Hybrid Pension Schemes:
Risk Allocation and Asset Liability Optimization

DISSertation
of the University of St. Gallen
Graduate School of Business Administration,
Economics, Law and Social Sciences (HSG)
to obtain the title of Doctor of Economics

submitted by
Roger Titus Baumann
from
Aeschlen (Bern)

Approved on the application of
Prof. Dr. Heinz Müller
and
Prof. Dr. Heinz Zimmermann

Dissertation Nr. 3067

Difo-Druck GmbH, Bamberg 2005
The University of St. Gallen, Graduate School of Business Administration, Economics, Law and Social Sciences (HSG) hereby consents to the printing of the present dissertation, without hereby expressing any opinion on the views herein expressed.

St. Gallen, June 30, 2005

The President:

Prof. Ernst Mohr, PhD
Acknowledgements

Alles, was die Menschen in Bewegung setzt, muss durch ihren Kopf hindurch;
aber welche Gestalt es in diesem Kopf annimmt, hängt sehr von den Umständen ab.
– Friedrich Engels

In the sense of the above quote I would like to thank particularly Professor Heinz Müller for his support. I became interested in the combination of old age pension schemes, risk and portfolio optimization through our common projects. His flexibility to give me as much time as I needed, and his willingness to discuss subject issues at any time, have contributed to keep me grounded throughout the uncertain and bold venture of writing a thesis. I would also like to express my sincere thanks to Professor Heinz Zimmermann for his assistance. His dedication to this area has been truly inspiring, and his economic advice has proved invaluable.

Equally, I would like to thank Stephan Müller for the numerous constructive discussions on portfolio optimization. With his intuition and his analytical expertise he has pointed out both possibilities as well as limits. Thanks also go to Thomas Pfiffner, Reto Leibundgut, Evelyn Ribi and Alex Gulde for the proof reading of different parts of my dissertation, and David Schiess for thoroughly checking the intricate appendix of chapter five with the same generous and unselfish dedication for which I have always respected and esteemed him on the football ground. Additional thanks go to Christian Jaag and Bernhard Thöny for critical comments on my ideas as well as the refreshing debates about “God and the world” on air and during the breaks. Special thanks also go to Thomas Häussler, who has been a short head ahead in languages already in high school and who has given a last polish to my English.

Furthermore, I would also like to thank my mother and my late father, who have known to motivate me by giving me self-confidence throughout the whole of my school days without ever exerting pressure. Finally, I owe very much to Yvonne, who has been particularly understanding during the final stage of my dissertation and who gives me strength.

St. Gallen, July 2005                    Roger Baumann
# Contents

List of Figures viii  
Notation ix  

1 Introduction 1  
1.1 Premises 2  
1.2 Pension Schemes, Financial Risk and Yield Guarantee 6  
1.3 Literature Overview 8  
1.4 Contents and Results of Thesis 15  

