derived from a Value-at-Risk approach. Thus, rather than criticising the idea of stress scenarios, we admit its advantages and seek to qualify the discussion about what the stress scenarios should look like. Our study is general enough to help answer this question no matter if the unit is a contract or a portfolio.

The paper is organized as follows: In Section 4.2 we introduce the insurance market and get a representation of the probability weighted reserve. In Section 4.3 we obtain worst-case scenarios and reserves for a single policy and describe numerical methods which are needed for the numerical calculations of these quantities. In Section 4.4, we extend the theory to cover a portfolio of policyholders, and finally we present some numerical calculations for both single policies and portfolios in Section 4.5.

4.2 Modelling and valuation of an insurance policy

Let T be a fixed finite time horizon and $(\Omega, \mathcal{F}, (\mathcal{F}(t))_{0 \leq t \leq T}, P)$ a filtered probability space with filtration satisfying the usual conditions of right-continuity and completeness. Let X be a pure jump process defined on this probability space with finite state space $\mathcal{S}$ which we assume consists of n states. The process $(X(t))_{t \in [0, T]}$ represents the state of a policyholder. We do not assume a deterministic starting value for X but only a starting distribution, which we denote π . This enables us to easily calculate the worst-case reserve for a homogeneous portfolio of insurance contracts, where each policyholder can be in different initial states like “Active” and “Disabled”. This is a simple special case of the more general theory for portfolios presented in Section 4.4. We denote by $\mathcal{J}$ the transition space of X defined by $\mathcal{J} = \{(j, k) \in \mathcal{S}^2 | j \neq k\}$.

Following the lines of Norberg (1999), we assume that the interest rate intensity ϕ is piecewise continuous and that $\int_0^T \phi(s)ds$ is finite. We define a discounting function v by the following forward equation:

$$\frac{d}{dt}v(t) = -v(t)\phi(t), \quad v(t_0) = 1, \quad (4.2.1)$$

where t_0 is not necessarily the initiation time of the contract. One can intuitively think of t_0 as 0 but we introduce the notation t_0 in order to be able to accurately formulate the verification lemma. The solution to (4.2.1) is given by

$$v(t) = e^{-\int_{t_0}^t \phi(\tau) d\tau},$$

and the discounting factor for the time span from s to t is given by

$$v(s, t) = \frac{v(t)}{v(s)} = e^{-\int_s^t \phi(\tau) d\tau}.$$

We let the $n \times n$ transition matrix for X be denoted by p , which is a function with two arguments. That is, for $t_0 \leq s \leq t \leq T$ we have that

$$p(s, t) = (P(X(t) = k | X(s) = j))_{(j,k) \in \mathcal{S}^2}.$$

By μ we denote the corresponding $n \times n$ intensity matrix. It is well-known that p is determined by Kolmogorov's forward equation:

$$\frac{\partial}{\partial t} p(s, t) = p(s, t) \mu(t), \quad p(s, s) = \mathbf{I}_n,$$

where $\mathbf{I}_n$ is the n -dimensional identity matrix.

With a little abuse of notation we denote by p (a function of one variable only) the marginal distribution of the random pattern of states:

$$p(t) = \begin{pmatrix} p_1(t) \\ p_2(t) \\ \vdots \\ p_n(t) \end{pmatrix},$$

where $p_j(t) = P(X(t) = j)$. The forward equation for p is given by

$$\frac{d}{dt} p(t) = \mu^{\text{tr}}(t) p(t), \quad p(t_0) = \pi, \tag{4.2.2}$$

where π is a given initial distribution over the n states and $\mu^{\text{tr}}(t)$ is the transpose of $\mu(t)$. Equation (4.2.2) follows by Kolmogorov's forward equation by using that $p_j = \sum_{i \in \mathcal{S}} \pi_i p_{ij}$, which gives us that the dynamics of p_j are equal to p_i times the columns of μ . This is exactly equal to $\mu^{\text{tr}}(t) p(t)$.

We consider an insurance contract with the following type of payments:

1. $b_i(t)$ is the rate of payments in state i at time t .
2. $b_{ij}(t)$ is a lump sum payment payable upon transition from state i to state j at time t .

We denote by $B_i(t)$ the accumulated payments in state i up to time t and by $B(t)$ we denote the present value at time t of future payments of the contract. We assume that all the functions b_{ij} and B_i have bounded variation on $[0, T]$ and that the functions b_i and b_{ij} are C^1 on $[0, T]$. The latter assumption can easily be relaxed to piecewise C^1 but this will result in cumbersome notation, since we need to deal with extra boundary conditions.

To keep the notation simple, we assume that no lump sum payments are paid out while sojourning in a state. That is, for all $t \in [0, T]$ we have that

$$\Delta B_i(t) = B_i(t) - B_i(t-) = 0. \tag{4.2.3}$$

However, the results of the paper can easily be extended to the case, where assumption (4.2.3) does not hold.

We are now able to define the *statewise* prospective reserves. The statewise prospective reserve in state $j \in \mathcal{S}$ is given by

$$\begin{aligned} V_j(t) &= \mathbb{E}[B(t)|X(t) = j] \\ &= \sum_{k \in \mathcal{S}} \int_t^T v(t, u) p_{jk}(t, u) \left(b_k(u) + \sum_{l \in \mathcal{S}: l \neq k} \mu_{kl}(u) b_{kl}(u) \right) du. \end{aligned}$$

The standard Thiele's (backward) differential equation for the reserve is given by

$$\frac{d}{dt} V_j(t) = -b_j(t) + \phi(t) V_j(t) - \sum_{k \in \mathcal{S}: k \neq j} (b_{jk}(t) + V_k(t) - V_j(t)) \mu_{jk}(t), \quad V_j(T) = 0, \quad (4.2.4)$$

where the boundary condition follows because of assumption (4.2.3). Note that the integral form of (4.2.4) is known as Thiele's integral equation of type II. With defining a mapping $W : [t_0, T] \times (0, \infty) \times [0, 1]^{|\mathcal{S}|} \rightarrow \mathbb{R}$ by

$$W(t, v, p) := v \sum_{j \in \mathcal{S}} p_j V_j(t), \quad (4.2.5)$$

the expected present value at time t_0 of the future payments between time t and time T equals

$$\begin{aligned} W(t, v(t), p(t)) &= \sum_{j \in \mathcal{S}} \pi_j \mathbb{E}[v(t) B(t) | X(t_0) = j] \\ &= v(t) \sum_{j \in \mathcal{S}} p_j(t) V_j(t) \\ &= \sum_{k \in \mathcal{S}} \int_t^T p_k(u) v(u) \left(b_k(u) + \sum_{l \in \mathcal{S}: l \neq k} b_{kl}(u) \mu_{kl}(u) \right) du. \end{aligned} \quad (4.2.6)$$

Here, the last equality follows from the Chapman-Kolmogorov equation and by collecting the discounting terms. The functions p and v in (4.2.6) follow from (4.2.1) and (4.2.2) and the initial values $v(t)$ and $p(t)$.

4.3 Calculation of the worst-case reserve

In this section, we derive the worst-case scenario for the probability weighted reserve, $\pi^{\text{tr}} V(t_0)$, which is related to W by (4.2.5). First, we establish a verification lemma and give a heuristic argument for the main ingredients. Second, we show existence of a worst-case scenario. Third, we translate the verification result for the probability weighted reserve W into a corresponding result for the statewise reserves V_j in a corollary. Finally, we outline two numerical methods for calculation of the worst-case reserve.

4.3.1 Verification lemma

We note that $\pi^{\text{tr}} V(t_0) = W(t_0, v(t_0), p(t_0)) = W(t_0, 1, \pi)$. This is useful since dynamic programming applies to W and not to $\pi^{\text{tr}} V$. This allows us to attack our optimization problem by maximizing W for all future time points, whereas $V_j, j \in \mathcal{S}$ is not in itself maximized for $t > t_0$.

In the following, we search for the worst-case reserve (the optimal value function) with respect to a set M , where $M \subset L_1^{1+|\mathcal{J}|}([t_0, T])$ is a set of integrable interest rate and transition intensity

paths. A set M belongs to $L_1^{1+|\mathcal{J}|}([t_0, T])$ if $\sum_{i=1}^{1+|\mathcal{J}|} \int_{t_0}^T |f_i(s)| ds < \infty$ for all $f_i(t)$ in $M_i(t)$. The “slices” $M(t)$ of M ,

$$M(t) = \{(\phi(t), \mu(t)) | (\phi, \mu) \in M\},$$

describe the parameter space at time t . We start by establishing a classical verification lemma.

Proposition 4.3.1. (*Verification lemma*) *Let $\bar{W}$ be a solution to the partial differential equation*

$$\begin{aligned} 0 = & \frac{\partial}{\partial t} \bar{W}(t, v, p) - \bar{\phi}(t, v, p) v \frac{\partial}{\partial v} \bar{W}(t, v, p) + \sum_{k \in \mathcal{S}} p_k \left(v b_k(t) \right. \\ & \left. + \sum_{l \in \mathcal{S}: l \neq k} \bar{\mu}_{kl}(t, v, p) \left(v b_{kl}(t) + \frac{\partial}{\partial p_l} \bar{W}(t, v, p) - \frac{\partial}{\partial p_k} \bar{W}(t, v, p) \right) \right), \quad \bar{W}(T, v, p) = 0, \\ (\bar{\phi}(t, v, p), \bar{\mu}(t, v, p)) = & \operatorname{argmax}_{(f, m) \in M(t)} \left\{ -f v \frac{\partial}{\partial v} \bar{W}(t, v, p) + \sum_{k \in \mathcal{S}} p_k \left(v b_k(t) \right. \right. \\ & \left. \left. + \sum_{l \in \mathcal{S}: l \neq k} m_{kl} \left(v b_{kl}(t) + \frac{\partial}{\partial p_l} \bar{W}(t, v, p) - \frac{\partial}{\partial p_k} \bar{W}(t, v, p) \right) \right) \right\}. \end{aligned} \quad (4.3.1)$$

Furthermore, let the ordinary differential equation system

$$\begin{aligned} 0 = & \frac{d}{dt} \bar{v}(t) + \bar{\phi}(t, \bar{v}(t), \bar{p}(t)) \bar{v}(t), \\ 0 = & \frac{d}{dt} \bar{p}_j(t) - \sum_{k \in \mathcal{S}: k \neq j} (\bar{\mu}_{kj}(t, \bar{v}(t), \bar{p}(t)) \bar{p}_k(t) - \bar{\mu}_{jk}(t, \bar{v}(t), \bar{p}(t)) \bar{p}_j(t)), \quad j \in \mathcal{S}, \end{aligned} \quad (4.3.2)$$

have a unique solution on $[s, T]$ for any initial condition $(\bar{v}(s), \bar{p}(s)) = (v, p)$ at any time $s \in [t_0, T]$ and any pair $(v, p) \in (0, \infty) \times [0, 1]^{|\mathcal{S}|}$. Then

$$\bar{W}(s, v, p) = \sup_{(\phi, \mu) \in M} W(s, v, p; \phi, \mu) = \sup_{(\phi, \mu) \in M} v \sum_{j \in \mathcal{S}} p_j V_j(s; \phi, \mu) \quad (4.3.3)$$

for all triples (s, v, p) .

Proof. See Bertsekas (2005, Section 3.2). □

Note that the notation “ $;\phi, \mu$ ” in $W(s, v, p; \phi, \mu)$ and $V_j(s; \phi, \mu)$ emphasises that W and V_j depends on the entire processes ϕ and μ .

In the following, we heuristically derive the differential equations in Proposition 4.3.2. We start by combining two different differential equations for W for a given (ϕ, μ) . First, the derivative of W with respect to t , using (4.2.6), is given by

$$\frac{d}{dt} W(t, v(t), p(t)) = - \sum_{k \in \mathcal{S}} p_k(t) v(t) \left(b_k(t) + \sum_{l \in \mathcal{S}: l \neq k} b_{kl}(t) \mu_{kl}(t) \right). \quad (4.3.4)$$

Second, we can also consider W as a function of three variables with dynamics given by

$$\begin{aligned} \frac{d}{dt}W(t, v(t), p(t)) &= \frac{\partial}{\partial t}W(t, v(t), p(t)) + \frac{\partial}{\partial v}W(t, v(t), p(t)) \left(\frac{d}{dt}v(t) \right) \\ &\quad + \nabla_p W(t, v(t), p(t)) \left(\frac{d}{dt}p(t) \right), \end{aligned} \quad (4.3.5)$$

where

$$\nabla_p W = \left(\frac{\partial}{\partial p_1} W, \frac{\partial}{\partial p_2} W, \dots, \frac{\partial}{\partial p_n} W \right).$$

By inserting the differential equations for $v(t)$ and $p(t)$ (given by (4.2.1) and (4.2.2), respectively) into (4.3.5), combining with (4.3.4) and rearranging the terms, we obtain the following differential equation

$$\begin{aligned} 0 &= \frac{\partial}{\partial t}W(t, v(t), p(t)) - \phi(t)v(t) \frac{\partial}{\partial v}W(t, v(t), p(t)) \\ &\quad + \sum_{k \in \mathcal{S}} p_k(t) \left(\sum_{l \in \mathcal{S}: l \neq k} \mu_{kl}(t) \left(v(t)b_{kl}(t) + \frac{\partial}{\partial p_l}W(t, v(t), p(t)) - \frac{\partial}{\partial p_k}W(t, v(t), p(t)) \right) \right. \\ &\quad \left. + v(t)b_k(t) \right), \end{aligned} \quad (4.3.6)$$

with terminal condition given by $W(T, v, p) = 0$. Choosing (ϕ, μ) such that the time-derivative is as small as possible at each time point and also using this (ϕ, μ) in the differential equations for v and p give us the equations that characterize the worst-case reserve. We obtain the differential equations for $\bar{v}$ and $\bar{p}$ by inserting the worst-case interest rate into (4.2.1) and the worst-case intensities into a coordinate-wise version of (4.2.2). Hereby, we have heuristically derived the system of differential equations in Proposition 4.3.2.

4.3.2 Existence

We now turn to the question of existence of a worst-case reserve.

Proposition 4.3.2. (*Existence of a worst-case scenario*) *Let M be a compact subset of the space $L_1^{1+|\mathcal{J}|}([t_0, T])$ which contains only nonnegative interest rates ϕ and nonnegative transition intensities μ_{jk} , $(j, k) \in \mathcal{J}$. Then for each $(t, v, p) \in [t_0, T] \times (0, \infty) \times [0, 1]^{|\mathcal{S}|}$ there exists a maximizing argument $(\bar{\phi}, \bar{\mu}) \in M$ for which $W(t, v, p; \bar{\phi}, \bar{\mu}) = \bar{W}(t, v, p)$.*

Proof. Since ϕ and μ_{jk} , $(j, k) \in \mathcal{J}$ are nonnegative, we necessarily have $|v(t)| \leq 1$ and $|p_j(t)| \leq 1$ for all t and j . Therefore, the reserves $V_j(t)$ are uniformly bounded by

$$|V_j(t)| \leq \sum_{k \in \mathcal{S}} \int_t^T \left(|b_k(s)| + \sum_{k \in \mathcal{S}: l \neq k} |b_{kl}(s)| \mu_{kl}(s) \right) ds < C$$

for a finite constant C (which is independent of j , t and μ) since the functions $b_k(t)$ and $b_{kl}(t)$ are bounded and M is a compact subspace of $L_1^{1+|\mathcal{J}|}([t_0, T])$. Consequently, on the set M the mapping $W(t, v, p; \phi, \mu)$ has the uniform upper bound given by

$$W(t, v, p; \phi, \mu) \leq v \sum_{j \in \mathcal{S}} p_j C.$$

Furthermore, Christiansen (2008, Theorem 4.4) showed that the reserves $V_j(t)$ are Fréchet differentiable with respect to the cumulative intensities $t \mapsto \int_{t_0}^t \phi(u)du$ and $t \mapsto \int_{t_0}^t \mu(u)du$ in the total variation norm. Since the operator that maps the intensities to the cumulative intensities is continuous and since the L_1 -norm of the intensities equals the total variation norm of the cumulative intensities, we can conclude that $V_j(t; \phi, \mu)$ is continuous with respect to (ϕ, μ) . Because of the linear representation (4.2.5), the mapping $W(t, v, p; \phi, \mu)$ is continuous in (ϕ, μ) , as well. From the boundedness and continuity of $W(t, v, p; \cdot, \cdot)$ on M and the compactness and completeness of M we can finally conclude that there exists a maximizing argument in M . $\square$

Proposition 4.3.2 gives existence of a worst-case reserve but we do not say anything about uniqueness and existence of a solution to the system of differential equations given by (4.3.2). What type of solution we can possibly expect crucially depends on the set M . We do not dig further into this question in this exposition.

In some examples later on we will construct sets M by specifying the t -slices, and the following lemma will help that we obtain sets that are compact in $L_1^{1+|\mathcal{J}|}$.

Lemma 4.3.3. *Let B, S be subsets of $L_1^{1+|\mathcal{J}|}([t_0, T])$ where B is a bounded and closed set and*

$$S := \left\{ f \in L_1^{1+|\mathcal{J}|}([t_0, T]) : \right. \quad (4.3.7)$$

$$\left. \text{there exists a version of } f \text{ with variation norm not greater than } C \right\}$$

for some $C < \infty$. Then the closure of $B \cap S$ is a compact subset of B .

Proof. Because of Tychonoff's theorem, it suffices to show the lemma just for the space $L_1([t_0, T])$. We let $h \in \mathbb{R}$. Using that all elements in S have a version that has finite variation, we can, with the help of Fubini's theorem, show that

$$\begin{aligned} \sup_{f \in B \cap S} \int_{t_0}^T |f(t+h)\mathbf{1}_{t+h \leq T} - f(t)| dt &\leq \sup_{f \in B \cap S} \int_{t_0}^T \int_{[t, t+h]} d|f|(s) dt \\ &\leq \sup_{f \in B \cap S} \int_{[t_0, T+h]} \int_{s-h}^s dt d|f|(s) \\ &\leq \sup_{f \in B \cap S} Ch, \end{aligned}$$

where $f(t) := 0$ for t outside of $[t_0, T]$ and where $|f| := f_+ + f_-$ for minimal non-decreasing functions f_+, f_- with $f = f_+ - f_-$. Since Ch converges to zero for $h \rightarrow 0$ (uniformly in f), and since B is bounded, from the Kolmogorov-Riesz theorem we can conclude that $B \cap S$ is pre-compact. Hence, the closure of $B \cap S$ is compact. Since B is closed, the closure of $B \cap S$ must be a subset of B . $\square$

In the following, we approach the differential equations by numerical methods. The existence of a worst-case reserve given by Proposition 4.3.2 together with convergence in various numerical calculations indicates that we do approximate a worst-case reserve.

4.3.3 Results for statewise reserves

We have throughout the section worked with the probability weighted reserve W . In order to prepare for the numerical calculations, we reformulate the verification theorem in the following corollary in terms of the statewise reserves V_j . The result follows directly from Proposition 4.3.1 and (4.2.6).

Corollary 4.3.4. *Let the assumptions of Proposition 4.3.1 be fulfilled, and let the ordinary differential equation system*

$$\begin{aligned}
\frac{d}{dt}\bar{V}_k(t) &= -b_k(t) + \bar{V}_k(t)\bar{\phi}(t) - \sum_{l \in \mathcal{S}: l \neq k} (b_{kl}(t) + \bar{V}_l(t) - \bar{V}_k(t)) \bar{\mu}_{kl}(t), \quad \bar{V}_k(T) = 0, \\
\frac{d}{dt}\bar{v}(t) &= -\bar{v}(t)\bar{\phi}(t), \quad \bar{v}(t_0) = 1, \\
\frac{d}{dt}\bar{p}(t) &= -\bar{\mu}^{tr}(t)\bar{p}(t), \quad \bar{p}(t_0) = \pi, \\
(\bar{\phi}(t), \bar{\mu}(t)) &= \operatorname{argmax}_{(f, m) \in M(t)} \left\{ -f \sum_{k \in \mathcal{S}} \bar{p}_k(t) \bar{V}_k(t) \right. \\
&\quad \left. + \sum_{k \in \mathcal{S}} \bar{p}_k(t) \sum_{l \in \mathcal{S}: l \neq k} m_{kl} (b_{kl}(t) + \bar{V}_l(t) - \bar{V}_k(t)) \right\}
\end{aligned} \tag{4.3.8}$$

have a unique solution $\bar{V} = (\bar{V}_1, \dots, \bar{V}_n)^{tr}$. Then

$$\bar{v}(t)\bar{p}^{tr}(t)\bar{V}(t) = \sup_{(\phi, \mu) \in M} \bar{v}(t) \sum_{j \in \mathcal{S}} \bar{p}_j(t) V_j(t; \phi, \mu), \tag{4.3.9}$$

in particular

$$\pi^{tr}\bar{V}(t_0) = \sup_{(\phi, \mu) \in M} \sum_{j \in \mathcal{S}} \pi_j V_j(t_0; \phi, \mu). \tag{4.3.10}$$

By solving the system (4.3.8) we are also able to calculate $\bar{W}$ and $\bar{V}$. Note, we have implicitly assumed that the interest rate and the intensities are not allowed to depend on the current state of the Markov chain. We do not know $\bar{p}(T)$ and $\bar{v}(T)$ but if we make a guess of the values $\bar{p}(T)$ and $\bar{v}(T)$ (this is a guess of dimension $(n+1)$ in total) and the guess is correct, we get that $\bar{p}(t_0) = \pi$ and $\bar{v}(t_0) = 1$. The problem is that we do not know the boundary conditions for all the differential equations at the same time point. This implies that we cannot use standard iterative, numerical methods to solve the differential equations. We now outline two methods, which can be used to obtain results numerically.

Shooting method: One way to overcome problems with boundary conditions at different time points is to apply the shooting method, see e.g. Orava and Lautala (1976). Here, we need to guess terminal conditions for $\bar{v}$ and $\bar{p}$ and then “shoot” until we hit the “right” starting values for $\bar{v}$ and $\bar{p}$. The shooting method is a method aiming at updating these guesses in order to obtain the right starting values as fast as possible. The system of ordinary differential equations given by (4.3.8) consists in total of $2n+1$ equations; $n+1$ forward equations and n backward equations. The standard way to choose whether to guess for the missing starting or terminal values is to make a guess of the lowest possible dimension. In the present case, this approach implies that we should guess the starting values for V_j and solve the equation system forward. Note however, that the differences between the two choices for the present case are minimal. Assuming that we make initial guesses $x_{T,0}$ and $y_{T,0}$ for the terminal values of $\bar{p}(T)$ and $\bar{v}(T)$ we obtain results $\bar{p}^{x_{T,0}}(t_0)$ and $\bar{v}^{y_{T,0}}(t_0)$, which we hope are close to π and 1, respectively. One should of course choose $x_{T,0}$ in the set $[0, 1]^n$ and $y_{T,0}$ in the set $[0, 1]$ since they are transition probabilities and a discount factor. Now the aim is to find roots for the function f given by

$$f(x, y) = \begin{pmatrix} \pi \\ 1 \end{pmatrix} - \begin{pmatrix} \bar{p}^x(t_0) \\ \bar{v}^y(t_0) \end{pmatrix}.$$

We can find these roots by choosing some starting values and apply a standard algorithm like Newton's Method to update the guesses. Hereby, we obtain a series of terminal conditions $(x_{T,i}, y_{T,i})$, $i = 0, 1, \dots$. We stop the algorithm, when each entry of $f(x_{T,i}, y_{T,i})$ is below a given tolerance level ϵ .

Fixed point equation method: Another numerical method that can be used to solve the system of differential equations is the “fixed point equation method”, see e.g. Bailey et al. (1968). They study problems a bit different from ours but the iteration idea is the same. They denote the method “Picard iteration” instead of “fixed-point equation method”. We note that this method is the one that has been used to obtain the numerical results in Section 4.5. The method is an iterative algorithm, which aims at solving the equation system (4.3.8) within a given tolerance level. The approach is to apply the following algorithm:

1. Choose a reasonable starting interest rate and transition intensities $(\bar{\phi}^0$ and $\bar{\mu}^0)$.
2. Solve the equations for $\bar{v}$ and $\bar{p}$ (forwards) using $\bar{\phi}^0$ and $\bar{\mu}^0$ and denote the solutions $\bar{v}^0$ and $\bar{p}^0$.
3. Solve the system of equations for $\bar{V}_j$ and $(\bar{\phi}, \bar{\mu})$ (backwards) using the values obtained in the former steps. We denote the solutions $\bar{V}_j^0$ and $(\bar{\phi}^1, \bar{\mu}^1)$.
4. Repeat step 2 and 3 (and increase the numbers of the superscripts accordingly) $\bar{i}$ times, where $\bar{i}$ is defined by

$$\bar{i} = \operatorname{argmin}_{i \in \mathbb{N}} \sup_{t \in [t_0, T]} \max \left\{ \max_{j \in \mathcal{S}} \left\{ \bar{V}_j^i(t) - \bar{V}_j^{i-1}(t) \right\}, \max_{j \in \mathcal{S}} \left\{ \bar{p}_j^i(t) - \bar{p}_j^{i-1}(t) \right\}, \bar{v}^i(t) - \bar{v}^{i-1}(t), \right. \\ \left. \bar{\phi}^{i+1}(t) - \bar{\phi}^i(t), \max_{(j,k) \in \mathcal{J}} \left\{ \bar{\mu}_{jk}^{i+1}(t) - \bar{\mu}_{jk}^i(t) \right\} \right\} < \epsilon,$$

where ϵ is a given tolerance level. The fixed-point equation method is not necessarily converging, unless we make further model restrictions; in particular assumptions on the set M in (4.3.8). However, once we have found a sequence $(\bar{V}^i, \bar{p}^i, \bar{v}^i)_{i \in \mathbb{N}}$ that converges pointwise to a limit $(\bar{V}^*, \bar{p}^*, \bar{v}^*)$, we can, under mild conditions, conclude that the latter limit is a solution to (4.3.8): Given that the assumptions of Proposition 4.3.2 hold and that the argmax in (4.3.8) is continuous with respect to the parameters $\bar{V}^i(t)$ and $\bar{p}^i(t)$ at $\bar{V}^*(t)$ and $\bar{p}^*(t)$, we can show that $(\bar{V}^*(t), \bar{p}^*(t), \bar{v}^*(t))$ solves the integral version of (4.3.8). This is shown by applying the dominated convergence theorem and using the constant C in the proof of Proposition 4.3.2 as a uniform majorant for the functions $\bar{V}_j^i(t)$.

Remark 4.3.5. *If we assume that the argmax in (4.3.8) does not depend on any of the factors $\bar{p}_k(t)$ for all $k \in \mathcal{S}$ things get numerically simpler. Unfortunately, this only holds for a very limited type of models like a multiple causes of death model, see Christiansen and Steffensen (2013). The numerical simplification is that the calculation of $\bar{p}$ and $\bar{\mu}$ decouples from the calculation of $\bar{V}_k$. That is, one can first solve the ODEs for $\bar{V}_k$ with $\bar{v}(t)\bar{p}(t) = 1$ (backwards) to obtain values for $(\bar{\mu}, \bar{\phi})$, secondly use these results to calculate $(\bar{p}, \bar{v})$ (forwards), and lastly use the results of $(\bar{\mu}, \bar{\phi})$ and $(\bar{p}, \bar{v})$ to calculate $\bar{V}_k$ (backwards). In this case the use of the shooting method is not necessary. If the argmax in (4.3.8) is dependent on but constant in the factors $\bar{p}_k(t)$ one can do exactly the same as before by basing the first values $\bar{v}$ and $\bar{p}$ on some arbitrary values of μ and ϕ in M . These simplifications are for a fixed starting state covered by Christiansen and Steffensen (2013).*

4.3.4 Some notes about the set M

One crucial ingredient of the calculations in the present paper is the set M , over which we maximize the probability weighted reserve. To make the calculations useful for a company they need to know certain probabilistic properties of the set, e.g. the confidence level. One non-statistical way of obtaining a set is to use an expert opinion. However, it is hard to deduce any properties from sets based on an expert opinion.

An alternative and more attractive method is to deduce a set with a certain coverage from a stochastic model by either analytical or numerical methods. One could ask for the advantages of this approach compared to just simulating reserves and finding confidence intervals numerically. The answer to this is twofold. First, you gain some insight about the nature of your risk by calculating the sets. Second, for many life insurance companies this method will be computationally faster and for some easier to implement. The speedup relative to directly simulating the reserves occurs because you only need to solve the system of differential equations characterizing the reserve once, after having found the set. The alternative is to simulate scenarios and solve differential equations over and over again to get reliable results. This is particularly important when making calculations for complex products modeled in Markov chains with many states in which case the differential equations take a long time to solve. An additional advantage comes from the possibility of using the same calculated sets for different products (for the same policyholder) and for different policyholders, who have similar characteristics. We do not think that it is a big drawback to use a parametric approach, since many companies already deal with estimation in such models when e.g. forecasting the future mortality.

4.4 Worst-case calculations for portfolios

In this section we consider an *inhomogeneous* portfolio, whereby allowing for nonidentical distributions of the jump processes modelling the policyholders. We generally assume that the policyholders are *stochastically independent* and that the force of interest is the same for all contracts. Given this independence assumption and assuming that there exists an intensity matrix for each policyholder, we can conclude that two or more policyholders change state at the same point in time with probability zero.

If the same set M is used for all policyholders in the portfolio but no interactions between the worst cases of different contracts are taken into account, we can easily obtain the worst-case reserve for the portfolio by calculating the worst-case reserve for each policyholder separately and then summing over all the reserves. In practice, we think the mortalities across the population are closely connected (dependent), so this is the case we want to study. Such a study is not doable within the theory of Christiansen and Steffensen (2013), because the requirements of that paper are not fulfilled in this case. This is because dependence between policyholders in a portfolio implies that the argmax in Christiansen and Steffensen (2013, Proposition 4.1) is not constant with respect to all the discounted transition probabilities; see also Remark 4.3.5. The simple approach outlined above leads to a rough upper bound only, for the worst-case in a portfolio. We illustrate this with a numerical example in Section 4.5.

There are a few simple cases of inhomogeneous portfolios, where we can confine ourselves to the theory of Section 4.3. One of the cases, which was shortly mentioned in the beginning of Section 4.3, is where the policyholders are governed by the same intensities and have the same insurance product but have different starting distributions. This case is covered because it is equivalent to calculation of the worst-case scenario for a single policy with a specific starting distribution. Another simple case is where the policyholders of the portfolios are governed by the

same intensities and starting distributions but have different insurance products. In this case, the worst-case reserve of the portfolio can be found as the worst-case reserve of a single policy, where the policyholder has the sum of the products of all the policyholders in the portfolio. However, in general we need some extended results to cover portfolios.

4.4.1 A representation for the worst-case reserve

The aim of this subsection is to find a simple representation for the worst-case reserve for a portfolio of policyholders. To do so, we start by introducing some additional notation for the setting with multiple policyholders. For the terms where the notation from Section 4.3 is applicable, we do not repeat these.

- L : Number of policyholders in the portfolio.
- $(\mathbf{X}(t))_{t \in [0, T]} = (X_1(t), \dots, X_L(t))_{t \in [0, T]}$: Representation of the state of the policyholders of the portfolio.
- $\mathcal{S} = \{(\mathcal{S}_1, \dots, \mathcal{S}_L) | \mathcal{S}_i \subset \mathcal{S}\}$: State space for all policies.
- $\mathcal{J} = \left\{(\mathbf{j}, \mathbf{k}) \in \mathcal{S} \times \mathcal{S} | \mathbf{j} = (j_1, \dots, j_l, \dots, j_L), \mathbf{k} = (j_1, \dots, \tilde{j}_l, \dots, j_L), j_l \neq \tilde{j}_l, l \in \{1, \dots, L\}\right\}$:
Transition space for all policies. The condition in the above formula means, that the vectors $\mathbf{j}$ and $\mathbf{k}$ differ in exactly one coordinate.
- $V_i^l(t)$: Statewise reserve at time t for policyholder l in state i .
- $b_{jk}^l(t)$ and $b_j^l(t)$: Lump sum and continuous payments at time t for policyholder l .
- $\mu_{jk}^l(t)$: Transition intensities at time t for policyholder l .
- $p_{jk}^l(s, t)$: Transition probabilities between time t and s for policyholder l .

In the following, we show that the “big model” collapses to something simpler. By “big model” we mean that the portfolio is modeled as a single policy, such that we can use the theory of Section 4.3. In the “big model”, we have $|\mathcal{S}_1| \cdot \dots \cdot |\mathcal{S}_L|$ states that cover all the different combinations of policyholders and states. By defining

$$b_{\mathbf{j}}(t) = \sum_{l=1}^L b_{j_l}^l(t) \text{ and } b_{\mathbf{j}\mathbf{k}}(t) = b_{j_l \tilde{j}_l}^l(t)$$

for $\mathbf{j} = (j_1, \dots, j_L)$ and $\mathbf{k} = (j_1, \dots, \tilde{j}_l, \dots, j_L)$ (recall that two or more jumps at the same time occur with probability zero), we get that the present portfolio value $\mathbf{B}(t)$ equals the sum of the

single present values,

$$\begin{aligned}
\mathbf{B}(t) &= \sum_{\mathbf{j} \in \mathcal{S}} \int_t^T v(t, s) \mathbf{1}_{\{\mathbf{X}(s)=\mathbf{j}\}} b_{\mathbf{j}}(s) ds + \sum_{(\mathbf{j}, \mathbf{k}) \in \mathcal{J}} \int_t^T v(t, s) b_{\mathbf{j}\mathbf{k}}(s) dN_{\mathbf{j}\mathbf{k}}(s) \\
&= \sum_{\mathbf{j} \in \mathcal{S}} \int_t^T v(t, s) \mathbf{1}_{\{\mathbf{X}(s)=\mathbf{j}\}} \sum_{l=1}^L b_{j_l}^l(s) ds + \sum_{(\mathbf{j}, \mathbf{k}) \in \mathcal{J}} \int_t^T v(t, s) b_{j_l \tilde{j}_l}^l(s) dN_{j_l \tilde{j}_l}^l(s) \\
&= \sum_{l=1}^L \sum_{j_1 \in \mathcal{S}_1, \dots, j_L \in \mathcal{S}_L} \int_t^T v(t, s) \prod_{k=1}^L \mathbf{1}_{\{X_k(s)=j_k\}} b_{j_l}^l(s) ds \\
&\quad + \sum_{l=1}^L \sum_{(j, k) \in \mathcal{J}_l} \int_t^T v(t, s) b_{jk}^l(s) dN_{jk}^l(s) \\
&= \sum_{l=1}^L \left(\sum_{j \in \mathcal{S}_l} \int_t^T v(t, s) \mathbf{1}_{\{X_l(s)=j\}} b_j^l(s) ds + \sum_{(j, k) \in \mathcal{J}_l} \int_t^T v(t, s) b_{jk}^l(s) dN_{jk}^l(s) \right) \\
&= \sum_{l=1}^L B^l(t),
\end{aligned} \tag{4.4.1}$$

where B^l is the present value of future payments for policyholder l , and N_{jk}^l counts the number of jumps from state j to state k for policyholder l .

We have that the reserve for each policyholder follows a differential equation of the type (4.2.4) and that the transition probabilities, p_j^l , follow the Kolmogorov equations given by (4.2.2). We now want to find the worst-case scenario for the portfolio. For this purpose, we define a mapping $\mathbf{W} : [t_0, T] \times (0, \infty) \times [0, 1]^{\mathbf{S}} \rightarrow \mathbb{R}$ by

$$\mathbf{W}(t, v, \mathbf{p}) = v \sum_{\mathbf{j} \in \mathcal{S}} \mathbf{p}_{\mathbf{j}} \mathbb{E}[\mathbf{B}(t) | \mathbf{X}(t) = \mathbf{j}].$$

That is, we want to find the worst-case scenario for

$$\begin{aligned}
\mathbf{W}(t, v(t), \mathbf{p}(t)) &= v(t) \sum_{\mathbf{j} \in \mathcal{S}} \pi_{\mathbf{j}} \mathbb{E}[\mathbf{B}(t) | \mathbf{X}(t_0) = \mathbf{j}] \\
&= v(t) \sum_{\mathbf{j} \in \mathcal{S}} \mathbf{p}_{\mathbf{j}}(t) \mathbb{E}[\mathbf{B}(t) | \mathbf{X}(t) = \mathbf{j}],
\end{aligned}$$

where $\mathbf{p}(t) = (P(\mathbf{X}(t) = \mathbf{j}))_{\mathbf{j} \in \mathcal{S}}$ and $\pi_{\mathbf{j}} = \pi_{j_1}^1 \cdot \dots \cdot \pi_{j_L}^L$ because of the independence assumption. Using again the stochastic independence of the policyholders and the additive structure of (4.4.1), we can show that

$$\begin{aligned}
\mathbf{W}(t, v(t), \mathbf{p}(t)) &= v(t) \sum_{l=1}^L \sum_{j \in \mathcal{S}_l} \pi_j^l \mathbb{E}[B^l(t) | X_l(t_0) = j] \\
&= \sum_{l=1}^L \sum_{j \in \mathcal{S}_l} \int_t^T p_j^l(u) v(u) \left(b_j^l(u) + \sum_{k \in \mathcal{S}_l: k \neq j} b_{jk}^l(u) \mu_{jk}^l(u) \right) du.
\end{aligned}$$

We aim to find

$$\sup_{(\phi, \mu^1, \dots, \mu^L) \in M} \mathbf{W}(t_0, v(t_0), \mathbf{p}(t_0); \phi, \mu^1, \dots, \mu^L), \tag{4.4.2}$$

where $M \subset L_1^{1+|\mathcal{J}|}([t_0, T])$ is a set of integrable interest rate and transition intensity paths. The starting point t_0 is the same for all policyholders of the portfolio and should therefore be thought of as a calendar time point rather than an age in case of a heterogeneous portfolio. Considering this portfolio as a single contract on the state space $\mathcal{S}$, one can apply the theory of Section 4.3 and obtain the desired results. Because of the huge state space $\mathcal{S}$, the computational workload seems enormous. However, we can simplify the formulas in (4.3.8).

Proposition 4.4.1. *For the “big model” the solutions of the differential equation system given by (4.3.8) are equivalent to the solutions of*

$$\begin{aligned}
\frac{d}{dt} \bar{V}_j^l(t) &= -b_j^l(t) + \bar{V}_j^l(t) \bar{\phi}(t) - \sum_{k \in \mathcal{S}_l: k \neq j} \left(b_{jk}^l(t) + \bar{V}_k^l(t) - \bar{V}_j^l(t) \right) \bar{\mu}_{jk}^l(t), \\
\bar{V}_j^l(T) &= 0, \quad l = 1, \dots, L, \\
\frac{d}{dt} \bar{v}(t) &= -\bar{v}(t) \bar{\phi}(t), \quad \bar{\phi}(t_0) = 1, \\
\frac{d}{dt} \bar{p}^l(t) &= -\left(\bar{\mu}^l(t) \right)^{tr} \bar{p}^l(t), \quad \bar{p}(t_0) = \pi^l, \quad l = 1, \dots, L, \\
(\bar{\phi}(t), \bar{\mu}^1(t), \dots, \bar{\mu}^L(t)) &= \underset{(f, m^1, \dots, m^L) \in M(t)}{\operatorname{argmax}} \left\{ -f \sum_{l=1}^L \sum_{j_l \in \mathcal{S}_l} \bar{p}_{j_l}^l(t) \bar{V}_{j_l}^l(t) + \sum_{l=1}^L \sum_{(j_l, \tilde{j}_l) \in \mathcal{J}_l} \bar{p}_{j_l}^l(t) m_{j_l \tilde{j}_l}^l \right. \\
&\quad \left. \times \left(b_{j_l \tilde{j}_l}^l(t) + \bar{V}_{\tilde{j}_l}^l(t) - \bar{V}_{j_l}^l(t) \right) \right\}.
\end{aligned} \tag{4.4.3}$$

Proof. We start by considering the argmax given in (4.3.2) for the entire portfolio. The interior of the argmax is given by

$$\begin{aligned}
& -fv \frac{\partial}{\partial v} \bar{\mathbf{W}} + \sum_{\mathbf{j} \in \mathcal{S}} \mathbf{p}_{\mathbf{j}} \left(vb_{\mathbf{j}} + \sum_{\mathbf{k} \in \mathcal{S}: \mathbf{k} \neq \mathbf{j}} m_{\mathbf{j}\mathbf{k}} (vb_{\mathbf{j}\mathbf{k}} + \nabla_{\mathbf{p}_{\mathbf{k}}} \bar{\mathbf{W}} - \nabla_{\mathbf{p}_{\mathbf{j}}} \bar{\mathbf{W}}) \right) \\
&= -fv \frac{\partial}{\partial v} \bar{\mathbf{W}} + \sum_{\mathbf{j} \in \mathcal{S}} \mathbf{p}_{\mathbf{j}} vb_{\mathbf{j}} + \sum_{(\mathbf{j}, \mathbf{k}) \in \mathcal{J}} \mathbf{p}_{\mathbf{j}} m_{\mathbf{j}\mathbf{k}} (vb_{\mathbf{j}\mathbf{k}} + \nabla_{\mathbf{p}_{\mathbf{k}}} \bar{\mathbf{W}} - \nabla_{\mathbf{p}_{\mathbf{j}}} \bar{\mathbf{W}}).
\end{aligned} \tag{4.4.4}$$

In (4.4.4), $\mathbf{j}$ and $\mathbf{k}$ are vectors of dimension L . We use that the lives of the policyholders are independent conditional on the intensities and assume that $\mathbf{j} = (j_1, \dots, j_l, \dots, j_L)^{tr}$ and $\mathbf{k} = (j_1, \dots, \tilde{j}_l, \dots, j_L)^{tr}$ with $j_l \neq \tilde{j}_l$. This means that we obtain:

$$\begin{aligned}
(4.4.4) &= -fv \sum_{l=1}^L \sum_{j_l \in \mathcal{S}_l} p_{j_l}^l V_{j_l}^l + v \sum_{l=1}^L \sum_{j_l \in \mathcal{S}_l} p_{j_l}^l b_{j_l}^l \\
&\quad + \sum_{l=1}^L \sum_{(j_l, \tilde{j}_l) \in \mathcal{J}_l} p_{j_l}^l m_{j_l \tilde{j}_l} v \left(b_{j_l \tilde{j}_l}^l + V_{j_1}^1 + \dots + V_{\tilde{j}_l}^l + \dots + V_{j_L}^L - \sum_{i=1}^L V_{j_i}^i \right) \\
&= -fv \sum_{l=1}^L \sum_{j_l \in \mathcal{S}_l} p_{j_l}^l V_{j_l}^l + \sum_{l=1}^L \sum_{j_l \in \mathcal{S}_l} v p_{j_l}^l b_{j_l}^l + \sum_{l=1}^L \sum_{(j_l, \tilde{j}_l) \in \mathcal{J}_l} p_{j_l}^l m_{j_l \tilde{j}_l} v \left(b_{j_l \tilde{j}_l}^l + V_{\tilde{j}_l}^l - V_{j_l}^l \right).
\end{aligned} \tag{4.4.5}$$

We see that (4.4.5) splits into components relating to each of the policyholders and that the argmax does not depend on the second of the three terms. That is, we have proved the proposition. $\square$

In the following two subsections, we show examples of worst-case scenarios for specific models of a portfolio.

4.4.2 Example 1: Solvency II

In this subsection we illustrate how the general theory above is used to generate a mortality stress scenario for a portfolio of contracts. This is a way of constructing a stress scenario as for instance the mortality stress in Solvency II. Depending on whether the portfolio is a specific portfolio or a stylised market portfolio the generated scenario is internally based or input to a standard calculation, respectively.

We consider a simple two-state life-death model and assume that there is no payments in the state “dead”. We are maximizing the reserve with respect to compact sets of the form

$$M = \overline{S \cap B},$$

where S is given by (4.3.7) and B is defined via its slices

$$B(t) = \left\{ (\phi(t), \mu^1(t), \dots, \mu^L(t)) \in \mathbb{R}_+^{L+1} \mid \right. \\ \left. \phi(t) \in \Phi(t), \mu^1(t) = \hat{\mu}^1(t)\alpha(t), \dots, \mu^L(t) = \hat{\mu}^L(t)\alpha(t) \right\}.$$

We assume that $\alpha(t) \in [\alpha_l(t), \alpha_h(t)]$ and $\Phi(t) = [\phi_l(t), \phi_h(t)]$ and that $\alpha_l, \alpha_h, \phi_l, \phi_h$ and $\hat{\mu}^i$ are bounded functions. Note that M is compact in $L_1^{1+|J|}$ by Lemma 4.3.3. Assuming that the L policyholders have different ages $x_1, \dots, x_L$ at the current time, a natural choice would be to model $\hat{\mu}^i$ of the form

$$\hat{\mu}^i(t) = \mu^{\text{be}}(x_i + t)\Lambda(x_i, t),$$

where μ^{be} is a best estimate mortality intensity and Λ is a longevity factor meaning that Λ is decreasing in time. In this setup the interest rate is independent of the mortality intensities and the mortality intensities are linearly dependent.

The argmax in (4.4.3) becomes

$$(\bar{\phi}(t), \bar{\mu}^1(t), \dots, \bar{\mu}^L(t)) \\ = \operatorname{argmax}_{(f, m^1, \dots, m^L) \in M(t)} \left\{ -f \sum_{l=1}^L \bar{p}_a^l(t) \bar{V}_a^l(t) + \sum_{l=1}^L \bar{p}_a^l(t) m^l \left(b_{ad}^l(t) - \bar{V}_a^l(t) \right) \right\}.$$

Because of linearity, this argmax can be found by calculating

$$\operatorname{argmax}_{(f, \alpha) \in (\Phi(t) \times [\alpha_l(t), \alpha_h(t)])} \left\{ -f \sum_{l=1}^L \bar{p}_a^l(t) \bar{V}_a^l(t) + \alpha \sum_{l=1}^L \bar{p}_a^l(t) \hat{\mu}^l(t) \left(b_{ad}^l(t) - \bar{V}_a^l(t) \right) \right\}$$

and multiplying $\hat{\mu}^1, \dots, \hat{\mu}^L$ with α .