2 The Model Basics 19  
2.1 Assumptions and Definitions 19  
2.1.1 Hybrid Pension Scheme 19  
2.1.2 Financial Market 21  
2.1.3 Wealth 22  
2.1.4 Liability 22  
2.1.5 Funding Ratio 23  
2.1.6 Stopping Time 24  
2.1.7 Utility Functions 24  
2.1.8 The Expected Utility Hypothesis 25  
2.1.9 Pareto Optimality 26  
2.2 The Hamilton-Jacobi-Bellmann equation 26  
2.2.1 General HJB Equation 26  
2.2.2 Optimal Terminal Value 28  
2.2.3 Time-Homogeneous Setting 28  
2.2.4 Long-Term Growth Rate 29  
2.2.5 A Useful Theorem 31  
2.A Appendix 33  
2.A.1 Proof of Theorem 2.9 33  
2.A.2 Proof of Theorem 2.10 36
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Asset Liability Optimization with Exogenous Liability</td>
<td></td>
</tr>
<tr>
<td>3.1</td>
<td>Merton’s Problem Revisited</td>
<td>39</td>
</tr>
<tr>
<td>3.1.1</td>
<td>Optimal Consumption and Bequest Motive</td>
<td>39</td>
</tr>
<tr>
<td>3.1.2</td>
<td>Terminal Wealth</td>
<td>41</td>
</tr>
<tr>
<td>3.2</td>
<td>Optimal Funding Ratio</td>
<td>42</td>
</tr>
<tr>
<td>3.2.1</td>
<td>General Model</td>
<td>43</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Special Case: No Net Contributions</td>
<td>46</td>
</tr>
<tr>
<td>3.3</td>
<td>The Impact of the Yield Guarantee</td>
<td>48</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Underfunded Pension</td>
<td>50</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Overfunded Pension</td>
<td>51</td>
</tr>
<tr>
<td>3.4</td>
<td>Conclusions</td>
<td>53</td>
</tr>
<tr>
<td>3.A</td>
<td>Appendix</td>
<td></td>
</tr>
<tr>
<td>3.A.1</td>
<td>Proof of Proposition 3.1</td>
<td>56</td>
</tr>
<tr>
<td>3.A.2</td>
<td>Distribution Function of $\tau$</td>
<td>57</td>
</tr>
<tr>
<td>4</td>
<td>Minimum Yield Guarantee and Surplus</td>
<td>61</td>
</tr>
<tr>
<td>4.1</td>
<td>Low Minimum Yield Guarantee and Bounded Funding Ratio</td>
<td>62</td>
</tr>
<tr>
<td>4.1.1</td>
<td>Model Setting</td>
<td>62</td>
</tr>
<tr>
<td>4.1.2</td>
<td>Solution</td>
<td>64</td>
</tr>
<tr>
<td>4.2</td>
<td>High Minimum Yield Guarantee and Surplus</td>
<td>69</td>
</tr>
<tr>
<td>4.2.1</td>
<td>General Model</td>
<td>70</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Special Case: CIES-Utility</td>
<td>72</td>
</tr>
<tr>
<td>4.2.3</td>
<td>Special Case: CIES-Utility and Constraint on Funding Ratio</td>
<td>75</td>
</tr>
<tr>
<td>4.3</td>
<td>Conclusions</td>
<td>81</td>
</tr>
<tr>
<td>5</td>
<td>Individual Risk-Sharing Models</td>
<td>83</td>
</tr>
<tr>
<td>5.1</td>
<td>The Basics of the Settings</td>
<td>86</td>
</tr>
<tr>
<td>5.2</td>
<td>The Reference Yield Plan</td>
<td>89</td>
</tr>
<tr>
<td>5.2.1</td>
<td>General Model</td>
<td>89</td>
</tr>
<tr>
<td>5.2.2</td>
<td>Special Case: No Net Contributions and CRRA Utility</td>
<td>92</td>
</tr>
<tr>
<td>5.2.3</td>
<td>Impact of the Reference Yield</td>
<td>95</td>
</tr>
<tr>
<td>5.3</td>
<td>Free Pension Selection</td>
<td>105</td>
</tr>
<tr>
<td>5.3.1</td>
<td>General Model</td>
<td>105</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Special Case: No Net Contributions and CRRA Utility</td>
<td>106</td>
</tr>
</tbody>
</table>
List of Figures

3.1 Upper boundary $\hat{r}_{0,T}$ ................................................. 52
3.2 Upper boundary $\hat{r}_{0,p}$ .................................................. 54
3.3 Distribution function of $\tau$ ................................................. 60
4.1 Trajectories of $W_t$, $L_t$ (or $X_t$) and $F_t$ .......................... 64
4.2 Development of $X_t$ and $F_t$ ............................................... 68
4.3 Relation of $\hat{r}_0$ and $\bar{r}_0$ ............................................... 80
5.1 Reference yield plan or yield guarantee setting? The employee’s perspective. 94
5.2 Reference yield plan or yield guarantee setting? The plan sponsor’s perspective. 97
5.3 Reference yield setting: Contracts ........................................ 99
5.4 Reference yield plan with contract or yield guarantee setting? The employee’s perspective. 101
5.5 Reference yield plan with contract or yield guarantee setting? The plan sponsor’s perspective. 104
5.6 Optima for the plan sponsor and the employee for constant portfolio .......... 107
5.7 Free pension choice: The choice .......................................... 108
5.8 Free pension choice vs. autonomous investment and yield guarantee setting:
   Boundaries for $r_0$. .................................................. 111
5.9 Contract setting: Offers .................................................. 115
5.10 Contract setting: Contract ................................................ 116
5.11 Contract setting: Solutions ............................................... 120
5.12 Contract setting: Solutions3D .......................................... 121
5.13 Contract setting: Pareto boundaries $R_{e,Pareto}$ ....................... 122
5.14 Example for parameters where the contract setting is worse than the yield guarantee setting. ............................................. 123
5.15 Contract setting: Pareto optimality ..................................... 125
5.16 Contract setting: Possible $r_0^*$. ........................................ 128
Notation

Uppercase Letters

$B$ point, see proposition 5.1,
$C$ net contribution process
$D$ domain
$D_a$ diagonal matrix with elements $a$
$E$ expectation operator
$F$ funding ratio process
$I$ indirect utility function of the employee
$J$ indirect utility function of the plan sponsor, $J$-function
$L$ liability process
$M$ index for “Merton solution”; martingale process
$P$ probability operator
$P$ probability measure
$Q$ radicand of the solution of a cubic equation
$R$ risk aversion of the CRRA utility
$R_e$ employee’s risk aversion
$R_p$ plan sponsor’s risk aversion
$S$ price process
$T$ time horizon
$U$ utility function
$W$ wealth process
$X$ accrued retirement assets process
$Z$ Wiener process
Notation

**Lowercase Letters**

- $a, b, c, d, g, h, k, q$: real support parameters
- $b$: vector process
- $c_i$: individual contribution process
- $f(.), g(.), h(.), m(.), y(.)$: real-valued support functions
- $i, j$: integer indexes
- $k, m, n$: dimensions
- $l$: liability supplement process
- $p$: probability
- $r$: return of the riskless asset
- $r_0$: yield guarantee, minimum yield, reference yield, quasi-guarantee
- $s$: service costs process
- $t$: point of time
- $u(.)$: time independent part of the utility function
- $w$: process
- $x$: arguments
- $x$: argument; proportion of the growth optimal portfolio
- $z$: argument

**Uppercase Greek Letters**

- $\Gamma, \Theta, \Xi$: static objective functions
- $\Pi$: portfolio value process
- $\Sigma$: variance-covariance matrix
- $\Phi$: distribution function
- $\Psi$: constraint process
- $\Omega$: set of states
Lowercase Greek Letters

\(\alpha\) participation rate (process)
\(\beta\) \(\beta \in \mathbb{R} \setminus \{0\}\)
\(\gamma\) weight factor of the untradable part
\(\delta\) CIES-utility function parameter
\(\epsilon\) bequest parameter
\(\varepsilon\) small positive real number
\(\zeta\) \(\zeta \triangleq \frac{r - r_0}{\hat{\pi} \Sigma \hat{\pi}}\)
\(\eta\) constraint parameter
\(\theta\) weight parameter, \(\theta \in [0, 1]\)
\(\vartheta\) stochastic process
\(\kappa, \psi, \varphi\) parameters
\(\lambda\) weight factor of the untradable part
\(\mu\) expected returns
\(\mu\) drift; drift parameter
\(\nu\) surplus process
\(\zeta\) control variable
\(\pi\) portfolio process
\(\hat{\pi}\) growth-optimal portfolio \(\hat{\pi} \triangleq \Sigma^{-1}(\mu - r 1)\)
\(\rho\) stochastic retirement assets return
\(\sigma\) diffusion parameters
\(\sigma\) diffusion; diffusion parameter
\(\varsigma\) service cost parameter
\(\tau\) stopping time
\(\upsilon\) optimal long-term growth rate
\(\phi\) real-valued function
\(\omega\) state
\(\rho\) time parameter of the CIES-utility function
Script and Set Letters

\( A \) generator of the Itô diffusion
\( C^1 \) set of once differentiable functions
\( C^2 \) set of twice differentiable functions
\( \mathcal{E}, \mathcal{K}, \mathcal{L}, \mathcal{M} \) sets
\( \mathcal{F} \) \( \sigma \)-algebra
\( \mathcal{G}, \mathcal{H}, \mathcal{Y} \) sets of admissible strategies
\( \mathbb{R} \) set of real numbers