We can also consider a portfolio of disability contracts, see Figure 4.1. That is, we study contracts which can be modeled within a three state Markov chain where recovery from “disabled” to

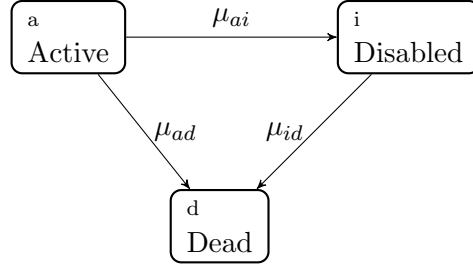


Figure 4.1: Disability model without recovery.

“active” is not possible. We assume no payments in the state “dead” and that μ_{ai} is fixed in the sense that it is a deterministic function which we are not maximizing the reserve with respect to. We want to find the worst-case reserve of this portfolio with respect to ϕ , μ_{ad} and μ_{id} with respect to the sets of the form

$$M = \overline{S \cap B},$$

where S is given by (4.3.7) and B is defined via its slices

$$\begin{aligned}
B(t) = \Big\{ & (\phi(t), \mu_{ad}^1(t), \mu_{id}^1(t), \dots, \mu_{ad}^L(t), \mu_{id}^L(t)) \in \mathbb{R}_+^{2L+1} \Big| \phi(t) \in \Phi(t), \\
& \mu_{ad}^1(t) = \hat{\mu}_{ad}^1(t)\alpha(t), \mu_{id}^1(t) = \hat{\mu}_{id}^1(t)\beta(t), \dots, \\
& \mu_{ad}^L(t) = \hat{\mu}_{ad}^L(t)\alpha(t), \mu_{id}^L(t) = \hat{\mu}_{id}^L(t)\beta(t) \Big\}.
\end{aligned}$$

We assume that $\Phi(t)$ is defined as above and that $\alpha(t) \in [\alpha_l(t), \alpha_h(t)]$ and $\beta(t) \in [\beta_l(t), \beta_h(t)]$ for bounded functions $\alpha_l, \alpha_h, \beta_l, \beta_h, \hat{\mu}_{ad}^i$ and $\hat{\mu}_{id}^i$. Note that M is compact in $L_1^{1+|J|}$ by Lemma 4.3.3.

Because of linearity, we can obtain $(\bar{\phi}(t), \bar{\mu}^1(t), \dots, \bar{\mu}^L(t))$ by calculating

$$\begin{aligned}
& \operatorname{argmax}_{(f, \alpha, \beta) \in (\Phi(t) \times [\alpha_l(t), \alpha_h(t)] \times [\beta_l(t), \beta_h(t)])} \Big\{ -f \sum_{l=1}^L \left(\bar{p}_a^l(t) \bar{V}_a^l(t) + \bar{p}_i^l(t) \bar{V}_i^l(t) \right) \\
& \quad + \alpha \sum_{l=1}^L \bar{p}_a^l(t) \hat{\mu}_{ad}^l(t) \left(b_{ad}^l(t) - \bar{V}_a^l(t) \right) \\
& \quad + \beta \sum_{l=1}^L \bar{p}_i^l(t) \hat{\mu}_{id}^l(t) \left(b_{id}^l(t) - \bar{V}_i^l(t) \right) \Big\}
\end{aligned}$$

and multiplying $\hat{\mu}_{ad}^1, \dots, \hat{\mu}_{ad}^L$ with α and $\hat{\mu}_{id}^1, \dots, \hat{\mu}_{id}^L$ with β .

In the above calculations the death and disability intensities are independent. Another possibility could be to make μ_{ad} and μ_{id} dependent. This is exactly what is described in Section 4.4.3. In the case $\alpha_l = \alpha_h$, we optimize over a singleton with respect to μ_{ad} for each time point and the argmax becomes trivial.

4.4.3 Example 2: Dependent version of the Solvency II example

This example is an extension of the result in Section 4.4.2 where we include sets of the form given by the case “Dependence” in Figure 4.2. Again, we assume no payments in the state “dead”. We

are maximizing the reserve with respect to sets of the form

$$M = \overline{S \cap B},$$

where S is given by (4.3.7) and B is defined via its slices

$$\begin{aligned} B(t) = \Big\{ & (\phi(t), \mu_{ad}^1(t), \mu_{id}^1(t), \dots, \mu_{ad}^L(t), \mu_{id}^L(t)) \in \mathbb{R}_+^{2L+1} \Big| \\ & \phi(t) \in \Phi(t), \mu_{ad}^1(t) = \hat{\mu}_{ad}^1(t)\alpha(t), \mu_{id}^1(t) = \hat{\mu}_{id}^1(t)\beta(t), \dots, \mu_{ad}^L(t) = \hat{\mu}_{ad}^L(t)\alpha(t), \\ & \mu_{id}^L(t) = \hat{\mu}_{id}^L(t)\beta(t), (\alpha(t), \beta(t)) \in \check{B}(t) \Big\}. \end{aligned}$$

We assume that $\Phi(t)$ is given as in Section 4.4.2 and $\check{B}(t)$ is given on the linear programming form

$$\check{B}(t) = \left\{ (x, y) \in \mathbb{R}_+^2 \left| \begin{pmatrix} -\frac{3(\beta_h(t)-\beta_l(t))}{\alpha_h(t)-\alpha_l(t)} & 1 \\ \frac{3(\beta_h(t)-\beta_l(t))}{\alpha_h(t)-\alpha_l(t)} & -1 \\ -\frac{\beta_h(t)-\beta_l(t)}{3(\alpha_h(t)-\alpha_l(t))} & 1 \\ \frac{\beta_h(t)-\beta_l(t)}{3(\alpha_h(t)-\alpha_l(t))} & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \leq \begin{pmatrix} -\frac{3(\beta_h(t)-\beta_l(t))}{\alpha_h(t)-\alpha_l(t)}\alpha_l(t) + \beta_l(t) \\ \frac{3(\beta_h(t)-\beta_l(t))}{\alpha_h(t)-\alpha_l(t)}\alpha_h(t) - \beta_h(t) \\ -\frac{\beta_h(t)-\beta_l(t)}{3(\alpha_h(t)-\alpha_l(t))}\alpha_l(t) + \beta_l(t) \\ \frac{\beta_h(t)-\beta_l(t)}{3(\alpha_h(t)-\alpha_l(t))}\alpha_h(t) - \beta_h(t) \end{pmatrix} \right\}.$$

We choose the functions $\alpha_h, \alpha_l, \beta_h, \beta_l$ in such a way that the slices $\check{B}(t)$ are closed and uniformly bounded (in t). Moreover, we assume that the functions $\hat{\mu}_{ad}^i$ and $\hat{\mu}_{id}^i$ are bounded. Note that M is compact in $L_1^{1+|J|}$ by Lemma 4.3.3. Because of linearity, we can find the argmax $(\bar{\phi}(t), \bar{\mu}^1(t), \dots, \bar{\mu}^L(t))$ in (4.4.3) by calculating

$$\begin{aligned} \operatorname{argmax}_{(f, \alpha, \beta) \in (\Phi(t) \times \check{B}(t))} \Big\{ & -f \sum_{l=1}^L \left(\bar{p}_a^l(t) \bar{V}_a^l(t) + \bar{p}_i^l(t) \bar{V}_i^l(t) \right) + \alpha \sum_{l=1}^L \bar{p}_a^l(t) \hat{\mu}_{ad}^l(t) \left(b_{ad}^l(t) - \bar{V}_a^l(t) \right) \\ & + \beta \sum_{l=1}^L \bar{p}_i^l(t) \hat{\mu}_{id}^l(t) \left(b_{id}^l(t) - \bar{V}_i^l(t) \right) \Big\} \end{aligned} \quad (4.4.6)$$

and multiplying $\hat{\mu}_{ad}^1, \dots, \hat{\mu}_{ad}^L$ with α and $\hat{\mu}_{id}^1, \dots, \hat{\mu}_{id}^L$ with β . To find (4.4.6), we must at each time point check each of the four extremal points of the set $\check{B}(t)$ combined with the extremal points of $\Phi(t)$.

4.5 Numerical calculations

We have performed the numerical calculations in this section by applying the “fixed point equation method” described in Section 4.3. In all our examples we obtained converging fixed-point sequences, and in a neighborhood of the limit the argmax in (4.3.8), seen as a mapping of reserves and transition probabilities, turned out to be continuous for almost all t . Taking into account the arguments in the lines before Remark 4.3.5, our numerical results are indeed approximations for the solutions of (4.3.8). First, we consider numerical calculations for a single a policy. Next, we consider similar calculations for an inhomogeneous portfolio.

4.5.1 Numerical calculations for a single policy

We consider the example of a simple disability policy described in Figure 4.1, where the payments are given by disability benefits in the state “Disabled” at a yearly rate $b_i = 1$ and lump sum payments paying out an amount of 3 upon transcription to the state “Dead” from either of the states “Active” or “Disabled”. For simplicity, we assume that no premiums are paid.

In the example we consider a person at the age of 35 and contract expiry at the age of 65. We let the short rate be 2% and let both the intensity from “Active” to “Disabled” (which we consider fixed) and the best estimate death intensity be given on a Gompertz-Makeham form. The exact intensity parameters are given in Table 4.1. We consider the same lower and upper

μ_{ad}^{be}	μ_{ai}
$0.0025 + 10^{5.804-10+0.038\chi}$	$0.00148 + 10^{4.97136-10+0.06\chi}$

Table 4.1: Best estimate intensities for a policyholder at age χ .

bounds for both the active-death and the disabled-death intensities. The lower bound is given by $U(t) = 0.8\mu_{ad}^{be}(t)$ whereas the upper bound is given by $L(t) = 1.15\mu_{ad}^{be}(t)$.

In the following, we find the worst-case scenario for different sets M using numerical methods. The argmax in (4.3.8) can be either easy or hard to obtain depending on the form of the slices of the set M . If we can formulate the optimization problem as a linear program, which is the case for all the sets presented in Figure 4.2, we know that we only need to search for the argmax in the extremal points of the sets. That is, for the two cases “Independence” and “Dependence”, we only need to evaluate the object function in four points, whereas for the case “Linear dependence”, we only need to evaluate the object function in two points. In the case of a linear program, the extremal points are quite obvious. In the more general case of a strictly convex set $M(t)$, Christiansen and Steffensen (2013, Appendix) outlines a way of obtaining the extremal points.

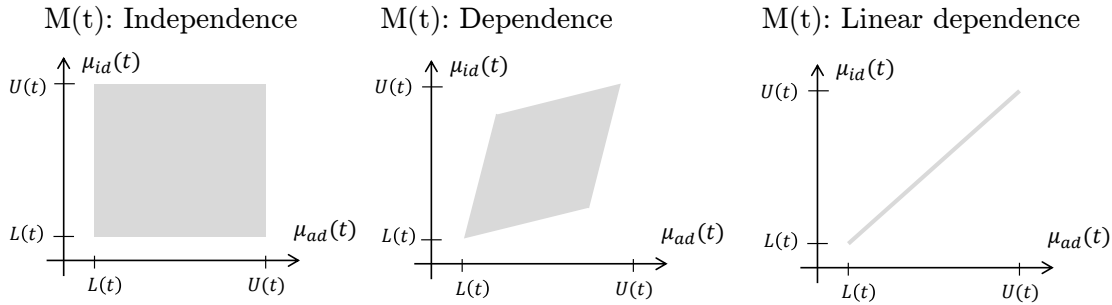


Figure 4.2: Three different trust regions.

The figures 4.3-4.5 show the worst-case bases (conditional on that the current state is “Active”) for the three different types of sets depicted in Figure 4.2. In the case “Dependence”, the four extremal points are

$$\{(L(t), L(t)), (L(t) + 0.25(U(t) - L(t)), L(t) + 0.75(U(t) - L(t))), (U(t), U(t)), (L(t) + 0.75(U(t) - L(t)), L(t) + 0.25(U(t) - L(t)))\}.$$

In the case of independence, the worst-case scenario is that the intensity μ_{ad} is as high as possible throughout the entire period, since the chances of getting disabled is not that high. On the other

hand, the intensity μ_{id} is only high at the very last part of the period of the contract, because there are no more disability benefits after the transition. Note that a bigger relative difference between b_i and $b_{ad} = b_{id}$ would have caused the shift from low to high intensity to happen earlier.

In the case of dependence, the situation is not equally simple. Here, the tradeoff between having a high intensity μ_{ad} and a low intensity μ_{id} results in the worst-case scenario for the middle of the time span of the contract becoming $\mu_{ad}(t) = L(t) + 0.75(U(t) - L(t))$ whereas $\mu_{id}(t) = L(t) + 0.25(U(t) - L(t))$. Note that these are not corner points of the marginals in contrast to the case of independence. The worst-case scenario occurs because the level of μ_{ad} is more important for the size of the reserve than the level of μ_{id} . This is because the level of μ_{id} is a second-order effect in the state “Active”, since μ_{id} only matters after transition to the state disabled. However, this second-order effect is so significant that the worst-case scenario is *not* to maximize both μ_{ad} and μ_{id} .

For the linear dependent case we, as in the dependent case, see that the impact of μ_{ad} is more significant than the impact of μ_{id} implying that both are maximized for the entire time span.

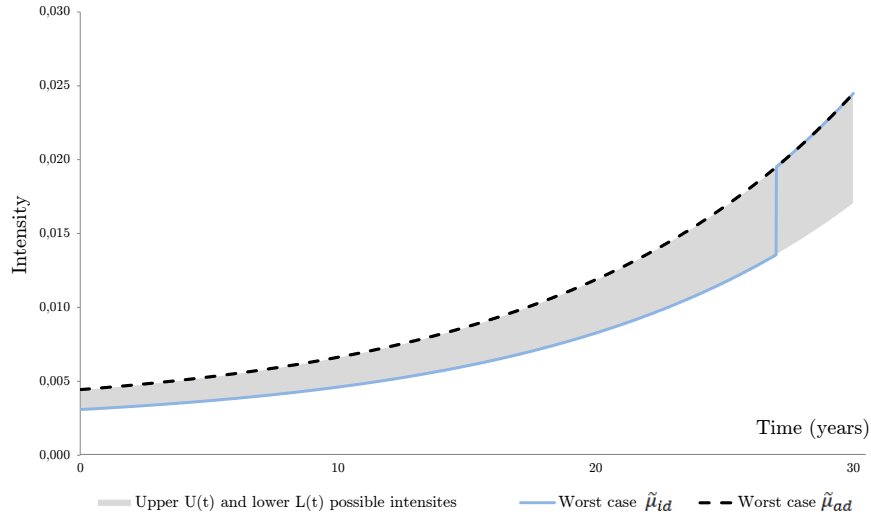


Figure 4.3: The worst-case death intensities in the case of independence.

In Figure 4.6 we see that the convergence to the fixed point is fairly fast: After only four iterations, we have obtained convergence.

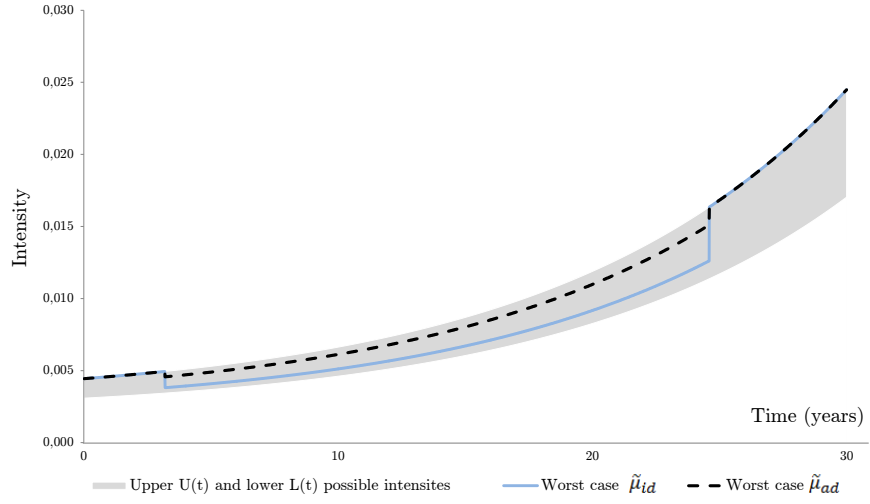


Figure 4.4: The worst-case death intensities in the case of dependence.

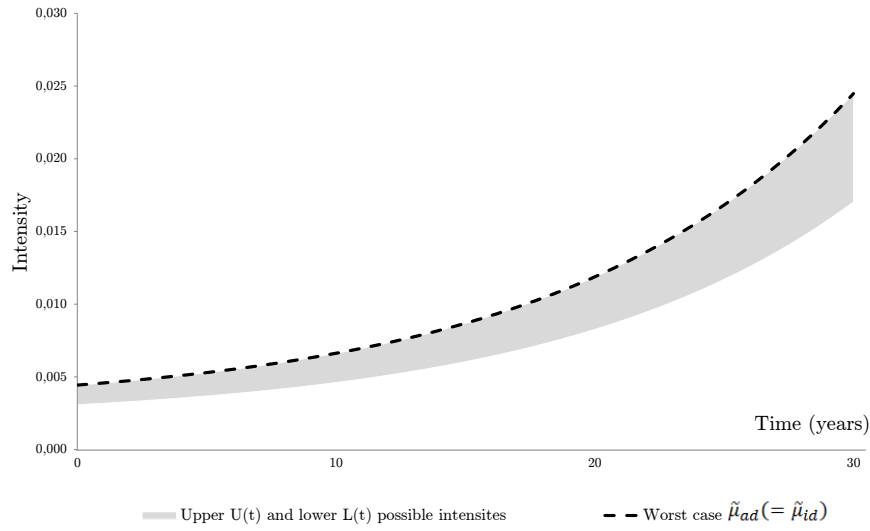


Figure 4.5: The worst-case death intensities in the case of linear dependence.

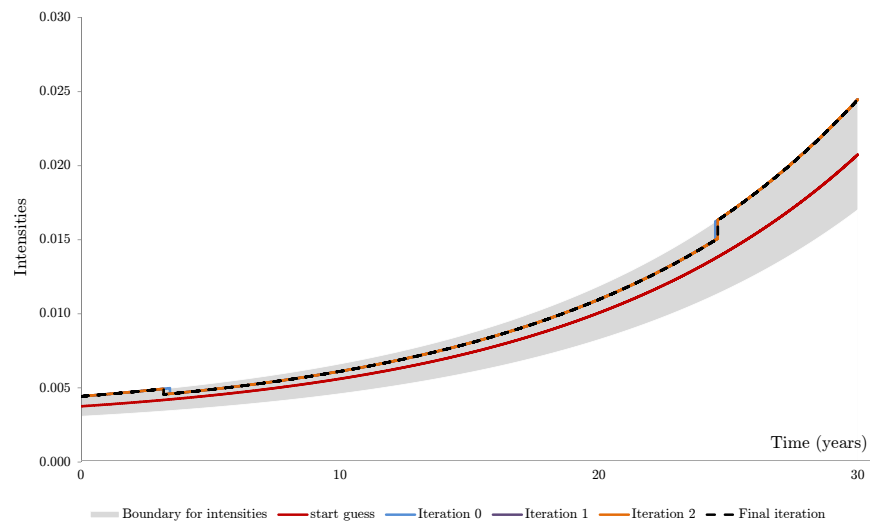


Figure 4.6: Convergence for the argmax of μ^{ad} .

4.5.2 Numerical calculations for a portfolio

In this section we illustrate the theory of Section 4.4 for a representative portfolio consisting of a young, a middle-aged, and a close-to-pension-aged policyholder. We are in the two state life-death model and are maximizing over the type of sets described in Section 4.4.2. We assume that all three representative policyholders have the same type of contract. That is, they have a term insurance paying 15 at death before retirement (age 67) and a life annuity paying a yearly rate of 1 starting at retirement. Their baseline death intensities are given as μ_{ad}^{be} in Table 4.1. For the present example the set of interest rates is $\Phi = \{0.02\}$, the lower multiplicative factor is $\alpha_l = 0.8$, and the upper multiplicative factor is $\alpha_h = 1.15$. The values of α_l and α_h are motivated by the mortality and longevity stresses from Solvency II, see EIOPA (2013).

We obtain the worst-case intensities illustrated in Figure 4.7 on a logarithmic scale. We denote by subscript 30 the youngest policyholder, by subscript 45 the middle-aged policyholder, and by subscript 60 the oldest policyholder. Moreover, we use “I” to indicate that quantities are calculated at an individual level, and “P” to indicate that quantities are calculated on portfolio level. We see from Figure 4.7 that the worst-case scenario for the oldest person is the same as the worst-case scenario at the portfolio. The scenario is that the intensity is as high as possible for the first seven years (until retirement of the oldest policyholder), and hereafter it is as low as possible. On the other hand, the worst case-scenarios for the two other policyholders are quite different compared to the worst-case scenario for the portfolio. The statewise worst-case reserves for the policyholders corresponding to the intensities in Figure 4.7 can be found in Figure 4.8.

Worst-case intensities

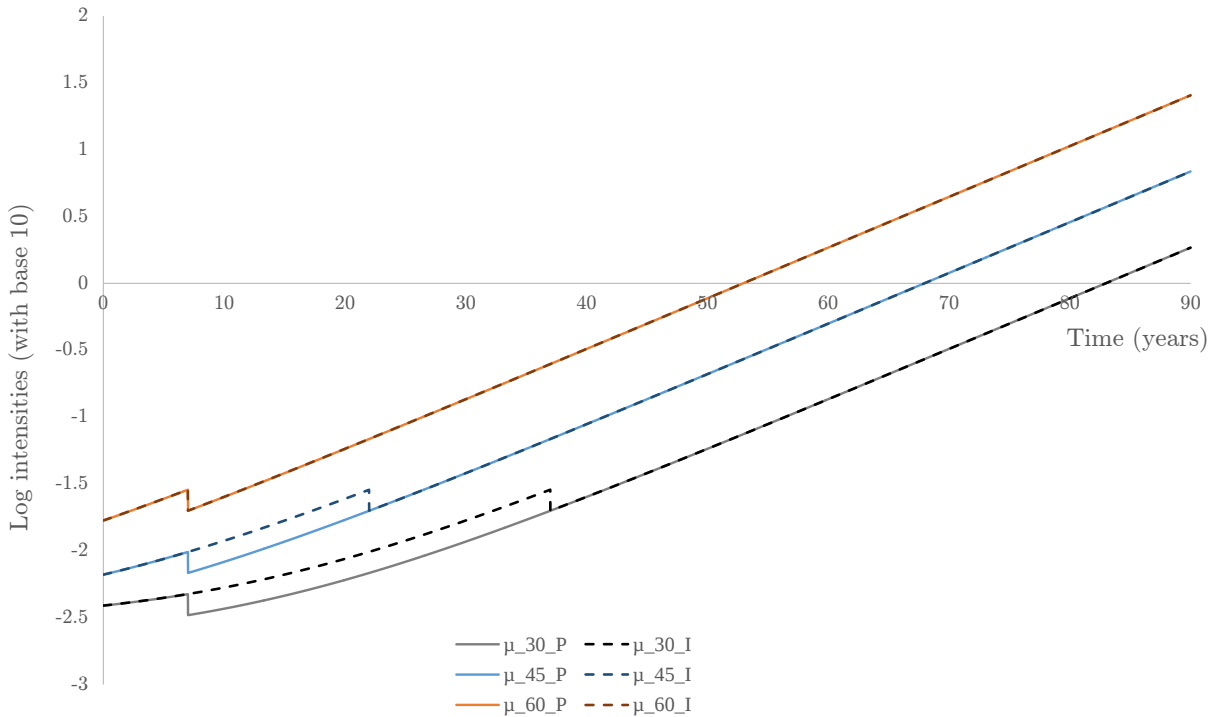


Figure 4.7: Worst-case intensities for the portfolio and for each individual policyholder.

In Table 4.2 we compare the reserves for the three policyholders in the portfolio calculated with different bases. The first is the best estimate basis, the second is α_h times the best estimate,

Statewise worst-case reserves

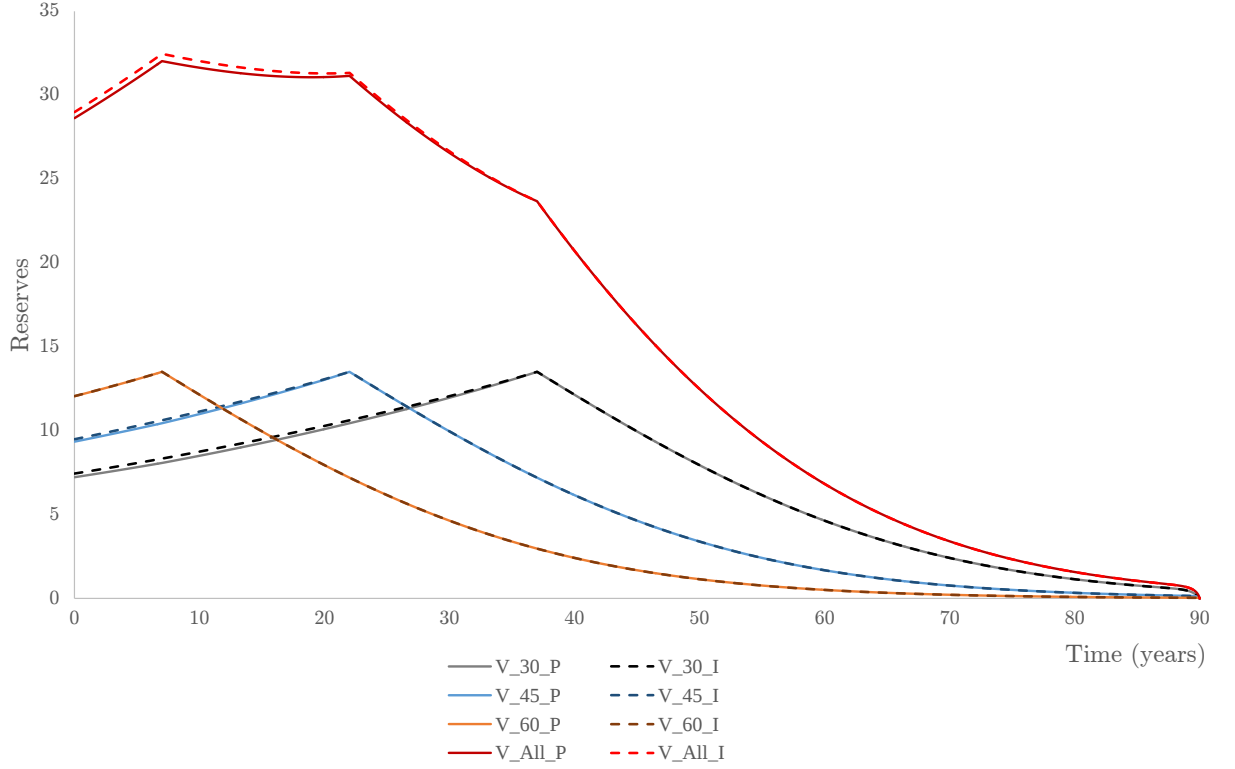


Figure 4.8: Statewise worst-case reserves calculated on basis of the worst-case scenarios for the individual policyholders and the worst-case scenario for the portfolio.

the third is α_l times the best estimate, and the fourth and fifth are the worst-case bases for the portfolio and the individual policyholders, respectively.

The numbers for “Solvency II (mortality)” and “Solvency II (longevity)” can be used to calculate the SCR in the “standard model”. Assuming that only mortality risk and longevity risk apply to our portfolio, the SCR is defined as

$$\text{SCR} = \sqrt{(\Delta V^{\text{mortality}})^2 + (\Delta V^{\text{longevity}})^2 - 2 \cdot 0.25 \cdot \Delta V^{\text{mortality}} \Delta V^{\text{longevity}}}, \quad (4.5.1)$$

where

$$\begin{aligned} \Delta V^{\text{mortality}} &= \sum_{l=1}^L \max \left(V^l \left(\mu^{\text{be}} \cdot 1.15 \right) - V^l \left(\mu^{\text{be}} \right), 0 \right), \\ \Delta V^{\text{longevity}} &= \sum_{l=1}^L \max \left(V^l \left(\mu^{\text{be}} \cdot 0.8 \right) - V^l \left(\mu^{\text{be}} \right), 0 \right). \end{aligned}$$

The result of this calculation together with the calculations of the worst-case reserves lead to three different “SCR-like” quantities presented in Table 4.3. We here see that the SCR for the entire portfolio is significantly smaller for the worst-case scenario for the portfolio ($\approx 6.8\%$ of the reserve) compared to the worst-case scenario for the individual policyholders ($\approx 8.2\%$ of the reserve). However, they are both bigger than the SCR calculated using (4.5.1) ($\approx 5.9\%$ of the reserve).

	PH 1	PH 2	PH 3	Sum
Best estimate	6.91	8.80	11.09	26.81
Solvency II (mortality)	6.81	8.57	10.60	25.97
Solvency II (longevity)	7.17	9.27	11.97	28.40
Worst-case (PF)	7.23	9.35	12.06	28.64
Worst-case (separate)	7.45	9.49	12.06	29.00

Table 4.2: Reserves calculated by different methods ($b_{ad} = 15$) for the three policyholders.

Solvency II	Worst-case (PF)	Worst-case (separate)
1.59	1.83	2.19

Table 4.3: SCR calculated by different methods ($b_{ad} = 15$).

4.5.2.1 Increasing the term insurance - making more shifts in the worst-case intensities

In the previous example, we saw that there was one shift for the worst-case intensity for the portfolio; the shift was from low to high intensity after seven years. This kind of structure is probably quite normal for a big portfolio. However, this is not necessarily the case. There can be many more shifts, as we illustrate in this section. The only difference compared to the former example is that the term insurance is increased from 15 to 32. This leads to the worst-case intensities in Figure 4.9, where we have jumps in the worst-case intensities after the retirement of each of the policyholders. We also note, as opposed to the example with $b_{ad} = 15$, that none of the individually worst-case intensities coincide with the worst-case intensity for the portfolio. A table equivalent to Table 4.3 can be found in Table 4.5. The results there illustrate that the relative differences between the results of the different calculation methods can be quite big.

A comparison of different types of SCR-like calculations can be found in Table 4.4. We note that the relative differences of the SCRs are much bigger than in the former example with $b_{ad} = 15$.

	PH 1	PH 2	PH 3	Sum
Best estimate	10.01	11.95	13.08	35.05
Solvency II (mortality)	10.30	12.12	12.86	35.28
Solvency II (longevity)	9.71	11.85	13.58	35.15
Worst-case (PF)	10.28	12.78	13.54	36.60
Worst-case (separate)	10.93	13.04	14.33	38.30

Table 4.4: Reserves calculated by different methods ($b_{ad} = 32$) for the three policyholders.

The three different calculation methods lead to the three different SCR presented in Table 4.5.

Solvency II	Worst-case (PF)	Worst-case (separate)
0.59	1.55	3.25

Table 4.5: SCR calculated by different methods ($b_{ad} = 32$).

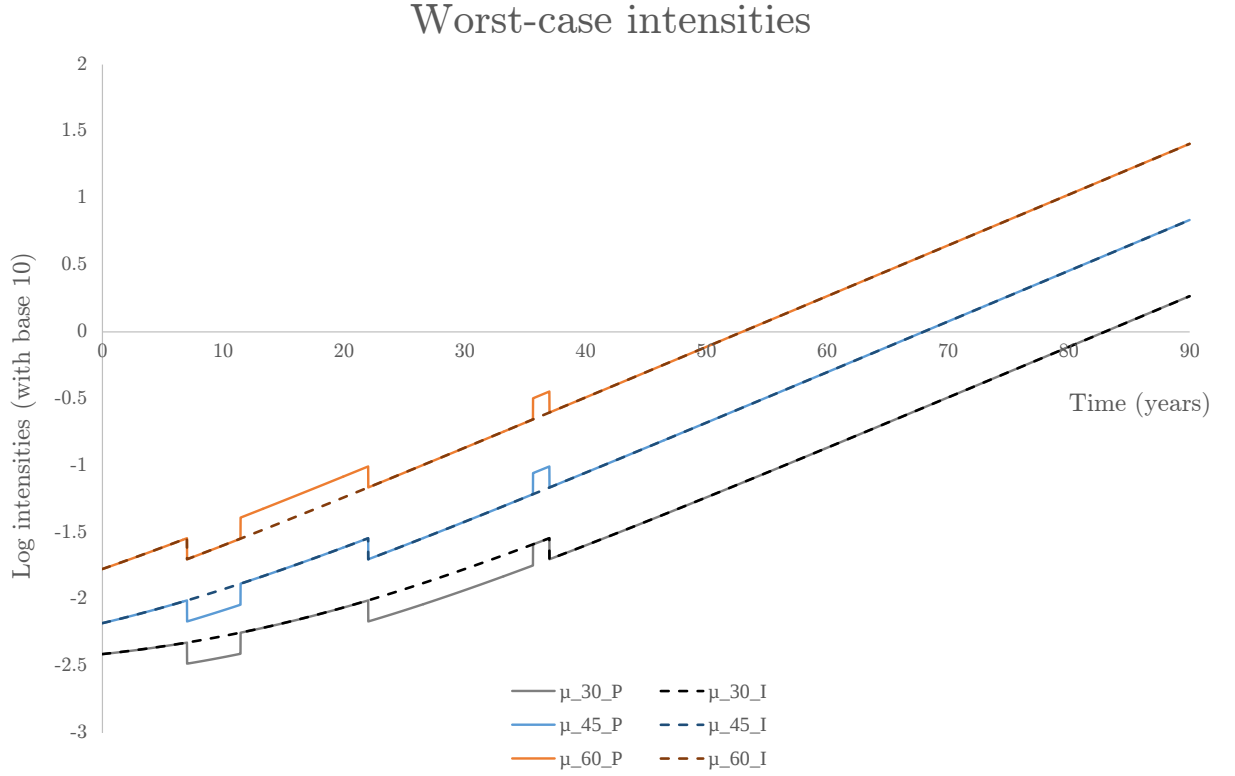


Figure 4.9: Worst-case intensities for the portfolio and for each individual policyholder ($b_{ad} = 32$).

4.5.3 Conclusion

First, what is meant by stressing a portfolio with respect to mortality and longevity by scaling the mortality rate by between 0.8 and 1.15? The standard formula (4.5.1) represents one interpretation. The negative correlation in (4.5.1) is a (normal) probabilistic formalization of the idea that experiencing future high mortality rates and future low mortality rates tend not to happen in the same realization. Our calculations do not assume such “restrictions”. We actually do find that the worst thing that can happen is that the mortality rate is high in the near future and low in the distant future and our worst-case approach really allows this realization to occur. Our numerical example shows that the standard formula based on a negative correlation of 0.25 leads to a too low capital requirement compared to the one obtained in our calculation. We do not claim that one can draw strong quantitative conclusions from this. What we do claim is that it is urgently important to understand exactly what is calculated and what is not. If one tends to believe that mortality rates actually can be high in the near future and low in the distant future, then calculations based on the standard formula may be dangerous.

Second, what is the intuition behind the high mortality rates in the near future and the low mortality rates in the distant future? This conforms with the basic understanding that in the near future, when policyholders are relatively young and hold positive sums at risk, high mortality rates are undesirable. Conversely, in the distant future, when policyholders are relatively old and hold negative sums at risk, low mortality rates are undesirable. The individual calculations in this section take into account these effects on a policy by policy basis in the sense what defines the near and distant future depends on the age of the individual policyholder. This is the simpler calculation giving a worst-case basis separately for each policy. The portfolio calculations deal with the situation where the same realized mortality rate counts for all policies, thinking of

uncertainty in the mortality rate as being at macro-level. Inhomogeneity in the portfolio now reduces the consequences of the worst case and the capital requirement goes down. It is important to understand that this has absolutely nothing to do with diversification but is related to portfolio inhomogeneity exclusively. The inhomogeneity in the portfolio with respect to age and products may be so involved that the worst-case “jumps” up and down before it finds its low when the distant future is finally met, as illustrated in subsection 4.5.2.1.

Third, what do these solvency calculations have to do with design and pricing, which was mentioned in the introduction? Here, it is important to remember that a policy is an object for safe-side calculation already upon pricing, before the contract is underwritten and goes into the solvency calculation. So the first safe-side calculations are part of the internal pricing and management procedures in contrast to solvency calculations where principles and restrictions are given from outside. Our results illustrate how one can ascertain a given level of prudence by setting the first order pricing basis. Individual and portfolio level calculations allow for setting the first order bases differently for different (groups of) policyholders. A more individual unit for pricing leads to a more prudent pricing basis and, thus, higher surplus contributions from the portfolio. This idea was maybe not relevant in the past due to technological limitations. We have then indirectly illustrated the prudence effects of micro-pricing in life insurance by different tailor-made first order bases used for a portfolio of inhomogeneous policies.

Fourth, what can we conclude in general about scenarios on the basis of our calculations? In contrast to some of the general references mentioned in the introduction, we do not criticize scenarios as a mean of solvency calculations and management for being too uninformative, too inaccurate or too simple. Rather, we push forward scenarios, exactly for being simple to work with and understand. They just have to be chosen such that information and accuracy is not lost in the translation between distributional aspects of intensities and reserves, respectively. We consider general policies and portfolios and find that the worst-case intensities are the ones that maximize the expected sum at risk. Since the intensities occur in the expectation itself, this is a delicate optimization and not just a check of the sign of a given sum at risk. However, once the calculational challenges are overcome, one is left with a stress calculation simple to implement, simple to interpret, and simple to communicate, while actually bounding the insolvency probability, see also the paragraph around (4.1.1).

Acknowledgments: We are grateful to an anonymous referee for many fruitful comments that greatly improved the paper.

Chapter 5

Affine processes and Markov chains: Interest Rate-Dependent Transition Rates in Life Insurance

Abstract: We obtain results for calculations of life insurance reserves where we model policyholder behavior governed by stochastic intensities within hierarchical Markov chain models. We model the interest rate and the transition rates as affine processes, where we allow for dependence between the interest rate and the transition rates. For many years, affine processes have been popular choices when modelling interest rates and mortality rates due to mathematical and computational tractability. Our main example of dependence is interest rate-dependent surrender rates, which are widely acknowledged in the literature, but the presented framework also allows for many other types of dependencies. The paper extends some results of transforms of affine processes presented in Duffie et al. (2000). We find two representations of transforms of affine processes needed for calculation of life insurance reserves and discuss their differences.

Keywords: Affine processes, Solvency II, expected policy holder behavior.

5.1 Introduction

This paper considers valuation of life insurance contracts within a Markov chain model in the case where the interest rate and the transition rates are dependent stochastic processes. We are working in an affine setup such that we are able to obtain results without needing to solve partial differential equations but only systems of ordinary differential equations as long as we consider models without cycles. Affine processes have for some time been used to model the interest rate and transition rates in life insurance, see e.g. Dahl and Møller (2006) and Biffis (2005). The results in this paper can be seen as a continuation of the work in Duffie et al. (2000). The recent paper Buchardt (2012) addresses some of the same questions as this paper. However, the focus and methods for proving the results in the two papers are quite different. Moreover, this paper contains a generalization to an arbitrary number of transitions within hierarchical Markov chain models and two different representations of transforms of affine processes and discussions of subjects in continuation to these. By a transform of an affine process we mean the conditional expected value of the product of the exponential of the integral of an affine transformation of the process, the exponential of an affine transformation of the process and a polynomial transform of

the process.

The surrender option is embedded in many life insurance contracts and the attention to the modelling of surrender rates and valuation of life insurance contracts where one takes into account the surrender option, has highly increased the last years. This is, in particular, motivated by the forthcoming Solvency II legislation, where the lapse risk plays an important role. One way to model the surrender rate is to make it dependent of the interest rate, which is a fairly natural approach as for example shown in Kuo et al. (2003). In De Giovanni (2010) a model for insurance liabilities is applied, where the surrender intensity is split into a rational part and an irrational part. The rational part is modelled as a function of the short rate squared, whereas the rate of the irrational part is assumed to be constant. The interest rate dependent surrender rates form our main example of our general results on dependent interest and transition rates.

The affine processes are widely used for interest rate modelling, credit risk modelling and modelling of mortality intensities because of the calculational tractability. That is, one only needs to solve ordinary differential equations (ODEs) instead of partial differential equations (PDEs) which in general is a much easier task. When we use affine processes to model transition intensities in a multistate Markov chain model the tractability can be a little reduced due to the fact that one needs to solve quite a lot of ODEs. The number of calculations necessary to obtain a result depends highly on the number of possible transitions (between different states) before reaching an absorbing state. Secondly, it also depends on the dimension of the underlying processes.

The paper is organized as follows: In Section 5.2 and 5.3 we motivate the work and introduce the basic insurance Markov chain model, which we are going to consider throughout the rest of the paper. Moreover, we introduce the class of affine processes and state conditions to ensure positivity of these processes. In Section 5.4 we present the main results which allow us to calculate statewise reserves in general hierarchical Markov models. Section 5.5 contains a small selection of different subsections relating to the results in Section 5.4. These are a description of the workload relating to the results in Section 5.4, a presentation of a special case, where one can allow for cycles in the Markov chain and still be able to benefit from the same results as in the case with no cycles, and a description of the class of processes called linear-quadratic processes.

5.2 Motivating example

The forthcoming Solvency II legislation demands that the life insurance companies take into account future policyholder behavior such as the likelihood of lapse during the remaining period of an insurance contract. In general, lapse refers to the exercise of policyholder options. In this paper we will focus on the two most important policyholder options. These are the surrender option and the free policy option, which are quite common features in many life insurance contracts. Moreover, lapse risk is a significant risk for many life insurance companies, and recognizing lapse can change the cash flow and reserves of a company significantly, see e.g. Henriksen et al. (2014) and Buchardt et al. (2013). In the following we will consider valuation of a general life insurance contract including both disability and lapse. The state space for such a contract is shown in Figure 5.1, where we have disregarded premium resumption, i.e. the possibility of a transition from state fa to state a .

In comparison with a standard disability model we have added four extra states to allow for lapse. It is not important for the results in this paper that e.g. the intensity from state a to state i is the same as from state fa to state fi , but for most modelling purposes this is the case. Note that even though the pure jump process, which determines the state of the insurance contract, is

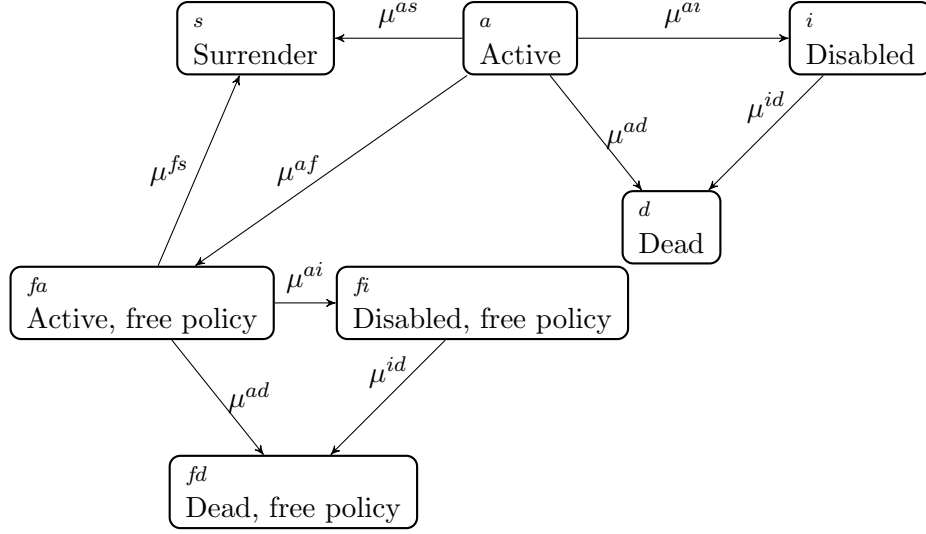


Figure 5.1: State space for a disability model with lapse.

Markovian, i.e. the intensities only depend on the current state of the process, the payments do not only depend on the state of the pure jump process. The reason for this is that the payments, after the contract has entered the free policy state, depend on the time of this jump. That is, the introduction of a free policy state gives rise to duration depend payments in the model. The multistate model given in Figure 5.1 is a hierarchical Markov model, i.e. a model without cycles. Within this model, the state process of the Markov chain can jump at most 3 times before reaching an absorbing state. Since the model has no cycles, the results for the statewise reserves can be obtained by integration of solutions to ODEs in the case, where the interest rate and transition intensities are dependent affine processes. We show this in Section 5.4.

For the purpose of illustrating the general challenge when evaluating contracts in this type of models and justified by wanting to keep the terms as simple as possible, we consider the (not very realistic) case where the only payment of the insurance contract is a death sum $b_{id}(t)$ triggered by the transition from state fi to state fd . We allow for the discount factor, r , and the transition intensities to be affine dependent on an underlying affine stochastic process X . Under the assumption that the insurance contract only includes one payment, the statewise reserve in state a at time t , denoted $V_a(t)$, is given by (assuming that we are allowed to interchange the expectation and the integrals):

$$\begin{aligned}
V_a(t) &= \mathbb{E} \left[\int_t^T \int_s^T \int_u^T e^{-\int_t^v r(\tau, X(\tau)) d\tau} e^{-\int_t^s \mu^{aa}(\tau, X(\tau)) d\tau} e^{-\int_s^u \mu^{ff}(\tau, X(\tau)) d\tau} e^{-\int_u^v \mu^{id}(\tau, X(\tau)) d\tau} \right. \\
&\quad \times f(s) \mu^{af}(s, X(s)) \mu^{ai}(u, X(u)) \mu^{id}(v, X(v)) b_{id}(v) dv du ds \Big| \mathcal{F}(t) \Big] \\
&= \int_t^T \int_s^T \int_u^T \mathbb{E} \left[e^{-\int_t^v r(\tau, X(\tau)) d\tau} e^{-\int_t^s \mu^{aa}(\tau, X(\tau)) d\tau} e^{-\int_s^u \mu^{ff}(\tau, X(\tau)) d\tau} e^{-\int_u^v \mu^{id}(\tau, X(\tau)) d\tau} \right. \\
&\quad \times \mu^{af}(s, X(s)) \mu^{ai}(u, X(u)) \mu^{id}(v, X(v)) \Big| \mathcal{F}(t) \Big] f(s) b_{id}(v) dv du ds,
\end{aligned} \tag{5.2.1}$$

where $\mu^{aa} = \mu^{ai} + \mu^{ad} + \mu^{af} + \mu^{as}$, $\mu^{ff} = \mu^{ai} + \mu^{ad} + \mu^{fs}$ and f is the free policy factor. The free policy factor is determined at time s and is equal to 1 in this example, since the example does not

include any premium payments. More generally, the free policy factor is given by $f(\tau) = \frac{V^*(\tau)}{V^{*+}(\tau)}$, where V^* is the technical reserve, and V^{*+} is the technical reserve where we only take into account the benefits of the contract. This means, that the free policy factor is a time-dependent, deterministic function.