Various symbols

1 vector containing ones
a vector; matrix
a' transpose of vector/matrix
a* optimal value
\( \bar{a}, \tilde{a} \) transformation or alternative value of a
\( \hat{a} \) upper boundary of a; except \( \hat{\pi}, \hat{\upsilon} \) and \( \hat{\alpha} \)
\( \check{a} \) lower boundary of a
\( \breve{a} \) zero point
a−, a− negative value; converges from left
a+, a+ positive value; converges from right
a' (.) first derivative
a'' (.) second derivative
|a| determinant of a
Chapter 1

Introduction

People are subjected to a life-cycle with periods where they are more and periods where they are less productive. Private savings provide one possibility to prepare for the non-income producing period after retirement. Most notably, life insurers offer a variety of different retirement plans and thus are able to consider the individual risk preferences of the client so as to help to find an optimal life-cycle planning.

However, since it is not possible for everybody to prepare individually for the unproductive period after retirement, modern societies need an old-age insurance system that provides minimal protection against impoverishment. More precisely, economically advanced societies need an insurance system that allows them to redistribute wealth from the rich to the poor in order to increase social security. Therefore, in most industrial countries there is a governmental old-age insurance system.

Old-age insurance has increasingly been supplemented by pension funds, which have grown fast in recent decades. Pension assets relative to GDP, for example, amounted to about 50% in the United States, 100% in the United Kingdom and 130% in Switzerland at the end of 2000. The main reason for this proportion is that in most of the countries the government encourages people to join a pension plan. The most common method is through tax incentives; in Switzerland, for instance, joining is even mandatory.

Considering that there already exists public old-age insurance for redistribution of wealth and a variety of private plans for individual purposes, the question arises as to why the government should be interested in pension funds. The most common argument is that the personal saving rate is too low. Indeed, some studies argue that there is a significant portion of households with saving rates too low to be explained by conventional life-cycle models, and a growing body of literature has focused on behavioral explanations as to why some households fail to save.¹ However, dismissing low savings as individual mistakes by people is problematic from our perspective. It would mean that people are not rational, which is inconsistent since we intend to develop a positive theory requiring

¹see e.g. Lusardi (1999), Lusardi, Skinner, and Venti (2001), Bernheim (1995) or Yakoboski and Dickemper (1997).
people to be rational. Other reasons for insufficient savings could be, e.g., that people do not have enough information, face high planning costs or lack financial literacy.

Or it could be argued that the saving rate is too low from the government perspective. The government might believe that a low personal saving rate causes negative externalities. For instance, it might believe that the domestic saving rate should be increased even though the personal attitudes towards saving are efficient from an individual perspective.\(^2\)

Taking everything into account, encouraging retirement planning per se could be part of a program to increase savings in general.\(^3\) This might be part of the reason that has lead some countries to support retirement arrangements, as for instance the United States, where the personal saving rate dropped from 10.6 per cent of disposable personal income in 1984 to -1 percent in 2001.\(^4\) Additionally, in the United States the domestic saving rate is, at about 5 per cent,\(^5\) among the lowest in the world.

However, to increase savings might not be the only purpose for encouraging retirement plans, particularly for countries with a high saving rate. In Switzerland for instance, the net domestic saving rate is about 15 per cent, i.e. roughly 10 per cent higher than in the USA\(^6\) and the personal saving rate is about 10 per cent. Why then is joining a pension plan even mandatory in Switzerland? Apparently, there must be still other arguments in favour of a pension fund as we will see in the next section.

### 1.1 Premises

Within a public old-age insurance system, the government faces a problem of moral hazard. Because people count on public old-age insurance they have no incentive to care

---

\(^2\)The question whether the domestic saving rate is too low and if so, how to encourage savings best, are topics of economic growth theory. Extensive textbooks about this issue are Barro and Sala-I-Martin (1995) or Romer (1996).