From equation (5.2.1) one sees that the important thing is to be able to calculate terms of the form

$$\begin{aligned} \mathbb{E} \left[e^{-\int_t^v r(\tau, X(\tau)) d\tau} e^{-\int_t^s \mu^{aa}(\tau, X(\tau)) d\tau} e^{-\int_s^u \mu^{ff}(\tau, X(\tau)) d\tau} e^{-\int_u^v \mu^{id}(\tau, X(\tau)) d\tau} \right. \\ \left. \times \mu^{af}(s, X(s)) \mu^{ai}(u, X(u)) \mu^{id}(v, X(v)) \middle| \mathcal{F}(t) \right] \end{aligned} \quad (5.2.2)$$

for different assumptions regarding the interest rate and transition intensities. The rest of the paper is about how to calculate terms like this. Note that for the general case with continuous and lump sum payments, we are still able to calculate statewise reserves, if we are able to calculate terms of the form (5.2.2). That is, being able to calculate terms of the form (5.2.2) and integrating the results makes it possible to obtain the statewise reserves for all states in Figure 5.1. More generally, the results make us able to calculate transition probabilities and statewise reserves in arbitrarily large hierarchical Markov chain models.

5.3 Insurance models and affine processes

Let T be a fixed finite time horizon and $(\Omega, \mathcal{F}, P)$ a probability space equipped with a filtration $\mathbb{F} = (\mathcal{F}(t))_{0 \leq t \leq T}$ satisfying the usual conditions of right-continuity and completeness. In the following we consider a class of stochastic processes called affine processes. Affine processes are stochastic processes with the property that the conditional characteristic functions are exponentially affine. One normally divides affine processes into diffusion processes and jump diffusion processes, where affine diffusion processes are the continuous subgroup of affine jump diffusion processes. For detailed descriptions of affine processes, see e.g. Duffie et al. (2003) or Filipović (2009, Chapter 10).

The dynamics of a d -dimensional time-inhomogeneous affine jump diffusion are given by

$$dX(t) = \mu(t, X(t))dt + \sigma(t, X(t))dW(t) + dJ(t), \quad X(0) = x, \quad (5.3.1)$$

where $\mu(t, x) = K_0(t) + K_1(t)x$ and $(\sigma(t, x)\sigma^{\text{tr}}(t, x))_{ij} = H_{0,ij}(t) + \sum_{k=1}^d H_{1,ij}^k(t)x_k$. Moreover, W is a d -dimensional standard Brownian motion and J is a pure jump process with arrival intensity $\lambda(t, x) = \lambda_0(t) + \lambda_1^{\text{tr}}(t)x$ and jump distribution ν . We assume that the interest rate has the form $r(t, x) = \rho_0(t) + \rho_1^{\text{tr}}(t)x$. In the above equations we have that $\lambda_0, \rho_0 \in \mathbb{R}$, $K_0, \lambda_1, \rho_1 \in \mathbb{R}^d$, $K_1, H_0 \in \mathbb{R}^{d \times d}$ and $H_1 \in \mathbb{R}^{d \times d \times d}$.

The class of processes given by (5.3.1) is fairly general and is living on the statespace $\mathbb{R}^d$. The purpose of this paper is to use affine processes to model the interest rate and the transition intensities of the Markov chain. That is, we want to restrict the class of processes in such a way, that the interest rate is non-negative and such that the transition intensities are *strictly* positive. To obtain non-negativity (in the case of time-homogeneous affine processes without jumps), Filipović (2009, Theorem 10.2) states the following conditions:

- $K_0 \in \mathbb{R}_+^d$.
- K_1 has non-negative off-diagonal elements.
- $H_0 = 0$.

- The only non-zero entries of H_1^k are $H_{1,kk}^k$, which are positive.

This result can be generalized to time-inhomogeneous affine processes (the conditions are the same), see Buchardt (2012). Positive jumps of X do not destroy the positivity, whereas negative jumps in general will. For the rest of the paper, we disregard jumps to keep the notation as simple as possible. Note, however, that the results presented here can easily be generalized from affine diffusion processes to affine jump diffusion processes.

To obtain the strict positivity for the intensities we need an extra condition to be fulfilled. The condition is known as the multivariate Feller condition. The condition is that $2K_{0,k}(t) \geq H_{1,kk}^k(t)$. That is, the structure of the non-negative affine processes is quite restricted compared to the general case and the structure leads to simplifications of the ODEs for the transforms of the affine processes. For example, the ODEs relating to the standard transform in Duffie et al. (2000) given by

$$\mathbb{E} \left[e^{-\int_t^T \rho_0(s) + \rho_1^{\text{tr}}(s) X(s) ds + u^{\text{tr}} X(T)} \middle| \mathcal{F}(t) \right] = e^{\alpha(t,T) + \beta^{\text{tr}}(t,T) X(t)}$$

simplify to:

$$\begin{aligned} \frac{d}{dt} \beta(t, T) &= \rho_1(t) - K_1^{\text{tr}}(t) \beta(t, T) - \frac{1}{2} \beta^{\text{tr}}(t, T) \text{diag} \left(H_{1,11}^1(t), \dots, H_{1,dd}^d(t) \right) \beta(t, T), \quad \beta(T, T) = u, \\ \frac{d}{dt} \alpha(t, T) &= \rho_0(t) - K_0^{\text{tr}}(t) \beta(t, T), \quad \alpha(T, T) = 0. \end{aligned}$$

However, for the rest of the paper, we consider, for the sake of completeness, the case where for example H_0 can be different from 0. This also allows us to use e.g. a Vasiček process, which can be negative, to model the interest rate.

5.4 Main result

In this section we present results enabling us to calculate transforms like (5.2.2) and similar transforms for more general hierarchical Markov chain models. We give two different representations of the results; a *dense* and a *non-dense* representation. We make these two concepts clear in Section 5.4.1. In both cases we rely on conditioning by use of the tower property.

The structure of this section is as follows: First, we present the results for a one-dimensional underlying stochastic process. Afterwards, we state some of the results for the multidimensional setting without giving additional proofs.

5.4.1 The conditional approach

The literature has for many years investigated the dependence between the interest rate and policy holder behavior. This possible dependence is called the interest rate hypothesis, see e.g. Dar and Dodds (1989) and references therein. A natural way of modelling this dependence is to model the lapse rates as functions of a one-dimensional interest rate process like in De Giovanni (2010). Due to this reasoning and because we are not gaining much more insight by extending the approach to a multidimensional setting, we introduce what we call “the conditional approach” in a one-dimensional setting. This also makes the results and expressions much more readable. We consider three intermediate time points t_1, t_2 and t_3 fulfilling that $t < t_1 < t_2 < t_3 < T$. Assuming that all the intensities in Figure 5.1 are given as affine transformations of an underlying

one-dimensional stochastic process, we need to calculate a quantity of the form¹

$$\begin{aligned} & \mathbb{E} \left[e^{-\int_{t_1}^{t_2} f_1(\tau, X(\tau)) d\tau} e^{-\int_{t_2}^{t_3} f_2(\tau, X(\tau)) d\tau} e^{-\int_{t_3}^T f_3(\tau, X(\tau)) d\tau} e^{-\int_{t_3}^T f_4(\tau, X(\tau)) d\tau} \right. \\ & \quad \left. \times \mu_1(t_1, X(t_1)) \mu_2(t_2, X(t_2)) \mu_3(t_3, X(t_3)) \right] \Big| \mathcal{F}(t), \end{aligned} \quad (5.4.1)$$

where μ_i and f_i , $i = 1, 2, 3$ are affine functions of X . Note that the term (5.4.1) is closely related to the calculation of transition probabilities in the Markov chain model, since the triple integral of (5.4.1) for $f_i = \mu_i$ is exactly a transition probability. By successively using the tower property we obtain that (5.4.1) equals

$$\begin{aligned} & \mathbb{E} \left[e^{-\int_{t_1}^{t_2} f_1(\tau, X(\tau)) d\tau} \mu_1(t_1, X(t_1)) \right] \mathbb{E} \left[e^{-\int_{t_2}^{t_3} f_2(\tau, X(\tau)) d\tau} \mu_2(t_2, X(t_2)) \right. \\ & \quad \left. \times \mathbb{E} \left[e^{-\int_{t_3}^T f_3(\tau, X(\tau)) d\tau} \mu_3(t_3, X(t_3)) \right] \Big| \mathcal{F}(t_3) \right] \Big| \mathcal{F}(t_2) \Big| \mathcal{F}(t_1) \Big| \mathcal{F}(t). \end{aligned} \quad (5.4.2)$$

Towards the end of this section we obtain closed form solutions for (5.4.2) by first calculating the innermost conditional expected value and then using this result for calculating the next conditional expected value and so on. To do so we need the results of the following two lemmas. In the following, we let ρ_i , u_i , $i = 1, 2$ and g_j , $j = 0, \dots, n$ be deterministic functions. To shorten the notation, we define by Υ_X the term

$$\Upsilon_X(s, v) := e^{u_0(v) + u_1(v)X(v) - \int_s^v \rho_0(\tau) + \rho_1(\tau)X(\tau) d\tau}.$$

Lemma 5.4.1 (Dense representation). *Let $s, v \in [t, T]$ with $s < v$. Assume that the system of ODEs stated in the lemma is uniquely solved by the functions $\beta, C_i, i = 0, \dots, n$, and that the following integrability conditions are fulfilled:*

$$E[|\Phi(v)|] < \infty$$

and

$$E \left[(\eta^2(t))^{\frac{1}{2}} \right] < \infty \text{ for } \eta(t) = \left(\Phi(t)\beta(t, v) + \Upsilon_X(t, v) \sum_{i=1}^n C_i(t, v) i x^{i-1} \right) \sigma(X(t)),$$

where $\Phi(t) = e^{-\int_0^t \rho_0(\tau) + \rho_1(\tau)X(\tau) d\tau} e^{\beta(t, v)X(t)} \sum_{i=0}^n C_i(t, v) X^i(t)$.

Then for $n \in \mathbb{N}$ there exist functions β and C_i for $i = 0, \dots, n$ given as solutions to a system of ODEs such that,

$$E \left[\Upsilon_X(s, v) \sum_{i=0}^n g_i(v) X^i(v) \right] \Big| \mathcal{F}(s) = e^{\beta(s, v)X(s)} \sum_{i=0}^n C_i(s, v) X^i(s). \quad (5.4.3)$$

The functions β and $C_i, i = 0, \dots, n$ are given by the following systems of ODEs (suppressing the

¹We can e.g. think of the term with the integral from t_3 to T as giving us the possibility of modelling a contract where the death sum is due at time T .

time arguments of the functions)

$$\begin{aligned}\frac{\partial}{\partial s}\beta &= \rho_1 - \left(K_1 + \frac{1}{2}H_1\beta\right)\beta, \\ \frac{\partial}{\partial s}C_0 &= \rho_0C_0 - (\beta C_0 + C_1)K_0 - (\beta^2C_0 + 2\beta C_1 + 2C_2)\frac{1}{2}H_0, \\ \frac{\partial}{\partial s}C_1 &= C_0\left(\frac{\partial}{\partial s}\beta - \rho_1\right) + \rho_0C_1 - 2C_2K_0 - 2C_2\beta H_0 - C_1K_1 - 3C_3H_0 - C_2H_1 \\ &\quad - \left(K_0 + \frac{1}{2}H_0\beta + H_1\right)C_1\beta - \left(K_1 + \frac{1}{2}H_1\beta\right)C_0\beta.\end{aligned}$$

For $n \geq 4$ and $i = 2, \dots, n-2$ we have the ODEs given by

$$\begin{aligned}\frac{\partial}{\partial s}C_i &= -C_{i-1}\left(\frac{\partial}{\partial s}\beta - \rho_1\right) + \rho_0C_i - (i+1)K_0C_{i+1} - (i+1)H_0C_{i+1}\beta \\ &\quad - iK_1C_i - iH_1C_i\beta - \frac{1}{2}(i+2)(i+1)C_{i+2}H_0 - (i+1)iC_{i+1}\frac{1}{2}H_1 \\ &\quad - \left(K_0 + \frac{1}{2}H_0\beta\right)\beta C_i - \left(K_1 + \frac{1}{2}H_1\beta\right)C_{i-1}\beta.\end{aligned}$$

For $n \geq 3$ we have the ODE given by

$$\begin{aligned}\frac{\partial}{\partial s}C_{n-1} &= -C_{n-2}\left(\frac{\partial}{\partial s}\beta - \rho_1\right) + \rho_0C_{n-1} - n(K_0 + \beta H_0)C_n - (n-1)(K_1 + \beta H_1)C_{n-1} \\ &\quad - n(n-1)C_n\frac{1}{2}H_1 - \left(K_0 + \frac{1}{2}H_0\beta\right)C_{n-1}\beta - \left(K_1 + \frac{1}{2}H_1\beta\right)C_{n-2}\beta.\end{aligned}$$

Finally, for $n \geq 2$ we have the ODE given by

$$\begin{aligned}\frac{\partial}{\partial s}C_n &= -C_{n-1}\left(\frac{\partial}{\partial s}\beta - \rho_1\right) + \rho_0C_n - n(K_1 + \beta H_1)C_n - \left(K_0 + \frac{1}{2}H_0\beta\right)C_n\beta \\ &\quad - \left(K_1 + \frac{1}{2}H_1\beta\right)C_{n-1}\beta.\end{aligned}$$

The corresponding boundary conditions are

$$\beta(v, v) = u_1(v), \quad C_i(v, v) = e^{u_0(v)}g_i(v), \quad i = 0, \dots, n.$$

Proof of dense representation. The proof of the dense representation follows the lines of the proof of Duffie et al. (2000, Proposition 1). We are going to show that

$$\mathcal{H}(s, X(s); v) := \mathbb{E} \left[\Upsilon_X(s, v) \sum_{i=0}^n g_i(v) X^i(v) \middle| \mathcal{F}(s) \right] = e^{\beta(s, v)X(s)} \sum_{i=0}^n C_i(s, v) X^i(s),$$

where β and C_i are fulfilling the ODEs stated in the lemma.

We denote by $\tilde{\mathcal{H}}$ the martingale corresponding to $\mathcal{H}$. The martingale is given by

$$\tilde{\mathcal{H}}(s, X(s); v) = e^{-\int_0^s \rho_0(\tau) + \rho_1(\tau)X(\tau) d\tau} \mathcal{H}(s, X(s); v).$$

By utilizing that the drift of the martingale $\tilde{\mathcal{H}}(s, X(s); v)$ is 0 we get by Itô's formula that $\tilde{\mathcal{H}}$ is given by the following PDE:

$$0 = \frac{\partial}{\partial s} \tilde{\mathcal{H}}(s, x; v) + (K_0(s) + K_1(s)x) \frac{\partial}{\partial x} \tilde{\mathcal{H}}(s, x; v) + \frac{1}{2} (H_0(s) + H_1(s)x) \frac{\partial^2}{\partial x^2} \tilde{\mathcal{H}}(s, x; v). \quad (5.4.4)$$

We guess that the solution to this PDE is given by the function $\mathcal{G}$:

$$\begin{aligned}\mathcal{G}(s, x; v) &= e^{-\int_0^s \rho_0(\tau) + \rho_1(\tau) X(\tau) d\tau} e^{\beta(s, v)x} \sum_{i=0}^n C_i(s, v) x^i \\ &:= \mathcal{E}(s, x; v) \sum_{i=0}^n C_i(s, v) x^i,\end{aligned}\tag{5.4.5}$$

where we have introduced the shorthand notation $\mathcal{E}$ for the exponential terms. We want to find functions β and C_i , such that $\mathcal{G}$ is a martingale. If $\mathcal{G}$ is a martingale, we have that

$$\mathbb{E}[\mathcal{G}(v, x; v) | \mathcal{F}(s)] = \mathcal{G}(s, x; v)$$

and we can get the result by multiplying both sides of (5.4.5) by $e^{\int_0^s \rho_0(\tau) + \rho_1(\tau) X(\tau) d\tau}$.

Inserting $\mathcal{G}$ in the place of $\tilde{\mathcal{H}}$ in the right-hand side of (5.4.4) gives us

$$\begin{aligned}& \frac{\partial}{\partial s} \mathcal{G}(s, x; v) + \frac{\partial}{\partial x} \mathcal{G}(s, x; v) (K_0(s) + K_1(s)x) + \frac{1}{2} \frac{\partial^2}{\partial x^2} \mathcal{G}(s, x; v) (H_0(s) + H_1(s)x) \\ &= \mathcal{E}(s, x; v) \left(\left(-\rho_0(s) - \rho_1(s)x + \frac{\partial}{\partial s} \beta(s, v)x \right) \sum_{i=0}^n C_i(s, v) x^i + \sum_{i=0}^n \frac{\partial}{\partial s} C_i(s, v) x^i \right) \\ &+ \mathcal{E}(s, x; v) \left(\beta(s, v) \sum_{i=0}^n C_i(s, v) x^i + \sum_{i=1}^n C_i(s, v) i x^{i-1} \right) (K_0(s) + K_1(s)x) \\ &+ \frac{1}{2} \mathcal{E}(s, x; v) \left(\beta^2(s, v) \sum_{i=0}^n C_i(s, v) x^i + 2\beta(s, v) \sum_{i=1}^n i C_i(s, v) x^{i-1} \right. \\ &\quad \left. + \sum_{i=2}^n C_i(s, v) i(i-1) x^{i-2} \right) (H_0(s) + H_1(s)x).\end{aligned}$$

We divide the right-hand side of the above term by $\mathcal{E}(s, x; v)$, suppress the arguments of all the functions, insert indicator functions to take care of small numbers of n , collect terms wrt. x , and

obtain that the above term is equal to

$$\begin{aligned}
& -\rho_0 C_0 + \frac{\partial}{\partial s} C_0 + (\beta C_0 + C_1) K_0 + (\beta^2 C_0 + 2\beta C_1 + 2C_2) \frac{1}{2} H_0 \\
& + x \left(C_0 \left(\frac{\partial}{\partial s} \beta - \rho_1 \right) - \rho_0 C_1 + \frac{\partial}{\partial s} C_1 + 2C_2 K_0 + 2C_2 \beta H_0 + C_1 K_1 + 3C_3 H_0 + C_2 H_1 \right. \\
& \quad \left. + \left(K_0 + \frac{1}{2} H_0 \beta + H_1 \right) C_1 \beta + \left(K_1 + \frac{1}{2} H_1 \beta \right) C_0 \beta \right) \\
& + \mathbf{1}_{\{n \geq 4\}} \sum_{i=2}^{n-2} x^i \left(C_{i-1} \left(\frac{\partial}{\partial s} \beta - \rho_1 \right) - \rho_0 C_i + \frac{\partial}{\partial s} C_i + (i+1) K_0 C_{i+1} + (i+1) H_0 C_{i+1} \beta \right. \\
& \quad \left. + i K_1 C_i + i H_1 C_i \beta + \frac{1}{2} (i+2)(i+1) C_{i+2} H_0 + (i+1) i C_{i+1} \frac{1}{2} H_1 \right. \\
& \quad \left. + \left(K_0 + \frac{1}{2} H_0 \beta \right) \beta C_i + \left(K_1 + \frac{1}{2} H_1 \beta \right) C_{i-1} \beta \right) \\
& + \mathbf{1}_{\{n \geq 3\}} x^{n-1} \left(C_{n-2} \left(\frac{\partial}{\partial s} \beta - \rho_1 \right) - \rho_0 C_{n-1} + \frac{\partial}{\partial s} C_{n-1} + n(K_0 + \beta H_0) C_n \right. \\
& \quad \left. + (n-1)(K_1 + \beta H_1) C_{n-1} \right. \\
& \quad \left. + n(n-1) C_n \frac{1}{2} H_1 + \left(K_0 + \frac{1}{2} H_0 \beta \right) C_{n-1} \beta + \left(K_1 + \frac{1}{2} H_1 \beta \right) C_{n-2} \beta \right) \\
& + \mathbf{1}_{\{n \geq 2\}} x^n \left(C_{n-1} \left(\frac{\partial}{\partial s} \beta - \rho_1 \right) - \rho_0 C_n + \frac{\partial}{\partial s} C_n + n(K_1 + \beta H_1) C_n + \left(K_0 + \frac{1}{2} H_0 \beta \right) C_n \beta \right. \\
& \quad \left. + \left(K_1 + \frac{1}{2} H_1 \beta \right) C_{n-1} \beta \right) \\
& + x^{n+1} C_n \left(\frac{\partial}{\partial s} \beta - \rho_1 + \left(K_1 + \frac{1}{2} H_1 \beta \right) \beta \right).
\end{aligned} \tag{5.4.6}$$

From (5.4.6) we can see, that we can make the term become zero by setting each of the $n+2$ terms equal to zero. In total this gives us a system of $(n+2)$ ODEs that β and $C_0, \dots, C_n$ need to fulfill for $\mathcal{G}$ to be a martingale. These ODEs are the ones stated in the lemma. The boundary values are the obvious ones: $\beta(v, v) = u_1(v)$ and $C_i(v, v) = e^{u_0(v)} g_i(v)$, $i = 0, \dots, n$.

Given the integrability conditions in the lemma (ensuring that the integral wrt. to the Brownian motion is a martingale) and functions $\beta, C_0, \dots, C_n$ fulfilling the system of ODEs, we have

$$\begin{aligned}
e^{\int_0^s \rho_0(\tau) + \rho_1(\tau) X(\tau) d\tau} \mathbb{E} [\mathcal{G}(v, x; v) | \mathcal{F}(s)] &= e^{\int_0^s \rho_0(\tau) + \rho_1(\tau) X(\tau) d\tau} \mathcal{G}(s, x; v) \\
&= e^{\beta(s, v) X(s)} \sum_{i=0}^n C_i(s, v) X^i(s).
\end{aligned} \tag{5.4.7}$$

On the other hand we have that

$$\begin{aligned}
& e^{\int_0^s \rho_0(\tau) + \rho_1(\tau) X(\tau) d\tau} \mathbb{E} [\mathcal{G}(v, x; v) | \mathcal{F}(s)] \\
&= \mathbb{E} \left[e^{-\int_s^v \rho_0(\tau) + \rho_1(\tau) X(\tau) d\tau} e^{\beta(v, v) X(v)} \sum_{i=0}^n C_i(v, v) X^i(v) \middle| \mathcal{F}(s) \right] \\
&= \mathbb{E} \left[\Upsilon_X(s, v) \sum_{i=0}^n g_i(v) X^i(v) \middle| \mathcal{F}(s) \right] \\
&= \mathcal{H}(s, X(s); v).
\end{aligned} \tag{5.4.8}$$

Since the left-hand sides of (5.4.7) and (5.4.8) are the same, we have that

$$\mathbb{E} \left[\Upsilon_X(s, v) \sum_{i=0}^n g_i(v) X^i(v) \middle| \mathcal{F}(s) \right] = e^{\beta(s, v) X(s)} \sum_{i=0}^n C_i(s, v) X^i(s),$$

and we have proved the lemma. $\square$

Next, we give a non-dense representation of the transform. In the lemma below, f_1 and f_2 are deterministic functions.

Lemma 5.4.2 (Non-dense representation). *Let $n \in \mathbb{N}$ be fixed and let $s, v \in [t, T]$ with $s < v$. Assume that the system of ODEs stated in the lemma is uniquely solved by the functions $\alpha, \beta, A_j, B_j, j = 1, \dots, n$ and that we have enough integrability to interchange differentiation and integration. That is, there exists a stochastic variable Y with finite expectation such that*

$$|\Upsilon_X(s, v) (f_0(v) + f_1(v) X(v))^n| \leq Y \text{ for all } \omega \in \Omega.$$

Then for $n \in \mathbb{N}$ there exist functions α, β, A_j and B_j for $j = 1, \dots, n$ given as solutions to a system of ODEs such that,

$$\begin{aligned}
& E[\Upsilon_X(s, v) (f_0(v) + f_1(v) X(v))^n | \mathcal{F}(s)] \\
&= e^{\alpha(s, v) + \beta(s, v) X(s)} \sum_{(i_1, \dots, i_n) \in \mathcal{C}_n} \kappa^n(i_1, \dots, i_n) \prod_{j=1}^n (A_j(s, v) + B_j(s, v) X(s))^{i_j},
\end{aligned} \tag{5.4.9}$$

where the set $\mathcal{C}_n$ is given by

$$\mathcal{C}_n = \left\{ v \in \mathbb{N}_0^n \middle| \sum_{i=1}^n i v_i = n \right\},$$

and $\kappa^n \in \mathbb{N}_0$ is given recursively by

$$\begin{aligned}
\kappa^n(i_1, \dots, i_n) &= \mathbf{1}_{\{i_n=1\}} + \mathbf{1}_{\{i_n \neq 1\}} \left(\kappa^{n-1}(i_1 - 1, i_2, \dots, i_{n-1}) \right. \\
&\quad \left. + \sum_{k=1}^{n-2} (i_k + 1) \kappa^{n-1}(i_1, \dots, i_k + 1, i_{k+1} - 1, \dots, i_{n-1}) \right),
\end{aligned}$$

where $\kappa^{n-1}(v_1, \dots, v_{n-1}) = 0$ for $\min(v_1, \dots, v_{n-1}) < 0$.

The functions α, β and $A_i, B_i, i = 1, \dots, n$ are given by the following systems of ODEs (suppressing the time arguments of the functions), where we use the notation $B_0 := \beta$:

$$\begin{aligned}\frac{\partial}{\partial s}\beta &= \rho_1 - \left(K_1 + \frac{1}{2}H_1\beta\right)\beta, \\ \frac{\partial}{\partial s}\alpha &= \rho_0 - \left(K_0 + \frac{1}{2}H_0\beta\right)\beta, \\ \frac{\partial}{\partial s}B_m &= -K_1B_m - \frac{1}{2}\sum_{i=0}^m \binom{m}{i} B_i H_1 B_{m-i}, \quad m = 1, \dots, n, \\ \frac{\partial}{\partial s}A_m &= -K_0B_m - \frac{1}{2}\sum_{i=0}^m \binom{m}{i} B_i H_0 B_{m-i}, \quad m = 1, \dots, n.\end{aligned}$$

The boundary conditions are given by $\beta(v, v) = u_1(v)$, $\alpha(v, v) = u_0(v)$, $B_1(v, v) = f_1(v)$, $A_1(v, v) = f_0(v)$ and $A_i(v, v) = B_i(v, v) = 0$ for $i > 1$.

Remark 5.4.3. We note that the class of transforms we are able to calculate with Lemma 5.4.1 is much broader than the class of transforms we are able to calculate with Lemma 5.4.2. However, this is not the entire story about the non-dense representation. By following the lines of the proof of Lemma 5.4.2 one can see, that we can calculate the same type of transforms with the two lemmas. The price to achieve this, we pay in terms of less simplicity for the non-dense representation. This means, that the results are not as simple as the ones given by (5.4.9), nor are we able to have so simple representations of the ODEs as the ones stated in Lemma 5.4.2. This is due to the fact, that in the general case, there are different “versions” of the functions A_i and B_i , since they have different boundary values. This also implies different ODEs for the different “versions” of A_i and B_i , since their dynamics are mutually dependent. Later in this section, we make these statements clear, and the results in the general case for $n = 3$ are outlined.

Remark 5.4.4. For a fixed n the values of the functions κ^n in Lemma 5.4.2 are given by the recursive formula. For $n = 4$ we for example get the coefficients in Table 5.1.

$\kappa^4(0, 0, 0, 1)$	$\kappa^4(1, 0, 1, 0)$	$\kappa^4(0, 2, 0, 0)$	$\kappa^4(2, 1, 0, 0)$	$\kappa^4(4, 0, 0, 0)$
1	4	3	6	1

Table 5.1: Example of the coefficients denoted by $\kappa^4()$.

By comparison, we see that Lemma 5.4.1 and Lemma 5.4.2 give us two representations for similar quantities. The reason to give two different representations is, that each of them have advantages in different situations. We refer to the representation (5.4.3) as the *dense representation*, whereas we refer to the representation (5.4.9) as the *non-dense representation*. The reason for this is that the dense representation contains at maximum $(n + 2)$ functions, which can be found by solving $n + 2$ ODEs, whereas the non-dense representation is overparameterized in the sense that the maximum number of functions evolve in an exponential manner. For some real numbers, see Table 5.2 on page 108.

The dense representation is interesting because it minimizes the number of ODEs we need to solve. On the other hand, the reason why the non-dense representation is interesting, even though it is

overparameterized, is due to that we can use the structure of the problem to obtain a nice representation, where we can reuse the results of the former calculation steps by “gluing” solutions to differential equations together. This is explained in Section 5.4.2, where the resulting system of ODEs is stated in the case of 3 jumps of the pure jump process. This “gluing” method is not available for the dense representation, since we collect all terms no matter where they come from to end up solving as few ODEs as possible. Moreover, when using this non-dense representation, results are easily extended to a multidimensional setting, whereas higher dimensions add significantly extra calculational workload to the calculation of the dense representation.

Now we give the proof of the non-dense representation.

Proof of non-dense representation. We are going to show that

$$\begin{aligned} & \mathbb{E} [\Upsilon_X(s, v) (f_0(v) + f_1(v)X(v))^n | \mathcal{F}(s)] \\ &= e^{\alpha(s, v) + \beta(s, v)X(s)} \sum_{(i_1, \dots, i_n) \in \mathcal{C}_n} \kappa^n(i_1, \dots, i_n) \prod_{j=1}^n (A_j(s, v) + B_j(s, v)X(s))^{i_j}, \end{aligned} \quad (5.4.10)$$

where the functions incorporated in the right-hand side of the equation are fulfilling a system of ODEs. We prove the lemma by an induction proof.

The induction basis: ($n = 1$)

The result follows in the same way as the “extended transform” in Duffie et al. (2000), where one includes time-dependence of the functions of the transform.

The induction step:

We assume that (5.4.10) holds for n and show that this implies, that it also holds for $n + 1$. To make the notation simpler, we define $\mathcal{B}_n$ by

$$\mathcal{B}_n(s, v) := \Upsilon_X(s, v) (f_0(v) + f_1(v)X(v))^n.$$

We take the derivative of the left-hand side of (5.4.10) wrt. v and obtain

$$\begin{aligned} & \frac{d}{dv} \mathbb{E} [\mathcal{B}_n(s, v) | \mathcal{F}(s)] \\ &= \mathbb{E} \left[\left(-\rho_0(v) - \rho_1(v)X(v) + \frac{d}{dv}u_0(v) + \frac{d}{dv}(u_1(v))X(v) \right. \right. \\ & \quad \left. \left. + u_1(v)(K_0(v) + K_1(v)X(v)) \right) \mathcal{B}_n(s, v) \middle| \mathcal{F}(s) \right] \\ &+ \mathbb{E} \left[\Upsilon_X(s, v)(n-1)(f_0(v) + f_1(v)X(v))^{n-1} \right. \\ & \quad \left. \times \left(\frac{d}{dv}f_0(v) + \frac{d}{dv}f_1(v)X(v) + f_1(v)(K_0(v) + K_1(v)X(v)) \right) \middle| \mathcal{F}(s) \right] \\ &= \mathbb{E} \left[\Upsilon_X(s, v) \left((f_0(v) + f_1(v)X(v))^{n-1} (L_0(v) + L_1(v)X(v)) \right. \right. \\ & \quad \left. \left. + (f_0(v) + f_1(v)X(v))^n (L_2(v) + L_3(v)X(v)) \right) \middle| \mathcal{F}(s) \right] \end{aligned} \quad (5.4.11)$$

for some functions L_i .

On the other hand, we can take the derivative of the right-hand side of (5.4.10) wrt. v and obtain:

$$\begin{aligned}
& \frac{d}{dv} \left(e^{\alpha(s,v)+\beta(s,v)X(s)} \sum_{(i_1, \dots, i_n) \in \mathcal{C}_n} \kappa^n(i_1, \dots, i_n) \prod_{j=1}^n (A_j(s, v) + B_j(s, v)X(s))^{i_j} \right) \\
&= e^{\alpha(s,v)+\beta(s,v)X(s)} \sum_{(i_1, \dots, i_n) \in \mathcal{C}_n} \left(\kappa^n(i_1, \dots, i_n) (A_1(s, v) + B_1(s, v)X(s))^{i_1+1} \right. \\
&\quad \times \prod_{j=2}^n (A_j(s, v) + B_j(s, v)X(s))^{i_j} \\
&\quad + \kappa^n(i_1, \dots, i_n) \sum_{k=1}^n \prod_{j=1, j \neq k}^n ((A_j(s, v) + B_j(s, v)X(s))^{i_j}) \\
&\quad \times (i_k (A_k(s, v) + B_k(s, v)X(s))^{i_k-1} (A_{k+1}(s, v) + B_{k+1}(s, v)X(s))) \Big).
\end{aligned}$$

We collect the terms and obtain:

$$\begin{aligned}
& e^{\alpha(s,v)+\beta(s,v)X(s)} \sum_{(i_1, \dots, i_n) \in \mathcal{C}_n} \left(\kappa^n(i_1, \dots, i_n) \prod_{j=1}^n (A_j(s, v) + B_j(s, v)X(s))^{i_j+1_{\{j=1\}}} \right. \\
&\quad \left. + \sum_{k=1}^n i_k \kappa^n(i_1, \dots, i_n) \prod_{j=1}^{n+1} ((A_j(s, v) + B_j(s, v)X(s))^{i_{j,k}}) \right) \quad (5.4.12) \\
&= e^{\alpha(s,v)+\beta(s,v)X(s)} \sum_{(i_1, \dots, i_{n+1}) \in \mathcal{C}_{n+1}} \kappa^{n+1}(i_1, \dots, i_{n+1}) \prod_{j=1}^{n+1} (A_j(s, v) + B_j(s, v)X(s))^{i_j}.
\end{aligned}$$

Here, $i_{j,k} = i_j \mathbf{1}_{\{j \leq n\}} - \mathbf{1}_{\{j=k\}} + \mathbf{1}_{\{j=k+1\}}$ and

$$\begin{aligned}
& \kappa^{n+1}(i_1, \dots, i_{n+1}) = \mathbf{1}_{\{i_{n+1}=1\}} \\
& + \mathbf{1}_{\{i_{n+1} \neq 1\}} \left(\kappa^n(i_1 - 1, i_2, \dots, i_n) + \sum_{k=1}^{n-1} (i_k + 1) \kappa^n(i_1, \dots, i_k + 1, i_{k+1} - 1, \dots, i_n) \right),
\end{aligned}$$

where $\kappa^n(v_1, \dots, v_n) = 0$ for $\min(v_1, \dots, v_n) < 0$. We see, that also the right-hand side has the form stated in the lemma for $n + 1$.

We note that neither the left-hand side nor the right-hand side have the exact form stated in the lemma for $n + 1$. Here, we outline why the results also hold, if we change some boundary conditions. We want to replace the boundary conditions $L_2(v)$ and $L_3(v)$ for $f_0(v)$ and $f_1(v)$ and the boundary conditions $L_0(v)$ and $L_1(v)$ for 0 and 0. The argument relies on the well-behaved structure of the ODEs. We have that the ODEs for both B_i and A_i , $i > 0$ are linear ODEs with closed form solutions. As we show later in the proof, the ODE for B_j is given by

$$\frac{\partial}{\partial s} B_j(s, v) = - (K_1(s) + \beta(s, v)) B_j(s, v) - \frac{1}{2} \sum_{i=1}^{j-1} \binom{j}{i} B_i(s, v) H_1(s) B_{j-i}(s, v). \quad (5.4.13)$$

Given that $B(v, v) = \varsigma$, the solution to (5.4.13) is given by

$$B_j(s, v) = e^{\int_s^v K_1(\tau) + \beta(\tau, v) H_1(\tau) d\tau} \times \left(\int_s^v \frac{1}{2} \sum_{i=1}^{j-1} \binom{j}{i} B_i(u, v) H_1(u) B_{j-i}(u, v) e^{\int_s^u -K_1(\tau) - \beta(\tau, v) H_1(\tau) d\tau} du + \varsigma \right).$$

Moreover, given $A_j(v, v) = \varsigma$, A_j is given by

$$A_j(s, v) = \int_s^v K_0(\tau) B_j(\tau, v) + \frac{1}{2} \sum_{i=0}^j \binom{j}{i} B_i(\tau, v) H_0(\tau) B_{j-i}(\tau, v) d\tau + \varsigma.$$

We claim that the solutions of A_j and B_j in conjunction with the form of the partial derivative of the right-hand side of (5.4.10) given by (5.4.12) yields, that we can change the boundary conditions and obtain the result given by (5.4.9).

Regarding the form of the set $\mathcal{C}_n$: Again, we give an induction proof, where the induction basis is obvious since it follows from Duffie et al. (2000), that the only element in $\mathcal{C}_1$ is the singleton $\{1\}$.

Now, the induction step. We consider (5.4.12) for a fixed element $(i_1, \dots, i_n) \in \mathcal{C}_n$. We note, that we in total have $n + 1$ additive terms. By the induction hypothesis for n we have that $\sum_{j=1}^n j i_j = n$. This gives us, that the $1 + \sum_{j=1}^n j i_j = n + 1$ for the first term of the left-hand side of (5.4.12). For term k of the last n terms the sums are given by $n - k + (k + 1) = n + 1$. That is, for a vector v in $\mathcal{C}_{n+1}$, we have that $\sum_{i=1}^{n+1} i v_i = n + 1$ and we have proved the form of $\mathcal{C}_i$, $i \in \mathbb{N}$.

Regarding the system of ODEs. We describe the procedure for the terms β and B_i . The calculations for terms α and A_i follow in exactly the same manner. We know from Duffie et al. (2000), that β fulfills the ODE:

$$\frac{\partial}{\partial s} \beta(s, v) = \rho_1(s) - K_1(s) \beta(s, v) - \frac{1}{2} \beta^2(s, v) H_1(s). \quad (5.4.14)$$

Taken the derivative of (5.4.14) wrt. v and setting $\frac{\partial}{\partial v} \beta(s, v) = B_1(s, v)$ we get

$$\frac{\partial}{\partial s} B_1(s, v) = -K_1(s) B_1(s, v) - \beta(s, v) H_1(s) B_1(s, v). \quad (5.4.15)$$

Taken the derivative of (5.4.15) wrt. v and setting $\frac{\partial}{\partial v} B_1(s, v) = B_2(s, v)$ we get

$$\frac{\partial}{\partial s} B_2(s, v) = -K_1(s) B_2(s, v) - \beta(s, v) H_1(s) B_2(s, v) - H_1(s) B_1^2(s, v). \quad (5.4.16)$$

The general representation of the ODEs can be proven by an induction proof:

The induction basis: ($n = 1$)

We use (5.4.15), the notation $B_0 := \beta$ and obtain:

$$\begin{aligned} \frac{\partial}{\partial s} B_1(s, v) &= -K_1(s) B_1(s, v) - 2 \frac{1}{2} B_0(s, v) H_1(s) B_1(s, v) \\ &= -K_1(s) B_1(s, v) - \frac{1}{2} \sum_{i=0}^1 \binom{1}{i} B_i(s, v) H_1(s) B_{1-i}(s, v). \end{aligned}$$

That is, the induction basis is correct.

The induction step:

We assume that equation (5.4.17) holds for a natural number n :

$$\frac{\partial}{\partial s} B_n(s, v) = -K_1(s)B_n(s, v) - \frac{1}{2} \sum_{i=0}^n \binom{n}{i} B_i(s, v) H_1(s) B_{n-i}(s, v). \quad (5.4.17)$$

We want to show, that (5.4.17) also holds for $n+1$. Differentiating the equation wrt. v yields:

$$\begin{aligned} \frac{\partial}{\partial v} \frac{\partial}{\partial s} B_n(s, v) &= -K_1(s) \frac{\partial}{\partial v} B_n(s, v) - \frac{1}{2} \sum_{i=0}^n \binom{n}{i} \left(\frac{\partial}{\partial v} (B_i(s, v)) H_1(s) B_{n-i}(s, v) \right. \\ &\quad \left. + B_i(s, v) H_1(s) \frac{\partial}{\partial v} (B_{n-i}(s, v)) \right) \\ \Rightarrow \frac{\partial}{\partial s} B_{n+1}(s, v) &= -K_1(s) B_{n+1}(s, v) - \frac{1}{2} \sum_{i=0}^n \binom{n}{i} \left(B_{i+1}(s, v) H_1(s) B_{n-i}(s, v) \right. \\ &\quad \left. + B_i(s, v) H_1(s) B_{n+1-i}(s, v) \right) \\ &= -K_1(s) B_{n+1}(s, v) - \frac{1}{2} \sum_{i=0}^{n+1} \binom{n+1}{i} B_i(s, v) H_1(s) B_{n+1-i}(s, v). \end{aligned}$$

This concludes the proof. □

We now wrap up the quantity given by (5.4.2) from inside out using the results of Lemma 5.4.1 and Lemma 5.4.2. Note, we cannot directly apply the results of the two lemmas, since there is some time-dependence in (5.4.2) which the lemma does not allow for. For the non-dense representation it means, that we in general are not able to represent the term (5.4.1) on this form:

$$e^{\alpha(s,v)+\beta(s,v)X(s)} \sum_{(i_1, \dots, i_n) \in \mathcal{C}_n} \kappa^n(i_1, \dots, i_n) \prod_{j=1}^n (A_j(s, v) + B_j(s, v)X(s))^{i_j}.$$

As we show in the following, we still have the same overall structure, but instead of a term like $(A_1(t, t_1) + B_n(t, t_1)x)^n$ we have a term $\prod_{k=1}^n (A_1(t, t_1; t_k) + B_1(t, t_1; t_k)x)$ reflecting that the jumps of the Markov chain happen at different points in time as opposed to the assumptions in Lemma 5.4.2. The dynamics of B_1 are the same as stated in Lemma 5.4.2, but the boundary values are different. For the functions $B_i, i \geq 2$ things get even messier because different “versions” both have different boundary conditions and different dynamics. The structure of the dynamics are the same, but the dynamics are different, since the dynamics of different versions of B_i depend on different versions of $B_j, j < i$ resulting in different versions of B_i . For three jumps, which include the model represented in Figure 5.1, this system is stated explicitly in Section 5.4.2. Otherwise, the approach is the same as in Lemma 5.4.1 and Lemma 5.4.2. In the terms below, the functions $\alpha_i, \beta_i, C_{ij}, A_{ij}$ and B_{ij} are solutions to systems of ODEs. The exact form of these ODEs are specified in Section 5.4.2. We let the affine functions $\mu_i, i = 1, 2, 3$ be given by $\mu_i(t, x) = \mu_{i0}(t) + \mu_{i1}(t)x$.

To exemplify this difference due to the time-dependence, we calculate the term (5.4.2) from inside and out by using Lemma 5.4.1 and Lemma 5.4.2. We do so by taking each step (each jump time)

at a time. This both covers the dense and the non-dense representation. In the notation we suppress that e.g. C_{30} implicitly depends on time T .

First we consider the **time point** $\mathbf{t}_3$, i.e. the time point when no jumps have occurred. The result follows directly from Duffie et al. (2000, Proposition 1):

$$\begin{aligned} \mathbb{E} \left[e^{-\int_{t_3}^T f_4(\tau, X(\tau)) d\tau} \middle| \mathcal{F}(t_3) \right] &= e^{\alpha_4(t_3, T) + \beta_4(t_3, T) X(t_3)}, \\ \alpha_4(T, T) &= \beta_4(T, T) = 0. \end{aligned} \quad (5.4.18)$$

Using the result (5.4.18), Lemma 5.4.1 and Lemma 5.4.2, we get that the **time** $\mathbf{t}_2$ value, where one jump has occurred, is given by:

Dense representation:

$$\begin{aligned} \mathbb{E} \left[e^{-\int_{t_2}^{t_3} f_3(\tau, X(\tau)) d\tau} e^{\alpha_4(t_3, T) + \beta_4(t_3, T) X(t_3)} \mu_3(t_3, X(t_3)) \middle| \mathcal{F}(t_2) \right] \\ = e^{\beta_3(t_2, t_3) X(t_2)} (C_{30}(t_2, t_3) + C_{31}(t_2, t_3) X(t_2)), \\ \beta_3(t_3, t_3) = \beta_4(t_3, T), C_{30}(t_3, t_3) = e^{\alpha_4(t_3, T)} \mu_{30}(t_3), C_{31}(t_3, t_3) = e^{\alpha_4(t_3, T)} \mu_{31}(t_3). \end{aligned} \quad (5.4.19)$$

Non-dense representation:

$$\begin{aligned} \mathbb{E} \left[e^{-\int_{t_2}^{t_3} f_3(\tau, X(\tau)) d\tau} e^{\alpha_4(t_3, T) + \beta_4(t_3, T) X(t_3)} \mu_3(t_3, X(t_3)) \middle| \mathcal{F}(t_2) \right] \\ = e^{\alpha_3(t_2, t_3) + \beta_3(t_2, t_3) X(t_2)} (A_{31}(t_2, t_3) + B_{31}(t_2, t_3) X(t_2)), \\ \alpha_3(t_3, t_3) = \alpha_4(t_3, T), \beta_3(t_3, t_3) = \beta_4(t_3, T), A_{31}(t_3, t_3) = \mu_{30}(t_3), B_{31}(t_3, t_3) = \mu_{31}(t_3). \end{aligned} \quad (5.4.20)$$

To get the value at time **time** $\mathbf{t}_1$, where two jumps occurred, we use (5.4.19) and (5.4.20).