\(^3\)Whether retirement saving programs increase saving has been a matter of much debate (see Eugen, Gale, and Scholz (1996); Hupbard and Skinner (1996); Poterba, Venti, and Wise (1996)). However, there is a consensus that pension plans and 401(k) plans are effective at stimulating savings, particulary among low-income households (see e.g. Gustman and Steinmeier (2000) and Eugen and Gale (2000)).


\(^5\)The exact value depends on the source. The most widely reported saving rates for the United States are the National Income and Product Accounts (NIPA) saving rate and the Flow of Funds Accounts (FFA) saving rate compiled by the Federal Reserve Board, see Lusardi, Skinner, and Venti (2001).

\(^6\)The Swiss domestic saving rate is among the highest in the world. Source: OECD
about retirement savings themselves. Furthermore, the extensive the redistributions are in order to increase the level of social security, the more it causes a moral hazard problem. However, contrary to the first impression it is possible to incorporate an implicit redistribution argument in the pension funds in order to increase social security. In funding-based retirement arrangements such as pension plans, the redistribution argument is expressed by a “redistribution of risk”. For the purpose of our context we define a redistribution of risk as a transfer of risk to support the ones at the expense of the others within at least one of the following two dimensions:

Firstly, the inter-generational dimension. If the financial market is prospering, the plan sponsor\(^7\) will build up funding reserves. Therefore, the benefits will not be adequate. In bad times he will reverse the accruals in order to afford the benefits. Thus, employees\(^8\) who retire right after a boom will lose relative to opportunities in individual investments, whereas the opposite is true for employees who retire after a crash.\(^9\)

Secondly, the intra-generational dimension. People with a low risk aversion are bounded and, therefore, bear opportunity costs. Their expected utility in a pension plan is smaller than within their private saving opportunities. The opposite is true for those employees who are highly risk averse.

In Switzerland, for instance, the social security argument is very strong as the federal constitution states that the public old-age insurance combined with retirement arrangement have to enable the accustomed living standard.\(^10\) As the contribution of public old-age insurance is rather small\(^11\), pension funds provide a substantial part of redistribution. Therefore, joining a pension plan is mandatory.

A redistribution of risk is usually done by appointing a generally binding minimum of benefits in retirement, although, admittedly, an inter-generational redistribution causes a shortfall\(^12\) risk. The crucial question we have to ask is: Who bears the shortfall risk? If it is...
a public pension fund, then the state and, consequently, the public bears the shortfall risk. This leads to a moral hazard problem regarding the plan manager since he is not liable. If the pension plan is offered by a firm, then the firm will bear the shortfall risk. However, the same moral hazard is likely to appear also in private pension plans if the government is the lender-of-last-resort. This leads to the question of why one should include a social security argument in the pension schemes at all if, at last consequence, it inevitably just leads to another moral hazard problem? Or put differently: why encourage or even stipulate retirement savings including a redistribution argument and not just enhance the public old-age insurance system? There are three arguments why pension schemes should bear at least a part of the redistribution:

1. The moral hazard problem localized by the plan sponsor is easier to control than the moral hazard problem localized by every single employee who does not appear in a pension fund. Also, the former is a repeated game and the latter is not, thus if a moral hazard problem is once localized by a plan sponsor, it will be controlled more rigorously. Additionally, a firm is exposed to competition. In the long run, a nonhazardous pension plan leads the firm to a better position in the labour market. Therefore, an increase of the guaranteed benefits by a public old-age insurance will lead to a more serious moral hazard problem than an increasing of guaranteed benefits by a pension scheme combined with a commitment by the government.

2. Generally, pension funds are subject to different risks than the public old-age insurance (specified e.g. by the government) and forces a payment/punishment of the party that is liable. In the USA the Pension Benefit Guarantee Corporation (PBGC) was established by the enactment of the Employee Retirement Income Security Act (ERISA) in order to insure private pension benefits. In Switzerland the “Sicherheitsfond” is economically a similar institution to the PBGC. But, even without a similar institution, it is obvious that if a big pension fund collapses, then the government will react.

3. Note that in countries where joining a pension plan is mandatory, there implicitly exists an even greater responsibility of the government. Therefore, in those countries the moral hazard problem might be even more serious. On the other hand, if attendance is free for the employees (consequently, no attendance would correspond with a higher wage) but firms are obligated to offer a pension plan, then the firm will face the problem of adverse selection. Finally, if it is also free to offer a pension plan, then the redistribution effect will decrease, since fewer firms will offer a plan containing high guarantees if bad times are expected.

An example of a state with a very much enhanced public old-age insurance is Germany. The benefits depend on several parameters. On the other hand, in the U.K. “contracting out” is possible, i.e. the employee can substitute the public insurance with a private pension plan. In the U.S. the full benefits vary between about US$ 620 and US$ 1660 per month; in Switzerland, as already mentioned, between about US$ 840 and US$ 1690.

In theory, in repeated games moral hazard disappears, see LaFont and Martimort (2002).
1.1 Premises

... which is often based on a pay-as-you-go system\textsuperscript{17}. The former, most notably, depends on the financial markets, the latter on the growth of the wage sum. Of course, these quantities are statistically dependent but there might be a diversification potential that may be tapped by a mix of the two systems.\textsuperscript{18}

3. If the public old-age insurance is a pay-as-you-go system, then it will be politically very difficult to reverse it since members of the currently retired generation accumulated the higher benefits when they were active and are therefore not willing to accept lower benefits paid by the currently active generation. This circumstance is called the inter-generation contract.

Taking everything into account we arrive at the following premise which provides the basis for the thesis:

I. It is possible to increase social security while including a redistribution of risk in pension funds. If people expect a high standard of social security then there are serious arguments to include a redistribution mechanism in pension funds instead of increasing the public old-age insurance.

The crucial points are: Firstly, how much risk should be redistributed between the employees within the pension plan and, secondly, how much risk should or can be borne by the plan sponsor, i.e. by the firm or the public? Or put differently: How much individuality or decentralization could still be realised with a requested level of redistribution? We will discuss these questions throughout the thesis. But in order to do this we need some further assumptions partly introduced before. We will assume that;

II. the plan sponsor aims to accomplish the postulated benefits even though there might be a moral hazard problem and

III. in order to have the essential transparency to provide premise II, the balance sheet of the pension fund must be separated from the balance sheet of the firm even though it might be the firm which bears the short-fall risk.

\textsuperscript{17}In a pay-as-you-go system, each year the system pays out roughly the same amount as it receives, i.e. within the same financial period the amount it receives in contributions is paid out in benefits. E.g. in the U.S. the Old Age and Survivors Insurance (OASI), in Switzerland the “Alters- und Hinterbliebenen Versicherung” (AHV) and the public old-age insurance in Germany are financed this way.

\textsuperscript{18}For the relation between economic growth and capital markets see e.g. Fama (1981) or Chen (1991).
For premises II and III to be true we require either an effective control mechanism (by the government, by the employees or through competition) or “prudent” plan sponsors. In most countries there actually exists a “prudent person rule”\(^{19}\) which can be deduced directly or indirectly from the law.\(^{20}\)

Additionally, we assume that;

IV. the supply of pension plans is ensured even though it might be a non-profit institution.

Thus, we assume that either for labour market reasons supply is induced by a broad demand, e.g. encouraged by tax incentives, or else attendance is even mandatory.

In the next section we will give a short introduction to different pension schemes. We will define a “hybrid” pension scheme and briefly discuss the impact of a yield guarantee on such a plan relative to the financial risk.

### 1.2 Pension Schemes, Financial Risk and Yield Guarantee

The differences between the pension schemes are determined by the benefits obtained by the retired member and the contributions paid by the active member. Generally, we distinguish between two main categories of funded pension schemes. If the future benefits are stipulated independently of the available pension asset the pension plan is referred to as a *defined benefit scheme* (DB plan). These objectives are somewhat different from those of *defined contribution schemes* (DC plan) that do not make a defined pension promise. Here, the value of the pension is determined by the value of the annuity which the pool of assets accumulated in the fund can purchase at retirement. Therefore, a notable advantage of DB plans is the potential they offer to provide a stable replacement rate of final income to employees. Of course, protection offered to employees is a risk borne by the