Dense representation:

$$\begin{aligned} \mathbb{E} \left[e^{-\int_{t_1}^{t_2} f_2(\tau, X(\tau)) d\tau} e^{\beta_3(t_2, t_3) X(t_2)} (C_{30}(t_2, t_3) + C_{31}(t_2, t_3) X(t_2)) \mu_2(t_2, X(t_2)) \middle| \mathcal{F}(t_1) \right] \\ = e^{\beta_2(t_1, t_2) X(t_1)} \left(\sum_{i=0}^2 C_{2i}(t_1, t_2) X^i(t_1) \right), \\ \beta_2(t_2, t_2) = \beta_3(t_2, t_3), C_{20}(t_2, t_2) = C_{30}(t_2, t_3) \mu_{20}(t_2), \\ C_{21}(t_2, t_2) = C_{31}(t_2, t_3) \mu_{20}(t_2) + C_{30}(t_2, t_3) \mu_{21}(t_2), C_{22}(t_2, t_2) = C_{31}(t_2, t_3) \mu_{21}(t_2). \end{aligned} \quad (5.4.21)$$

Non-dense representation:

$$\begin{aligned} \mathbb{E} \left[e^{-\int_{t_1}^{t_2} f_2(\tau, X(\tau)) d\tau} e^{\alpha_3(t_2, t_3) + \beta_3(t_2, t_3) X(t_2)} \right. \\ \left. \times (A_{31}(t_2, t_3) + B_{31}(t_2, t_3) X(t_2)) \mu_2(t_2, X(t_2)) \middle| \mathcal{F}(t_1) \right] \\ = e^{\alpha_2(t_1, t_2) + \beta_2(t_1, t_2) X(t_1)} \left((A_{21}(t_1, t_2) + B_{21}(t_1, t_2) X(t_1)) (A_{22}(t_1, t_2) + B_{22}(t_1, t_2) X(t_1)) \right. \\ \left. + A_{23}(t_1, t_2) + B_{23}(t_1, t_2) X(t_1) \right), \\ \alpha_2(t_2, t_2) = \alpha_3(t_2, t_3), \beta_2(t_2, t_2) = \beta_3(t_2, t_3), A_{21}(t_2, t_2) = A_{31}(t_2, t_3), B_{21}(t_2, t_2) = B_{31}(t_2, t_3), \\ A_{22}(t_2, t_2) = \mu_{20}(t_2), B_{22}(t_2, t_2) = \mu_{21}(t_2), A_{23}(t_2, t_2) = B_{23}(t_2, t_2) = 0. \end{aligned} \quad (5.4.22)$$

In the same manner, we can also obtain results at **time** $\mathbf{t}$ of the quantities

$$\begin{aligned} & \mathbb{E} \left[e^{-\int_t^{t_1} f_1(\tau, X(\tau)) d\tau} e^{\beta_2(t_1, t_2) X(t_1)} \sum_{i=0}^2 (C_{2i}(t_1, t_2) X^i(t_1, t_2)) \mu_1(t_1, X(t_1)) \middle| \mathcal{F}(t_1) \right] \\ &= e^{\beta(t, t_1) X(s)} \left(\sum_{i=0}^3 C_i(t, t_1) X^i(t) \right). \end{aligned} \quad (5.4.23)$$

We can also calculate the non-dense representation

$$\begin{aligned} & \mathbb{E} \left[e^{-\int_t^{t_1} f_1(\tau, X(\tau)) d\tau} e^{\alpha_2(t_1, t_2) + \beta_2(t_1, t_2) X(t_1)} \left((A_{21}(t_1, t_2) + B_{21}(t_1, t_2) X(t_1)) \right. \right. \\ & \quad \left. \left. \times (A_{22}(t_1, t_2) + B_{22}(t_1, t_2) X(t_1)) + (A_{23}(t_1, t_2) + B_{23}(t_1, t_2) X(t_1)) \right) \mu_1(t_1, X(t_1)) \middle| \mathcal{F}(t) \right]. \end{aligned} \quad (5.4.24)$$

The result is quite lengthy and since it is stated in details in Section 5.4.2, we skip it here.

5.4.2 ODEs for the non-dense transforms

As mentioned in the former subsection we cannot get the ODEs for the terms (5.4.18)-(5.4.24) directly from Lemma 5.4.2. This is why we state the ODEs in this subsection. When one wants to calculate a statewise reserve of an insurance contract or the transition probabilities in a Markov chain model, the approach is to do the calculations backwards starting from the last time point and obtain the first systems of ODEs. Then we use these values as input for the next system of ODEs and so on. After discretizing the entire time interval, we need to calculate the transforms for all different time points and at last integrate.

Because of the dependence of the different jump times, it is not doable for the non-dense representation to give a representation as neat as the one in Lemma 5.4.2. By neat we mean a representation which holds for fixed time points s and v and does not depend on anything after time v . Therefore, the ODEs for a model with up to 3 jumps are stated here. That is, one is able to calculate all statewise reserves in the model given by Figure 5.1. In practice, it is not too interesting to do calculations for a very high n due to the computational workload one encounters, see Section 5.5.1. The workload does not come from solving the equations for a single transform, but rather from the massive number of transforms needed in order to obtain the value for a reserve.

Before stating the precise results of (5.4.20), (5.4.22) and (5.4.24), one should note, that due to the structure of the transforms, it is possible to “glue” together the solutions to some of the ODEs. Among others, it holds for the pairs $\{\alpha_4, \alpha_3\}$, $\{\beta_4, \beta_3\}$, $\{A_{31}, A_{21}\}$ and $\{B_{31}, B_{21}\}$. That is, we can specify some of the ODEs which have boundary conditions at time s wrt. boundary conditions for time points after time s . This works because for a given function f it is equivalent to specify a function ε either by

$$\frac{d}{dt} \varepsilon(t) = f(t), \quad \varepsilon(v) = K, \quad (5.4.25)$$

or by

$$\frac{d}{dt} \varepsilon(t) = f(t), \quad \varepsilon(s) = \tilde{\varepsilon}(s), \quad \frac{d}{dt} \tilde{\varepsilon}(t) = f(t), \quad \tilde{\varepsilon}(v) = K, \quad (5.4.26)$$

where $t < s < v$ and K is a constant. So by “gluing” we mean, that we instead of the representation (5.4.25) which arises in a natural way from the conditional approach, use the equivalent representation (5.4.26). The latter representation highlights that the function ε has its natural boundary condition at time v .

If we “glue” the various differential equations, we obtain the following results for the different points in time. To shorten notation, we write $(A_1 + B_1 X)(t, t_3, T)$ rather than $A_1(t, t_3, T) + B_1(t, t_3, T)X(t)$. We mark by **red** text the time point for which the boundary condition is specified, and by **blue** text we mark the time points in which the functions (excluding the two functions for which we are specifying the dynamics), that the ODEs are depending on, have boundary conditions. For example the arguments of $A_1(t, t_3, T)$ follow in this way: The T is caused by the dynamics being dependent on $\beta(t, T)$ with boundary condition in T . Moreover, the dynamics also depend on $B_1(t, t_3, T)$ but as stated previously, we suppress this. The argument t_3 arise due to that $A_1(t, t_3, T)$ has a boundary condition at time t_3 .

Time t_2 :

$$(5.4.20) = e^{\alpha(t_2, T) + \beta(t_2, T)X(t_2)} (A_1 + B_1 X)(t_2, t_3, T).$$

Time t_1 :

$$(5.4.22) = e^{\alpha(t_1, T) + \beta(t_1, T)X(t_1)} ((A_1 + B_1 X)(t_1, t_3, T) (A_1 + B_1 X)(t_1, t_2, T) + (A_2 + B_2 X)(t_1, t_2, t_3, T)).$$

Time t :

$$(5.4.24) = e^{\alpha(t, T) + \beta(t, T)X(t)} ((A_1 + B_1 X)(t, t_3, T) (A_1 + B_1 X)(t, t_2, T) \times (A_1 + B_1 X)(t, t_1, T) + (A_2 + B_2 X)(t, t_2, t_3, T) (A_1 + B_1 X)(t, t_1, T) + (A_2 + B_2 X)(t, t_1, t_3, T) (A_1 + B_1 X)(t, t_2, T) + (A_1 + B_1 X)(t, t_3, T) (A_2 + B_2 X)(t, t_1, t_2, T) + (A_3 + B_3 X)(t, t_1, t_1, t_2, t_3, T)).$$

The above functions are solutions to the following systems of ODEs (we only state the ODEs for the β and the B_i functions, since the ODEs for α and the A_i functions are the same except that K_0 is substituted for K_1 , H_0 is substituted for H_1 , and μ_0 is substituted for μ_1 in the boundary

conditions):

$$\begin{aligned}
\frac{\partial}{\partial t} B_1(t, \mathbf{t}_3, T) &= -K_1(t)B_1(t, \mathbf{t}_3, T) - \beta(t, T)H_1(t)B_1(t, \mathbf{t}_3, T), & B_1(t_3, \mathbf{t}_3, T) &= \mu_1(t_3). \\
\frac{\partial}{\partial t} B_1(t, \mathbf{t}_2, T) &= -K_1(t)B_1(t, \mathbf{t}_2, T) - \beta(t, T)H_1(t)B_1(t, \mathbf{t}_2, T), & B_1(t_2, \mathbf{t}_2, T) &= \mu_1(t_2). \\
\frac{\partial}{\partial t} B_1(t, \mathbf{t}_1, T) &= -K_1(t)B_1(t, \mathbf{t}_1, T) - \beta(t, T)H_1(t)B_1(t, \mathbf{t}_1, T), & B_1(t_1, \mathbf{t}_1, T) &= \mu_1(t_1). \\
\frac{\partial}{\partial t} B_2(t, \mathbf{t}_2, \mathbf{t}_3, T) &= -K_1(t)B_2(t, \mathbf{t}_2, \mathbf{t}_3, T) - \beta(t, T)H_1(t)B_2(t, \mathbf{t}_2, \mathbf{t}_3, T) \\
&\quad - H_1(t)B_1(t, \mathbf{t}_2, T)B_1(t, \mathbf{t}_3, T), & B_2(t_2, \mathbf{t}_2, \mathbf{t}_3, T) &= 0. \\
\frac{\partial}{\partial t} B_2(t, \mathbf{t}_1, \mathbf{t}_2, T) &= -K_1(t)B_2(t, \mathbf{t}_1, \mathbf{t}_2, T) - B_1(t, \mathbf{t}_1, T)H_1(t)B_1(t, \mathbf{t}_2, T) \\
&\quad - \beta(t, T)H_1(t)B_2(t, \mathbf{t}_1, \mathbf{t}_2, T), & B_2(t_1, \mathbf{t}_1, \mathbf{t}_2, T) &= 0. \\
\frac{\partial}{\partial t} B_2(t, \mathbf{t}_1, \mathbf{t}_3, T) &= -K_1(t)B_2(t, \mathbf{t}_1, \mathbf{t}_3, T) - B_1(t, \mathbf{t}_1, T)H_1(t)B_1(t, \mathbf{t}_3, T) \\
&\quad - \beta(t, T)H_1(t)B_2(t, \mathbf{t}_1, \mathbf{t}_3, T), & B_2(t_1, \mathbf{t}_1, \mathbf{t}_3, T) &= 0. \\
\frac{\partial}{\partial t} B_3(t, \mathbf{t}_1, \mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3, T) &= -K_1(t)B_3(t, \mathbf{t}_1, \mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3, T) - \beta(t, T)H_1(t)B_3(t, \mathbf{t}_1, \mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3, T) \\
&\quad - B_1(t, \mathbf{t}_1, T)H_1(t)B_2(t, \mathbf{t}_2, \mathbf{t}_2, \mathbf{t}_3, T) \\
&\quad - H_1(t)B_2(t, \mathbf{t}_1, \mathbf{t}_3, T)B_1(t, \mathbf{t}_2, T) - H_1(t)B_2(t, \mathbf{t}_1, \mathbf{t}_2, T)B_1(t, \mathbf{t}_3, T), \\
&\quad B_3(t_1, \mathbf{t}_1, \mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3, T) = 0.
\end{aligned}$$

To illustrate how one obtains the above ODEs, we show how to obtain the ODE for $B_2(t, \mathbf{t}_1, \mathbf{t}_3, T)$ by taking the derivative of the ODE for $B_1(t, \mathbf{t}_3, T)$ wrt. t_1 .

$$\begin{aligned}
\frac{\partial}{\partial t_1} \frac{\partial}{\partial t} B_1(t, \mathbf{t}_3, T) &= \frac{\partial}{\partial t_1} (-K_1(t)B_1(t, \mathbf{t}_3, T) - \beta(t, T)H_1(t)B_1(t, \mathbf{t}_3, T)) \\
\Rightarrow \frac{\partial}{\partial t} \frac{\partial}{\partial t_1} B_1(t, \mathbf{t}_3, T) &= -K_1(t) \frac{\partial}{\partial t_1} B_1(t, \mathbf{t}_3, T) - \left(\frac{\partial}{\partial t_1} \beta(t, T) \right) H_1(t)B_1(t, \mathbf{t}_3, T) \\
&\quad - \beta(t, T)H_1(t) \frac{\partial}{\partial t_1} B_1(t, \mathbf{t}_3, T) \\
\Rightarrow \frac{\partial}{\partial t} B_2(t, \mathbf{t}_1, \mathbf{t}_3, T) &= -K_1(t)B_2(t, \mathbf{t}_1, \mathbf{t}_3, T) - B_1(t, \mathbf{t}_1, T)H_1(t)B_1(t, \mathbf{t}_3, T) \\
&\quad - \beta(t, T)H_1(t)B_2(t, \mathbf{t}_1, \mathbf{t}_3, T).
\end{aligned}$$

5.4.3 Generalizing to higher dimensions

The results of the previous subsection can easily be generalized to the case where the underlying process is a multidimensional affine process. That is, we assume that X is given by the stochastic differential equation (5.3.1) but without jumps. Unfortunately, it is rather hard to state the dense representation for a general dimension (d) of the affine process and for a general number of jumps (n), even though it is doable in practice for fixed numbers n and d by following the approach in the proof of Lemma 5.4.1. It is especially hard to state a general result for a dense representation, since one gets all different kinds of mixed terms. For the non-dense approach, the result is exactly the same except that the dimension of the ODEs is d and not 1. Here, we state a multidimensional version of Lemma 5.4.2 without giving any proof.

Theorem 5.4.5 (Non-dense representation). *Let $s, v \in [t, T]$ with $s < v$. Assuming the same integrability conditions as in Lemma 5.4.2 (in a multidimensional setting), we have that*

$$\begin{aligned} & E \left[e^{-\int_s^v \rho_0(\tau) + \rho_1^{tr}(\tau) X(\tau) d\tau} e^{u_0(v) + u_1^{tr}(v) X(v)} (f_0(v) + f_1^{tr}(v) X(v))^n \middle| \mathcal{F}(s) \right] \\ &= e^{\alpha(s,v) + \beta^{tr}(s,v)x} \sum_{(i_1, \dots, i_n) \in \mathcal{C}_n} \kappa^n(i_1, \dots, i_n) \prod_{j=1}^n (A_j(s, v) + B_j^{tr}(s, v)x)^{i_j}, \end{aligned}$$

where $\kappa^n(i_1, \dots, i_n) \in \mathbb{N}_0$ and the set $\mathcal{C}_n$ is given by

$$\mathcal{C}_n = \left\{ v \in \mathbb{N}_0^n \middle| \sum_{i=1}^n i v_i = n \right\}.$$

Moreover, $\alpha, \beta, A_i, B_i, i = 1, \dots, n$ fulfill the ODEs

$$\begin{aligned} \frac{\partial}{\partial s} \beta(s, v) &= \rho_1(s) - K_1^{tr}(s) \beta(s, v) - \frac{1}{2} \beta^{tr}(s, v) H_1(s) \beta(s, v), \\ \frac{\partial}{\partial s} \alpha(s, v) &= \rho_0(s) - K_0^{tr}(s) \beta(s, v) - \frac{1}{2} \beta^{tr}(s, v) H_1(s) \beta(s, v), \\ \frac{\partial}{\partial s} B_n(s, v) &= -K_1^{tr}(s) B_n(s, v) + \frac{1}{2} \sum_{i=0}^n \binom{n}{i} B_i^{tr}(s, v) H_1(s) B_{n-i}(s, v), \\ \frac{\partial}{\partial s} A_n(s, v) &= -K_0^{tr}(s) B_n(s, v) + \frac{1}{2} \sum_{i=0}^n \binom{n}{i} B_i^{tr}(s, v) H_0(s) B_{n-i}(s, v) \end{aligned}$$

with terminal conditions

$$\begin{aligned} \beta(v, v) &= u_1(v), \quad \alpha(v, v) = u_0(v), \quad B_1(v, v) = f_1(v), \quad A_1(v, v) = f_0(v), \\ B_j(v, v) &= 0, \quad A_j(v, v) = 0, \quad j > 1. \end{aligned}$$

By similar arguments as in the one-dimensional case, we can “glue” the ODEs and obtain exactly the same results as in Section 5.4.2. We just need to replace the one-dimensional notation with the multidimensional vector notation.

5.5 Different subtopics relating to affine processes and reserves in Markov chain models

In this section we present different topics relating to the results in Section 5.4. These topics are a comparison of the dense and the non-dense representation and a description of a special case, where one can obtain simple results even in the case of cycles in the Markov chain model. Lastly, we comment on that one, under some additional assumptions, can include quadratic terms in the dynamics of the underlying stochastic process and quadratic terms in the transforms of the underlying stochastic processes and still be able to obtain results of the same type as the ones in Section 5.4.

5.5.1 Comparison of representations/computational efficiency

Affine processes have gained their popularity in areas as economics, life insurance and credit risk modelling because of their computational tractability. This tractability for some modelling purposes is caused by the opportunity to solve ODEs instead of PDEs. In general, it is much simpler

and computationally faster to solve ODEs instead of PDEs. Since computational efficiency is one of the main advantages of the class of affine processes, it is relevant to consider the computational workload. Our main example in this paper is to calculate life insurance reserves, why we in this section consider the computational workload of calculating these reserves. Moreover, this section includes some comments about the differences between the dense and the non-dense representation. In short, the main difference is that the dense representation leads to a minimum number of ODEs to solve, whereas the non-dense representation is overparameterized. On the other hand, the ODEs for the dense representation are in general more complicated and harder to obtain than the ODEs for the non-dense representation.

In Table 5.2 we show a comparison between the dense and the non-dense representation in terms of the number of ODEs needed to be solved to calculate the “standard transforms” given in Lemma 5.4.1 for the one-dimensional case. This comparison covers different numbers of jumps in the case where the underlying stochastic process is of dimension 1 and 2, respectively. When we state the number of differential equations that needs to be solved, we count two-dimensional ODE as 2 equations.

Dimension: 1			Dimension: 2		
Jumps	Dense	Non-dense	Jumps	Dense	Non-dense
0	2	2	0	3	3
1	3	4	1	5	6
2	4	8	2	8	12
3	5	16	3	12	24

Table 5.2: Comparison of the number of ODEs in case of one- and two-dimensional affine processes.

Regardless of whether we consider the dense or the non-dense representation of the transform, the main problem with either approach is that one needs to calculate lots of ODEs to obtain a reserve. The computational challenge is not so much the dimension of the underlying stochastic process as it is the number of jumps. It does for example hold for the non-dense representation, that the number of ODEs is linear in the dimension in the sense, that the number of ODEs is given as a multiple of $(d + 1)$, where d is the dimension of the affine stochastic process. This is not too bad and the greater problems arise when calculating the value of a reserve. In that case, one needs to integrate the solutions of the ODEs. Moreover, the different ODEs depend on each other through the fact that the solutions to some ODEs are needed as boundary values for other ODEs. To exemplify this, we consider the case with an underlying one-dimensional stochastic process and a calculation of the quantity given by (5.2.1). If the length of the contract $(T - t)$ is 30 years, and we assume that we discretize each year into 10 intervals, we in total need to calculate ODEs in the magnitude of 1 billion time points.

5.5.2 Affine processes and models with cycles

Usually when dealing with Markov chains, affine processes are almost exclusively used for models *without* cycles, which are also called hierarchical Markov chain models. The reason for this is that transition probabilities in such models can be expressed as integrals of solutions to ODEs as shown in the previous sections. Nonetheless, various products in the life insurance industry would more naturally be valued in Markov chain models with cycles. The most prominent of these examples is the case of disability insurance, where it is naturally to assume, that the insur-

ance policy can “jump” between the states “active/paying premiums” and “disabled/receiving disability benefits”.

In models with cycles the affine structure is in the general case not of any help. However, there is a class of interesting examples, where the affine structure is still useful, which we outline here. The affine structure is useful in the case, where one can diagonalize an intensity matrix Q (also known as a transition rate matrix) of a continuous time Markov chain such that

$$Q(t, X(t)) = VD(t, X(t))V^{-1},$$

where V is time-homogeneous and D is a diagonal matrix whose non-zero entries are affine in the underlying affine process.

The most common case where one is able to diagonalize the intensity matrix is in the case of proportional intensities. An example of such a matrix could be

$$Q(t, x) = \begin{pmatrix} -(\kappa + \rho)x & \kappa x & \rho x \\ \varrho x & -(\varrho + \phi)x & \phi x \\ 0 & 0 & 0 \end{pmatrix},$$

which corresponds to Figure 5.2.

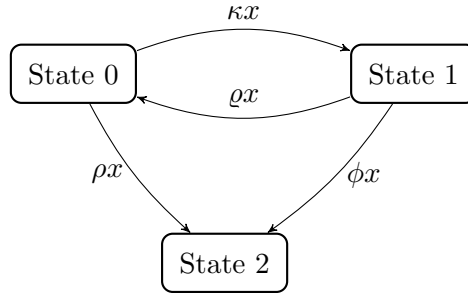


Figure 5.2: Diagonalizable three state model.

In such a model we can easily obtain results for the transition probabilities. We denote by Z the state process of the Markov chain and by τ_T the integral $\int_t^T X(s)ds$. By the tower property and by making a time change in the matrix exponential, we get that for example the probability of going from “State z ” to “State j ” from time t to time T is given by

$$\begin{aligned}
P_{Z(t)=z, X(t)}(Z(T) = j) &= \mathbb{E} [\mathbf{1}_{\{Z(T)=j\}} \mid Z(t) = z, X(t)] \\
&= \mathbb{E} [\mathbb{E} [\mathbf{1}_{\{Z(T)=j\}} \mid Z(t) = z, X(s), s \leq T,] \mid Z(t) = z, X(t)] \\
&= \mathbb{E} [(e^{Q\tau_T})_{zj} \mid Z(t) = z, X(t)] \\
&= \mathbb{E} [(V e^{D\tau_T} V^{-1})_{zj} \mid Z(t) = z, X(t)] \\
&= \sum_{i=0}^2 \left(v_{zi} \tilde{v}_{ij} \mathbb{E} \left[e^{d_i \int_t^T X(s)ds} \mid \mathcal{F}(t) \right] \right),
\end{aligned} \tag{5.5.1}$$

where v_{ij} are the entries of V , $\tilde{v}_{ij}$ are the entries of V^{-1} , $D = \text{diag}(d_1 X, d_2 X, d_3 X)$ and the term $\mathbb{E} [e^{d_i \int_t^T X(s)ds} \mid \mathcal{F}(t)]$ is a standard transform considered in this paper. That is, one can obtain transition probabilities by diagonalizing a matrix, calculating simple transforms of affine processes and summing the terms.

5.5.3 The linear quadratic class

In the literature one sometimes sees examples of transition rates modelled as linear quadratic functions. One example is De Giovanni (2010), where the surrender rate γ is given by

$$\gamma(t) = L_0 r^2(t) + L_1.$$

Here, r is the short rate and L_0 and L_1 are constants. That is, in general it could be interesting to be able to consider other transforms than the previously described affine ones. Moreover, it would also make the modelling possibilities more flexible, if we were able to have other types of dynamics than the affine ones. This section is about including these cases and still be able to obtain results by solving ODEs instead of PDEs.

If we add some more structure, we can extend the class of stochastic processes from the affine to the linear quadratic case. This requires some extra conditions to be fulfilled (for the coordinates, where one has quadratic terms). In one dimension, we can calculate transforms of the form

$$\mathbb{E} \left[e^{-\int_t^T \rho_0(s) + \rho_1(s)X(s) + \frac{1}{2}\rho_2(s)X^2(s)ds} \middle| \mathcal{F}(t) \right] = e^{c(t) + b(t)X(t) + \frac{1}{2}\Gamma(t)X^2(t)}, \quad (5.5.2)$$

where the dynamics of the underlying stochastic process are allowed to contain the following terms:

$$dX(t) = (K_0(t) + K_1(t)X(t))dt + \sqrt{H_0(t)}dW(t). \quad (5.5.3)$$

For a two-dimensional stochastic process, where there is a quadratic term in one of the dimensions but not in the other, we can calculate transforms of the form

$$\mathbb{E} \left[e^{-\int_t^T \rho_0(s) + \rho_1(s)X_1(s) + \rho_2(s)X_2(s) + \rho_3(s)X_1^2(s)ds} \middle| \mathcal{F}(t) \right] = e^{a(t) + b(t)X_1(t) + b_2(t)X_2(t) + c(t)X_1^2(t)},$$

where the dynamics are allowed to contain the following terms:

$$d \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} = \begin{pmatrix} K_{10}(t) + K_{11}(t)X_1(t) \\ K_{20}(t) + K_{21}^{\text{tr}}(t)X(t) + K_{22}(t)X_1^2(t) \end{pmatrix} dt + \sigma(t)dW(t), \quad (5.5.4)$$

where $\sigma\sigma^{\text{tr}}$ is given by

$$\sigma(t)\sigma^{\text{tr}}(t) = \begin{pmatrix} H_{10}(t) & 0 \\ 0 & H_{20}(t)X_2(t) \end{pmatrix}.$$

This setup still covers many of the well known interest rate models e.g. the Vasiček and the Hull-White model. It can be shown in general, that the linear quadratic class is equivalent to the affine class, which can be proven by an expansion of the statespace for the affine processes. For a proof and more details about this equivalence, see Cheng and Scaillet (2007).

Chapter 6

Optimal surplus distribution problems for regulated funds with assets and liabilities

Abstract: This paper studies optimal surplus distribution strategies for an infinite time horizon within a model where asset and liability values are modeled by correlated, geometric Brownian motions. The controls considered are, that we either increase liabilities or decrease assets. The increase of liabilities could be used in the modelling of non-for-profit mutual funds or pension funds. On the other hand, the decrease in assets could be used for modelling of for-profit companies. In the first part of the paper, we prove optimality of a barrier strategy under the simple solvency constraint that no distribution can be made if the funding ratio is below a certain level. In the second and third part of the paper, we study barrier strategies in a model with more advanced solvency constraints and a model where ruin must be prevented. The advanced solvency constraint is that there is an interval $[\alpha_1, \alpha_2]$ in which the allowance of controlling the process depends on which of the two barriers α_1 and α_2 that has been crossed last. The interpretation of the advanced solvency constraint is that an institution which have been in an emergency state should not be allowed to distribute before having achieved a given level of consolidation. To prevent ruin, we either lower the liabilities (in the pension fund case) or we inject capital (in the for-profit case). This leads to different levels for the optimal upper barrier for the case where we control liabilities and for the case where we control assets, respectively. This is different to the case where ruin is a possibility. The formalistic results are exemplified in a series of numerical studies including cost/benefit analysis of the different solvency constraints and the effect of making the price of raising capital dependent on the funding ratio.

Keywords: Optimal dividends, capital injection, funding ratio, stochastic control, regulation, solvency.

6.1 Introduction

6.1.1 Motivation and main contributions

In actuarial risk theory, the *stability problem* is about modelling the dynamics of risky businesses in a stylised fashion in order to help them make decisions about how to manage their risk; see (Bühlmann, 1970) for a classical reference. Over the past century, a variety of (decision)

criteria were considered, including the probability of ruin (see Asmussen and Albrecher, 2010, for a recent comprehensive review) or the expected present value of dividends (see Albrecher and Thonhauser, 2009; Avanzi, 2009, for recent reviews). In their purest form, these criteria have various shortcomings, that researchers have tried to address over time. The criteria are also sometimes modified or augmented to better fit some specific contexts.

One of the criticisms formulated against the expected present value of dividends criterion (see Gerber, 1974, for instance) as introduced in de Finetti (1957) was the lack of explicit focus (or consideration) of solvency in the criterion, and in its optimal definition. In this paper, we consider a profitable, risky setting with two separate, correlated asset and liability processes (see Section 6.1.2). The company that is considered is allowed to distribute excess profits (traditionally referred to as *dividends* in the literature), but is regulated and is subject to particular regulatory (solvency) constraints (the nature of which, inspired by Paulsen, 2003; Avanzi and Wong, 2012, is further developed in Section 6.1.4). Importantly, because of its bivariate nature, such *distributions* of excess profits can take two alternative forms. These can originate from a reduction of assets (and hence a payment to owners), but also from an increase of liabilities (when these represent the wealth of owners, such as in pension funds). The latter is particularly relevant if leakages do not make sense because of the context, such as in pension funds where assets are locked until retirement.

Mathematically, both distribution avenues are treated in a very similar way (although there are material differences in some cases). For the sake of brevity we will provide full details only for one case, and only results for the other. We elected to focus primarily on the ‘increase of liabilities’ case, as we believe this is the most innovative in this context.

Introducing solvency constraints improve the longevity of the regulated fund, see Sections 6.2 and 6.3, but may not prevent it. In such cases where the company must stop its activities, it may be profitable to rescue it. We consider this in Section 6.4. The form of the rescue measures—injections of assets, or decrease of liabilities—will depend on the context; this is briefly discussed in Section 6.1.5.

A bivariate geometric Brownian motion was introduced in Gerber and Shiu (2003). They considered two problems: (a) to keep the funding ratio (ratio of assets to liabilities) within a band, equalising inflows and outflows. They conjectured a fund “should” do so; (b) to maximise (in absence of inflows) the expected present value of outflows (dividends). They conjectured that a barrier dividend strategy should be optimal. Decamps et al. (2006) extended (a) to finite time horizon, while Decamps et al. (2009) proved that the conjecture in (b) is correct. Also, Chen and Yang (2010) extended the results of Gerber and Shiu (2003) to a regime-switching environment.

In this paper, we extend the results of Gerber and Shiu (2003) and Decamps et al. (2009) by including simple and advanced solvency constraints. We also consider the case when the company is rescued to prevent ruin. More specifically, we obtain closed form expressions for the expected present value of distributions (asset decrements or liability increments), penalised by rescue measures when those are considered (asset increments or liability decrements), when a distribution barrier is used, and under a range of solvency constraints. When the company cannot be rescued, optimal barrier levels are the same whether assets or liabilities are controlled, but these differ when the company is rescued to prevent ruin. Furthermore, we show that a barrier strategy is optimal amongst all admissible strategies when a simple solvency constraint is applied.

6.1.2 A bivariate asset and liability process

As mentioned earlier, we consider a bivariate surplus process, where assets and liabilities are modeled as correlated geometric Brownian motions. The dynamics of the assets, which we denote by $X_1 := A$, and the liabilities, which we denote by $X_2 := L$, are given by

$$\begin{aligned} d\vec{X}(t) &= d \begin{pmatrix} A(t) \\ L(t) \end{pmatrix} = \begin{pmatrix} \mu_A & 0 \\ 0 & \mu_L \end{pmatrix} \begin{pmatrix} A(t) \\ L(t) \end{pmatrix} dt + \begin{pmatrix} \sigma_A A(t) & 0 \\ \rho \sigma_L L(t) & \sqrt{1 - \rho^2} \sigma_L L(t) \end{pmatrix} dW(t) \\ &= \mu(\vec{X}(t))dt + \sigma(\vec{X}(t))dW(t), \end{aligned} \quad (6.1.1)$$

where $A(0) = A_0$, $L(0) = L_0$, $\rho \in]-1, 1[$ is the correlation factor, and where W is a standard two-dimensional Brownian motion. Following the lines of Gerber and Shiu (2003), we assume that the discount rate is greater than the drift of the assets, which is greater than the drift of the liabilities. That is,

$$\delta > \mu_A \quad (6.1.2)$$

and

$$\mu_A > \mu_L. \quad (6.1.3)$$

Equation (6.1.2) makes us avoid that the supremum of the expected present value of dividends is infinity and (6.1.3) makes us avoid the case, where it is optimal to pay out all dividends immediately; see Gerber and Shiu (2003, Section 9). The funding ratio, Y , is defined as the ratio of assets to liabilities. That is,

$$Y(t) = \frac{A(t)}{L(t)}, \quad t \geq 0.$$

This model setting is identical to that of Gerber and Shiu (2003). Note that Sethi and Taksar (2002) considered dividends and capital injections for a company whose surplus is modeled by a (univariate) geometric Brownian motion, which is a more traditional, uni-dimensional formulation.

6.1.3 Distribution of profits

Since $\mu_A > \mu_L$, the funding ratio will increase to infinity on average, and it makes sense that excess profits would be distributed in some way. In what follows, we will consider two ways of doing so:

- A. Increase liabilities;
- B. Decrease assets.

For the rest of the paper, we will refer to both cases as to “Case A” and “Case B”, respectively. The result of either will be referred to as a “distribution”.

The increase of liabilities (Case A) could be used in the modelling of non-for-profit mutual funds or pension funds. On the other hand, the decrease of assets (Case B) could be used in the modelling of for-profit companies. The latter case is equivalent to paying out dividends, which is the standard assumption in the actuarial dividend literature.

In this paper, we consider barrier type distributions, which we prove optimal in some cases. Note that here the barrier is defined on the funding ratio (distributions are made if the funding ratio is beyond a certain barrier level, so that the funding ratio is brought back to that particular

level). The model is illustrated in Figure 6.1, where assets, liabilities and funding ratio both for the uncontrolled (black lines) and the controlled (grey lines) processes are depicted for a given sample path of the two-dimensional Brownian motion. Moreover, the dotted line in the figures show the positive distribution processes D_A^π and D_B^π defined below. Note that the controlled funding ratio is the same for both cases, whereas the expected present values will be different.

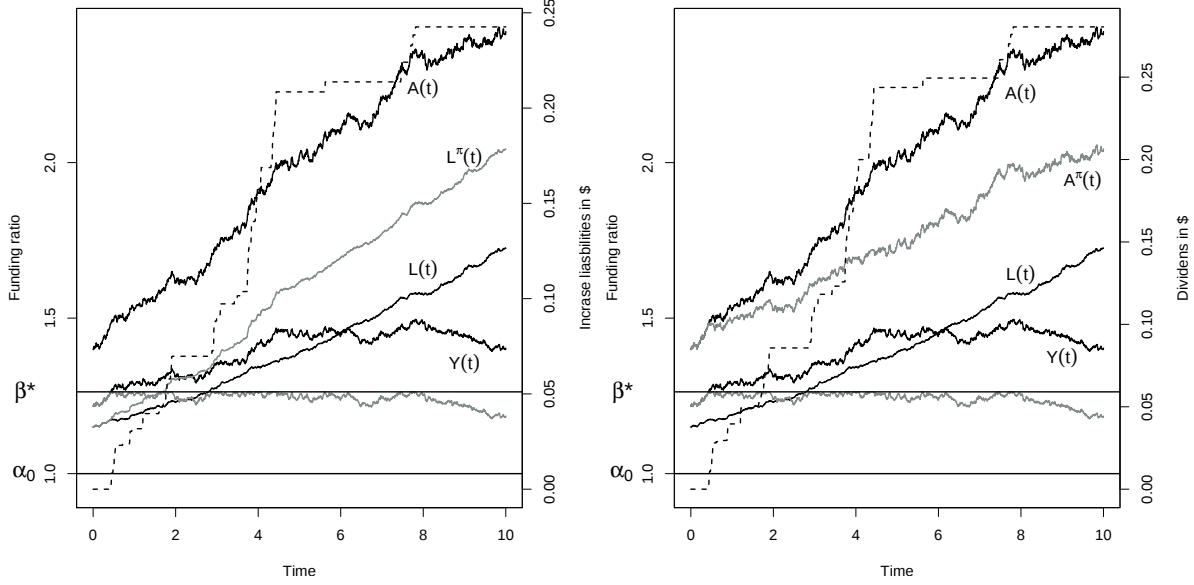


Figure 6.1: Figure illustrating the model. Left: Case A. Right: Case B. The uncontrolled processes are in black and the controlled processes are in grey. The dotted lines depict the undiscounted, aggregated payment processes.

Distributions will either translate into increasing liabilities (Case A) or decreasing assets (Case B). The controlled processes, which we denote $\vec{X}_A^\pi$ and $\vec{X}_B^\pi$, have dynamics

$$d\vec{X}_A^\pi(t) = \mu \left(\vec{X}_A^\pi(t) \right) dt + \sigma \left(\vec{X}_A^\pi(t) \right) dW(t) + \begin{pmatrix} 0 \\ dD_A^\pi(t) \end{pmatrix}$$

and

$$d\vec{X}_B^\pi(t) = \mu \left(\vec{X}_B^\pi(t) \right) dt + \sigma \left(\vec{X}_B^\pi(t) \right) dW(t) - \begin{pmatrix} dD_B^\pi(t) \\ 0 \end{pmatrix},$$

where D_A^π is the payment stream of increases of the liabilities (Case A) and D_B^π is the dividend payment stream. We denote by L^π liabilities after addition of D_A^π and by A^π assets after subtraction of D_B^π . The funding ratios of the controlled processes are then given by

$$Y_A^\pi(t) = \frac{A(t)}{L^\pi(t)} \text{ and } Y_B^\pi(t) = \frac{A^\pi(t)}{L(t)}, \quad t \geq 0,$$

respectively.

Remark 6.1.1. *The model allows for deterministic (multiplicative) increase of the assets and liabilities, plus (correlated) random variations. Because of the nature of the processes, these variations are continuous, and one might argue that abrupt changes in assets and liabilities (jumps) should be allowed in order to reflect the random nature of the businesses, and/or expected changes in scale. Beyond the fact that these would require developments beyond the scope of a single paper, we believe our model is still reasonable for the following reasons:*

1. *Case A: the formulation of our model means that we consider an accumulation scheme in equilibrium, that is, where contributions are continuously offset by payouts. This is an approximation, but we believe it is good enough for our analysis. If significant assets and liabilities were to enter or leave the fund, this typically would lead to a specific procedure and distribution rule (partial liquidation).*
2. *Case B: additional contributions to the company can be made from time to time without affecting the conclusions as long as these are made so as not to make existing shareholders richer or poorer. In terms of our model this means that they would be made at the existing funding ratio. We will see later that a change of scale that does not impact the funding ratio has no impact on our conclusions (how to control the process).*

6.1.4 Bankruptcy and advanced solvency constraints

Of course, the fund may become bankrupt. This will occur as soon as the funding ratio reaches a given level α_0 . For either of the cases A and B we denote by τ_{α_0} the time of ruin, which is the stopping time defined as the first time the funding ratio of the controlled processes equals α_0 . For the rest of the paper we will use the notation $\varpi \in \{A, B\}$ to simplify notation where possible. Using this notation, the bankruptcy time for the two cases is given by

$$\tau_{\alpha_0} = \inf \{t \geq 0 | Y_{\varpi}^{\pi}(t) = \alpha_0\}.$$

For Case A, α_0 could be below or above 1, depending on the nature of the fund (partially funded public or fully funded private, for instance). For Case B, the level α_0 would typically be 1.

We now introduce a simple solvency constraint. Because one might want to restrict admissible distributions so as not to bring the funding ratio too close to α_0 , we set a level $\alpha_1 \geq \alpha_0$ under which distributions cannot bring the funding ratio. For Case A, it is obvious that a regulator would not allow distributions to be made all the way to the bankruptcy level. In this context, α_1 could, for instance, include some statutory reserves dedicated to buffer fluctuations of assets. For B, such a restriction (unless it is exogenously imposed) does not make much sense in terms of pure maximisation of dividends as, if it did, this would be taken into account in the maximisation process. However, pure maximisation of the expected present value of dividends until ruin is hardly a perfect objective, and this mechanism has a significant impact on the stability of the process as illustrated later in Section 6.2.3. Hence, the controller might still want to forego some expected return in favour of some more stability. Of course, α_1 could also be exogenously determined by a regulator. For example, considering a regulated insurance fund, α_0 and α_1 could correspond to the Minimum Capital Requirement and Solvency Capital Requirement, respectively, under the framework of the new European regulatory system Solvency II.

So far we have two areas: between α_0 and α_1 , where no distribution is allowed, and beyond α_1 , where distributions are allowed. Because of diffusion, the fund might fluctuate very often between both areas if it is in the vicinity of α_1 (which is likely to happen in reality, especially if the barrier level is close to α_1). We introduce a third level $\alpha_2 \geq \alpha_1$ with the following mechanism (see also Figure 6.2). When the process is in the region between α_1 and α_2 and coming from α_2 —that is, when the last visit at either α_1 or α_2 was α_2 —distributions are allowed. When the last visit was α_1 we consider that the process is in recovery and no distribution is allowed. This has the following interpretation:

- A. Under A, we assume that the fund is not fully funded if its funding ratio is below α_1 (although it is not severe enough to declare bankruptcy). Hence it is not realistic to assume that a fund that has just recovered and just reached α_1 could immediately start distributing some of its

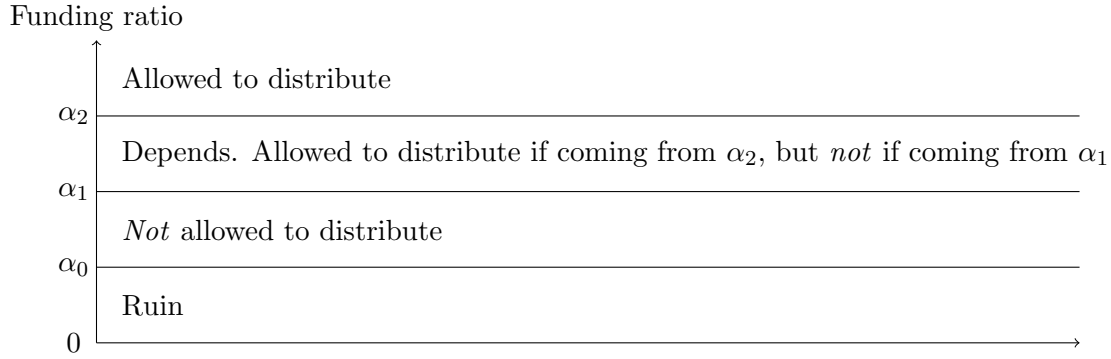


Figure 6.2: Graphical illustration of a model with advanced solvency constraints $\{\alpha_0, \alpha_1, \alpha_2\}$.

excess profits. In other words, $\alpha_2 > \alpha_1$ describes a situation where downcrossing α_1 would trigger some alarm and put the fund in an emergency state under which no distribution is allowed (and perhaps, under which the fund is closely monitored by the regulator). This state would revert back to normal when the process upcrosses the level $\alpha_2 > \alpha_1$ again.

- B. Under B, $\alpha_1 = \alpha_2$ may lead to erratic periods of dividend payments if the barrier is equal to α_1 , which is unrealistic; see also Avanzi and Wong (2012) for a discussion of this. In particular, erratic dividend payments was another criticism formulated in Gerber (1974).

Remark 6.1.2. *The introduction of advanced solvency constraints is in some way similar in spirit to the introduction of “Parisian implementation delays”; see, for instance, Dassios and Wu (2009). In their paper, dividends are not payable immediately when the barrier is crossed, but only if the surplus stays above that barrier for a certain period of time.*

6.1.5 Rescue measures

Distributing profits usually means that ruin is certain. If profits are allowed to be distributed, it makes sense that the reverse operation could also be allowed to prevent ruin, and could indeed be profitable in certain cases. This idea goes back to Borch (1974, Chapter 20) and Porteus (1977).

Here, ruin will be prevented in different ways, depending on the context. In Case A, it will be done by decreasing liabilities. Indeed, when in financial distress, rescue measures in pension funds usually materialise as reductions of benefits, as it may not be possible to ask for additional money from members. Furthermore, additional contributions from members would normally lead to increases of liabilities at the same time, which would improve the funding ratio, but rescue would require a higher amount of money than a reduction in benefits, and may even be unable to bring the ratio to an acceptable level (if the minimum is higher than or equal to 1). In Case B, we will be injecting capital, just as in the first problem considered in Gerber and Shiu (2003), but with an aim at maximizing the difference between dividends and capital injections, rather than matching them.

Lest the problem becomes trivial, rescue measures will attract transaction costs. With proportional transaction costs, it turns out (unsurprisingly) that if such measures are warranted (have positive expected value), these will be made only at level α_0 to avoid bankruptcy. If a more complex transaction cost structure is considered (one that would depend on the target funding ratio, for instance), this may no longer be the case. We discuss this in Section 6.4.3, but a formal treatment of this is outside the scope of this paper.

6.2 With simple solvency constraints

In this section we consider a *simple solvency constraint* setting $\{\alpha_0, \alpha_1, \alpha_2\}$ where $\alpha_1 > \alpha_0$, but where $\alpha_2 = \alpha_1$. This is equivalent to the constraint introduced in Paulsen (2003, in a univariate, pure diffusion setting). In mathematical terms, this constraint translates into the condition

$$\int_0^{\tau_{\alpha_0}} \mathbf{1}_{\{Y_{\varpi}^{\pi}(s) < \alpha_1\}} dD_{\varpi}^{\pi}(s) = 0. \quad (6.2.1)$$

In this framework, we want to determine the optimal control strategy π that maximizes the expected present value of the distribution denoted by

$$J_{\varpi,1}(\vec{x}; \pi) = \mathbb{E}^{\vec{x}} \left[\limsup_{t \rightarrow \infty} \int_0^{t \wedge \tau_{\alpha_0}} e^{-\delta s} dD_{\varpi}^{\pi}(s) \right].$$

where $\mathbb{E}^{\vec{x}}[\cdot]$ is the expected value given $X(0) = \vec{x}$. Note that the subscript “1” indicates that we consider a model with simple solvency constraint. Likewise, we will throughout the paper use the subscript “2” in the model with advanced solvency constraints and the subscript “3” in the model where ruin must be prevented. The optimal strategy π^* fulfills that

$$V_{\varpi,1}(\vec{x}; \pi^*) = \sup_{\pi \in \Pi} J_{\varpi,1}(\vec{x}; \pi), \quad (6.2.2)$$

where Π is the set of admissible strategies. A strategy is said to be admissible if D_{ϖ}^{π} is a non-decreasing, $\mathcal{F}$ -adapted process with $D_{\varpi}^{\pi}(0-) = 0$ and satisfies (6.2.1). Moreover, we assume that D_{ϖ}^{π} has càdlàg sample paths. The solvency restriction (6.2.1) implies, that we do not allow for any liquidation (“take your money and run”) strategy, whereby a single dividend payment could bring the company to bankruptcy.