\(^{19}\)The “prudent person rule” can generally be stated as follows: “A fiduciary must discharge his or her duties with the care, skill, prudence and diligence that a prudent person acting in a like capacity would use in the conduct of an enterprise of like character and aims” (Source: International Network of Pensions Regulators and Supervisors: Note by the OECD Secretariat)

\(^{20}\)e.g. in Switzerland BVG Art. 65, BVV2 Art. 50 Abs. 2. In the UK and US the prudent person rule as applied to pensions is based on common law, but adapted to the pension environment, primarily in the Pensions Act 1995 (UK) and the Employee Retirement Income Security Act of 1974 (“ERISA”) (US).
1.2 Pension Schemes, Financial Risk and Yield Guarantee

plan sponsors. With a DC plan in its pure form the investment risk is fully borne by the employees. This might be the main reason why there has been a significant trend away from DB plans to DC plans in order to reduce the risk borne by the plan sponsors in recent years. In the U.S. the number of DC plans grew 100% between 1980 and 1999 while the number of DB plans had its peak in 1983 and fell 71% between 1983 and 1999\textsuperscript{21}. But apart from advantages for the plan sponsors, DC plans also have some benefits for the employees. Firstly, employees always know exactly the value of the funds in their retirement account and secondly, only a DC plan allows a design within which the employees have the freedom to make investment choices from the array of options available in their company’s plan\textsuperscript{22}.

In the purest form of a DC plan the employees bear the whole financial risk but they can select a risk profile consistent with their own preferences and circumstances. Such a pension plan includes no redistribution of risk. The closest form to this scenario in practice are the 401k-plans in the U.S., which have grown rapidly in the last 20 years\textsuperscript{23}. The employee can choose her preferred investment policy out of a variety offered by the plan sponsor. The more different investment options that are offered by the plan sponsor and the more options are open for contributions, the closer this scheme comes to the individual optimization in a life-cycle-model. However, in a framework such as the 401k-plan the risk is fully externalized and, therefore, there is no opportunity for asset liability optimization by the plan sponsor.

Within the framework of the DC plan it is possible to design hybrid plans combining features of the DB plan with those of a pure DC plan, e.g. a model where the plan sponsor offers a DC plan including a shift of risk to or from the employee to the plan sponsor. Therefore, we define a hybrid plan as a pension plan i) in which the benefits depend on the accumulation of either predetermined or free contributions, and ii) which imposes a redistribution of risk. The most common tool to implement such a plan is a generally binding yield guarantee. The downside of the strategy is that the higher the guarantee, the higher the risk of slipping into a funding gap. If an already underfunded pension

\textsuperscript{21}Source: “Private Pension Plan Bulletin,” U.S. Department of Labor, Number 12, Summer 2004

\textsuperscript{22}There exist a variety of other tradeoffs between DC plans and DB plans, e.g. concerning inflation risks, wage growth uncertainty, operating costs or portability losses. We will focus on the financial market risk. For a more extensive comparison of DB plans and DC plans see Bodie, Marcus, and Merton (1988).

\textsuperscript{23}Between 1983 and 1999 the number of plans increased from 17,303 to 335,121, the participants from 7.5 m to 38.6 m and the assets from 92 bn to 1790 bn (Source: “Private Pension Plan Bulletin,” U.S. Department of Labor, Number 12, Summer 2004). A good survey of the practical relevance of 401k-plans is given in Poterba, Venti, and Wise (1998).
fund is confronted with a high minimum yield it will hardly ever reach a debit balance. Thus, if the minimum yield is high, the plan sponsor has to build up substantial funding reserves and, consequently, a minimum yield guarantee turns into a fixed yield guarantee. Within the discussion about a yield guarantee, we have to distinguish between the case with a low minimum yield, which is smaller than the riskless yield, and the case with a high minimum yield, which is higher than the riskless yield. In the first case it is the funding ratio that is expected to increase, even if the plan sponsor invests fully in the riskless asset. This scheme is, for example, practiced by life insurances in the EU where the minimum yield mustn’t exceed 60% of the yield of government bonds. In the second case the minimum yield is higher than the riskless yield. Therefore, the plan sponsor has to take substantial risk in his investment to avoid a shortfall. This situation, for example, is apparent in Switzerland where the government sets a minimum yield which is mandatory for all plan sponsors. In the last decade this minimum yield exceeded the yield of the riskless asset most of the time.