6.2.1 Value of distributions when liabilities are increased (Case A)

For a given value of $\vec{x}$, we denote by $F(\vec{x})$ the value function for the optimal strategy π^* . We allow for singular control strategies and by standard methods we get that F admits the following HJB equation:

$$0 = \max \left\{ (\mathcal{A} - \delta)F(\vec{x}), 1 + \frac{\partial}{\partial x_2} F(\vec{x}) \right\}, \quad F(\alpha_0 x_2, x_2) = 0, \quad (6.2.3)$$

where

$$\begin{aligned} \mathcal{A}f(\vec{x}) = & \mu_A x_1 \frac{\partial}{\partial x_1} f(\vec{x}) + \mu_L x_2 \frac{\partial}{\partial x_2} f(\vec{x}) + \frac{1}{2} \sigma_A^2 x_1^2 \frac{\partial^2}{\partial x_1^2} f(\vec{x}) \\ & + \frac{1}{2} \sigma_L^2 x_2^2 \frac{\partial^2}{\partial x_2^2} f(\vec{x}) + \rho \sigma_A \sigma_L x_1 x_2 \frac{\partial^2}{\partial x_1 \partial x_2} f(\vec{x}). \end{aligned} \quad (6.2.4)$$

Based on the results in Shreve et al. (1984), Decamps et al. (2009) and Paulsen (2003), a qualified guess is that the optimal strategy for increasing the liabilities is of the barrier type. By barrier, we refer to a constant level of funding ratio. We denote the optimal barrier in an unconstrained model (such as considered in Gerber and Shiu, 2003) by β_0^* and the optimal barrier in the model with (simple) solvency constraint by β_1^* . We use the notation β_0^* and β_1^* without any subscript

A to simplify notation. Besides, we show later that the optimal barrier is the same for cases A and B. We conjecture that the optimal barrier for the assets is given by $\beta_1^* L$, where

$$\beta_1^* = \begin{cases} \beta_0^*, & \beta_0^* \geq \alpha_1, \\ \alpha_1, & \beta_0^* < \alpha_1. \end{cases}$$

With techniques similar to the ones in Decamps et al. (2009) (who also derived β_0^* but had no solvency constraint) we will show that β_1^* is the optimal barrier. To make the paper self contained, we also give a proof for the case $\beta_0^* \geq \alpha_1$ and we make the proof a bit more detailed compared to Decamps et al. (2009).

First, we consider the value function of a barrier strategy with an arbitrary barrier level β . The corresponding value function is denoted by

$$G_A^\beta(\vec{x}) = \mathbb{E}^{\vec{x}} \left[\int_{0-}^{\tau_{\alpha_0}} e^{-\delta s} dC_\beta(s) \right],$$

where C_β is notation for a strategy, where we increase liabilities in order to keep the funding ratio below β . The value of G_A^β is given by

$$G_A^\beta(\vec{x}) = \begin{cases} G_A(\vec{x}; \beta), & x_1 \in [\alpha_0 x_2, \beta x_2], \\ \frac{x_1}{\beta} - x_2 + G_A\left(x_1, \frac{x_1}{\beta}; \beta\right), & x_1 > \beta x_2, \end{cases} \quad (6.2.5)$$

where $G_A(\cdot; \beta)$ is given by the differential equation

$$(\mathcal{A} - \delta)G_A(\vec{x}; \beta) = 0 \text{ for } \alpha_0 \leq \frac{x_1}{x_2} \leq \beta. \quad (6.2.6)$$

The differential equation (6.2.6) can be obtained by the following heuristic reasoning. Let $\frac{x_1}{x_2} \leq \beta \Leftrightarrow x_1 \leq \beta x_2$ and consider an infinitesimal time interval of length dt . We get that

$$G_A(x_1, x_2; \beta) = e^{-\delta dt} \mathbb{E} \left[G_A(x_1 + \mu_A x_1 dt + \sigma_A x_1 W_1(dt), \right. \\ \left. x_2 + \mu_L x_2 dt + \rho \sigma_L x_2 W_1(dt) + \sqrt{1 - \rho^2} \sigma_L x_2 W_2(dt); \beta \right]. \quad (6.2.7)$$

Developing the expectation using Taylor series, subtracting $G_A(x_1, x_2; \beta)$ and dividing by dt on both sides yields

$$G_A(x_1, x_2; \beta) \left(\frac{e^{\delta dt} - 1}{dt} \right) = \mathcal{A}G_A(x_1, x_2; \beta).$$

We let $dt \rightarrow 0$ and using l'Hôpital's rule we obtain (6.2.6).

The boundary conditions for the value function, which hold for all levels of barrier β , are given by

$$G_A(\alpha_0 x_2, x_2; \beta) = 0, \quad (6.2.8)$$

$$\frac{\partial}{\partial x_2} G_A(x_1, x_2; \beta) \Big|_{x_2 = \frac{x_1}{\beta}} = -1. \quad (6.2.9)$$

Moreover, we have the boundary condition

$$\frac{\partial^2}{\partial x_2^2} G_A(x_1, x_2; \beta_0^*) \Big|_{x_2 = \frac{x_1}{\beta_0^*}} = 0, \quad (6.2.10)$$

which only holds for the optimal barrier β_0^* . Condition (6.2.8) follows directly from the definition of ruin. Condition (6.2.9) is similar to (2.6) in Avanzi and Gerber (2008) and can be obtained in a similar way. Lastly, we get condition (6.2.10) by taking the derivative of (6.2.9) with respect to β . This gives us that (using the chain rule for partial derivatives)

$$\begin{aligned} & \frac{\partial}{\partial \beta} \left(\frac{\partial}{\partial x_2} G_A(x_1, x_2; \beta) \Big|_{x_2 = \frac{x_1 -}{\beta}} \right) \\ &= - \frac{(x_2 -)}{\beta^2} \frac{\partial^2}{\partial x_2^2} G_A(x_1, x_2; \beta) \Big|_{x_2 = \frac{x_1 -}{\beta}} + \frac{\partial^2}{\partial x_2 \partial \beta} G_A(x_1, x_2; \beta) \Big|_{x_2 = \frac{x_1 -}{\beta}} = \frac{\partial}{\partial \beta} (-1) = 0. \end{aligned} \quad (6.2.11)$$

Since the term $\frac{\partial^2}{\partial x_2 \partial \beta} G_A(x_1, x_2; \beta) \Big|_{x_2 = \frac{x_1 -}{\beta}}$ in (6.2.11) is 0 for the optimal barrier, we get (6.2.10) by inserting β_0^* and dividing the equation by $(x_2 -)/\beta^2$. To find the solution to equation (6.2.6), we take advantage of the fact that G_A^β (and G_A) are homogeneous functions of degree 1, which follows since the dynamics of both assets and liabilities are linear in the assets and liabilities, respectively. That is, for a given constant ϱ it holds that $G_A^\beta(\varrho x_1, \varrho x_2; \beta) = \varrho G_A^\beta(x_1, x_2; \beta)$ so that the important quantity (up to a scaling factor) is the ratio x_1/x_2 . The solution to the system of equations is given in the following lemma:

Lemma 6.2.1. *The solution to differential equation (6.2.6) with boundary conditions (6.2.8) and (6.2.9) is given by*

$$G_A(x_1, x_2; \beta) = x_2 \frac{\left(\frac{x_1}{\alpha_0 x_2} \right)^{\zeta_1} - \left(\frac{x_1}{\alpha_0 x_2} \right)^{\zeta_2}}{(1 - \zeta_2) \left(\frac{\beta}{\alpha_0} \right)^{\zeta_2} - (1 - \zeta_1) \left(\frac{\beta}{\alpha_0} \right)^{\zeta_1}}, \quad (6.2.12)$$

where

$$\begin{aligned} \tilde{\sigma}^2 &= \sigma_A^2 + \sigma_L^2 - 2\rho\sigma_A\sigma_L, \\ \zeta_1 &= \frac{\frac{1}{2}\tilde{\sigma}^2 - (\mu_A - \mu_L) - \sqrt{\frac{1}{4}\tilde{\sigma}^4 + (\mu_A - \mu_L)^2 - \tilde{\sigma}^2(\mu_A + \mu_L - 2\delta)}}{\tilde{\sigma}^2}, \\ \zeta_2 &= \frac{\frac{1}{2}\tilde{\sigma}^2 - (\mu_A - \mu_L) + \sqrt{\frac{1}{4}\tilde{\sigma}^4 + (\mu_A - \mu_L)^2 - \tilde{\sigma}^2(\mu_A + \mu_L - 2\delta)}}{\tilde{\sigma}^2}. \end{aligned} \quad (6.2.13)$$

Proof: See Appendix 6.A.1.

We see from (6.2.12), that only the denominator depends on β and that both the numerator and the denominator are negative. Furthermore, it is interesting to note that, apart from a scaling factor of x_2 , the function is only expressed in terms of ratios of x_1 to $\alpha_0 x_2$. Hence, the shape of the value function is unaffected by the scale of the two processes.

We now take the derivative of the denominator and set the derivative equal to 0 to find the maximum of the denominator (and the maximum of $G_A(\cdot; \beta)$). The resulting optimal barrier level is

$$\beta_0^* = \alpha_0 \left(\frac{\zeta_2(\zeta_2 - 1)}{\zeta_1(\zeta_1 - 1)} \right)^{\frac{1}{\zeta_1 - \zeta_2}} = \alpha_0 \left(\frac{\zeta_1(\zeta_1 - 1)}{\zeta_2(\zeta_2 - 1)} \right)^{\frac{1}{\zeta_2 - \zeta_1}}. \quad (6.2.14)$$

Because of the assumption that $\mu_A > \mu_L$, we have that $\beta_0^* > \alpha_0$. We obtain this result by using the representation given by (6.A.3) for both the numerator and the denominator of (6.2.14)

$$\frac{\zeta_1(\zeta_1 - 1)}{\zeta_2(\zeta_2 - 1)} = \frac{\delta - \mu_L - (\mu_A - \mu_L)\zeta_1}{\delta - \mu_L - (\mu_A - \mu_L)\zeta_2} > 1, \quad (6.2.15)$$

and using the fact that $\frac{1}{\zeta_2 - \zeta_1} > 0$; see (6.2.13).

We note that the optimal barrier β_0^* exists and is unique. This is a result of the following argument: We let f denote the denominator of (6.2.12) as a function of β . Using the assumptions (6.1.2) and (6.1.3) and the result (6.2.15), we get that $f'(\alpha_0) > 0$ and $f''(\beta) < 0$ for all $\beta > \alpha_0$. Thus we know the optimal barrier exists and is unique. Also, note that for $\delta = \mu_A$ we get that $\zeta_2 = 1$ which implies that $\beta_0^* \rightarrow \infty$ for $\mu_A \rightarrow \delta$ (if we become very patient (δ is getting closer to μ_A) it is optimal to wait more before distributing profits) and for $\mu_A \leq \mu_L$ we get $\beta_0^* = \alpha_0$ (if the company is not profitable we should liquidate it immediately).

In order for the candidate solution to be useful, we need it to fulfill the HJB equation. That is the case, as stated in the following lemma:

Lemma 6.2.2. *The candidate solution $G_A^{\beta_0^*}$ given in (6.2.5) with barrier β_0^* given by (6.2.14) fulfills all the conditions of the HJB equation (6.2.3) and the partial derivatives of $G_A^{\beta_0^*}$ are bounded.*

Proof: See appendix 6.A.2.

Next, the verification lemma:

Lemma 6.2.3. *If a non-negative function $G \in C^1(\mathbb{R}_+)$ is also twice continuously differentiable except at countably many points and satisfies for $x \geq 0$ that*

$$(\mathcal{A} - \delta)G(\vec{x}) \leq 0, \quad (6.2.16)$$

$$\frac{\partial^2}{\partial x_1^2} G(\vec{x}) \leq 0, \quad (6.2.17)$$

$$K_1 \geq \frac{\partial}{\partial x_1} G(\vec{x}) \geq H_1, \quad (6.2.18)$$

$$\frac{\partial^2}{\partial x_2^2} G(\vec{x}) \leq 0, \quad (6.2.19)$$

$$K_2 \leq \frac{\partial}{\partial x_2} G(\vec{x}) \leq -1 \quad (6.2.20)$$

for some finite constants K_1 , H_1 and K_2 , then

$$G(\vec{x}) \geq V_{A,1}(\vec{x}; \pi^*).$$

Moreover, if there exists a point $\beta^* \in \mathbb{R}_+$ such that $G \in C^1(\mathbb{R}_+) \cap C^2(\mathbb{R}_+ \setminus \{\beta^*\})$ satisfies

$$(\mathcal{A} - \delta)G(\vec{x}) = 0 \text{ and } \frac{\partial}{\partial x_2} G(\vec{x}) \leq -1 \text{ for } x_2 \in \left[\frac{x_1}{\beta^*}, \frac{x_1}{\alpha_0} \right], \quad (6.2.21)$$

$$(\mathcal{A} - \delta)G(\vec{x}) < 0 \text{ and } G(\vec{x}) = \frac{x_1}{\beta^*} - x_2 + G\left(x_1, \frac{x_1}{\beta^*}\right) \text{ for } x_2 \in \left(0, \frac{x_1}{\beta^*}\right), \quad (6.2.22)$$

then it holds that

$$G(\vec{x}) = V_{A,1}(\vec{x}; \pi^*),$$

and for the case $\beta^* \geq \alpha_1$, the optimal strategy is a barrier strategy with barrier β^* , whereas for the case $\beta^* < \alpha_1$, the optimal strategy is a barrier strategy with barrier α_1 .

Proof: See appendix 6.A.3.

The proof of Lemma 6.2.3 extends that of Decamps et al. (2009) to include solvency constraints and an explicit proof of concavity. The solvency constraint requires a second part of proof, where we use some subresults of the first part of the proof.

6.2.2 Value of distributions when assets are decreased (Case B)

In this section we consider the optimal dividend strategy under a simple solvency constraint. The proofs of the results are similar to the ones in Section 6.2.1 we therefore omit these. The structure of the section resembles the one of Section 6.2.1, and we state the results without a long introduction.

The HJB equation is given by

$$0 = \max \left\{ (\mathcal{A} - \delta)F(\vec{x}), 1 - \frac{\partial}{\partial x_1}F(\vec{x}) \right\}, \quad F(\alpha_0 x_2, x_2) = 0, \quad (6.2.23)$$

where $\mathcal{A}$ is given by (6.2.4). The value function for a barrier strategy is then given by

$$G_B^\beta(\vec{x}) = \begin{cases} G_B(\vec{x}; \beta), & x_1 \in [\alpha_0 x_2, \beta x_2], \\ x_1 - \beta x_2 + G_B(\beta x_2, x_2; \beta), & x_1 > \beta x_2. \end{cases} \quad (6.2.24)$$

The boundary conditions for the value function, which hold for all levels of barrier β , are given by

$$G_B(\alpha_0 x_2, x_2; \beta) = 0, \quad (6.2.25)$$

$$\frac{\partial}{\partial x_1} G_B(x_1, x_2; \beta) \big|_{x_1 = \beta x_2 -} = 1. \quad (6.2.26)$$

Moreover we have the boundary condition

$$\frac{\partial^2}{\partial x_1^2} G_B(x_1, x_2; \beta_0^*) \big|_{x_1 = \beta_0^* x_2 -} = 0, \quad (6.2.27)$$

which only holds for the optimal barrier β_0^* .

The resulting value function is

$$G_B(x_1, x_2; \beta) = \beta x_2 \frac{\left(\frac{x_1}{\alpha_0 x_2}\right)^{\zeta_1} - \left(\frac{x_1}{\alpha_0 x_2}\right)^{\zeta_2}}{\zeta_1 \left(\frac{\beta}{\alpha_0}\right)^{\zeta_1} - \zeta_2 \left(\frac{\beta}{\alpha_0}\right)^{\zeta_2}}, \quad (6.2.28)$$

where ζ_1 and ζ_2 are given by (6.2.13). Note that for $x_2 = 1$ this result is also given by Gerber and Shiu (2003, equation (9.6)).

The optimal barrier is the same as the one in Case A, such that β_0^* is given by (6.2.14). As in Case A, one can show that the optimal barrier exists and is unique.

The verification lemma for Case B is given below:

Lemma 6.2.4. *If a non-negative function $G \in C^1(\mathbb{R}_+)$ is also twice continuously differentiable except at countably many points and satisfies for $x \geq 0$ that*

$$(\mathcal{A} - \delta)G(\vec{x}) \leq 0, \quad (6.2.29)$$

$$\frac{\partial^2}{\partial x_1^2} G(\vec{x}) \leq 0, \quad (6.2.30)$$

$$K_1 \geq \frac{\partial}{\partial x_1} G(\vec{x}) \geq 1, \quad (6.2.31)$$

$$\frac{\partial^2}{\partial x_2^2} G(\vec{x}) \leq 0, \quad (6.2.32)$$

$$K_2 \leq \frac{\partial}{\partial x_2} G(\vec{x}) \leq H_2 \quad (6.2.33)$$

for some finite constants K_1 , K_2 and H_2 then

$$G(\vec{x}) \geq V_{B,1}(\vec{x}; \pi^*).$$

Moreover, if there exists a point $\beta^* \in \mathbb{R}_+$ such that $G \in C^1(\mathbb{R}_+) \cap C^2(\mathbb{R}_+ \setminus \{\beta^*\})$ with

$$(\mathcal{A} - \delta)G(\vec{x}) = 0 \text{ and } \frac{\partial}{\partial x_1}G(\vec{x}) \geq 1 \text{ for } x_1 \in [\alpha_0 x_2, \beta^* x_2], \quad (6.2.34)$$

$$(\mathcal{A} - \delta)G(\vec{x}) < 0 \text{ and } G(\vec{x}) = x_1 - \beta^* x_2 + G(\beta^* x_2, x_2) \text{ for } x_1 \in (\beta^* x_2, \infty), \quad (6.2.35)$$

then it holds that

$$G(\vec{x}) = V_{B,1}(\vec{x}; \pi^*),$$

and for the case $\beta^* \geq \alpha_1$, the optimal strategy is a barrier strategy with barrier β^* , whereas for the case $\beta^* < \alpha_1$, the optimal strategy is a barrier strategy with barrier α_1 .

The structure of the proof is the same as in the proof of Lemma 6.2.3 though the calculations and comparisons of the different terms are somewhat different.

6.2.3 The impact of the simple solvency constraint

In this section, we consider the impact of the simple solvency constraint on the stability of operations. In Figure 6.3, we compare outcomes of 10,000 simulations for the aggregate distributed amount in cases A (left) and B (right), as well as the time to bankruptcy (middle—note bankruptcy times are the same for both cases A and B since the funding ratio for the two cases are the same), when a simple solvency constraint is applied (horizontal axis) or not (vertical axis). In all three scatterplots, values in the top left triangle are those where the outcome in absence of constraints beats that in presence of a simple solvency constraint, whereas outcomes in the bottom right triangle are those where the constraints beats the base case. In terms of dividends, we know that the absence of constraints will lead to a higher expected present value—*on average*. What the three graphs in Figure 6.3 teach us is that, when there is a substantial difference in terms of distributions, it is in favour of the solvency constraint (especially in Case A). Also, the coefficient of variation of the average aggregate distribution amount is in both cases lower with the solvency constraint than without. In terms of solvency itself, the case with solvency constraint clearly dominates the one without. Therefore, regulators could argue that the solvency constraint in this case is really effective and comes at a relatively small cost.

Note that the bankruptcy times are the same for both cases A and B since the funding ratio for the two cases are the same. The plots are for 10.000 simulations performed according to a simple Euler scheme. The simulations were censored at time $T = 15.000$ (unless declared bankrupt before). For the following mean values and variances, the first element in the vector corresponds to the data on the first axis and the second element corresponds to the data on the second axis.

Left plot: Mean = (0.77395, 0.77806), Variance = (0.04563, 0.04709).

Middle plot: Mean = (5040.92724, 2420.87073), Variance = (4348.931², 2403.101²).

Right plot: Mean = (1.81492, 1.84436), Variance = (0.74447, 0.81895).

While these results would be quantitatively different with other parameters, conclusions would be qualitatively similar.

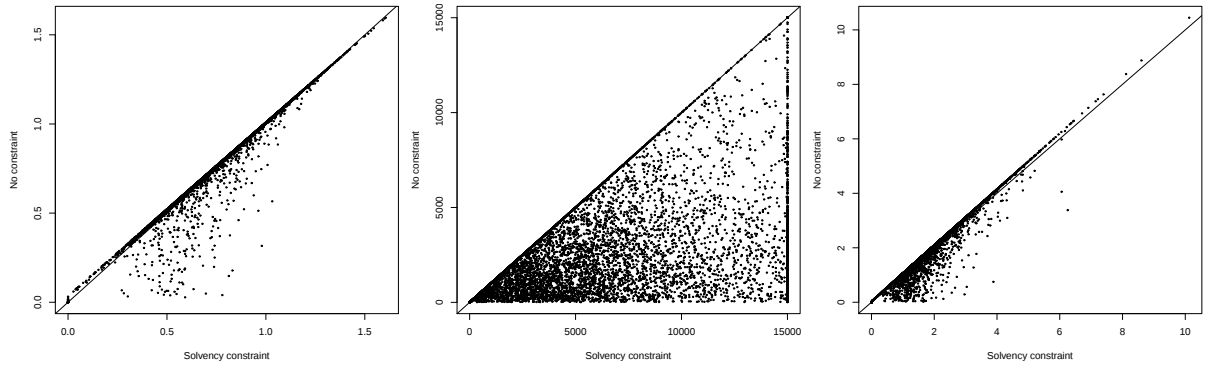


Figure 6.3: Scatterplots for no solvency constraints and simple solvency constraint, respectively. The left figure shows the distributions for Case A, the middle figure shows ruin times (which are the same for both Case A and Case B), and the right figure shows dividends for Case B. For parameters used, see Table 6.1 in Appendix 6.B, set no. 1.

6.3 Advanced solvency constraints

In this section, we consider the advanced type of solvency constraint with barriers $\{\alpha_0, \alpha_1, \alpha_2\}$ introduced and described in Section 6.1.4; see in particular Figure 6.2. Here, we are allowed to control the process when the funding ratio is in the interval (α_1, α_2) if and only if the funding ratio last crossed α_2 , rather than α_1 . Note, that this model is a generalization of that developed in the previous section. This follows, since this model simplifies to the model in Section 6.2 for $\alpha_2 = \alpha_1$. The main challenges in the current model are how to formulate the constraint mathematically, how to obtain a value function explicitly, and how to determine the optimal barrier level.

We start by defining a set of stopping times $\tau_n^\pi, n \in \mathbb{N}_0$ given by

$$\begin{aligned}\tau_{\varpi,0}^\pi &= \inf\{t \geq 0 : Y_\varpi^\pi(t) \leq \alpha_1 \vee Y_\varpi^\pi(t) \geq \alpha_2\}, \\ \tau_{\varpi,i}^\pi &= \inf\{t > \tau_{\varpi,i-1}^\pi : (Y_\varpi^\pi(t) \leq \alpha_1 \wedge Y_\varpi^\pi(\tau_{\varpi,i-1}^\pi) \geq \alpha_2) \vee (Y_\varpi^\pi(t) \geq \alpha_2 \wedge Y_\varpi^\pi(\tau_{\varpi,i-1}^\pi) \leq \alpha_1)\}, \\ i &= 1, 2, \dots\end{aligned}$$

Now that we are equipped with those stopping times, we define a 0-1 process ϕ_ϖ such that

$$\phi_\varpi(t) = \mathbf{1}_{\{t < \tau_{\varpi,0}^\pi \cup Y_\varpi^\pi(t) \geq \alpha_1\}} + \sum_{i=1}^{\infty} \left(\mathbf{1}_{\{t \in [\tau_{\varpi,i-1}^\pi, \tau_{\varpi,i}^\pi]\}} \mathbf{1}_{\{Y_\varpi^\pi(\tau_{\varpi,i-1}^\pi) \geq \alpha_2\}} \right). \quad (6.3.1)$$

The process ϕ_ϖ is 1 when we are allowed to control the funding ratio and 0, when we are not. That is, we can formulate the new solvency constraint as

$$\int_0^{\tau_{\alpha_0}} (1 - \phi_\varpi(s)) dD_\varpi^\pi(s) = 0. \quad (6.3.2)$$

Due to the form of the function ϕ_ϖ given by (6.3.1) we assume that distributions are initially allowed if $\alpha_1 < Y_\varpi^\pi(0) < \alpha_2$ (until the first stopping time $\tau_{\varpi,0}^\pi$).

As in the former section, we denote the optimal barrier in the model without any solvency constraints by β_0^* , and we denote the optimal barrier in the model with advanced solvency constraints by β_2^* . In order to conjecture an optimal barrier strategy, we consider different possibilities for the relationship between β_0^* and $\{\alpha_0, \alpha_1, \alpha_2\}$. We know that $\beta_0^* > \alpha_0$, so we only need to consider the following cases:

Case 1: $\alpha_0 < \beta_0^* < \alpha_1$

Case 2: $\alpha_1 \leq \beta_0^* < \alpha_2$

Case 3: $\beta_0^* \geq \alpha_2$

We want to consider each of the three cases separately:

1. The conjecture is *not* that you distribute as much as you are allowed to (down to α_1). Instead the conjecture is, that the optimal strategy is a barrier strategy with level $\Lambda > \alpha_1$ (strictly higher), which enables you to keep paying dividends for some time. If the barrier was α_1 we would lose any opportunity to pay dividends until we reach α_2 again.
2. Following the lines of the previous point, we conjecture that the optimal strategy is a barrier strategy with barrier $\Lambda \geq \beta_0^*$. The supplement $\Lambda - \beta_0^*$ would be larger as β_0^* is close to α_1 . This leads to a mix of lump sum distributions of size $\left(\frac{1}{\Lambda} - \frac{1}{\alpha_2}\right) A$ for Case A and $(\alpha_2 - \Lambda)L$ for Case B (when ϕ switches from 0 to 1) and payments at the barrier according to the oscillation of the Brownian motion (when the process ϕ is in state 1). This control can be seen both as alternating between a non-singular and an impulse control, or it can be seen as a singular control.
3. In this case the solvency constraint is no constraint at all in terms of optimal dividend strategies, and we get the same result as in the unconstrained case. That is, $\beta_2^* = \beta_0^*$.

An example of a sample path for “Case 2” is found in Figure 6.4 where β_2^* is the optimal barrier (which will be determined later). The dotted lines show the distributions in Case A and Case B, respectively. The grey parts of the funding ratio process illustrates time spans where you are *not* allowed to pay out dividends ($\phi = 0$) and the black parts of the line illustrates time spans where you are allowed to pay out dividends ($\phi = 1$). Note that $\beta_0^* = 1.3722$. We observe that distributions consist of infinitesimal payments at the barrier β_2^* (when $\phi = 1$) and lump sum payments of size $\left(\frac{1}{\beta_2^*} - \frac{1}{\alpha_2}\right) A$ (Case A) and $(\alpha_2 - \beta_2^*)L$ (Case B), when the funding ratio hits α_2 .

The objective here is the same as in Section 6.2, where we considered a simple solvency constrain. The only difference is, that we now consider the solvency constraint (6.3.2) instead of the solvency constraint (6.2.1).

In what follows, we conjecture that the optimal strategy is as described here, and determine the associated value function and optimal barrier level β_2^* .

6.3.1 Value of distributions when liabilities are increased (Case A)

We assume that $\alpha_1 \leq \beta$, since a control strategy within the present solvency regime is not well defined/allowed for $\beta < \alpha_1$. We denote by $V_{A,2}^\beta(\vec{x})$ the value function for a barrier strategy with barrier β for the funding ratio. By a barrier strategy we mean exactly the same as described in Section 6.2. The value function is then given by

$$V_{A,2}^\beta(\vec{x}) = \begin{cases} V_{A,2}^0(\vec{x}; \beta), & x_1 \in [\alpha_0 x_2, \alpha_2 x_2] \wedge \phi = 0, \\ V_{A,2}^1(\vec{x}; \beta), & x_1 \in [\alpha_1 x_2, \beta x_2] \wedge \phi = 1, \\ \frac{x_1}{\beta} - x_2 + V_{A,2}^1\left(x_1, \frac{x_1}{\beta}; \beta\right), & x_1 > \beta x_2 \wedge \phi = 1, \end{cases} \quad (6.3.3)$$

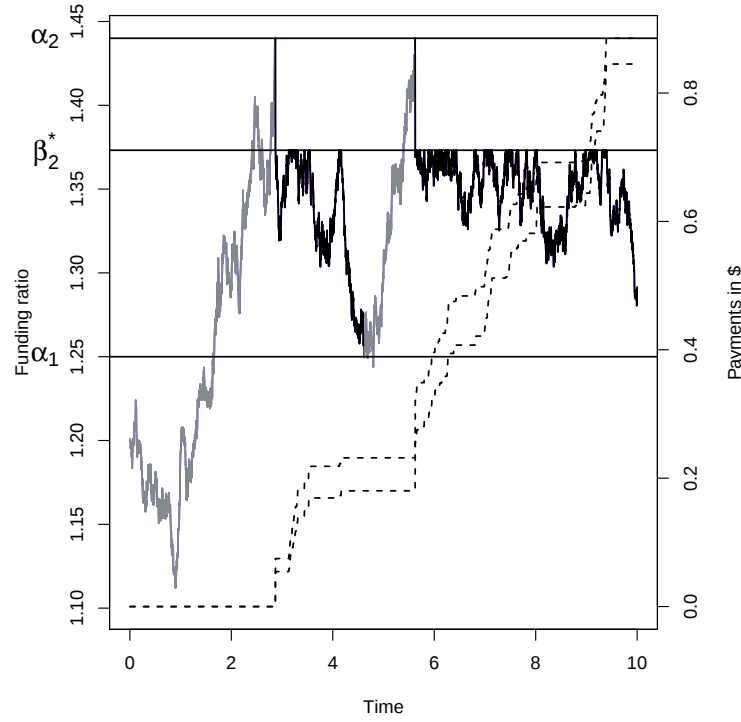


Figure 6.4: Illustration of the funding ratio and distributions in case of advanced solvency constraints.

where $V_{A,2}^0(\cdot; \beta)$ and $V_{A,2}^1(\cdot; \beta)$ fulfill the following systems of PDEs:

$$\begin{aligned}
 (\mathcal{A} - \delta)V_{A,2}^0(\vec{x}; \beta) &= 0 \text{ for } \alpha_0 \leq \frac{x_1}{x_2} \leq \alpha_2, & V_{A,2}^0(\alpha_0 x_2, x_2; \beta) &= 0, \\
 V_{A,2}^0\left(x_1, \frac{x_1}{\alpha_2}; \beta\right) &= V_{A,2}^1\left(x_1, \frac{x_1}{\alpha_2}; \beta\right), \\
 (\mathcal{A} - \delta)V_{A,2}^1(\vec{x}; \beta) &= 0 \text{ for } \alpha_1 \leq \frac{x_1}{x_2} \leq \beta, & V_{A,2}^1\left(x_1, \frac{x_1}{\alpha_1}; \beta\right) &= V_{A,2}^0\left(x_1, \frac{x_1}{\alpha_1}; \beta\right), \\
 \frac{\partial}{\partial x_2} V_{A,2}^1(x_1, x_2; \beta)|_{x_2=\frac{x_1}{\beta}} &= -1.
 \end{aligned} \tag{6.3.4}$$

Note that for $x_2 \leq \frac{x_1}{\alpha_2}$ we automatically have that $\phi = 1$. The notation $V_{A,2}^0$ highlights that this is the value function for $\phi = 0$ and the notation $V_{A,2}^1$ highlights that this is the value function for $\phi = 1$. Moreover, for the optimal level of the barrier, β_2^* , we have the smooth fit condition that

$$\frac{\partial^2}{\partial x_2^2} V_{A,2}^1(x_1, x_2; \beta_2^*)|_{x_2=\frac{x_1}{\beta_2^*}} = 0. \tag{6.3.5}$$

The boundary conditions $V_{A,2}^1(\alpha_1) = V_{A,2}^0(\alpha_1)$ and $V_{A,2}^0(\alpha_2) = V_{A,2}^1(\alpha_2)$ are due to the continuity of the diffusion term. Since $V_{A,2}^0$ and $V_{A,2}^1$ depend on each other, the system of PDEs needs to be solved simultaneously. A graphical representation of the value function for a given barrier β is illustrated in Figure 6.5. Note that in the figure we have omitted some arguments for clarity. The left rectangle illustrates the domain of $V_{A,2}^0$ and the right “open” rectangle illustrates the domain of $V_{A,2}^1$.

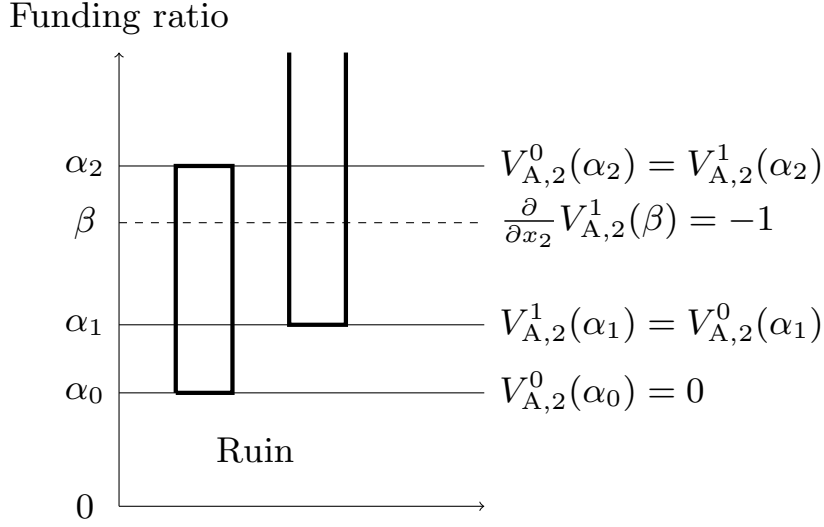


Figure 6.5: Illustration of the domains and boundary conditions in the case of advanced solvency constraints.

We know from Section 6.2, that $V_{A,2}^i$, $i = 0, 1$, are given by

$$V_{A,2}^i(x_1, x_2; \beta) = K_{1,i} x_1^{\zeta_1} x_2^{1-\zeta_1} + K_{2,i} x_1^{\zeta_2} x_2^{1-\zeta_2}, \quad (6.3.6)$$

where ζ_1 and ζ_2 are given by (6.2.13) and $K_{1,i}$ and $K_{2,i}$ are some constants, that fulfills the equation $(\mathcal{A} - \delta)V_{A,2}^i = 0$. That is, including the boundary conditions, we get that the value function is specified by

$$V_{A,2}^0(x_1, x_2; \beta) = C_1 x_1^{\zeta_1} x_2^{1-\zeta_1} + C_2 x_1^{\zeta_2} x_2^{1-\zeta_2}, \quad (x_1 \in [\alpha_0 x_2, \alpha_2 x_2]), \quad (6.3.7)$$

$$V_{A,2}^1(x_1, x_2; \beta) = \tilde{C}_1 x_1^{\zeta_1} x_2^{1-\zeta_1} + \tilde{C}_2 x_1^{\zeta_2} x_2^{1-\zeta_2}, \quad (x_1 \in [\alpha_1 x_2, \beta x_2]), \quad (6.3.8)$$

$$V_{A,2}^0(\alpha_0 x_2, x_2; \beta) = 0, \quad (6.3.9)$$

$$V_{A,2}^0\left(x_1, \frac{x_1}{\alpha_2}; \beta\right) = V_{A,2}^1\left(x_1, \frac{x_1}{\alpha_2}; \beta\right), \quad (6.3.10)$$

$$V_{A,2}^1\left(x_1, \frac{x_1}{\alpha_1}; \beta\right) = V_{A,2}^0\left(x_1, \frac{x_1}{\alpha_1}; \beta\right), \quad (6.3.11)$$

$$\frac{\partial}{\partial x_2} V_{A,2}^1(x_1, x_2; \beta)|_{x_2=\frac{x_1}{\beta}} = -1, \quad z \geq \beta. \quad (6.3.12)$$

The solution to this system is provided in Theorem 6.3.1.

Theorem 6.3.1. *The solutions $V_{A,2}^0$ and $V_{A,2}^1$ to the system of equations given by (6.3.7)-(6.3.12) are given by*

$$\begin{aligned} V_{A,2}^0(x_1, x_2; \beta) &= C_1 \left(x_1^{\zeta_1} x_2^{1-\zeta_1} - \alpha_0^{\zeta_1-\zeta_2} x_1^{\zeta_2} x_2^{1-\zeta_2} \right) \text{ and} \\ V_{A,2}^1(x_1, x_2; \beta) &= \left(C_1 \left(1 - \alpha_0^{\zeta_1-\zeta_2} \alpha_1^{\zeta_2-\zeta_1} \right) - \tilde{C}_2 \alpha_1^{\zeta_2-\zeta_1} \right) x_1^{\zeta_1} x_2^{1-\zeta_1} + \tilde{C}_2 x_1^{\zeta_2} x_2^{1-\zeta_2}, \end{aligned}$$

where

$$\begin{aligned}\xi &= \frac{\min(\beta, \alpha_2)^{\zeta_2-1} - \alpha_1^{\zeta_2-\zeta_1} \min(\beta, \alpha_2)^{\zeta_1-1}}{(1-\zeta_2)\beta^{\zeta_2} - (1-\zeta_1)\alpha_1^{\zeta_2-\zeta_1}\beta^{\zeta_1}}, \\ C_1 &= \frac{\left(\frac{1}{\beta} - \frac{1}{\alpha_2}\right)^+ - \xi}{\alpha_2^{\zeta_1-1} - \alpha_0^{\zeta_1-\zeta_2}\alpha_2^{\zeta_2-1} + (\alpha_0^{\zeta_1-\zeta_2}\alpha_1^{\zeta_2-\zeta_1} - 1) \min(\beta, \alpha_2)^{\zeta_1-1} + \xi(1-\zeta_1)(1-\alpha_0^{\zeta_1-\zeta_2}\alpha_1^{\zeta_2-\zeta_1})\beta^{\zeta_1}}, \\ \tilde{C}_2 &= \frac{-1 - (1-\zeta_1)(C_1 - C_1\alpha_0^{\zeta_1-\zeta_2}\alpha_1^{\zeta_2-\zeta_1})\beta^{\zeta_1}}{(1-\zeta_2)\beta^{\zeta_2} - (1-\zeta_1)\alpha_1^{\zeta_2-\zeta_1}\beta^{\zeta_1}}.\end{aligned}$$

Proof: See appendix 6.A.4.

The optimal barrier β_2^* is obtained by maximizing the value function with respect to β . As for β_0^* and β_1^* we omit the subscript A for β_2^* to simplify notation, because the optimal barrier should, again, be the same for Case A and Case B. The fact that the funding ratio after distribution is the same in both cases supports this claim, but to prove this formally is surprisingly challenging. However, based on numerical studies, β_2^* is indeed the same in both cases. Furthermore, the optimal barrier seems to behave nicely and we did not encounter any problems with existence or uniqueness.

6.3.2 Value of distributions when assets are decreased (Case B)

For Case B, the value function is given by

$$V_{B,2}^\beta(\vec{x}) = \begin{cases} V_{B,2}^0(\vec{x}; \beta), & x_1 \in [\alpha_0 x_2, \alpha_2 x_2] \wedge \phi = 0, \\ V_{B,2}^1(\vec{x}; \beta), & x_1 \in [\alpha_1 x_2, \beta x_2] \wedge \phi = 1, \\ x_1 - \beta x_2 + V_{B,2}^1(\beta x_2, x_2; \beta), & x_1 > \beta x_2 \wedge \phi = 1, \end{cases} \quad (6.3.13)$$

where $V_{B,2}^0$ and $V_{B,2}^1$ are given by the following specification:

$$V_{B,2}^0(x_1, x_2; \beta) = C_1 x_1^{\zeta_1} x_2^{1-\zeta_1} + C_2 x_1^{\zeta_2} x_2^{1-\zeta_2}, \quad (x_1 \in [\alpha_0 x_2, \alpha_2 x_2]), \quad (6.3.14)$$

$$V_{B,2}^1(x_1, x_2; \beta) = \tilde{C}_1 x_1^{\zeta_1} x_2^{1-\zeta_1} + \tilde{C}_2 x_1^{\zeta_2} x_2^{1-\zeta_2}, \quad (x_1 \in [\alpha_1 x_2, \beta x_2]), \quad (6.3.15)$$

$$V_{B,2}^0(\alpha_0 x_2, x_2; \beta) = 0, \quad (6.3.16)$$

$$V_{B,2}^0(\alpha_2 x_2, x_2; \beta) = V_{B,2}^1(\alpha_2 x_2, x_2; \beta), \quad (6.3.17)$$

$$V_{B,2}^1(\alpha_1 x_2, x_2; \beta) = V_{B,2}^0(\alpha_1 x_2, x_2; \beta), \quad (6.3.18)$$

$$\frac{\partial}{\partial x_1} V_{B,2}^1(x_1, x_2; \beta)|_{x_1=zx_2} = 1, \quad z \geq \beta. \quad (6.3.19)$$

Theorem 6.3.2. *The solutions $V_{B,2}^0$ and $V_{B,2}^1$ to the system of equations given by (6.3.14)-(6.3.19) are given by*

$$\begin{aligned}V_{B,2}^0(x_1, x_2; \beta) &= C_1 \left(x_1^{\zeta_1} x_2^{1-\zeta_1} - \alpha_0^{\zeta_1-\zeta_2} x_1^{\zeta_2} x_2^{1-\zeta_2} \right) \text{ and} \\ V_{B,2}^1(x_1, x_2; \beta) &= \left(C_1 \left(1 - \alpha_0^{\zeta_1-\zeta_2} \alpha_1^{\zeta_2-\zeta_1} \right) - \tilde{C}_2 \alpha_1^{\zeta_2-\zeta_1} \right) x_1^{\zeta_1} x_2^{1-\zeta_1} + \tilde{C}_2 x_1^{\zeta_2} x_2^{1-\zeta_2},\end{aligned}$$

where

$$\xi = \frac{\min(\beta, \alpha_2)^{\zeta_2} - \alpha_1^{\zeta_2 - \zeta_1} \min(\beta, \alpha_2)^{\zeta_1}}{\zeta_2 \beta^{\zeta_2 - 1} - \zeta_1 \alpha_1^{\zeta_2 - \zeta_1} \beta^{\zeta_1 - 1}},$$

$$C_1 = \frac{(\alpha_2 - \beta)^+ + \xi}{\alpha_2^{\zeta_1} - \alpha_0^{\zeta_1 - \zeta_2} \alpha_2^{\zeta_2} + (\alpha_0^{\zeta_1 - \zeta_2} \alpha_1^{\zeta_2 - \zeta_1} - 1) \min(\beta, \alpha_2)^{\zeta_1} + \xi (1 - \alpha_0^{\zeta_1 - \zeta_2} \alpha_1^{\zeta_2 - \zeta_1}) \zeta_1 \beta^{\zeta_1 - 1}},$$

$$\tilde{C}_2 = \frac{1 - \zeta_1 (C_1 - C_1 \alpha_0^{\zeta_1 - \zeta_2} \alpha_1^{\zeta_2 - \zeta_1}) \beta^{\zeta_1 - 1}}{\zeta_2 \beta^{\zeta_2 - 1} - \zeta_1 \alpha_1^{\zeta_2 - \zeta_1} \beta^{\zeta_1 - 1}}.$$

As in Case A, it is easy to obtain the optimal barrier by maximizing the value function with respect to β .

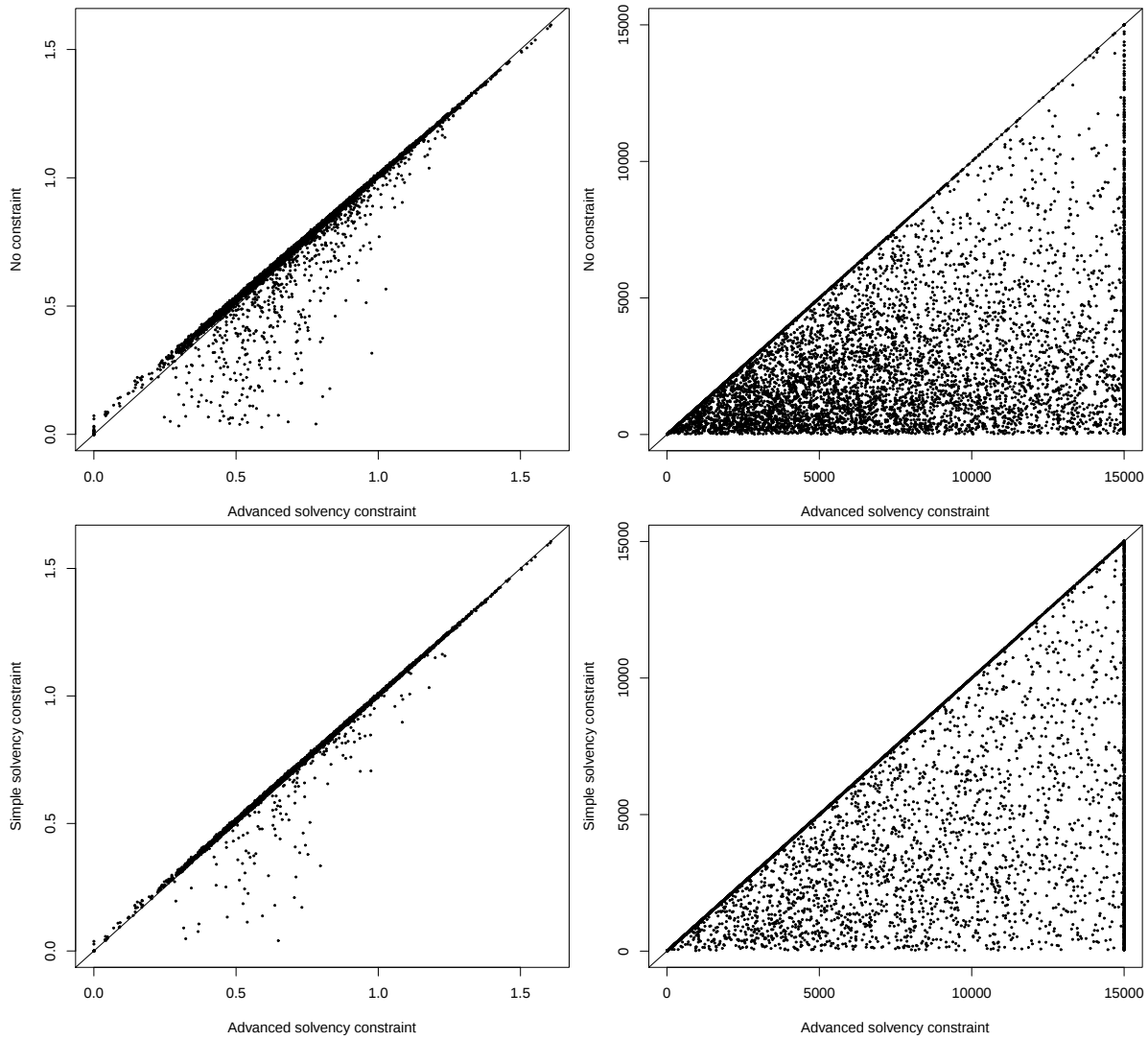


Figure 6.6: Scatterplots for unconstrained and advanced solvency constraints (above) and for simple and advanced solvency constraint (below), respectively. For parameters used, see Table 6.1 in Appendix 6.B, set no. 1.

6.3.3 Numerical studies

In this section, we illustrate numerically the impact of imposing advanced solvency constraints.

6.3.3.1 Moving from simple to advanced solvency constraints

We begin by extending our discussion started in Section 6.2.3. Figure 6.6 shows similar results as the ones in Figure 6.3, but consists of comparisons between no solvency constraint and advanced solvency constraints, and simple solvency constraints versus advanced solvency constraints. The left plots show the aggregate distributions, whereas the right plots show the ruin times. Here, we see that advanced solvency constraints lead to significant improvements wrt. ruin times compared to the simple constraint, of a nature that is qualitatively the same as what was discussed in Section 6.2.3. In other words, there is an additional, substantial difference between advanced and simple solvency constraints.

The plots are for 10.000 simulations and a censoring time $T = 15.000$. For the following vectors of mean values and variances, the first element in the vector corresponds to the data on the first axis and the second element corresponds to the data on the second axis.

Left plot above: Mean = (0.76932, 0.77806), Variance = (0.04582, 0.04709).

Right plot above: Mean = (6903.53093, 2420.87073), Variance = (5132.906², 2403.101²).

Left plot below: Mean = (0.76932, 0.77395), Variance = (0.04582, 0.04563).

Right plot below: Mean = (6903.53093, 5040.92724), Variance = (5132.906², 4348.931²) .

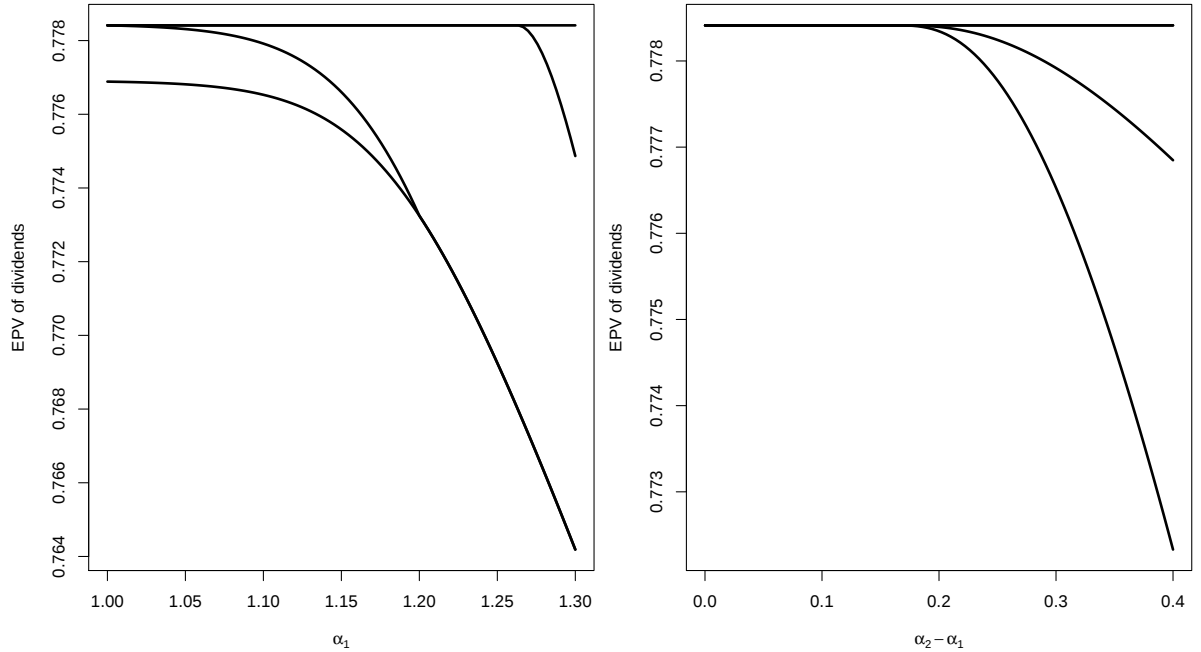


Figure 6.7: Value functions for Case B for the unconstrained case, simple solvency constraint and advanced solvency constraint ($\phi \in \{0, 1\}$) as a function of α_1 and α_2 , respectively. For parameters used, see Table 6.1 in Appendix 6.B, set no. 1.

Next, we investigate whether most of the differences between simple and advanced solvency constraints occur when α_2 moves a bit away from α_1 , or whether they occur slowly as α_2 moves away from α_1 . We also investigate this as α_1 moves away from α_0 in the simple framework. This

is illustrated in Figure 6.7 for case B. We see that small spacing between the α 's have marginal impact initially, even though they can have a large impact on stability as discussed above.

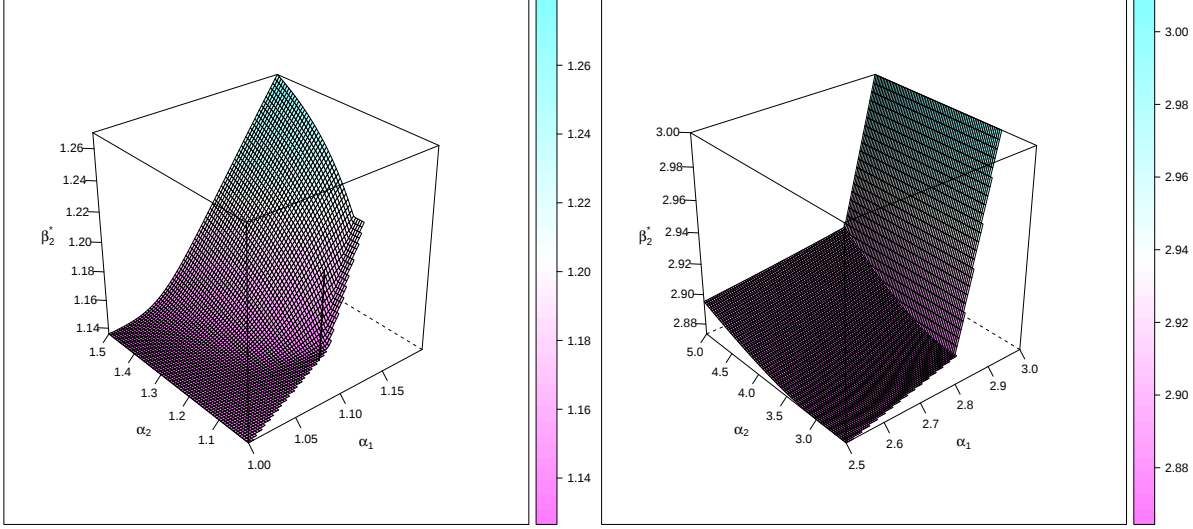


Figure 6.8: Plot of optimal barrier level for different combinations of α_1 and α_2 . Left: $\sigma_A = 0.02$ and $\beta_0^* = 1.13588$. Right: $\sigma_A = 0.3$ and $\beta_0^* = 2.87326$. For parameters used, see Table 6.1 in Appendix 6.B, set no. 2 (left plot) and set no. 3 (right plot).

6.3.3.2 Moving from β_0^* to β_2^*

The relationship between β_1^* and β_0^* is trivial, but the relationship between β_2^* and β_0^* is not, as explained early in this section. Figure 6.8 compares the optimal barrier level in a model without solvency constraints with the optimal barrier level in the model with advanced solvency constraints, and shows how the optimal barrier without constraints, β_0^* , is no longer optimal. Instead we get the optimal barrier β_2^* (adjusted compared to β_0^*) when advanced solvency constraints are introduced.

The right edge of the surfaces correspond to the simple solvency constraint ($\alpha_1 = \alpha_2$, so that $\beta_2^* = \beta_1^*$). There we can see the trivial, linear relationship between β_1^* and β_0^* , which is flat as long as $\beta_0^* > \alpha_1$, and then increases linearly such that $\beta_1^* = \alpha_1$.

When the volatility is rather low (as on the left plot), and we move towards the left on the surface (α_2 -wards), for given low α_1 , the optimal barrier does not change and is very close to β_0^* . This is because β_0^* is far enough from α_1 , and the process is very stable. When we increase volatility (moving to the right plot), β_0^* increases and even for low values of α_1 , β_2^* increases with α_2 .

Now, if we move towards the right of the surface for given α_2 , we can observe an increase of the optimal barrier even before the kink. This is because moving α_1 towards the barrier level makes periods when no distributions are allowed more likely, which is a problem particularly for low volatility (a low volatility means that the process can get stuck in a non-distribution state for a very long period of time). This effect seems to dominate the ‘kink’ effect especially for low volatilities.

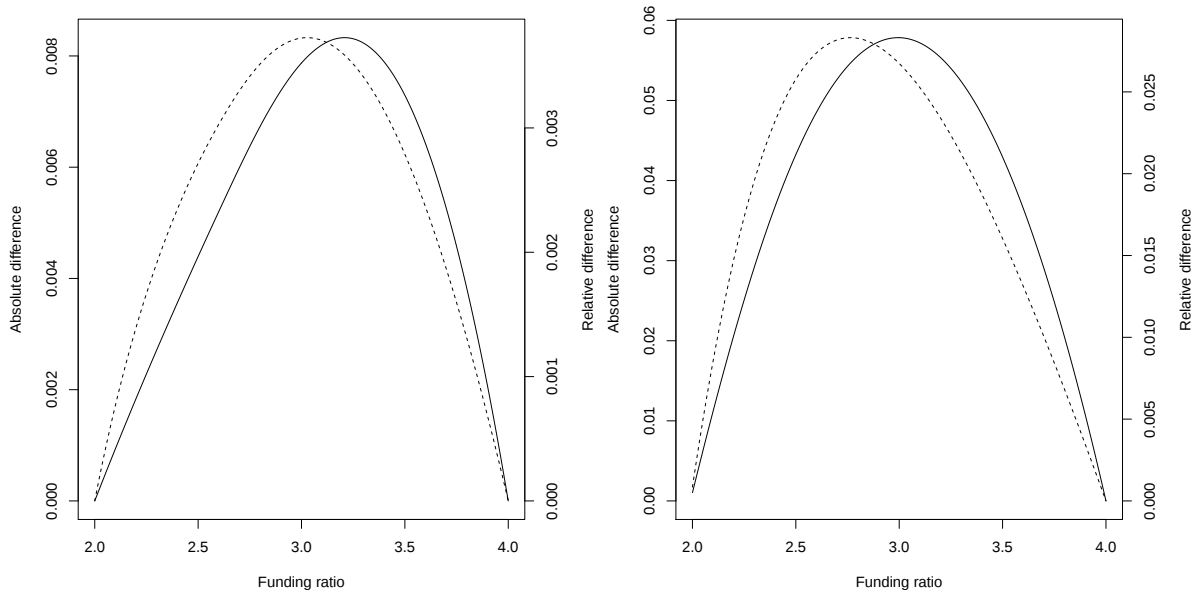


Figure 6.9: Differences between $V_{B,2}^0$ and $V_{B,2}^1$. Note that $V_{B,2}^0(\alpha_1) = V_{B,2}^1(\alpha_1)$ and $V_{B,2}^0(\alpha_2) = V_{B,2}^1(\alpha_2)$. The differences illustrated with the dotted line are relative to the values of $V_{B,2}^0$. For parameters used, see Table 6.1 in Appendix 6.B, set no. 4 (left plot) and set no. 5 (right plot).

6.3.3.3 The cost of not being able to distribute

Under the advanced solvency constraint, we have two different value functions when the funding ratio is in the interval between α_1 and α_2 . One can interpret the differences between these two value functions as the cost of being in the undesirable *no distribution* state. The difference between the value functions for $\phi = 0$ (the “undesirable” state) and $\phi = 1$ (the “good” state) in Case B is illustrated in Figure 6.9. The left plot is for a high value of σ_A (0.25) whereas the right figure is for a low value of σ_A (0.01). The reason that the differences are smallest for the most volatile model is that higher volatility leads to more switches between both environments, decreasing the influence of whether $\phi = 0$ or $\phi = 1$.

6.3.3.4 Sensitivity analysis for the volatility and correlation

Figure 6.10 shows the impact of volatility (first row) and correlation (second row) on the optimal barrier in absence of solvency constraint (first column), or with advanced solvency constraints (second column). The immediate observation is that effects that were trivial before the introduction of advanced solvency constraints are not trivial anymore. It becomes hard to describe which forces actually drive this sensitivity analysis.

In terms of correlation one can observe that higher correlation levels will lead to lower levels of the optimal barrier. This is because high correlation makes the funding ratio evolve in a (relatively) stable manner, such that we can choose a barrier that is not too far away from α_1 . On the other hand, the funding ratio goes wild for low values of ρ which pushes the β_2^* up.

In terms of volatility, absence of solvency constraints means that higher volatility levels will generally lead to higher barrier levels (except when σ_A is really close to 0, where σ_L is driving the volatility and the lack of correlation pushes the optimal barrier up slightly). For a σ_A that is not too high, there is an initial advantage in increasing volatility as it will make it easier to

leave state $\phi = 0$ if you fall into a no distribution period. However, as σ_A is going towards 0 or towards infinity, the effects present in absence of solvency constraint seem to dominate this marginal effect.

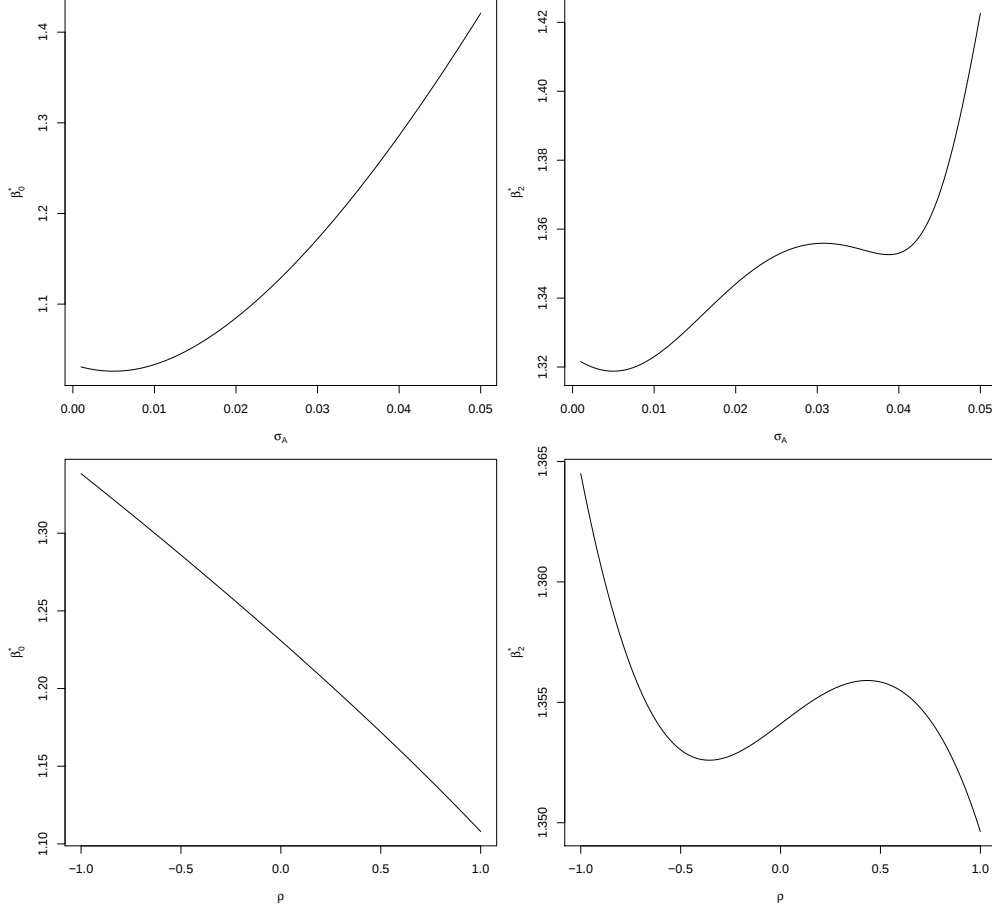


Figure 6.10: Sensitivity plots with respect to σ_A (above) and ρ (below). For parameters used, see Table 6.1, set no. 6.

6.4 When ruin must be prevented

In this section, we assume that the fund under consideration has the opportunity to implement rescue measures in order to prevent ruin; see Section 6.1.5. In this case, we consider the net distributions

$$J_{\varpi,3}(\vec{x};\pi) = \mathbb{E}^{\vec{x}} \left[\limsup_{t \rightarrow \infty} \int_0^t e^{-\delta s} dD_{\varpi}^{\pi}(s) - \int_0^t \kappa e^{-\delta s} dE_{\varpi}^{\pi}(s) \right], \quad (6.4.1)$$

where κ is a constant strictly greater than 1 which recognizes that there is a price of lowering liabilities or raising capital. The way ruin is prevented will depend on the context, as discussed in Section 6.1.5: liabilities are decreased in Case A, and assets increased in Case B. In (6.4.1), D_{ϖ}^{π} is the control process at the upper (distribution) barrier β , and E_{ϖ}^{π} is the control process at the lower (rescue) barrier γ . This means that the controlled processes for the assets and liabilities in

the present section have the form

$$d\vec{X}_A^\pi(t) = \mu \left(\vec{X}_A^\pi(t) \right) dt + \sigma \left(\vec{X}_A^\pi(t) \right) dW(t) + \begin{pmatrix} 0 \\ dD_A^\pi(t) - dE_A^\pi(t) \end{pmatrix}$$

and

$$d\vec{X}_B^\pi(t) = \mu \left(\vec{X}_B^\pi(t) \right) dt + \sigma \left(\vec{X}_B^\pi(t) \right) dW(t) - \begin{pmatrix} dD_B^\pi(t) - dE_B^\pi(t) \\ 0 \end{pmatrix}.$$

Because of transaction costs it makes sense that the optimal rescue level is the lowest possible (such that $\gamma^* = \alpha_0$), which is what we will assume in what follows. Note that this may not hold any more if transaction costs are not constant; see also Section 6.4.3.3. Furthermore, we conjecture that the form of the optimal distribution strategy is as before, that is, of the barrier type.

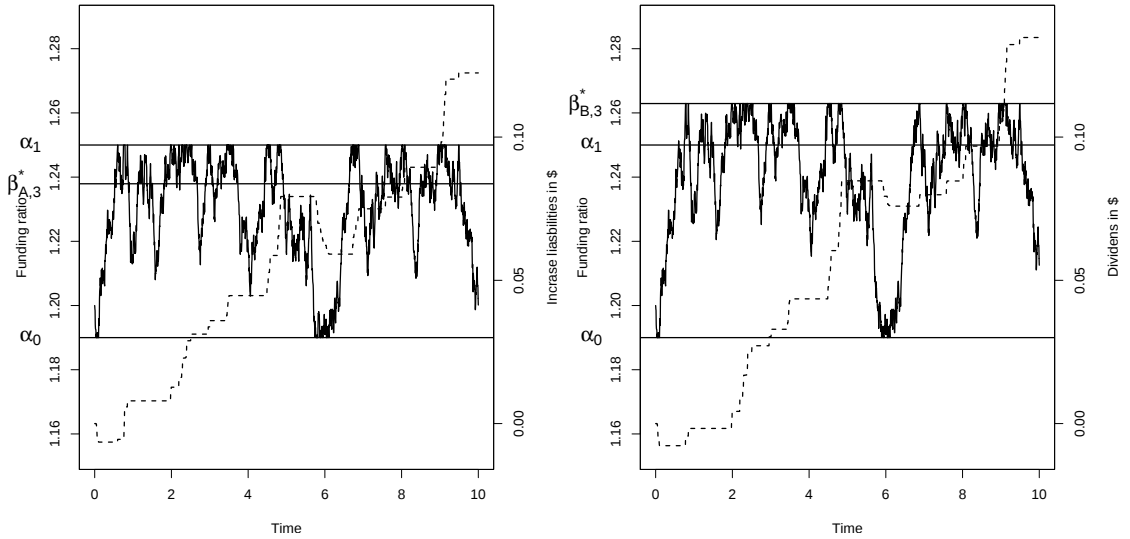


Figure 6.11: Simulation of the controlled funding ratio in the model where ruin must be prevented. The dotted lines are the distributions. Left: Case A. Right: Case B.

The model is illustrated in Figure 6.11, where the funding ratios are depicted by solid lines and net distributions (distributions, minus rescue measures) by dotted lines for a given sample path of the two-dimensional Brownian motion. The two plots in Figure 6.11 also contain the optimal upper barriers, which we find in the following two subsections. Note that the sample paths of the controlled funding ratios are different, even for the same level of distribution barrier. The optimal barriers in Figure 6.11 are different in Case A and Case B. Note also that the time to ruin in the preceding sections will be the time of the first lowering of liabilities (Case A) or first capital injection (Case B).

Remark 6.4.1. *In this section, we assume that rescue is forced. If we relaxed that assumption (but that is outside the scope of this paper), we would need to determine whether such rescue measures are worth taking or not. It makes sense that this would be the case if, and only if, $V_{\varpi,3}(\gamma x_2, x_2; \beta, \gamma)$ was greater than or equal to zero. This conjecture is supported by previous results in the literature (see, for instance Avanzi et al., 2011). This is explored in Section 6.4.3.4.*

6.4.1 In Case A, where rescue measures are to reduce benefits (decrease liabilities)

Preventing ruin by decreasing liabilities at the lower barrier adds an extra branch in the differential equation characterising the value function given by (6.2.3). The differential equation becomes

$$0 = \max \left\{ (\mathcal{A} - \delta)F(\vec{x}), 1 + \frac{\partial}{\partial x_2}F(\vec{x}), -\frac{\partial}{\partial x_2}F(\vec{x}) - \kappa \right\}. \quad (6.4.2)$$

We assume that $\gamma < \beta$ to make the strategy admissible. The expected present value of net distributions for a barrier strategy with upper barrier β and lower barrier γ is denoted by $G_A^{\beta, \gamma}$ and is given by

$$G_A^{\beta, \gamma}(\vec{x}) = \begin{cases} G_A(\vec{x}; \beta, \gamma), & x_1 \in [\gamma x_2, \beta x_2], \\ \frac{x_1}{\beta} - x_2 + G_A\left(x_1, \frac{x_1}{\beta}; \beta, \gamma\right), & x_1 > \beta x_2, \\ -\kappa \left(x_2 - \frac{x_1}{\gamma}\right) + G_A\left(x_1, \frac{x_1}{\gamma}; \beta, \gamma\right), & x_1 < \gamma x_2, \end{cases} \quad (6.4.3)$$

where $G_A(x_1, x_2; \beta, \gamma)$ is given by the differential equation

$$(\mathcal{A} - \delta)G_A(\vec{x}; \beta, \gamma) = 0 \text{ for } \gamma \leq \frac{x_1}{x_2} \leq \beta.$$

Since we have forced decreases of liabilities at the lower barrier, we do not have a boundary condition at the ruin level like (6.2.8). Instead, we have a boundary condition at the lower barrier. In total, the boundary conditions are

$$\frac{\partial}{\partial x_2}G_A(x_1, x_2; \beta, \gamma) \Big|_{x_2 = \frac{x_1}{\beta}} = -1, \quad (6.4.4)$$

$$\frac{\partial}{\partial x_2}G_A(x_1, x_2; \beta, \gamma) \Big|_{x_2 = \frac{x_1}{\gamma}} = -\kappa. \quad (6.4.5)$$

Moreover, for the optimal upper barrier, we also have the boundary condition

$$\frac{\partial^2}{\partial x_2^2}(x_1, x_2; \beta_{A,3}^*, \gamma) \Big|_{x_2 = \frac{x_1}{\beta_{A,3}^*}} = 0 \quad (6.4.6)$$

where $\beta_{A,3}^*$ is the optimal barrier level for increasing liabilities. For a heuristic derivation of condition (6.4.5), see Avanzi et al. (2011, Section 3.1).

By the calculations in Appendix 6.A.1, we get that $G_A(\cdot; \beta, \gamma)$ has the form given by (6.A.5). Using condition (6.4.5) we obtain that C_1 and C_2 fulfill the equation

$$C_1(1 - \zeta_1)x_1^{\zeta_1} \left(\frac{x_1}{\gamma}\right)^{-\zeta_1} + C_2(1 - \zeta_2)x_1^{\zeta_2} \left(\frac{x_1}{\gamma}\right)^{-\zeta_2} = -\kappa. \quad (6.4.7)$$

This means that

$$C_2 = \frac{-\kappa - C_1(1 - \zeta_1)\gamma^{\zeta_1}}{(1 - \zeta_2)\gamma^{\zeta_2}} = \frac{-\kappa\gamma^{-\zeta_2} - C_1(1 - \zeta_1)\gamma^{\zeta_1 - \zeta_2}}{1 - \zeta_2},$$

for a given level γ . Moreover, using condition (6.4.4), we obtain that

$$C_1(1 - \zeta_1)x_1^{\zeta_1} \left(\frac{x_1}{\beta}\right)^{-\zeta_1} + \frac{-\kappa - C_1(1 - \zeta_1)\gamma^{\zeta_1}}{(1 - \zeta_2)\gamma^{\zeta_2}}(1 - \zeta_2)x_1^{\zeta_2} \left(\frac{x_1}{\beta}\right)^{-\zeta_2} = -1.$$

This implies that C_1 is given by

$$C_1 = \frac{-1 + \kappa \left(\frac{\beta}{\gamma}\right)^{\zeta_2}}{(1 - \zeta_1) \left(\beta^{\zeta_1} - \gamma^{\zeta_1} \left(\frac{\beta}{\gamma}\right)^{\zeta_2}\right)} \quad (6.4.8)$$

That is, in total we get that the value function for a strategy with dividend barrier β and capital injection barrier γ with $\beta > \gamma$ is

$$G_A(x_1, x_2; \beta, \gamma) = C_1 x_1^{\zeta_1} x_2^{1-\zeta_1} + \frac{-\kappa - C_1(1 - \zeta_1)\gamma^{\zeta_1}}{(1 - \zeta_2)\gamma^{\zeta_2}} x_1^{\zeta_2} x_2^{1-\zeta_2}, \quad (6.4.9)$$

where C_1 is given by (6.4.8). To find the optimal barriers, $\beta_{A,3}^*$ and γ_A^* , we need to maximize (6.4.9) for $\frac{x_1}{x_2} \in [\gamma, \beta]$. For general values of x_1 and x_2 the objective function is given by

$$\begin{aligned} \mathcal{O}_A(x_1, x_2, \gamma, \beta) &= \left(\frac{x_1}{\beta} - x_2\right)^+ - \kappa \left(x_2 - \frac{x_1}{\gamma}\right)^+ \\ &+ G_A\left(x_1, x_2 + \left(\frac{x_1}{\beta} - x_2\right)^+ - \left(x_2 - \frac{x_1}{\gamma}\right)^+; \beta, \gamma\right). \end{aligned} \quad (6.4.10)$$

We aim at finding the value

$$\sup_{(\gamma, \beta) \in \Pi} \mathcal{O}_A(x_1, x_2, \gamma, \beta).$$

We assume that the optimal lower barrier is equal to α_0 and hence consider γ to be a fixed value in (6.4.10). We want to maximize (6.4.9) wrt. β . We rearrange the terms and see that

$$G_A(x_1, x_2; \beta, \gamma) = \frac{-\kappa x_1^{\zeta_2} x_2^{1-\zeta_2}}{(1 - \zeta_2)\gamma^{\zeta_2}} + C_1 x_1^{\zeta_1} x_2^{1-\zeta_1} + C_1 \frac{-\zeta_1 \gamma^{\zeta_1 - \zeta_2}}{\zeta_2} x_1^{\zeta_2} x_2^{1-\zeta_2},$$

where the first term of the RHS does not depend on β and $x_1^{\zeta_1} x_2^{1-\zeta_1}$, $x_1^{\zeta_2} x_2^{1-\zeta_2}$ and $\frac{-\zeta_1 \gamma^{\zeta_1 - \zeta_2}}{\zeta_2}$ are all positive terms. That is, to maximize (6.4.9), we need to maximize C_1 as a function of β . The partial derivative of C_1 wrt. β is given by

$$\begin{aligned} \frac{\partial}{\partial \beta} C_1 &= \frac{\partial}{\partial \beta} \left(\frac{-1 + \kappa \left(\frac{\beta}{\gamma}\right)^{\zeta_2}}{(1 - \zeta_1) \left(\beta^{\zeta_1} - \gamma^{\zeta_1} \left(\frac{\beta}{\gamma}\right)^{\zeta_2}\right)} \right) \\ &= \frac{1}{(1 - \zeta_1) \left(\beta^{\zeta_1} - \gamma^{\zeta_1} \left(\frac{\beta}{\gamma}\right)^{\zeta_2}\right)^2} \left(\zeta_2 \kappa \gamma^{-\zeta_2} \beta^{\zeta_2-1} \left(\left(\beta^{\zeta_1} - \gamma^{\zeta_1} \left(\frac{\beta}{\gamma}\right)^{\zeta_2}\right) \right) \right. \\ &\quad \left. + \left(1 - \kappa \left(\frac{\beta}{\gamma}\right)^{\zeta_2} \right) \left(\zeta_1 \beta^{\zeta_1-1} - \gamma^{\zeta_1-\zeta_2} \zeta_2 \beta^{\zeta_2-1} \right) \right) \\ &= \frac{\zeta_1 \beta^{\zeta_1-1} - \zeta_2 \gamma^{\zeta_1-\zeta_2} \beta^{\zeta_2-1} + \zeta_2 \kappa \gamma^{-\zeta_2} \beta^{\zeta_2+\zeta_1-1} - \zeta_1 \kappa \gamma^{-\zeta_2} \beta^{\zeta_2+\zeta_1-1}}{(1 - \zeta_1) \left(\beta^{\zeta_1} - \gamma^{\zeta_1} \left(\frac{\beta}{\gamma}\right)^{\zeta_2}\right)^2}. \end{aligned}$$

Setting this term equal to zero gives us an equation for the optimal barrier:

$$\zeta_1 \beta^{\zeta_1-1} - \zeta_2 \gamma^{\zeta_1-\zeta_2} \beta^{\zeta_2-1} = \beta^{\zeta_2+\zeta_1-1} (\zeta_1 \kappa \gamma^{-\zeta_2} - \zeta_2 \kappa \gamma^{-\zeta_2}). \quad (6.4.11)$$

Based on numerical studies, we conjecture that the optimal barrier $\beta_{A,3}^*$ exists and is unique. However, it has turned out to be surprisingly difficult to formally prove this.

6.4.2 In case B, where rescue measures are to inject capital (increase assets)

Gerber and Shiu (2003) match inflow and outflow in the situation where $\kappa = 1$. Here, we set $\kappa > 1$ and obtain the expected present value of outflow minus inflow if a barrier strategy is applied. The distribution barrier that maximises this difference exists and is unique.

Introducing capital injections adds an extra branch in the differential equation characterizing the value function given by (6.2.23). The differential equation in the case of capital injections has the following form:

$$0 = \max \left\{ (\mathcal{A} - \delta)F(\vec{x}), 1 - \frac{\partial}{\partial x_1} F(\vec{x}), \frac{\partial}{\partial x_1} F(\vec{x}) - \kappa \right\}. \quad (6.4.12)$$

We conjecture that the optimal dividend and capital injection strategies are barrier strategies with dividend barrier β and capital injection barrier γ . We assume that $\gamma < \beta$ to make the strategy admissible. The expected present value of dividends for a barrier strategy with barrier β and capital injection barrier γ is denoted by $G_B^{\beta, \gamma}$ and is given by

$$G_B^{\beta, \gamma}(\vec{x}) = \begin{cases} G_B(\vec{x}; \beta, \gamma), & x_1 \in [\gamma x_2, \beta x_2], \\ (x_1 - \beta x_2) + G_B(\beta x_2, x_2; \beta, \gamma), & x_1 > \beta x_2, \\ -\kappa(x_1 - \gamma x_2) + G_B(\gamma x_2, x_2; \beta, \gamma), & x_1 < \gamma x_2, \end{cases} \quad (6.4.13)$$

where $G_B(x_1, x_2; \beta, \gamma)$ is given by the differential equation

$$(\mathcal{A} - \delta)G_B(\vec{x}; \beta, \gamma) = 0 \text{ for } \gamma \leq \frac{x_1}{x_2} \leq \beta.$$

Since we have forced capital injections we do not have a boundary condition at the ruin level as in (6.2.25). Instead, we have a boundary condition at the capital injection level. In total, the boundary conditions are

$$\frac{\partial}{\partial x_1} G_B(x_1, x_2; \beta, \gamma) \big|_{x_1 = \beta x_2 -} = 1, \quad (6.4.14)$$

$$\frac{\partial}{\partial x_1} G_B(x_1, x_2; \beta, \gamma) \big|_{x_1 = \gamma x_2 -} = \kappa, \quad (6.4.15)$$

Moreover, for the optimal upper barrier, we also have the boundary condition

$$\frac{\partial^2}{\partial x_1^2} G_B(x_1, x_2; \beta_{B,3}^*, \gamma) \big|_{x_1 = \beta_{B,3}^* x_2 -} = 0, \quad (6.4.16)$$

where $\beta_{B,3}^*$ is the optimal level for dividend payments.

By the calculations in Appendix 6.A.1, we get that $G_B(\cdot; \beta, \gamma)$ has the form given by (6.A.5). Using condition (6.4.15) we obtain that C_1 and C_2 fulfill the equation

$$C_1 \zeta_1 (\gamma x_2)^{\zeta_1 - 1} x_2^{1 - \zeta_1} + C_2 \zeta_2 (\gamma x_2)^{\zeta_2 - 1} x_2^{1 - \zeta_2} = \kappa. \quad (6.4.17)$$

This means that

$$C_2 = \frac{\kappa - C_1 \zeta_1 \gamma^{\zeta_1 - 1}}{\zeta_2 \gamma^{\zeta_2 - 1}} = \frac{\kappa \gamma^{1 - \zeta_2} - C_1 \zeta_1 \gamma^{\zeta_1 - \zeta_2}}{\zeta_2},$$

for a given level γ . Moreover, using condition (6.4.14), we obtain that

$$C_1 \zeta_1 (x_2 \beta)^{\zeta_1 - 1} x_2^{1 - \zeta_1} + \frac{\kappa - C_1 \zeta_1 \gamma^{\zeta_1 - 1}}{\zeta_2 \gamma^{\zeta_2 - 1}} \zeta_2 (x_2 \beta)^{\zeta_2 - 1} x_2^{1 - \zeta_2} = 1.$$

This implies that C_1 is given by

$$C_1 = \frac{\gamma^{\zeta_2-1} - \kappa\beta^{\zeta_2-1}}{\zeta_1(\beta^{\zeta_1-1}\gamma^{\zeta_2-1} - \gamma^{\zeta_1-1}\beta^{\zeta_2-1})} = \frac{1 - \kappa\beta^{\zeta_2-1}\gamma^{1-\zeta_2}}{\zeta_1(\beta^{\zeta_1-1} - \gamma^{\zeta_1-\zeta_2}\beta^{\zeta_2-1})}. \quad (6.4.18)$$

That is, in total we get that the value function for a strategy with dividend barrier β and capital injection barrier γ with $\beta > \gamma$ is

$$G_B(x_1, x_2; \beta, \gamma) = C_1 x_1^{\zeta_1} x_2^{1-\zeta_1} + \frac{\kappa\gamma^{1-\zeta_2} - C_1\zeta_1\gamma^{\zeta_1-\zeta_2}}{\zeta_2} x_1^{\zeta_2} x_2^{1-\zeta_2}, \quad (6.4.19)$$

where C_1 is given by (6.4.18).

To find the optimal barriers, $\beta_{B,3}^*$ and γ_B^* , we need to maximize (6.4.19) for $\frac{x_1}{x_2} \in [\gamma, \beta]$. For general values of x_1 and x_2 the objective function is given by

$$\mathcal{O}_B(x_1, x_2, \gamma, \beta) = (x_1 - \beta x_2)^+ - \kappa(\gamma x_2 - x_1)^+ + G_B(x_1 - (x_1 - \beta x_2)^+ + (\gamma x_2 - x_1)^+, x_2; \beta, \gamma). \quad (6.4.20)$$

We aim at finding the value

$$\sup_{(\gamma, \beta) \in \Pi} \mathcal{O}_B(x_1, x_2, \gamma, \beta).$$

Again, we assume that the optimal lower barrier is equal to α_0 and hence consider γ to be a fixed value in (6.4.20). We want to maximize (6.4.19) wrt. β . We rearrange the terms and see that

$$G_B(x_1, x_2; \beta, \gamma) = \frac{\kappa\gamma^{1-\zeta_2} x_1^{\zeta_2} x_2^{1-\zeta_2}}{\zeta_2} + C_1 x_1^{\zeta_1} x_2^{1-\zeta_1} + C_1 \frac{-\zeta_1\gamma^{\zeta_1-\zeta_2}}{\zeta_2} x_1^{\zeta_2} x_2^{1-\zeta_2},$$

where the first term of the RHS does not depend on β and $x_1^{\zeta_1} x_2^{1-\zeta_1}$, $x_1^{\zeta_2} x_2^{1-\zeta_2}$ and $\frac{-\zeta_1\gamma^{\zeta_1-\zeta_2}}{\zeta_2}$ are all positive terms. That is, to maximize (6.4.19) we need to maximize C_1 as a function of β . The partial derivative of C_1 wrt. β is given by

$$\begin{aligned} \frac{\partial}{\partial \beta} C_1 &= \frac{\partial}{\partial \beta} \left(\frac{1 - \kappa\beta^{\zeta_2-1}\gamma^{1-\zeta_2}}{\zeta_1(\beta^{\zeta_1-1} - \gamma^{\zeta_1-\zeta_2}\beta^{\zeta_2-1})} \right) \\ &= \frac{1}{\zeta_1^2(\beta^{\zeta_1-1} - \gamma^{\zeta_1-\zeta_2}\beta^{\zeta_2-1})^2} \left(-(\zeta_2 - 1)\kappa\beta^{\zeta_2-2}\gamma^{1-\zeta_2}\zeta_1(\beta^{\zeta_1-1} - \gamma^{\zeta_1-\zeta_2}\beta^{\zeta_2-1}) \right. \\ &\quad \left. - (1 - \kappa\beta^{\zeta_2-1}\gamma^{1-\zeta_2})\zeta_1((\zeta_1 - 1)\beta^{\zeta_1-2} - (\zeta_2 - 1)\gamma^{\zeta_1-\zeta_2}\beta^{\zeta_2-2}) \right). \end{aligned}$$

Setting this term equal to zero leads to the equation

$$\begin{aligned} &-(\zeta_2 - 1)\kappa\beta^{\zeta_2-2}\gamma^{1-\zeta_2}\zeta_1(\beta^{\zeta_1-1} - \gamma^{\zeta_1-\zeta_2}\beta^{\zeta_2-1}) \\ &= (1 - \kappa\beta^{\zeta_2-1}\gamma^{1-\zeta_2})\zeta_1((\zeta_1 - 1)\beta^{\zeta_1-2} - (\zeta_2 - 1)\gamma^{\zeta_1-\zeta_2}\beta^{\zeta_2-2}). \end{aligned}$$

This gives us an equation of the form (dividing by β^{ζ_2-3})

$$\zeta_1 \left(\gamma^{1-\zeta_2}\kappa(\zeta_1 - \zeta_2)\beta^{\zeta_1} + (\zeta_2 - 1)\gamma^{\zeta_1-\zeta_2}\beta + (1 - \zeta_1)\beta^{1-\zeta_2+\zeta_1} \right) = 0. \quad (6.4.21)$$

The optimal dividend barrier level $\beta_{B,3}^*$ is given as the solution to (6.4.21). The existence and uniqueness of the optimal barrier level is given by the following lemma.

Lemma 6.4.2. *An optimal barrier level $\beta_{B,3}^*$ exists and is unique.*

Proof: See appendix 6.A.5.

6.4.3 Numerical studies

In this section, we discuss the impact of adding rescue measures in the objective.

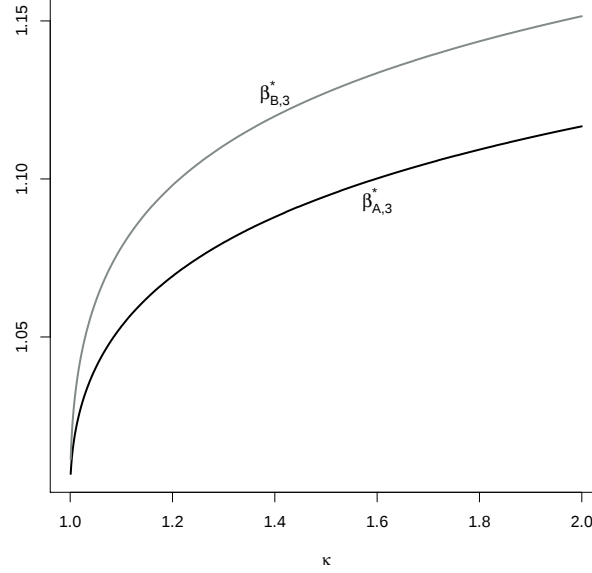


Figure 6.12: The optimal barriers $\beta_{\varpi,3}^*$ as a function of κ . In black $\beta_{A,3}^*$ and in grey $\beta_{B,3}^*$. For parameters used, see Table 6.1 in Appendix 6.B, set no. 1.

6.4.3.1 The impact of transaction costs on the optimal distribution barrier

Whichever value κ has, the optimal rescue barrier will always be α_0 . On the other hand, the optimal distribution barrier is affected by different values of κ . This is illustrated in Figure 6.12. For $\kappa = 1$, the optimal distribution barrier is α_0 since there is no reason to hold an extra buffer when additional capital comes at no cost. It then increases as κ increases.

6.4.3.2 The impact of the rescue and distribution barriers on the value function

In Figure 6.13, we depict surface plots for γ and β of the value function for the model in the case where ruin must be prevented. We see that the worst choice one can make is to have both β and γ close to one.

6.4.3.3 Insights into surplus-dependent transaction costs

So far, we have both theoretically and numerically only considered a constant value of κ . However, one could argue that a fixed cost for raising capital in Case B is not too realistic. A more realistic approach could be to make the cost dependent on the funding ratio such that the lower the funding ratio the higher the cost (in other words, it is easier to raise capital when the fund is better funded). Arbitrarily and for illustration purposes, we choose to model this relationship by

$$\kappa(x) = 1.01 + e^{(K-1.5)-Kx}, \quad (6.4.22)$$

where the argument is the funding ratio and K is a given constant. This means, for instance, that the price of raising capital for the funding ratio at the ruin level is $\kappa(\alpha_0) \approx 1.23$ for $\alpha_0 = 1$.

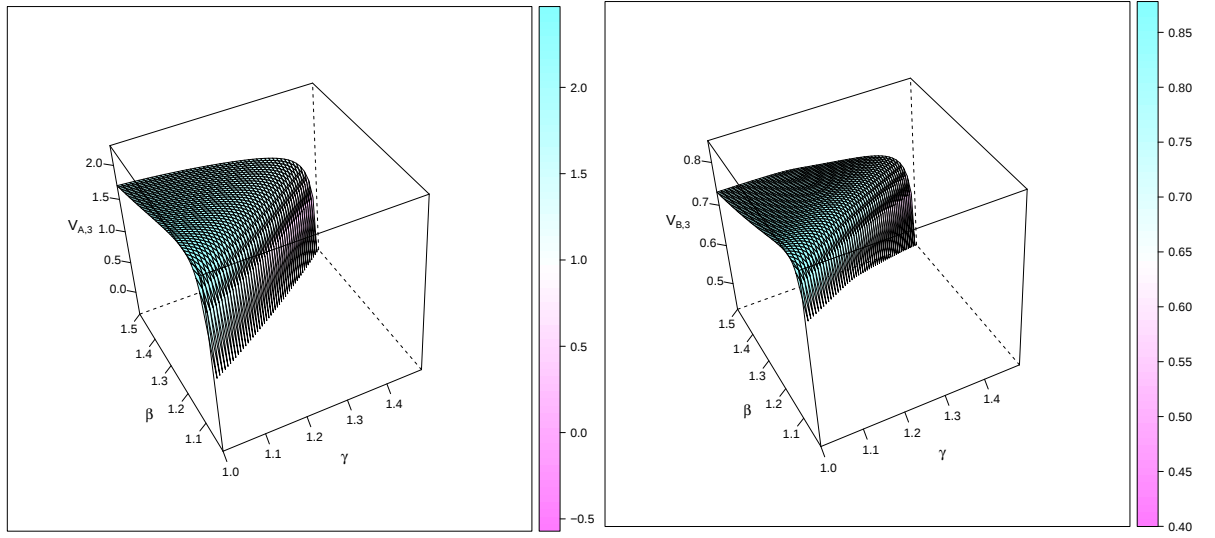


Figure 6.13: Surface plots of the value functions as functions of γ and β with $\kappa = 1.05$. Left: Case A. Right: Case B. For parameters used, see Table 6.1 in Appendix 6.B, set no. 7 (left plot) and set no. 1 (right plot).

For higher values of the funding ratio, κ is exponentially decreasing to 1.01 as the funding ratio tends to infinity. In Figure 6.14, we see the optimal barriers for different penalty functions of the form (6.4.22). The modelling of the penalty functions implies that, independent of the parameter K , the price of capital is the same for the funding ratio equal to α_0 . This type of price structure has pushed the optimal lower barrier far away from $\alpha_0 = 1$. As the steepness (the speed of the exponential decay of $\kappa(x)$) of the penalty functions decrease, one can choose an increasingly lower barrier without getting penalised too much. This also implies that the expected present value of dividends increases dramatically.

Importantly, these results hold for relatively high levels of volatility. Indeed, if we decreased σ_A in Figure 6.14 from 0.2 to 0.05, it would always be optimal to have the barrier as low as possible (that is, equal to α_0). In the right plot of Figure 6.14, we see a plot of the same type as the right plot in Figure 6.13 but with the difference that transaction costs are according to (6.4.22) with $K = 6$. Note that this makes the value function highly negative for many combinations of the lower and upper bounds and also that it changes the shape of the figure a bit compared to the right plot in Figure 6.13. The optimal values are given by $\gamma = 1.4225$ and $\beta = 2.1448$.

6.4.3.4 Should we rescue?

In this section, we explore the conjecture spelt out in Remark 6.4.1, which asserts that, in fact, rescue measures should be taken only if the value function at the rescue barrier is nonnegative. In Figure 6.15, we consider Case A and plot the values of κ which makes the value function exactly equal to 0 for the optimal choice of the upper barrier. The left plot is with a simple solvency constraint, the right is without. One observes from the figure that lower risk levels will allow for higher levels of transaction costs κ .

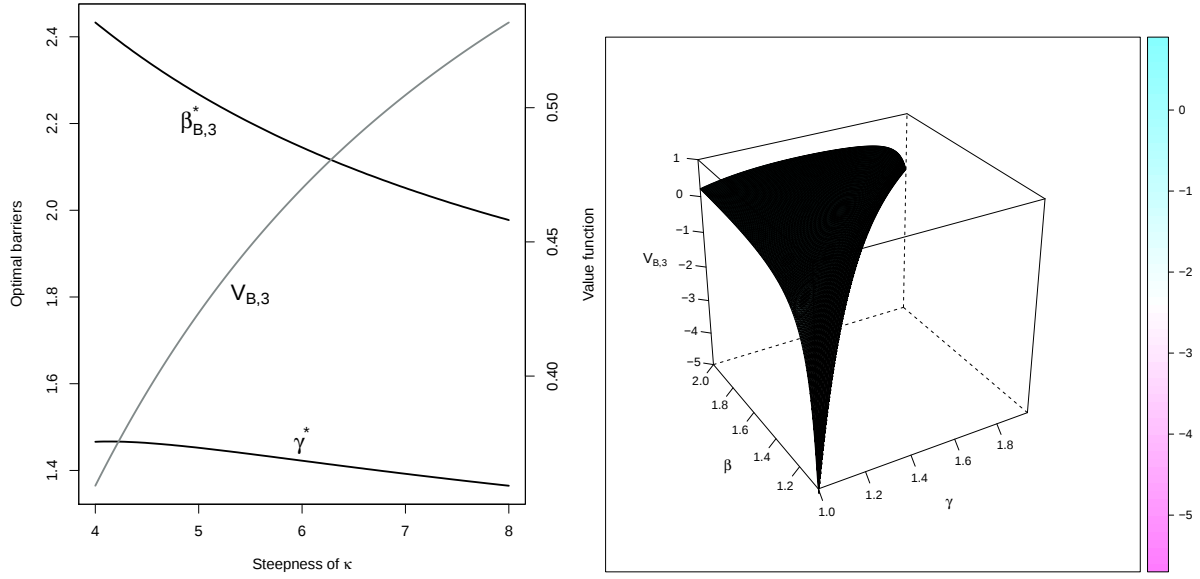


Figure 6.14: Left: Optimal capital injection and dividend barrier (in black) and corresponding value functions (in grey) for different steepness of κ . The values on the first axis refers to values K in (6.4.22). Right: The value function for different combinations of lower and upper barriers. For parameters used, see Table 6.1 in Appendix 6.B, set no. 8.

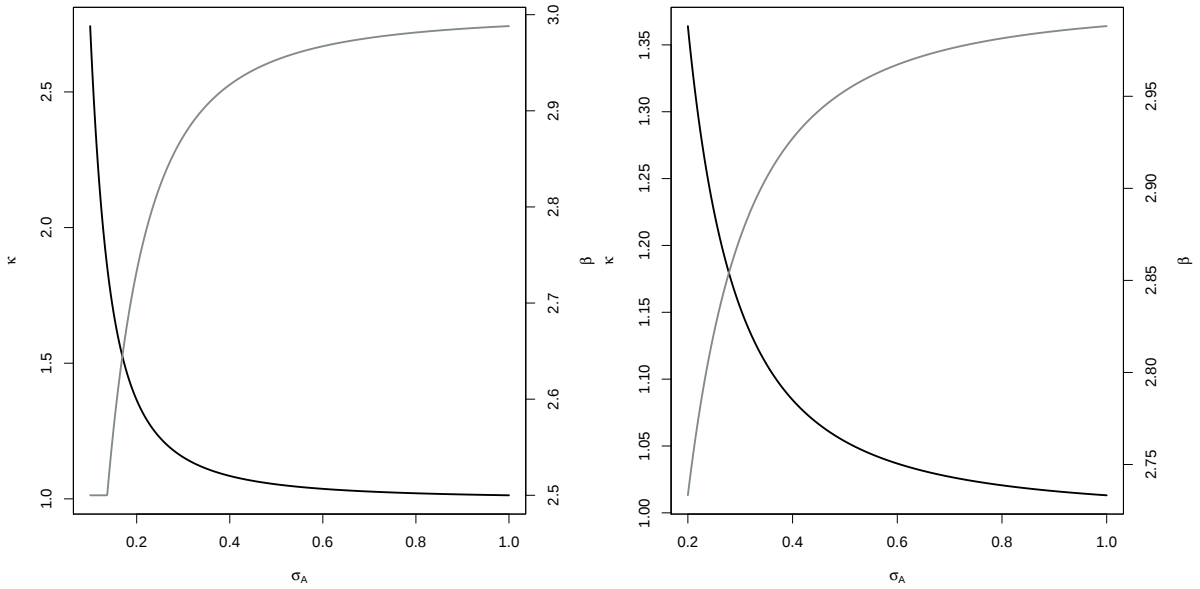


Figure 6.15: Illustration of which values of κ (black line) that makes the value function equal to 0 at the funding ratio α_0 . The corresponding optimal values for β are in grey. The left plot is with solvency constraint ($\alpha_1 = 2.5$) and the right plot is without solvency constraint. For parameters used, see Table 6.1 in Appendix 6.B, set no. 9.

Acknowledgments: Part of the paper was written while Henriksen was visiting the other authors at the UNSW Business School and the Department of Mathematics and Statistics of the University of Montreal. Henriksen would like to thank both Universities for their hospitality.

Appendix

6.A Proofs

6.A.1 Proof of Lemma 6.2.1 (Value function without constraints)

Proof of Lemma 6.2.1. We introduce the notation $\tilde{G}(\cdot; \beta)$ by

$$\begin{aligned} G_A(x_1, x_2; \beta) &= (x_1 + x_2) G_A\left(\frac{x_1}{x_1 + x_2}, \frac{x_2}{x_1 + x_2}; \beta\right) = (x_1 + x_2) G_A(y, 1 - y; \beta) \\ &= (x_1 + x_2) \tilde{G}(y; \beta), \end{aligned}$$

where $y = \frac{x_1}{x_1 + x_2}$. For reformulation of the HJB equation, we need the following derivatives

$$\frac{\partial}{\partial x_i} G_A(\vec{x}; \beta) \text{ and } \frac{\partial^2}{\partial x_i \partial x_j} G_A(\vec{x}; \beta), \quad i, j = 1, 2.$$

We get

$$\begin{aligned} \frac{\partial}{\partial x_1} G_A(x_1, x_2; \beta) &= \frac{\partial}{\partial x_1} \left((x_1 + x_2) \tilde{G}\left(\frac{x_1}{x_1 + x_2}; \beta\right) \right) \\ &= \tilde{G}\left(\frac{x_1}{x_1 + x_2}; \beta\right) + (x_1 + x_2) \tilde{G}'\left(\frac{x_1}{x_1 + x_2}; \beta\right) \left(\frac{1}{x_1 + x_2} - \frac{x_1}{(x_1 + x_2)^2} \right) \\ &= \tilde{G}(y; \beta) + \tilde{G}'(y; \beta)(1 - y), \\ \frac{\partial}{\partial x_2} G_A(x_1, x_2; \beta) &= \frac{\partial}{\partial x_2} \left((x_1 + x_2) \tilde{G}\left(\frac{x_1}{x_1 + x_2}; \beta\right) \right) \\ &= \tilde{G}\left(\frac{x_1}{x_1 + x_2}; \beta\right) + (x_1 + x_2) \tilde{G}'\left(\frac{x_1}{x_1 + x_2}; \beta\right) \left(-\frac{x_1}{(x_1 + x_2)^2} \right) \\ &= \tilde{G}(y; \beta) + \tilde{G}'(y; \beta)(-y), \\ \frac{\partial^2}{\partial x_1^2} G_A(x_1, x_2; \beta) &= \frac{\partial^2}{\partial x_1^2} \left((x_1 + x_2) \tilde{G}\left(\frac{x_1}{x_1 + x_2}; \beta\right) \right) \\ &= \frac{\partial}{\partial x_1} \left(\tilde{G}\left(\frac{x_1}{x_1 + x_2}; \beta\right) + \frac{x_2}{x_1 + x_2} \tilde{G}'\left(\frac{x_1}{x_1 + x_2}; \beta\right) \right) \\ &= \tilde{G}'\left(\frac{x_1}{x_1 + x_2}; \beta\right) \left(\frac{1}{x_1 + x_2} - \frac{x_1}{(x_1 + x_2)^2} - \frac{x_2}{(x_1 + x_2)^2} \right) \\ &\quad + \frac{x_2}{x_1 + x_2} \tilde{G}''\left(\frac{x_1}{x_1 + x_2}; \beta\right) \frac{x_2}{(x_1 + x_2)^2} \\ &= \tilde{G}''(y; \beta) \frac{x_2^2}{(x_1 + x_2)^3}. \end{aligned}$$

Likewise,

$$\begin{aligned}\frac{\partial^2}{\partial x_2^2} G_A(x_1, x_2; \beta) &= \tilde{G}''(y; \beta) \frac{x_1^2}{(x_1 + x_2)^3}, \\ \frac{\partial^2}{\partial x_1 \partial x_2} G_A(x_1, x_2; \beta) &= -\tilde{G}''(y; \beta) \frac{x_1 x_2}{(x_1 + x_2)^3}.\end{aligned}$$

In total we get that

$$\begin{aligned}\frac{\mathcal{A}G_A(\vec{x}; \beta) - \delta G_A(\vec{x}; \beta)}{x_1 + x_2} &= \frac{1}{2} (\sigma_A^2 + \sigma_L^2 - 2\rho\sigma_A\sigma_L) y^2(1-y)^2 \tilde{G}''(y; \beta) + (\mu_A - \mu_L) y(1-y) \tilde{G}'(y; \beta) \\ &\quad + (\mu_A y + \mu_L(1-y) - \delta) \tilde{G}(y; \beta),\end{aligned}\tag{6.A.1}$$

and the second branch of the HJB equation has the following form

$$1 + \frac{\partial}{\partial x_2} G_A(\vec{x}; \beta) = 1 + \tilde{G}(y; \beta) - y \tilde{G}'(y; \beta).$$

The boundary condition for $\tilde{G}(\cdot; \beta)$ (which holds for all values of β) is given by (6.2.8):

$$G_A(\alpha_0 x_2, x_2; \beta) = 0 \Rightarrow (\alpha_0 + 1) x_2 G_A\left(\frac{\alpha_0 x_2}{(\alpha_0 + 1) x_2}, \frac{x_2}{(\alpha_0 + 1) x_2}; \beta\right) = 0 \Rightarrow \tilde{G}\left(\frac{\alpha_0}{1 + \alpha_0}; \beta\right) = 0.\tag{6.A.2}$$

We guess that the solution to the right-hand side of equation (6.A.1) equal to 0 has a solution of the form:

$$\tilde{G}(y; \beta) = y^\vartheta (1-y)^\varphi.$$

Inserting this in (6.A.1) gives:

$$\begin{aligned}&\frac{\mathcal{A}G_A(\vec{x}; \beta) - \delta G_A(\vec{x}; \beta)}{x_1 + x_2} \\ &= \frac{1}{2} (\sigma_A^2 + \sigma_L^2 - 2\rho\sigma_A\sigma_L) y^2(1-y)^2 \\ &\quad \times \left(y^{\vartheta-2} (\vartheta^2 - \vartheta) (1-y)^\varphi - 2 y^{\vartheta-1} \vartheta (1-y)^{\varphi-1} \varphi + y^\vartheta (1-y)^{\varphi-2} (\varphi^2 - \varphi) \right) \\ &\quad + (\mu_A - \mu_L) y(1-y) \left(\vartheta y^{\vartheta-1} (1-y)^\varphi - \varphi y^\vartheta (1-y)^{\varphi-1} \right) + (\mu_A y + \mu_L(1-y) - \delta) y^\vartheta (1-y)^\varphi.\end{aligned}$$

Dividing the above equation with $y^\vartheta (1-y)^\varphi = \tilde{G}(y)$ and setting $\tilde{\sigma}^2 = \sigma_A^2 + \sigma_L^2 - 2\rho\sigma_A\sigma_L$ yields that the right-hand side is equal to

$$\frac{1}{2} (\vartheta^2 - \vartheta) \tilde{\sigma}^2 + \vartheta(\mu_A - \mu_L) + \mu_L - \delta + (\vartheta + \varphi - 1) \left((-\vartheta \tilde{\sigma}^2 + \mu_A - \mu_L) y + \frac{1}{2} \tilde{\sigma}^2 (\vartheta + \varphi) y^2 \right).$$

Setting $(*) = 0$ gives us a quadratic equation for ϑ :

$$\frac{1}{2} \tilde{\sigma}^2 \vartheta^2 + \left(\mu_A - \mu_L - \frac{1}{2} \tilde{\sigma}^2 \right) \vartheta + \mu_L - \delta = 0,\tag{6.A.3}$$

and by setting $(**) = 0$ we get that $\varphi = 1 - \vartheta$. We denote by ζ_1 and ζ_2 the two solutions to the quadratic equation. The solutions are given by

$$\begin{aligned}
& -(\mu_A - \mu_L - \frac{1}{2}\tilde{\sigma}^2) \pm \sqrt{(\mu_A - \mu_L - \frac{1}{2}\tilde{\sigma}^2)^2 - 4\frac{1}{2}\tilde{\sigma}^2(\mu_L - \delta)} \\
& \quad \quad \quad \tilde{\sigma}^2 \\
& = \frac{\frac{1}{2}\tilde{\sigma}^2 - (\mu_A - \mu_L) \pm \sqrt{\frac{1}{4}\tilde{\sigma}^4 + (\mu_A - \mu_L)^2 - (\mu_A - \mu_L)\tilde{\sigma}^2 - 2\tilde{\sigma}^2(\mu_L - \delta)}}{\tilde{\sigma}^2} \\
& = \frac{\frac{1}{2}\tilde{\sigma}^2 - (\mu_A - \mu_L) \pm \sqrt{\frac{1}{4}\tilde{\sigma}^4 + (\mu_A - \mu_L)^2 - \tilde{\sigma}^2(\mu_A + \mu_L - 2\delta)}}{\tilde{\sigma}^2}. \tag{6.A.4}
\end{aligned}$$

Because the coefficient of the quadratic term of (6.A.3), $\frac{1}{2}\tilde{\sigma}^2$, is greater than 0 and because the left-hand side of (6.A.3) is negative for $\vartheta = 0$ by (6.1.2) and (6.1.3) the quadratic equation (6.A.3) has a positive solution, which we denote ζ_2 , and negative solution, which we denote ζ_1 . Because the left-hand side of (6.A.3) is equal to $\mu_A - \delta < 0$ for $\vartheta = 1$, we get that $\zeta_2 > 1$.

By using that $\tilde{G}(y; \beta) = G_A(y, 1 - y; \beta)$ we get that a general solution is given by

$$G_A(x_1, x_2; \beta) = C_1 x_1^{\zeta_1} x_2^{1-\zeta_1} + C_2 x_1^{\zeta_2} x_2^{1-\zeta_2} \tag{6.A.5}$$

Using (6.2.8), we get the following relation between the coefficients C_1 and C_2 :

$$C_1(\alpha_0 x_2)^{\zeta_1} x_2^{1-\zeta_1} + C_2(\alpha_0 x_2)^{\zeta_2} x_2^{1-\zeta_2} = 0 \Rightarrow -C_1 \alpha_0^{\zeta_1} = C_2 \alpha_0^{\zeta_2} \Rightarrow C_2 = -C_1 \alpha_0^{\zeta_1 - \zeta_2}.$$

Another boundary condition, given by (6.2.9), is that at the upper barrier at level β , we have that

$$\frac{\partial}{\partial x_2} G_A(x_1, x_2; \beta) \Big|_{x_2 = \frac{x_1 -}{\beta}} = -1.$$

We use this to determine C_1 :

$$\begin{aligned}
\frac{\partial}{\partial x_2} G_A(x_1, x_2; \beta) &= \frac{\partial}{\partial x_2} \left(C_1 x_1^{\zeta_1} x_2^{1-\zeta_1} - C_1 \alpha_0^{\zeta_1 - \zeta_2} x_1^{\zeta_2} x_2^{1-\zeta_2} \right) \\
&= C_1 \left((1 - \zeta_1) x_1^{\zeta_1} x_2^{-\zeta_1} - \alpha_0^{\zeta_1 - \zeta_2} (1 - \zeta_2) x_1^{\zeta_2} x_2^{-\zeta_2} \right). \tag{6.A.6}
\end{aligned}$$

That is,

$$\begin{aligned}
& \frac{\partial}{\partial x_2} G_A(x_1, x_2; \beta) \Big|_{x_2 = \frac{x_1 -}{\beta}} = -1, \\
& \Rightarrow C_1 \left((1 - \zeta_1) x_1^{\zeta_1} \left(\frac{x_1 -}{\beta} \right)^{-\zeta_1} - \alpha_0^{\zeta_1 - \zeta_2} (1 - \zeta_2) x_1^{\zeta_2} \left(\frac{x_1 -}{\beta} \right)^{-\zeta_2} \right) = -1, \\
& \Rightarrow C_1 = -\frac{1}{(1 - \zeta_1)\beta^{\zeta_1} - \alpha_0^{\zeta_1 - \zeta_2}(1 - \zeta_2)\beta^{\zeta_2}}.
\end{aligned}$$

In total, this yields that $G_A(\cdot; \beta)$ for a given barrier strategy β is given by

$$G_A(x_1, x_2; \beta) = \frac{x_1^{\zeta_1} x_2^{1-\zeta_1} - \alpha_0^{\zeta_1 - \zeta_2} x_1^{\zeta_2} x_2^{1-\zeta_2}}{\alpha_0^{\zeta_1 - \zeta_2}(1 - \zeta_2)\beta^{\zeta_2} - (1 - \zeta_1)\beta^{\zeta_1}} \tag{6.A.7}$$

$$\begin{aligned}
& = x_2 \frac{\left(\frac{x_1}{\alpha_0 x_2} \right)^{\zeta_1} - \left(\frac{x_1}{\alpha_0 x_2} \right)^{\zeta_2}}{(1 - \zeta_2) \left(\frac{\beta}{\alpha_0} \right)^{\zeta_2} - (1 - \zeta_1) \left(\frac{\beta}{\alpha_0} \right)^{\zeta_1}}. \tag{6.A.8}
\end{aligned}$$

□

6.A.2 Proof of Lemma 6.2.2 (Fulfillment of conditions in the HJB equation)

Proof of Lemma 6.2.2. We prove the lemma by showing that

$$\begin{aligned} (\mathcal{A} - \delta)G_A^{\beta_0^*}(\vec{x}) &\begin{cases} = 0, & x_2 \in \left[\frac{x_1}{\beta_0^*}, \frac{x_1}{\alpha_0}\right], \\ < 0, & x_2 < \frac{x_1}{\beta_0^*}, \end{cases} \\ \frac{\partial}{\partial x_2}G_A^{\beta_0^*}(\vec{x}) &\begin{cases} \leq -1, & x_2 \in \left[\frac{x_1}{\beta_0^*}, \frac{x_1}{\alpha_0}\right], \\ = -1, & x_2 < \frac{x_1}{\beta_0^*}. \end{cases} \end{aligned} \quad (6.A.9)$$

To prove (6.A.9) we need that $G_A^{\beta_0^*}$ is concave in its first and second variable, respectively. That is, we start by showing that

$$\frac{\partial^2}{\partial x_1^2}G_A^{\beta_0^*}(\vec{x}) \leq 0 \text{ and } \frac{\partial^2}{\partial x_2^2}G_A^{\beta_0^*}(\vec{x}) \leq 0. \quad (6.A.10)$$

First, we consider the case $x_2 \geq \frac{x_1}{\beta_0^*}$. The second-order derivative with respect to x_1 is given by

$$\begin{aligned} \frac{\partial^2}{\partial x_1^2}G_A(x_1, x_2; \beta_0^*) &= \frac{\zeta_1(\zeta_1 - 1)x_1^{\zeta_1-2}x_2^{1-\zeta_1} - \alpha_0^{\zeta_1-\zeta_2}\zeta_2(\zeta_2 - 1)x_1^{\zeta_2-2}x_2^{1-\zeta_2}}{\alpha_0^{\zeta_1-\zeta_2}(1 - \zeta_2)(\beta_0^*)^{\zeta_2} - (1 - \zeta_1)(\beta_0^*)^{\zeta_1}} \\ &= \frac{-\zeta_1(\zeta_1 - 1)\left(\frac{x_1}{x_2}\right)^{\zeta_1-2} + \alpha_0^{\zeta_1-\zeta_2}\zeta_2(\zeta_2 - 1)\left(\frac{x_1}{x_2}\right)^{\zeta_2-2}}{x_2\left((1 - \zeta_1)(\beta_0^*)^{\zeta_1} - \alpha_0^{\zeta_1-\zeta_2}(1 - \zeta_2)(\beta_0^*)^{\zeta_2}\right)}. \end{aligned} \quad (6.A.11)$$

Here, the denominator is positive. That is, one needs to show that the numerator is non-positive:

$$\begin{aligned} &-\zeta_1(\zeta_1 - 1)\left(\frac{x_1}{x_2}\right)^{\zeta_1-2} + \alpha_0^{\zeta_1-\zeta_2}\zeta_2(\zeta_2 - 1)\left(\frac{x_1}{x_2}\right)^{\zeta_2-2} \leq 0 \\ \Leftrightarrow &\left(\frac{x_1}{x_2\alpha_0}\right)^{\zeta_2-\zeta_1} \leq \frac{\zeta_1(\zeta_1 - 1)}{\zeta_2(\zeta_2 - 1)} \\ \Leftrightarrow &\frac{x_1}{x_2\alpha_0} \leq \left(\frac{\zeta_1(\zeta_1 - 1)}{\zeta_2(\zeta_2 - 1)}\right)^{\frac{1}{\zeta_2-\zeta_1}} = \frac{\beta_0^*}{\alpha_0}, \end{aligned} \quad (6.A.12)$$

which is exactly the case since we consider the case $x_2 \geq \frac{x_1}{\beta_0^*}$. In the last step in the above calculation we have used (6.2.14). Moreover, we have that $\frac{\partial^2}{\partial x_1^2}G_A^{\beta_0^*}(\vec{x}) = 0$ for $x_2 \geq \frac{x_1}{\beta_0^*}$. That is, the first half of equation (6.A.10) holds.

For the case $x_2 \geq \frac{x_1}{\beta_0^*}$, the second-order derivative with respect to x_2 is given by

$$\begin{aligned} \frac{\partial^2}{\partial x_2^2}G_A(x_1, x_2; \beta_0^*) &= \frac{\zeta_1(\zeta_1 - 1)x_1^{\zeta_1}x_2^{-\zeta_1-1} - \alpha_0^{\zeta_1-\zeta_2}\zeta_2(\zeta_2 - 1)x_1^{\zeta_2}x_2^{-\zeta_2-1}}{\alpha_0^{\zeta_1-\zeta_2}(1 - \zeta_2)(\beta_0^*)^{\zeta_2} - (1 - \zeta_1)(\beta_0^*)^{\zeta_1}} \\ &= \frac{-\zeta_1(\zeta_1 - 1)\left(\frac{x_1}{x_2}\right)^{\zeta_1+1} + \alpha_0^{\zeta_1-\zeta_2}\zeta_2(\zeta_2 - 1)\left(\frac{x_1}{x_2}\right)^{\zeta_2+1}}{x_1\left((1 - \zeta_1)(\beta_0^*)^{\zeta_1} - \alpha_0^{\zeta_1-\zeta_2}(1 - \zeta_2)(\beta_0^*)^{\zeta_2}\right)}. \end{aligned} \quad (6.A.13)$$

Using again that the denominator is positive and by similar calculations as in (6.A.12) we get that the numerator is non-positive. Together with the observation that $\frac{\partial^2}{\partial x_2^2}G_A^{\beta_0^*}(\vec{x}) = 0$ for $x_2 \geq \frac{x_1}{\beta_0^*}$

we have that the second half of equation (6.A.10) also holds.

Also note that the second-order derivative with respect to x_2 vanishes at the optimal barrier: The numerator of (6.A.13) divided by $\alpha_0^{\zeta_1+1}$ is

$$-\zeta_1(\zeta_1 - 1) \left(\frac{x_1}{\alpha_0 x_2} \right)^{\zeta_1+1} + \zeta_2(\zeta_2 - 1) \left(\frac{x_1}{\alpha_0 x_2} \right)^{\zeta_2+1}.$$

For $x_2 = \frac{x_1}{\beta_0^*}$ and inserting the value of β_0^* given by (6.2.15) we get that

$$\begin{aligned} & -\zeta_1(\zeta_1 - 1) \left(\frac{\alpha_0 \left(\frac{\zeta_2(\zeta_2-1)}{\zeta_1(\zeta_1-1)} \right)^{\frac{1}{\zeta_1-\zeta_2}}}{\alpha_0} \right)^{\zeta_1+1} + \zeta_2(\zeta_2 - 1) \left(\frac{\alpha_0 \left(\frac{\zeta_2(\zeta_2-1)}{\zeta_1(\zeta_1-1)} \right)^{\frac{1}{\zeta_1-\zeta_2}}}{\alpha_0} \right)^{\zeta_2+1} \\ &= -\zeta_1(\zeta_1 - 1) \left(\frac{\zeta_2(\zeta_2 - 1)}{\zeta_1(\zeta_1 - 1)} \right)^{\frac{\zeta_1+1}{\zeta_1-\zeta_2}} + \zeta_2(\zeta_2 - 1) \left(\frac{\zeta_2(\zeta_2 - 1)}{\zeta_1(\zeta_1 - 1)} \right)^{\frac{\zeta_2+1}{\zeta_1-\zeta_2}} \\ &= \zeta_1(\zeta_1 - 1) \left(- \left(\frac{\zeta_2(\zeta_2 - 1)}{\zeta_1(\zeta_1 - 1)} \right)^{\frac{\zeta_1+1}{\zeta_1-\zeta_2}} + \left(\frac{\zeta_2(\zeta_2 - 1)}{\zeta_1(\zeta_1 - 1)} \right)^{\frac{\zeta_2+1}{\zeta_1-\zeta_2}+1} \right) = 0. \end{aligned}$$

That is, we have the smooth fit condition, $\frac{\partial^2}{\partial x_1^2} G_A(\vec{x}; \beta_0^*) \Big|_{x_2 = \frac{x_1}{\beta_0^*}} = 0$ holds. Hereafter, we are ready to show the four different conditions in (6.A.9) one by one:

1. Let $x_2 \in [\frac{x_1}{\beta_0^*}, \frac{x_1}{\alpha_0}]$: By construction of the candidate, we have that $(\mathcal{A} - \delta)G_A(\vec{x}) = 0$.
2. Let $x_2 < \frac{x_1}{\beta_0^*}$. Since $G_A(\cdot; \beta_0^*)$ is homogeneous we have that $G_A(x_1, \frac{x_1}{\beta_0^*}; \beta_0^*) = x_1 G_A(1, \frac{1}{\beta_0^*}; \beta_0^*)$. We use this and get that

$$\begin{aligned} (\mathcal{A} - \delta)G_A^{\beta_0^*}(\vec{x}) &= \mu_A x_1 \frac{\partial}{\partial x_1} G_A^{\beta_0^*}(\vec{x}) + \mu_L x_2 \frac{\partial}{\partial x_2} G_A^{\beta_0^*}(\vec{x}) + \frac{1}{2} \sigma_A^2 x_1^2 \frac{\partial^2}{\partial x_1^2} G_A^{\beta_0^*}(\vec{x}) \\ &\quad + \frac{1}{2} \sigma_L^2 x_2^2 \frac{\partial^2}{\partial x_2^2} G_A^{\beta_0^*}(\vec{x}) + \rho \sigma_A \sigma_L x_1 x_2 \frac{\partial^2}{\partial x_1 \partial x_2} G_A^{\beta_0^*}(\vec{x}) - \delta G_A^{\beta_0^*}(\vec{x}) \\ &= \mu_A \frac{x_1}{\beta_0^*} + \mu_A G_A \left(x_1, \frac{x_1}{\beta_0^*}; \beta_0^* \right) - \mu_L x_2 - \delta \left(\frac{x_1}{\beta_0^*} - x_2 + G_A \left(x_1, \frac{x_1}{\beta_0^*}; \beta_0^* \right) \right). \end{aligned} \tag{6.A.14}$$

We have by (6.A.8) that

$$G_A \left(x_1, \frac{x_1}{\beta_0^*}; \beta_0^* \right) = \frac{x_1}{\beta_0^*} \frac{\left(\frac{\beta_0^*}{\alpha_0} \right)^{\zeta_1} - \left(\frac{\beta_0^*}{\alpha_0} \right)^{\zeta_2}}{(1 - \zeta_2) \left(\frac{\beta_0^*}{\alpha_0} \right)^{\zeta_2} - (1 - \zeta_1) \left(\frac{\beta_0^*}{\alpha_0} \right)^{\zeta_1}}. \tag{6.A.15}$$

By (6.2.14) the optimal barrier, β_0^* fulfills the equation:

$$\zeta_1(\zeta_1 - 1) \left(\frac{\beta_0^*}{\alpha_0} \right)^{\zeta_1} = \zeta_2(\zeta_2 - 1) \left(\frac{\beta_0^*}{\alpha_0} \right)^{\zeta_2}. \tag{6.A.16}$$

By dividing each of the four terms of the big fraction in (6.A.15) with either the left or the right-hand side of (6.A.16) we obtain

$$\begin{aligned} G_A \left(x_1, \frac{x_1}{\beta_0^*}; \beta_0^* \right) &= \frac{x_1 \frac{1}{\zeta_1(\zeta_1-1)} - \frac{1}{\zeta_2(\zeta_2-1)}}{\beta_0^* \frac{1}{\zeta_1} - \frac{1}{\zeta_2}} \\ &= \frac{x_1}{\beta_0^*} \frac{\zeta_1 + \zeta_2 - 1}{\zeta_1 \zeta_2 - (\zeta_1 + \zeta_2 - 1)}. \end{aligned}$$

Using that ζ_1 and ζ_2 are given by (6.A.4), we get that $\zeta_1 + \zeta_2 - 1 = \frac{2(\mu_L - \mu_A)}{\bar{\sigma}^2}$ and that $\zeta_1 \zeta_2 = \frac{2(\mu_L - \delta)}{\bar{\sigma}^2}$. All in all,

$$G_A \left(x_1, \frac{x_1}{\beta_0^*}; \beta_0^* \right) = \frac{x_1}{\beta_0^*} \frac{\mu_A - \mu_L}{\delta - \mu_A}. \quad (6.A.17)$$

That is,

$$\begin{aligned} (\mathcal{A} - \delta) G_A^{\beta_0^*}(\vec{x}) &= \mu_A \frac{x_1}{\beta_0^*} - \mu_L x_2 + \mu_A \left(\frac{x_1}{\beta_0^*} \frac{\mu_A - \mu_L}{\delta - \mu_A} \right) - \delta \left(\frac{x_1}{\beta_0^*} - x_2 + \frac{x_1}{\beta_0^*} \frac{\mu_A - \mu_L}{\delta - \mu_A} \right) \\ &= (\mu_A - \delta) \frac{x_1}{\beta_0^*} + (\delta - \mu_L) x_2 - (\mu_A - \mu_L) \frac{x_1}{\beta_0^*} \\ &= (\mu_L - \delta) \frac{x_1}{\beta_0^*} + (\delta - \mu_L) x_2 \\ &= (\delta - \mu_L) \left(x_2 - \frac{x_1}{\beta_0^*} \right) \\ &< 0. \end{aligned} \quad (6.A.18)$$

3. Due to (6.2.9) and because $\frac{\partial^2}{\partial x_2^2} G_A^{\beta_0^*}(\vec{x})$ is non-positive for $x_2 \in \left[\frac{x_1}{\beta_0^*}, \frac{x_1}{\alpha_0} \right]$, we have that $\frac{\partial}{\partial x_2} G_A^{\beta_0^*}(\vec{x}) \leq -1$ for $x_2 \in \left[\frac{x_1}{\beta_0^*}, \frac{x_1}{\alpha_0} \right]$ as wanted.

4. Let $x_1 > \beta_0^* x_2$: By (6.2.5), we have that $\frac{\partial}{\partial x_2} G_A^{\beta_0^*}(\vec{x}) = -1$.

Finally, we show the boundedness of the partial derivatives of $G_A^{\beta_0^*}$.

Since $G_A^{\beta_0^*}$ is concave (in its first argument) we have for $x_1 \in [\alpha_0 x_2, \beta_0^* x_2]$ that

$$\frac{\partial}{\partial x_1} G_A^{\beta_0^*}(\vec{x}; \beta_0^*) \leq \frac{\zeta_1 \left(\frac{\alpha_0 x_2}{x_2} \right)^{\zeta_1 - 1} - \zeta_2 \alpha_0 \zeta_1 - \zeta_2 \left(\frac{\alpha_0 x_2}{x_2} \right)^{\zeta_2 - 1}}{\alpha_0 \zeta_1 - \zeta_2 (1 - \zeta_2) (\beta_0^*)^{\zeta_2} - (1 - \zeta_1) (\beta_0^*)^{\zeta_1}} \leq K_1 < \infty$$

for some constant K_1 . On the other hand, $\frac{\partial}{\partial x_1} G_A^{\beta_0^*}(\vec{x}; \beta_0^*) \geq \frac{1}{\beta_0^*} + G_A \left(1, \frac{1}{\beta_0^*}; \beta_0^* \right) := H_1 > -\infty$.

The partial derivative of $G_A^{\beta_0^*}$ with respect to x_2 is for $x_2 \in \left[\frac{x_1}{\beta_0^*}, \frac{x_1}{\alpha_0} \right]$ given by

$$\frac{\partial}{\partial x_2} G_A^{\beta_0^*}(\vec{x}; \beta_0^*) = \frac{(1 - \zeta_1) x_1^{\zeta_1} x_2^{-\zeta_1} - (1 - \zeta_2) \alpha_0^{\zeta_1 - \zeta_2} x_1^{\zeta_2} x_2^{-\zeta_2}}{\alpha_0 \zeta_1 - \zeta_2 (1 - \zeta_2) (\beta_0^*)^{\zeta_2} - (1 - \zeta_1) (\beta_0^*)^{\zeta_1}} < 0, \quad (6.A.19)$$

since the numerator is positive and the denominator is negative. Because $G_A^{\beta_0^*}$ is concave in the second argument, the minimum for the first derivative of $G_A^{\beta_0^*}(\vec{x}; \beta_0^*)$ with respect to x_2 is attained for the biggest possible value of x_2 , which is $x_2 = \frac{x_1}{\alpha_0}$. We insert in (6.A.19) and get

$$\left. \frac{\partial}{\partial x_2} G_A^{\beta_0^*}(\vec{x}; \beta_0^*) \right|_{x_2 = \frac{x_1}{\alpha_0}} = \frac{(1 - \zeta_1)\alpha_0^{\zeta_1} - (1 - \zeta_2)\alpha_0^{\zeta_1}}{\alpha_0^{\zeta_1 - \zeta_2}(1 - \zeta_2)(\beta_0^*)^{\zeta_2} - (1 - \zeta_1)(\beta_0^*)^{\zeta_1}} > K_2 > -\infty$$

for some constant K_2 . On the other hand we have previously proved that $G_A^{\beta_0^*}(\vec{x}; \beta_0^*) \leq -1$. In total we get that $K_2 < \frac{\partial}{\partial x_2} G_A^{\beta_0^*}(\vec{x}; \beta_0^*) \leq -1$. This concludes the proof. $\square$

Lemma 6.A.1. *Let $\alpha_1 > \beta^*$, where β^* is the optimal barrier. Then it holds*

$$G_A\left(x_1, \frac{x_1}{\alpha_1}; \alpha_1\right) \geq \left(\frac{\beta^*}{\alpha_1}\right) G_A\left(x_1, \frac{x_1}{\beta^*}; \beta^*\right).$$

Proof of Lemma 6.A.1. Using (6.A.7) and that $\frac{\alpha_1}{\beta^*} > 1$ we get that

$$\begin{aligned} G_A\left(x_1, \frac{x_1}{\alpha_1}; \alpha_1\right) &= \frac{x_1^{\zeta_1} \left(\frac{x_1}{\alpha_1}\right)^{1-\zeta_1} - \alpha_0^{\zeta_1 - \zeta_2} x_1^{\zeta_2} \left(\frac{x_1}{\alpha_1}\right)^{1-\zeta_2}}{\alpha_0^{\zeta_1 - \zeta_2}(1 - \zeta_2)\alpha_1^{\zeta_2} - (1 - \zeta_1)\alpha_1^{\zeta_1}} \\ &= x_1 \frac{\left(\beta^* \left(\frac{\alpha_1}{\beta^*}\right)\right)^{\zeta_1 - 1} - \alpha_0^{\zeta_1 - \zeta_2} \left(\beta^* \left(\frac{\alpha_1}{\beta^*}\right)\right)^{\zeta_2 - 1}}{\alpha_0^{\zeta_1 - \zeta_2}(1 - \zeta_2) \left(\beta^* \left(\frac{\alpha_1}{\beta^*}\right)\right)^{\zeta_2} - (1 - \zeta_1) \left(\beta^* \left(\frac{\alpha_1}{\beta^*}\right)\right)^{\zeta_1}} \\ &= x_1 \frac{(\beta^*)^{\zeta_1 - 1} \left(\frac{\alpha_1}{\beta^*}\right)^{\zeta_1 - \zeta_2} - \alpha_0^{\zeta_1 - \zeta_2} (\beta^*)^{\zeta_2 - 1}}{\alpha_0^{\zeta_1 - \zeta_2}(1 - \zeta_2) (\beta^*)^{\zeta_2} \left(\frac{\alpha_1}{\beta^*}\right)^{\zeta_1 - \zeta_2 + 1} - (1 - \zeta_1) (\beta^*)^{\zeta_1} \left(\frac{\alpha_1}{\beta^*}\right)^{\zeta_1 - \zeta_2 + 1}} \\ &= x_1 \left(\frac{\beta^*}{\alpha_1}\right) \frac{(\beta^*)^{\zeta_1 - 1} \left(\frac{\alpha_1}{\beta^*}\right)^{\zeta_1 - \zeta_2} - \alpha_0^{\zeta_1 - \zeta_2} (\beta^*)^{\zeta_2 - 1}}{\alpha_0^{\zeta_1 - \zeta_2}(1 - \zeta_2) (\beta^*)^{\zeta_2} - (1 - \zeta_1) (\beta^*)^{\zeta_1} \left(\frac{\alpha_1}{\beta^*}\right)^{\zeta_1 - \zeta_2}} \\ &\geq \left(\frac{\beta^*}{\alpha_1}\right) G_A\left(x_1, \frac{x_1}{\beta^*}; \beta^*\right). \end{aligned}$$

The above inequality follows in the following way: The big fraction above consists of a negative numerator and a negative denominator. The numerator consists of a positive term subtracted by a positive term, whereas the denominator consists of negative term subtracted by a positive term. Since $\left(\frac{\alpha_1}{\beta^*}\right)^{\zeta_1 - \zeta_2} < 1$, removing this term lead to a bigger (still negative) numerator and lower value for the denominator. That is, removing the term leads to a lower value. This concludes the proof. $\square$

6.A.3 Proof of Lemma 6.2.3 (Verification Lemma)

Proof of Verification Lemma (Lemma 6.2.3). We let D_A^π be a general control strategy (not limited to barrier strategies). For the sake of simplicity we write X^π instead of X_A^π in the proof.

Moreover, we denote the continuous part of X^π by $X^{\pi,c}$. By Itô's Lemma we get that

$$\begin{aligned}
& e^{-\delta(t \wedge \tau_{\alpha_0}-)} G(X_1(t \wedge \tau_{\alpha_0}-), X_2^\pi(t \wedge \tau_{\alpha_0}-)) \\
&= G(A_0, L_0) - \int_{0-}^{t \wedge \tau_{\alpha_0}-} \delta e^{-\delta s} G(X_1(s), X_2^\pi(s)) ds \\
& \quad + \sum_{i=1}^2 \int_{0-}^{t \wedge \tau_{\alpha_0}-} e^{-\delta s} \frac{\partial}{\partial x_i} G(X_1(s), X_2^\pi(s)) dX_i^{\pi,c}(s) \\
& \quad + \int_{0-}^{t \wedge \tau_{\alpha_0}-} e^{-\delta s} \frac{\sigma_A^2(X_1(s))^2}{2} \frac{\partial^2}{\partial x_1^2} G(X_1(s), X_2^\pi(s)) ds \\
& \quad + \int_{0-}^{t \wedge \tau_{\alpha_0}-} e^{-\delta s} \frac{\sigma_L^2(X_2^\pi(s))^2}{2} \frac{\partial^2}{\partial x_2^2} G(X_1(s), X_2^\pi(s)) ds \\
& \quad + \int_{0-}^{t \wedge \tau_{\alpha_0}-} e^{-\delta s} \rho \sigma_A \sigma_L X_1(s) X_2^\pi(s) \frac{\partial^2}{\partial x_1 \partial x_2} G(X_1(s), X_2^\pi(s)) ds \\
& \quad + \sum_{\substack{\Delta X_2^\pi(s) \neq 0 \\ s \leq t \wedge \tau_{\alpha_0}-}} e^{-\delta s} (G(X_1(s), X_2^\pi(s-)) + \Delta X_2^\pi(s) - G(X_1(s), X_2^\pi(s-))).
\end{aligned}$$

We insert the dynamics of X^π , collect terms and obtain:

$$\begin{aligned}
& e^{-\delta(t \wedge \tau_{\alpha_0}-)} G(X_1(t \wedge \tau_{\alpha_0}-), X_2^\pi(t \wedge \tau_{\alpha_0}-)) \\
&= G(A_0, L_0) + \int_{0-}^{t \wedge \tau_{\alpha_0}-} e^{-\delta s} (\mathcal{A} - \delta) G(X_1(s), X_2^\pi(s)) ds \\
& \quad + \int_{0-}^{t \wedge \tau_{\alpha_0}-} e^{-\delta s} \sigma_A X_1(s) \frac{\partial}{\partial x_1} G(X_1(s), X_2^\pi(s)) dW_1(s) \\
& \quad + \int_{0-}^{t \wedge \tau_{\alpha_0}-} e^{-\delta s} \sigma_L X_2^\pi(s) \frac{\partial}{\partial x_2} G(X_1(s), X_2^\pi(s)) (\rho dW_1(s) + \sqrt{1 - \rho^2} dW_2(s)) \\
& \quad + \int_{0-}^{t \wedge \tau_{\alpha_0}-} e^{-\delta s} \frac{\partial}{\partial x_2} G(X_1(s), X_2^\pi(s)) dD_A^{\pi,c}(s) \\
& \quad + \sum_{\substack{\Delta X_2^\pi(s) \neq 0 \\ s \leq t \wedge \tau_{\alpha_0}-}} e^{-\delta s} (G(X_1(s), X_2^\pi(s)) - G(X_1(s), X_2^\pi(s-))) \\
&= G(A_0, L_0) + \int_{0-}^{t \wedge \tau_{\alpha_0}-} e^{-\delta s} (\mathcal{A} - \delta) G(X_1(s), X_2^\pi(s)) ds \tag{6.A.20} \\
& \quad + \int_{0-}^{t \wedge \tau_{\alpha_0}-} e^{-\delta s} \sigma_A X_1(s) \frac{\partial}{\partial x_1} G(X_1(s), X_2^\pi(s)) dW_1(s) \\
& \quad + \int_{0-}^{t \wedge \tau_{\alpha_0}-} e^{-\delta s} \sigma_L X_2^\pi(s) \frac{\partial}{\partial x_2} G(X_1(s), X_2^\pi(s)) (\rho dW_1(s) + \sqrt{1 - \rho^2} dW_2(s)) \\
& \quad - \int_{0-}^{t \wedge \tau_{\alpha_0}-} e^{-\delta s} dD_A^\pi(s) - \int_{0-}^{t \wedge \tau_{\alpha_0}-} e^{-\delta s} \left(-1 - \frac{\partial}{\partial x_2} G(X_1(s), X_2^\pi(s-)) \right) dD_A^\pi(s) \\
& \quad + \sum_{\substack{\Delta X_2^\pi(s) \neq 0 \\ s \leq t \wedge \tau_{\alpha_0}-}} e^{-\delta s} \left(G(X_1(s), X_2^\pi(s)) - G(X_1(s), X_2^\pi(s-)) \right. \\
& \quad \quad \left. - \frac{\partial}{\partial x_2} G(X_1(s), X_2^\pi(s-)) (X_2^\pi(s) - X_2^\pi(s-)) \right),
\end{aligned}$$

where the operator $\mathcal{A}$ is given by (6.2.4).

Now, the aim is to prove that

$$G(A_0, L_0) \geq J_{A,1}(\vec{x}; \pi_d) = E^{A_0, L_0} \left[\int_{0-}^{\tau_{\pi_d}-} e^{-\delta s} dD_A^\pi(s) \right]. \quad (6.A.21)$$

We rearrange the terms and get that

$$G(A_0, L_0) \quad (6.A.22)$$

$$= \int_{0-}^{t \wedge \tau_{\alpha_0}-} e^{-\delta s} dD_A^\pi(s) + e^{-\delta(t \wedge \tau_{\alpha_0}-)} G(X_1(t \wedge \tau_{\alpha_0}-), X_2^\pi(t \wedge \tau_{\alpha_0}-)) \quad (6.A.23)$$

$$- \int_{0-}^{t \wedge \tau_{\alpha_0}-} e^{-\delta s} (\mathcal{A} - \delta) G(X_1(s), X_2^\pi(s)) ds \quad (6.A.24)$$

$$- \int_{0-}^{t \wedge \tau_{\alpha_0}-} e^{-\delta s} \sigma_A X_1(s) \frac{\partial}{\partial x_1} G(X_1(s), X_2^\pi(s)) dW_1(s) \quad (6.A.25)$$

$$- \int_{0-}^{t \wedge \tau_{\alpha_0}-} e^{-\delta s} \sigma_L X_2^\pi(s) \frac{\partial}{\partial x_2} G(X_1(s), X_2^\pi(s)) \left(\rho dW_1(s) + \sqrt{1 - \rho^2} dW_2(s) \right) \quad (6.A.26)$$

$$+ \int_{0-}^{t \wedge \tau_{\alpha_0}-} e^{-\delta s} \left(-1 - \frac{\partial}{\partial x_2} G(X_1(s), X_2^\pi(s-)) \right) dD_A^\pi(s) \quad (6.A.27)$$

$$- \sum_{\substack{\Delta X_2^\pi(s) \neq 0 \\ s \leq t \wedge \tau_{\alpha_0}-}} e^{-\delta s} \left(G(X_1(s), X_2^\pi(s)) - G(X_1(s), X_2^\pi(s-)) \right. \\ \left. - \frac{\partial}{\partial x_2} G(X_1(s), X_2^\pi(s-)) (X_2^\pi(s) - X_2^\pi(s-)) \right). \quad (6.A.28)$$

To prove the two stated results, we need to show that the limit of (6.A.23) is ≥ 0 , that

$$(\mathcal{A} - \delta) G(X_1(s), X_2^\pi(s)) \leq 0,$$

some integrability conditions such that the integrals with respect to the Brownian motions vanish when taking the expected values of (6.A.25) and (6.A.26). Moreover, we need to show that $\left(-1 - \frac{\partial}{\partial x_1} G(\vec{x}) \right) \geq 0$, such that (6.A.27) is positive, and that for $z \geq y$ it holds that

$$z \geq y : G(x_1, z) - G(x_1, y) - (z - y) \frac{\partial}{\partial x_2} G(\vec{x}) \Big|_{x_2=y} \leq 0,$$

such that also the term (6.A.28) is positive. If we are able to show these results, we have that

$G(A_0, L_0) \geq \mathbb{E}^{A_0, L_0} \left[\int_{0-}^{t \wedge \tau_{\alpha_0}-} e^{-\delta s} dD_A^\pi(s) \right]$. We consider each of the terms separately.

We have that

$$\begin{aligned}
& \lim_{t \rightarrow \infty} e^{-\delta(t \wedge \tau_{\alpha_0}-)} G(X_1(t \wedge \tau_{\alpha_0}-), X_2^\pi(t \wedge \tau_{\alpha_0}-)) \\
&= \lim_{t \rightarrow \infty} e^{-\delta(t \wedge \tau_{\alpha_0}-)} G(X_1(t \wedge \tau_{\alpha_0}-), X_2^\pi(t \wedge \tau_{\alpha_0}-)) \mathbf{1}_{\{\tau_{\alpha_0} < \infty\}} \\
&\quad + \lim_{t \rightarrow \infty} e^{-\delta(t \wedge \tau_{\alpha_0}-)} G(X_1(t \wedge \tau_{\alpha_0}-), X_2^\pi(t \wedge \tau_{\alpha_0}-)) \mathbf{1}_{\{\tau_{\alpha_0} = \infty\}} \\
&= e^{-\delta(\tau_{\alpha_0}-)} G(\alpha_0 X_2^\pi(\tau_{\alpha_0}-), X_2^\pi(\tau_{\alpha_0}-)) \mathbf{1}_{\{\tau_{\alpha_0} < \infty\}} \\
&\quad + \lim_{t \rightarrow \infty} e^{-\delta(t \wedge \tau_{\alpha_0}-)} G(X_1(t \wedge \tau_{\alpha_0}-), X_2^\pi(t \wedge \tau_{\alpha_0}-)) \mathbf{1}_{\{\tau_{\alpha_0} = \infty\}} \\
&\geq 0.
\end{aligned}$$

That $(\mathcal{A} - \delta) G(X_1(s), X_2^\pi(s)) \leq 0$ is exactly assumption (6.2.16). To show that the expected values of the integrals with respect to the two Brownian motions are zero, we define a monotone increasing sequence of stopping times by $\rho_n = \inf\{t : e^{-\delta t} X_1(t) \geq n\}$. We now consider the truncated integral given by

$$\int_{0-}^{t \wedge \tau_{\alpha_0} - \wedge \rho_n} e^{-\delta s} \sigma_A X_1(s) \frac{\partial}{\partial x_1} G(X_1(s), X_2^\pi(s)) dW_1(s). \quad (6.A.29)$$

We know that the expected value of (6.A.29) is zero, because the integrand is bounded. This means (taking the rest of the calculations into account) that

$$G(A_0, L_0) \geq \mathbb{E}^{A_0, L_0} \left[\int_{0-}^{t \wedge \tau_{\alpha_0} - \wedge \rho_n} e^{-\delta s} dD_A^\pi(s) \right].$$

By letting $t \rightarrow \infty$ and $n \rightarrow \infty$ we get by using Lebesgue's monotone convergence theorem that

$$G(A_0, L_0) \geq \mathbb{E}^{A_0, L_0} \left[\int_{0-}^{\tau_{\alpha_0}-} e^{-\delta s} dD_A^\pi(s) \right].$$

Using that the first derivative of the second variable of G is bounded by assumption (6.2.20), we get by calculations analogously to the ones for the term (6.A.25) that (6.A.26) = 0. That (6.A.27) is greater than or equal to 0 follows directly from assumption (6.2.18). Finally, because of the concavity of G in its first argument, which follows by assumption (6.2.17), we have that for $z > y$:

$$\left. \frac{\partial}{\partial x_2} G(x_1, x_2) \right|_{x_2=y} \geq \frac{G(x_1, z) - G(x_1, y)}{z - y} \Rightarrow 0 \geq G(x_1, z) - G(x_1, y) - (z - y) \left. \frac{\partial}{\partial x_2} G(x_1, x_2) \right|_{x_2=y},$$

this implies that (6.A.28) ≥ 0 . This ends the proof of the first part of the lemma for the case

$\alpha_1 \leq \beta^*$. The case, $\alpha_1 > \beta^*$ will be covered below.

For the second half of the proof we show that optimal strategy is in fact a barrier strategy.

We take expectations of (6.A.22), rearrange and obtain

$$\begin{aligned}
& \mathbb{E}^{A_0, L_0} \left[e^{-\delta(t \wedge \tau_{\alpha_0} -)} G(X_1(t \wedge \tau_{\alpha_0} -), X_2^\pi(t \wedge \tau_{\alpha_0} -)) \right] \\
&= G(A_0, L_0) + \mathbb{E}^{A_0, L_0} \left[\int_{0-}^{t \wedge \tau_{\alpha_0} -} e^{-\delta s} (\mathcal{A} - \delta) G(X_1(s), X_2^\pi(s)) ds \right] \\
&+ \mathbb{E}^{A_0, L_0} \left[\int_{0-}^{t \wedge \tau_{\alpha_0} -} e^{-\delta s} \frac{\partial}{\partial x_2} G(X_1(s), X_2^\pi(s-)) dD_A^\pi(s) \right] \\
&+ \mathbb{E}^{A_0, L_0} \left[\sum_{\substack{\Delta X_2^\pi(s) \neq 0 \\ s \leq t \wedge \tau_{\alpha_0} -}} e^{-\delta s} (G(X_1(s), X_2^\pi(s)) - G(X_1(s), X_2^\pi(s-))) \right. \\
&\quad \left. - \frac{\partial}{\partial x_2} G(X_1(s), X_2^\pi(s)) (X_2^\pi(s) - X_2^\pi(s-)) \right]. \tag{6.A.30}
\end{aligned}$$

We have that $(\mathcal{A} - \delta) G(X_1(s), X_2^\pi(s)) = 0$, since we consider a barrier strategy which means that $X_1 \leq \beta^* X_2^\pi$. The last term of (6.A.30) is zero due to the proposed dividend strategy. In total we get that,

$$\begin{aligned}
& \mathbb{E}^{A_0, L_0} \left[e^{-\delta(t \wedge \tau_{\alpha_0} -)} G(X_1(t \wedge \tau_{\alpha_0} -), X_2^\pi(t \wedge \tau_{\alpha_0} -)) \right] \\
&= G(A_0, L_0) + \mathbb{E}^{A_0, L_0} \left[\int_{0-}^{t \wedge \tau_{\alpha_0} -} e^{-\delta s} \frac{\partial}{\partial x_2} G(X_1(s), X_2^\pi(s-)) dD_A^\pi(s) \right] \\
&= G(A_0, L_0) + \mathbb{E}^{A_0, L_0} \left[\int_{0-}^{t \wedge \tau_{\alpha_0} -} e^{-\delta s} \frac{\partial}{\partial x_2} G(X_1(s), X_2^\pi(s-)) \mathbf{1}_{\{X_1(s) \geq \beta^* X_2^\pi(s)\}} dD_A^\pi(s) \right] \tag{6.A.31} \\
&= G(A_0, L_0) - \mathbb{E}^{A_0, L_0} \left[\int_{0-}^{t \wedge \tau_{\alpha_0} -} e^{-\delta s} dD_A^\pi(s) \right],
\end{aligned}$$

where we have used assumption (6.2.22). Letting $t \rightarrow \infty$ in (6.A.31) we get that

$$G(A_0, L_0) = \mathbb{E}^{A_0, L_0} \left[\lim_{t \rightarrow \infty} \int_{0-}^{t \wedge \tau_{\alpha_0} -} e^{-\delta s} dD_A^\pi(s) \right].$$

We now assume that $\beta^* < \alpha_1$. We denote by V_{α_1} the value function for a barrier strategy with barrier for the assets given by $\alpha_1 L$. The value function V_{α_1} is twice continuously differentiable in both arguments except in the point $(\alpha_1 x_2, x_2)$. However, since $V_{\alpha_1}(\alpha_1 x_2 -) = V_{\alpha_1}(\alpha_1 x_2 +) = 1$,

we can use Itô's Lemma and obtain

$$\begin{aligned}
& e^{-\delta(t \wedge \tau_{\alpha_0}-)} V_{\alpha_1}(X_1(t \wedge \tau_{\alpha_0}-), X_2^\pi(t \wedge \tau_{\alpha_0}-)) \\
&= \int_{0-}^{t \wedge \tau_{\alpha_0}-} e^{-\delta s} \left(\frac{\mu_A}{\alpha_1} X_1(s) - \mu_L X_2^\pi(s) + \mu_A G_A \left(X_1(s), \frac{X_1(s)}{\alpha_1}; \alpha_1 \right) \right. \\
&\quad \left. - \delta \left(\frac{X_1(s)}{\alpha_1} - X_2^\pi(s) + G_A \left(X_1(s), \frac{X_1(s)}{\alpha_1}; \alpha_1 \right) \right) \right) \mathbf{1}_{\{Y_A^\pi(s) > \alpha_1\}} ds \\
&\quad + V_{\alpha_1}(A_0, L_0) + \int_{0-}^{t \wedge \tau_{\alpha_0}-} e^{-\delta s} \sigma_A X_1(s) \frac{\partial}{\partial x_1} V_{\alpha_1}(X_1(s), X_2^\pi(s)) dW_1(s) \\
&\quad + \int_{0-}^{t \wedge \tau_{\alpha_0}-} e^{-\delta s} \sigma_L X_2^\pi(s) \frac{\partial}{\partial x_2} V_{\alpha_1}(X_1(s), X_2^\pi(s)) \left(\rho dW_1(s) + \sqrt{1 - \rho^2} dW_2(s) \right) \\
&\quad + \int_{0-}^{t \wedge \tau_{\alpha_0}-} \frac{\partial}{\partial x_2} V_{\alpha_1}(X_1, X_2^\pi(s-)) e^{-\delta s} dD_A^\pi(s),
\end{aligned} \tag{6.A.32}$$

where we have used that V_{α_1} is homogeneous. By the same arguments as for the case $\alpha_1 \leq \beta^*$, we can show that the derivatives are bounded, and that we ultimately get that the mean value of the integrals with respect to the Brownian motion is 0. We only increase liabilities for $X_2^\pi \leq \frac{X_1(s)}{\alpha_1}$ in which case $\frac{\partial}{\partial x_2} V_{\alpha_1}(X_1, X_2^\pi(s-)) = -1$. Moreover, we also have that the left-hand side of (6.A.32) is 0 in the limit. That is, we can show that

$$\lim_{t \rightarrow \infty} \mathbb{E} \left[\int_{0-}^{t \wedge \tau_{\alpha_0}-} e^{-\delta s} dD_A^\pi(s) \right] \leq V_{\alpha_1}(A_0, L_0)$$

by showing that for all positive values t it holds that

$$\begin{aligned}
& \mathbb{E} \left[\int_{0-}^{t \wedge \tau_{\alpha_0}-} e^{-\delta s} \left(\frac{\mu_A}{\alpha_1} X_1(s) - \mu_L X_2^\pi(s) + \mu_A G_A \left(X_1(s), \frac{X_1(s)}{\alpha_1}; \alpha_1 \right) \right. \right. \\
&\quad \left. \left. - \delta \left(\frac{X_1(s)}{\alpha_1} - X_2^\pi(s) + G_A \left(X_1(s), \frac{X_1(s)}{\alpha_1}; \alpha_1 \right) \right) \right) \mathbf{1}_{\{Y_A^\pi(s) > \alpha_1\}} ds \right] \leq 0.
\end{aligned}$$

To show this inequality, we show, that for $\frac{x_1}{\alpha_1} > x_2$ we have that

$$\frac{\mu_A}{\alpha_1} x_1 - \mu_L x_2 + \mu_A G_A \left(x_1, \frac{x_1}{\alpha_1}; \alpha_1 \right) - \delta \left(\frac{1}{\alpha_1} x_1 - x_2 + G_A \left(x_1, \frac{x_1}{\alpha_1}; \alpha_1 \right) \right) \leq 0.$$

Using (6.1.2), (6.1.3), Lemma 6.A.1 (p. 149), that $\frac{x_1}{\alpha_1} > x_2$ and that $G_A \left(x_1, \frac{x_1}{\beta^*}; \beta^* \right)$ is given by (6.A.17), we get that

$$\begin{aligned}
& \frac{\mu_A}{\alpha_1} x_1 - \mu_L x_2 + \mu_A G_A \left(x_1, \frac{x_1}{\alpha_1}; \alpha_1 \right) - \delta \left(\frac{x_1}{\alpha_1} - x_2 + G_A \left(x_1, \frac{x_1}{\alpha_1}; \alpha_1 \right) \right) \\
&\leq x_2 \left(\mu_A - \delta + \delta - \mu_L - (\delta - \mu_A) \frac{G_A \left(x_1, \frac{x_1}{\alpha_1}; \alpha_1 \right)}{x_2} \right) \\
&\leq x_2 \left(\mu_A - \mu_L - \frac{\delta - \mu_A}{x_2} \left(\frac{\beta^*}{\alpha_1} \right) \left(\frac{x_1}{\beta^*} \frac{\mu_A - \mu_L}{\delta - \mu_A} \right) \right) \\
&\leq x_2 (\mu_A - \mu_L - (\mu_A - \mu_L)) \\
&= 0.
\end{aligned}$$

That is, we get that (after taking limits)

$$V_{\alpha_1}(A_0, L_0) \geq E^{A_0, L_0} \left[\int_{0-}^{\tau_{\alpha_0}-} e^{-\delta s} dD_A^\pi(s) \right].$$

□

6.A.4 Proof of Theorem 6.3.1 (Value function under advanced solvency constraints)

Proof of Theorem 6.3.1. Using (6.3.7) and (6.3.9) we get:

$$C_2 = -C_1 \alpha_0^{\zeta_1 - \zeta_2}, \quad (6.A.33)$$

such that

$$V_{A,2}^0(x_1, x_2; \beta) = C_1 x_1^{\zeta_1} x_2^{1-\zeta_1} - C_1 \alpha_0^{\zeta_1 - \zeta_2} x_1^{\zeta_2} x_2^{1-\zeta_2}.$$

Condition (6.3.11) states that

$$C_1 x_1^{\zeta_1} \left(\frac{x_1}{\alpha_1} \right)^{1-\zeta_1} + C_2 x_1^{\zeta_2} \left(\frac{x_1}{\alpha_1} \right)^{1-\zeta_2} = \tilde{C}_1 x_1^{\zeta_1} \left(\frac{x_1}{\alpha_1} \right)^{1-\zeta_1} + \tilde{C}_2 x_1^{\zeta_2} \left(\frac{x_1}{\alpha_1} \right)^{1-\zeta_2}.$$

This is a binding condition for either $\tilde{C}_1$ or $\tilde{C}_2$, whereas the other parameter can vary freely. That is, we choose to represent $\tilde{C}_1$ as

$$\begin{aligned} \tilde{C}_1 &= \frac{C_1 x_1^{\zeta_1} \left(\frac{x_1}{\alpha_1} \right)^{1-\zeta_1} + C_2 x_1^{\zeta_2} \left(\frac{x_1}{\alpha_1} \right)^{1-\zeta_2} - \tilde{C}_2 x_1^{\zeta_2} \left(\frac{x_1}{\alpha_1} \right)^{1-\zeta_2}}{x_1^{\zeta_1} \left(\frac{x_1}{\alpha_1} \right)^{1-\zeta_1}} \\ &= C_1 + C_2 x_1^{\zeta_2 - \zeta_1} \left(\frac{x_1}{\alpha_1} \right)^{-\zeta_2 + \zeta_1} - \tilde{C}_2 x_1^{\zeta_2 - \zeta_1} \left(\frac{x_1}{\alpha_1} \right)^{-\zeta_2 + \zeta_1} \\ &= C_1 + C_2 \alpha_1^{\zeta_2 - \zeta_1} - \tilde{C}_2 \alpha_1^{\zeta_2 - \zeta_1}. \end{aligned} \quad (6.A.34)$$

Using (6.3.8), (6.A.34) and (6.A.33), we get that $V_{A,2}^1$ has the representation

$$V_{A,2}^1(x_1, x_2; \beta) = \left(C_1 - C_1 \alpha_0^{\zeta_1 - \zeta_2} \alpha_1^{\zeta_2 - \zeta_1} - \tilde{C}_2 \alpha_1^{\zeta_2 - \zeta_1} \right) x_1^{\zeta_1} x_2^{1-\zeta_1} + \tilde{C}_2 x_1^{\zeta_2} x_2^{1-\zeta_2}. \quad (6.A.35)$$

That is, we have the representations for $V_{A,2}^0$ and $V_{A,2}^1$ given in the theorem but we still need to determine C_1 and $\tilde{C}_2$. The partial derivative of $V_{A,2}^1$ with respect to x_2 in the point $\frac{x_1}{\beta}$ is given by

$$\begin{aligned} \frac{\partial}{\partial x_2} V_{A,2}^1(x_1, x_2; \beta) \Big|_{x_2 = \frac{x_1}{\beta}} &= (1 - \zeta_1) \left(C_1 - C_1 \alpha_0^{\zeta_1 - \zeta_2} \alpha_1^{\zeta_2 - \zeta_1} - \tilde{C}_2 \alpha_1^{\zeta_2 - \zeta_1} \right) x_1^{\zeta_1} \left(\frac{x_1}{\beta} \right)^{-\zeta_1} \\ &\quad + (1 - \zeta_2) \tilde{C}_2 x_1^{\zeta_2} \left(\frac{x_1}{\beta} \right)^{-\zeta_2} \\ &= (1 - \zeta_1) \left(C_1 - C_1 \alpha_0^{\zeta_1 - \zeta_2} \alpha_1^{\zeta_2 - \zeta_1} - \tilde{C}_2 \alpha_1^{\zeta_2 - \zeta_1} \right) \beta^{\zeta_1} \\ &\quad + (1 - \zeta_2) \tilde{C}_2 \beta^{\zeta_2}. \end{aligned}$$

This means, we get the following equation by using condition (6.3.12):

$$(1 - \zeta_1) \left(C_1 - C_1 \alpha_0^{\zeta_1 - \zeta_2} \alpha_1^{\zeta_2 - \zeta_1} - \tilde{C}_2 \alpha_1^{\zeta_2 - \zeta_1} \right) \beta^{\zeta_1} + (1 - \zeta_2) \tilde{C}_2 \beta^{\zeta_2} = -1.$$

From this it follows that

$$\tilde{C}_2 \left((1 - \zeta_2) \beta^{\zeta_2} - (1 - \zeta_1) \alpha_1^{\zeta_2 - \zeta_1} \beta^{\zeta_1} \right) + (1 - \zeta_1) \left(C_1 - C_1 \alpha_0^{\zeta_1 - \zeta_2} \alpha_1^{\zeta_2 - \zeta_1} \right) \beta^{\zeta_1} = -1,$$

such that $\tilde{C}_2$ is given by

$$\tilde{C}_2 = \frac{-1 - (1 - \zeta_1) (C_1 - C_1 \alpha_0^{\zeta_1 - \zeta_2} \alpha_1^{\zeta_2 - \zeta_1}) \beta^{\zeta_1}}{(1 - \zeta_2) \beta^{\zeta_2} - (1 - \zeta_1) \alpha_1^{\zeta_2 - \zeta_1} \beta^{\zeta_1}}. \quad (6.A.36)$$

We can represent $V_{A,2}^0 \left(x_1, \frac{x_1}{\alpha_2}; \beta \right)$ by

$$\begin{aligned} V_{A,2}^0 \left(x_1, \frac{x_1}{\alpha_2}; \beta \right) &= C_1 x_1^{\zeta_1} \left(\frac{x_1}{\alpha_2} \right)^{1 - \zeta_1} - C_1 \alpha_0^{\zeta_1 - \zeta_2} x_1^{\zeta_2} \left(\frac{x_1}{\alpha_2} \right)^{1 - \zeta_2} \\ &= \left(\frac{x_1}{\alpha_2} \right) C_1 \left(\alpha_2^{\zeta_1} - \alpha_0^{\zeta_1 - \zeta_2} \alpha_2^{\zeta_2} \right), \end{aligned} \quad (6.A.37)$$

and $V_{A,2}^1 \left(x_1, \frac{x_1}{\alpha_2}; \beta \right)$ by

$$\begin{aligned} V_{A,2}^1 \left(x_1, \frac{x_1}{\alpha_2}; \beta \right) &= \left(\frac{1}{\beta} - \frac{1}{\alpha_2} \right)^+ x_1 + V_{A,2}^1 \left(x_1, \left(\frac{x_1}{\min(\beta, \alpha_2)} \right); \beta \right) \\ &= x_1 \left(\left(\frac{1}{\beta} - \frac{1}{\alpha_2} \right)^+ + \left(C_1 - C_1 \alpha_0^{\zeta_1 - \zeta_2} \alpha_1^{\zeta_2 - \zeta_1} - \tilde{C}_2 \alpha_1^{\zeta_2 - \zeta_1} \right) \min(\beta, \alpha_2)^{\zeta_1 - 1} \right. \\ &\quad \left. + \tilde{C}_2 \min(\beta, \alpha_2)^{\zeta_2 - 1} \right), \end{aligned} \quad (6.A.38)$$

where we have used (6.A.35). By condition (6.3.10), we have that (6.A.37) equals (6.A.38). Using this and rearranging the terms give us that

$$\begin{aligned} &C_1 \left(\alpha_2^{\zeta_1 - 1} - \alpha_0^{\zeta_1 - \zeta_2} \alpha_2^{\zeta_2 - 1} + (\alpha_0^{\zeta_1 - \zeta_2} \alpha_1^{\zeta_2 - \zeta_1} - 1) \min(\beta, \alpha_2)^{\zeta_1 - 1} \right) \\ &= \left(\frac{1}{\beta} - \frac{1}{\alpha_2} \right)^+ - \tilde{C}_2 \left(\alpha_1^{\zeta_2 - \zeta_1} \min(\beta, \alpha_2)^{\zeta_1 - 1} - \min(\beta, \alpha_2)^{\zeta_2 - 1} \right). \end{aligned} \quad (6.A.39)$$

Inserting $\tilde{C}_2$ given by (6.A.36) in (6.A.39) leads to the equation

$$\begin{aligned} &C_1 \left(\alpha_2^{\zeta_1 - 1} - \alpha_0^{\zeta_1 - \zeta_2} \alpha_2^{\zeta_2 - 1} + (\alpha_0^{\zeta_1 - \zeta_2} \alpha_1^{\zeta_2 - \zeta_1} - 1) \min(\beta, \alpha_2)^{\zeta_1 - 1} \right) \\ &= \left(\frac{1}{\beta} - \frac{1}{\alpha_2} \right)^+ + \xi \left(-1 - (1 - \zeta_1) \left(C_1 - C_1 \alpha_0^{\zeta_1 - \zeta_2} \alpha_1^{\zeta_2 - \zeta_1} \right) \beta^{\zeta_1} \right), \end{aligned}$$

where

$$\xi = \frac{\min(\beta, \alpha_2)^{\zeta_2 - 1} - \alpha_1^{\zeta_2 - \zeta_1} \min(\beta, \alpha_2)^{\zeta_1 - 1}}{(1 - \zeta_2) \beta^{\zeta_2} - (1 - \zeta_1) \alpha_1^{\zeta_2 - \zeta_1} \beta^{\zeta_1}}. \quad (6.A.40)$$

Solving with respect to C_1 gives us that

$$C_1 = \frac{\left(\frac{1}{\beta} - \frac{1}{\alpha_2} \right)^+ - \xi}{\alpha_2^{\zeta_1 - 1} - \alpha_0^{\zeta_1 - \zeta_2} \alpha_2^{\zeta_2 - 1} + (\alpha_0^{\zeta_1 - \zeta_2} \alpha_1^{\zeta_2 - \zeta_1} - 1) \min(\beta, \alpha_2)^{\zeta_1 - 1} + \xi \left((1 - \zeta_1) (1 - \alpha_0^{\zeta_1 - \zeta_2} \alpha_1^{\zeta_2 - \zeta_1}) \beta^{\zeta_1} \right)}.$$

This concludes the proof. $\square$

6.A.5 Proof of Lemma 6.4.2 (Existence of optimal barrier with capital injections)

Proof of Lemma 6.4.2. The problem is to show that the equation (6.4.11) has a unique root. We denote by Ψ the function

$$\Psi(\beta) = \zeta_1 \left(\beta^{\zeta_1} \gamma^{1-\zeta_2} \kappa \zeta_1 - \beta^{\zeta_1} \gamma^{1-\zeta_2} \kappa \zeta_2 + \gamma^{\zeta_1-\zeta_2} \beta \zeta_2 - \beta^{1-\zeta_2+\zeta_1} \zeta_1 - \gamma^{\zeta_1-\zeta_2} \beta + \beta^{1-\zeta_2+\zeta_1} \right).$$

We want to prove that $\text{sgn}(\Psi(\gamma)) \neq \text{sgn}(\lim_{\beta \rightarrow \infty} \Psi(\beta))$, and that

$$\text{sgn}(\Psi'(x)) = \text{sgn}(\Psi'(y))$$

for $x, y > \gamma$.

1.

$$\Psi(\gamma) = \zeta_1 \gamma^{1-\zeta_2+\zeta_1} (\kappa \zeta_1 - \kappa \zeta_2 - \zeta_1 + \zeta_2) > 0.$$

2.

$$\lim_{\beta \rightarrow \infty} \Psi(\beta) = \lim_{\beta \rightarrow \infty} \zeta_1 \left(\gamma^{\zeta_1-\zeta_2} \beta \zeta_2 - \gamma^{\zeta_1-\zeta_2} \beta \right) = -\infty.$$

3.

$$\begin{aligned} \Psi'(\beta) = -\zeta_1 \Big(& -\beta^{\zeta_1-1} \zeta_1^2 \gamma^{1-\zeta_2} \kappa + \beta^{\zeta_1-1} \zeta_1 \gamma^{1-\zeta_2} \kappa \zeta_2 + \beta^{\zeta_1-\zeta_2} \zeta_1^2 - \beta^{\zeta_1-\zeta_2} \zeta_1 \zeta_2 \\ & + \beta^{\zeta_1-\zeta_2} \zeta_2 - \gamma^{\zeta_1-\zeta_2} \zeta_2 - \beta^{\zeta_1-\zeta_2} + \gamma^{\zeta_1-\zeta_2} \Big). \end{aligned}$$

We want to show that the inner part of the delimiters is negative. It is given by

$$\kappa \beta^{\zeta_1-1} \gamma^{1-\zeta_2} (-\zeta_1^2 + \zeta_1 \zeta_2) + \beta^{\zeta_1-\zeta_2} (\zeta_1^2 - \zeta_1 \zeta_2 + \zeta_2 - 1) - \gamma^{\zeta_1-\zeta_2} (\zeta_2 - 1). \quad (6.A.41)$$

To get that (6.A.41) is negative, we need the two following conditions to be fulfilled:

$$\beta^{\zeta_1-1} \gamma^{1-\zeta_2} \geq \beta^{\zeta_1-\zeta_2} \text{ and } \gamma^{\zeta_1-\zeta_2} \geq \beta^{\zeta_1-\zeta_2}. \quad (6.A.42)$$

Since $\left(\frac{\gamma}{\beta}\right)^{1-\zeta_2} \geq 1$ and $\left(\frac{\gamma}{\beta}\right)^{\zeta_1-\zeta_2} \geq 1$, both inequalities holds and $\Psi'(\beta) < 0$.

That is, the optimal barrier exists and is unique. $\square$

6.B Parameter values

No.	ρ	δ	μ_A	μ_L	σ_A	σ_L	α_0	α_1	α_2	κ	A_0	L_0
1	0.5	0.055	0.05	0.04	0.03	0.01	1	1.3	1.35	1.05	1.2	1
2	0.5	0.055	0.05	0.04	0.02	0.01	1	-	-	1.05	1.2	1
3	0.5	0.055	0.05	0.04	0.3	0.01	1	-	-	1.05	1.2	1
4	0.5	0.05	0.04	0.02	0.25	0.1	1	2	4	1.05	-	-
5	0.5	0.05	0.04	0.02	0.01	0.1	1	2	4	1.05	-	-
6	0.5	0.055	0.05	0.03	0.03	0.01	1	1.3	1.5	1.05	1.2	1
7	0.5	0.055	0.05	0.04	0.06	0.01	1	1.3	1.35	1.05	1.2	1
8	0.5	0.055	0.05	0.04	0.2	0.01	1	1.3	1.35	1.05	1.2	1
9	0.5	0.055	0.05	0.04	-	0.01	1	1.3	1.35	1.05	1	1
10	0.5	0.055	0.05	0.04	0.03	0.01	1	1.4	-	1.05	1.2	1

Table 6.1: Parameter values for numerical illustrations.

Bibliography

- Albrecher, H. and Thonhauser, S. (2009). Optimality results for dividend problems in insurance, *RACSAM Revista de la Real Academia de Ciencias; Serie A, Matemáticas* **100**(2), 295–320.
- Asmussen, S. and Albrecher, H. (2010). *Ruin Probabilities*, Vol. 14 of *Advanced Series on Statistical Science and Applied Probability*, 2 edn, World Scientific Singapore.
- Avanzi, B. (2009). Strategies for dividend distribution: A review, *North American Actuarial Journal* **13**(2), 217–251.
- Avanzi, B. and Gerber, H. U. (2008). Optimal dividends in the dual model with diffusion, *ASTIN Bulletin* **38**(2), 653–667.
- Avanzi, B., Shen, J. and Wong, B. (2011). Optimal dividends and capital injections in the dual model with diffusion, *ASTIN Bulletin* **41**(2), 611–644.
- Avanzi, B. and Wong, B. (2012). On a mean reverting dividend strategy with Brownian motion, *Insurance: Mathematics and Economics* **51**(2), 229–238.
- Bacinello, A. R. (2003). Fair valuation of a guaranteed life insurance participating contract embedding a surrender option, *Journal of Risk and Insurance* **70**(3), 461–487.
URL: <http://dx.doi.org/10.1111/1539-6975.t01-1-00060>
- Bailey, P. B., Shampine, L. F. and Waltman, P. E. (1968). *Nonlinear two point boundary value problems*, Academic Press, New York.
- Barbarin, J. (2007). Local Risk-Minimization For Payment Processes. Application to Insurance Contracts With Surrender Option or to Default Sensitive Contingent Claims. Working paper, Université Catholique de Louvain.
- Bertsekas, D. P. (2005). *Dynamic Programming and Optimal Control*, Vol. 1, 3rd edn, Athena Scientific, Belmont, MA.
- Biffis, E. (2005). Affine Processes for Dynamic Mortality and Actuarial Valuations, *Insurance: Mathematics and Economics*, 443–468.
- Blake, D. (2011). Welcome talk for longevity 7: Seventh international longevity risk and capital markets solutions conference. Available at http://www.cass.city.ac.uk/__data/assets/pdf_file/0016/112930/Pres_-Blake.pdf.
- Blake, D. and Biffis, E. (2012). Keeping some skin in the game: How to start a capital market in longevity risk transfers. Workingpaper available at <http://www.pensions-institute.org/workingpapers/wp1207.pdf>.
- Blake, D., Boardman, T. and Cairns, A. (2010). Sharing longevity risk: Why governments should issue longevity bonds, *MPRA Paper 34184*, University Library of Munich, Germany.

- Blake, D., Cairns, A. and Dowd, K. (2008). The birth of the life market, *Asia-Pacific Journal of Risk and Insurance* **3**(1), 1–32.
- Borch, K. (1974). *The Mathematical Theory of Insurance*, Lexington Books, D.C. Heath and Company, Lexington (Massachusetts), Toronto, London.
- Börger, M. (2010). Deterministic shock vs. stochastic value-at-risk - an analysis of the Solvency II standard model approach to longevity risk, *Blätter der DGVFM* **31**(2), 225–259.
URL: <http://dx.doi.org/10.1007/s11857-010-0125-z>
- Buchardt, K. (2012). Life insurance valuation with depend affine rates. Working paper.
- Buchardt, K. and Møller, T. (2013). Life insurance cash flows with policyholder behaviour. Preprint (submitted), available at <http://www.math.ku.dk/~buchardt/>.
- Buchardt, K., Møller, T. and Schmidt, K. B. (2013). Cash flows and policyholder behaviour in the semi-markov life insurance setup. To appear in *Scandinavian Actuarial Journal*.
- Bühlmann, H. (1970). *Mathematical Methods in Risk Theory*, Grundlehren der mathematischen Wissenschaften, Springer-Verlag, Berlin, Heidelberg, New York.
- Chen, P. and Yang, H. (2010). Pension funding problem with regime-switching geometric brownian motion assets and liabilities, *Applied Stochastic Models in Business and Industry* **26**(2), 125–141.
- Cheng, P. and Scaillet, O. (2007). Linear-quadratic jump-diffusion modeling with application to stochastic volatility, *Mathematical Finance* **17**(4), 575–598.
URL: <http://dx.doi.org/10.1111/j.1467-9965.2007.00316.x>
- Christiansen, M. C. (2008). A sensitivity analysis concept for life insurance with respect to a valuation basis of infinite dimension, *Insurance: Mathematics and Economics* **42**(2), 680–690.
URL: <http://ideas.repec.org/a/eee/insuma/v42y2008i2p680-690.html>
- Christiansen, M. C. and Steffensen, M. (2013). Safe-Side Scenarios for Financial and Biometrical Risk, *ASTIN Bulletin*, 1–35.
URL: http://www.journals.cambridge.org/abstract_S0515036113000160
- Colwell, D. B. and Elliott, R. J. (1993). Discontinuous asset prices and non-attainable contingent claims¹, *Mathematical Finance* **3**(3), 295–308.
- Dahl, M., Melchior, M. and Møller, T. (2008). On systematic mortality risk and risk-minimization with survivor swaps, *Scandinavian Actuarial Journal* **2008**(2–3), 114–146.
URL: <http://www.tandfonline.com/doi/abs/10.1080/03461230701795873>
- Dahl, M. and Møller, T. (2006). Valuation and Hedging of Life Insurance Liabilities with Systematic Mortality Risk, *Insurance: Mathematics and Economics* **39**, 193–217.
- Dar, A. and Dodds, C. (1989). Interest rates, the emergency fund hypothesis and saving through endowment policies: Some empirical evidence for the u.k., *The Journal of Risk and Insurance* **56**(3), pp. 415–pp. 433.
URL: <http://www.jstor.org/stable/253166>
- Dassios, A. and Wu, S. (2009). On barrier strategy dividends with parisian implementation delay for classical surplus processes, *Insurance: Mathematics and Economics* **45**(2), 195–202.

- de Finetti, B. (1957). Su un'impostazione alternativa della teoria collettiva del rischio, *Transactions of the XVth International Congress of Actuaries* **2**, 433–443.
- De Giovanni, D. (2010). Lapse rate modeling: a rational expectation approach, *Scandinavian Actuarial Journal* **2010**(1), 56–67.
URL: <http://www.tandfonline.com/doi/abs/10.1080/03461230802550649>
- Decamps, M., Schepper, A. D. and Goovaerts, M. (2006). A path integral approach to asset-liability management, *Physica A: Statistical Mechanics and its Applications* **363**(2), 404 – 416.
- Decamps, M., Schepper, A. D. and Goovaerts, M. (2009). Spectral decomposition of optimal asset-liability management, *Journal of Economic Dynamics and Control*. **33**(3), 710–724.
- Devineau, L. and Loisel, S. (2009). Risk aggregation in Solvency II: How to converge the approaches of the internal models and those of the standard formula?, *Post-Print hal-00403662*, HAL.
URL: <http://ideas.repec.org/p/hal/journal/hal-00403662.html>
- Doff, R. (2008). A Critical Analysis of the Solvency II Proposals, *The Geneva Papers on Risk and Insurance Issues and Practice* **33**, 193–206.
- Duffie, D., Filipović, D. and Schachermayer, W. (2003). Affine processes and applications in finance, *Annals of Applied Probability* **13**(3), 984–1053.
- Duffie, D., Pan, J. and Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions, *Econometrica* **68**(6), 1343–1376.
URL: <http://ideas.repec.org/a/ecm/emetrp/v68y2000i6p1343-1376.html>
- EIOPA (2011). EIOPA Report on the fifth Quantitative Impact Study (QIS5) for Solvency II.
- EIOPA (2013). Technical specification on the long term guarantee assessment (part i), *Technical report*, European Insurance and Occupational Pensions Authority.
- Eling, M. and Kiesenbauer, D. (2014). What policy features determine life insurance lapse? an analysis of the german market, *Journal of Risk and Insurance* **81**(2), 241–269.
URL: <http://dx.doi.org/10.1111/j.1539-6975.2012.01504.x>
- Filipović, D. (2009). *Term-Structure Models: A Graduate Course*, Springer Finance, Springer.
URL: <http://books.google.at/books?id=KqcSh6CavaAC>
- Föllmer, H. and Schweizer, M. (1990). Hedging of Contingent Claims under Incomplete Information, in M. H. A. Davis and R. J. Elliott (eds), *Applied Stochastic Analysis*, Stochastic Monographs, Vol. 5, Gordon and Breach, 389–414.
- Föllmer, H. and Schweizer, M. (2010). *Minimal Martingale Measure*, John Wiley & Sons, Ltd.
- Föllmer, H. and Sondermann, D. (1986). Hedging of Non-Redundant Contingent Claims, in W. Hildenbrand and A. Mas-Colell (eds), *Contributions to Mathematical Economics*, Elsevier Science Publishers B.V. (North-Holland), North-Holland, 205–223.
- Gerber, H. U. (1974). The dilemma between dividends and safety and a generalization of the Lundberg-Cramér formulas, *Scandinavian Actuarial Journal* **1974**, 46–57.
- Gerber, H. U. and Shiu, E. S. W. (2003). Geometric brownian motion models for assets and liabilities: From pension funding to optimal dividends, *North American Actuarial Journal* **7**(3), 37–56.

- Grosen, A. and Jørgensen, P. L. (2000). Fair valuation of life insurance liabilities: The impact of interest rate guarantees, surrender options, and bonus policies, *Insurance: Mathematics and Economics* **26**(1), 37 – 57.
URL: <http://www.sciencedirect.com/science/article/pii/S0167668799000414>
- Henriksen, L. F. B., Nielsen, J. W., Steffensen, M. and Svensson, C. (2014). Markov chain modeling of policyholder behavior in life insurance and pension, *European Actuarial Journal* **4**(1), 1–29.
URL: <http://dx.doi.org/10.1007/s13385-014-0091-2>
- Kim, C. (2005). Modeling surrender and lapse rates with economic variables, *North American Actuarial Journal* **9**(4), 56–70.
URL: <http://dx.doi.org/10.1080/10920277.2005.10596225>
- Kuo, W., Tsai, C. and Chen, W.-K. (2003). An empirical study on the lapse rate: The cointegration approach, *Journal of Risk and Insurance* **70**(3), 489–508.
URL: <http://dx.doi.org/10.1111/1539-6975.t01-1-00061>
- Li, J. and Szimayer, A. (2011). The uncertain mortality intensity framework: Pricing and hedging unit-linked life insurance contracts, *Insurance: Mathematics and Economics* **49**(3), 471 – 486.
- Li, J. and Szimayer, A. (2014). The effect of policyholders’ rationality on unit-linked life insurance contracts with surrender guarantees, *Quantitative Finance* **14**(2), 327–342.
URL: <http://www.tandfonline.com/doi/abs/10.1080/14697688.2013.825922>
- Loisel, S. and Milhaud, X. (2011). From deterministic to stochastic surrender risk models: Impact of correlation crises on economic capital, *European Journal of Operational Research* **214**(2), 348 – 357.
URL: <http://www.sciencedirect.com/science/article/pii/S0377221711003821>
- Milhaud, X. (2013). Exogenous and endogenous risk factors management to predict surrender behaviours, *ASTIN Bulletin* **43**, 373–398.
- Norberg, R. (1999). A theory of bonus in life insurance, *Finance and Stochastics* **3**(4), 373–390.
URL: <http://dx.doi.org/10.1007/s007800050067>
- Olivieri, A. and Pitacco, E. (2008). Assessing the cost of capital for longevity risk, *Insurance: Mathematics and Economics* **42**(3), 1013–1021.
URL: <http://ideas.repec.org/a/eee/insuma/v42y2008i3p1013-1021.html>
- Orava, P. and Lautala, P. (1976). Back-and-forth shooting method for solving two-point boundary-value problems, *Journal of Optimization Theory and Applications* **18**(4), 485–498.
URL: <http://dx.doi.org/10.1007/BF00932657>
- Paulsen, J. (2003). Optimal dividend payouts for diffusions with solvency constraints, *Finance and Stochastics* **7**(4), 457–473.
- Porteus, E. L. (1977). On optimal dividend, reinvestment, and liquidation policies for the firm, *Operations Research* **25**(5), 818–834.
- Protter, P. and Shimbo, K. (2008). No arbitrage and general semimartingales., *Markov processes and related topics: A Festschrift for Thomas G. Kurtz. Selected papers of the conference, Madison, WI, USA, July 10–13, 2006.*, Beachwood, OH: IMS, Institute of Mathematical Statistics, , 267–283.

- Riesner, M. (2006). Hedging Life Insurance Contracts in a Lévy Process Financial Market, *Insurance: Mathematics and Economics* **38**(3), 599–608.
- Schweizer, M. (1988). Hedging of Options in a General Semimartingale Model, Diss. ETHZ no. 8615, Zürich.
- Schweizer, M. (2001). A Guided Tour through Quadratic Hedging Approaches, in E. Jouini, J. Cvitanić and M. Musiela (eds), *Option Pricing, Interest Rates and Risk Management*, Cambridge University Press, 538–574.
- Schweizer, M. (2008). Local Risk-Minimization for Multidimensional Assets and Payment Streams, *Banach Center Publications* **83**, 213–229.
- Sethi, S. P. and Taksar, M. I. (2002). Optimal financing of a corporation subject to random returns, *Mathematical Finance* **12**(2), 155–172.
- Shreve, S. E., Lehoczky, J. P. and Gaver, D. P. (1984). Optimal consumption for general diffusions with absorbing and reflecting barriers, *SIAM Journal on Control and Optimization* **22**(1), 55–75.
- Siu, T. K. (2005). Fair valuation of participating policies with surrender options and regime switching, *Insurance: Mathematics and Economics* **37**(3), 533 – 552.
URL: <http://www.sciencedirect.com/science/article/pii/S0167668705000648>
- Steffen, T. (2008). Solvency II and the Work of CEIOPS, *The Geneva Papers on Risk and Insurance Issues and Practice* **33**, 60–65.
- Steffensen, M. (2000). A no arbitrage approach to thiele’s differential equation, *Insurance: Mathematics and Economics* **27**(2), 201 – 214.
URL: <http://www.sciencedirect.com/science/article/pii/S0167668700000482>
- Steffensen, M. (2002). Intervention options in life insurance, *Insurance: Mathematics and Economics* **31**(1), 71 – 85. Special Issue: Papers presented at the 5th {IME} Conference, Penn State University, University Park, PA, 23-25 July 2001.
URL: <http://www.sciencedirect.com/science/article/pii/S0167668702001270>
- Tsai, C., Kuo, W. and Chen, W.-K. (2002). Early surrender and the distribution of policy reserves, *Insurance: Mathematics and Economics* **31**(3), 429 – 445.
URL: <http://www.sciencedirect.com/science/article/pii/S0167668702001889>
- Valdez, E. A. (2001). Bivariate analysis of survivorship and persistency, *Insurance: Mathematics and Economics* **29**(3), 357 – 373. Papers presented at the 4th {IME} Conference, Universitat de Barcelona, Barcelona, 24-26 July 2001.
URL: <http://www.sciencedirect.com/science/article/pii/S0167668701000798>
- Vandaele, N. and Vanmaele, M. (2008). A Locally Risk-Minimizing Hedging Strategy for Unit-Linked Life Insurance Contracts in a Lévy Process Financial Market, *Insurance: Mathematics and Economics* **42**(3), 1128–1137.