In the next section we will give a survey of the corresponding literature pertinent to this thesis. In particular, we will discuss publications about the asset liability optimization method we will use throughout the thesis and about retirement plans, particularly DC plans and hybrid plans.

1.3 Literature Overview

Markowitz (1952) was among the first to analyze the trade-off between risk and re-

---

24 The riskless yield essentially stands for the return of the riskless asset minus other fixed costs.
25 92/96/EWG
26 Note that in this case, as ultimate consequence, the plan sponsor is not able to guarantee the minimum yield. He has to be reinjured which causes additional costs (in this thesis covered by the service costs). In Switzerland it is mandatory to contribute to the “Sicherheitsfond”. Of course it is economically clear that even the “Sicherheitsfond” could collapse since the financial risk cannot be eliminated. Thus, in the end the risk is borne by the public. The higher the guarantee, the higher the probability that the employees themselves have to rehabilitate the finances. Throughout this thesis we assume that this probability and the arisen access to the single employee are small enough such that the described effect is negligible.
27 BVG Art. 15 Abs. 2
28 Of course, it is not clear which asset we can call riskless. Our statement refers to the domestic 1-month- or 3-month-LIBOR
turning search for a model of how to optimize the return given a level of risk. His static one-period model was extended to a multi-period setting by Tobin (1965). In the same spirit Samuelson (1969) solved an optimal consumption and investment problem within a discrete-time multi-period model and Merton (1969) within a continuous-time model. Merton was the first to apply the idea of stochastic dynamic programming considering CRRA (constant relative risk aversion) utility functions and assuming log-normal distributed asset prices to provide explicit solutions. This so-called "Merton problem" is discussed in almost every textbook about financial theory. In Merton (1971), he showed that within this setting the optimization would lead to investment in mutual funds, i.e. the investor would hold stocks in a fixed proportion. Further, also in Merton (1971), he allowed for more general price processes and obtained a three-fund separation result. He stated that the investor will hold a risk free asset, a growth-optimal portfolio and a hedge portfolio to protect himself from adverse movements of the state variable. Merton (1990) contains a multitude of variations of optimal portfolio-problems, most notably an intertemporal equilibrium model (first published in Merton (1973)).

Cox and Huang (1989); Karatzas, Lehoczky, and Shreve (1987) supplemented this approach with general Itô-processes as underlying and used the martingale representation technique. This method splits the optimization problem into three subproblems:

Firstly, we need to find a martingale measure (which is unique if markets are complete). Roughly speaking, the martingale measure serves as a “discount factor” inasmuch as under this probability measure the expected value of future wealth (considering all in- and outflows) has to be at most as high as the initial wealth. Secondly, the problem is solved under this constraint as if it was a “static” one. Finally, based on the martingale-representation-theorem, the optimal wealth process can be replicated and the optimal portfolio strategy can be derived. The method is treated in several textbooks, e.g. see Karatzas and Shreve (1998) or Korn (1997). The martingale method is much more intuitive than the dynamic programming approach and avoids the partial differential equation known as the Hamilton-Jacobi-Bellman (HJB) equation. However, the approach also faces some difficulties. Firstly, we have to find a martingale measure. Within complete markets the martingale measure is unique. Therefore, a unique deflator process can be created. But if markets are incomplete (or if we include constraints), the martingale measure is not unique which in turn causes some problems, again see Karatzas and Shreve (1998) or Korn (1997). Secondly, if we solve concrete problems, the solution of the

---

29See e.g. Ingersoll Jr. (1987) or Björk (1998)
30see Bellmann (1957)
static problem can cause difficulties. In particular, it is not certain whether we find an explicit form of the Lagrangian multiplier. Finally, we have to replicate the wealth process, which in general is the most demanding part.

To summarize, if we i) intend to find an explicit solution to a specific optimization problem, ii) act within incomplete markets and iii) assume simple underlying processes (which is the case in this thesis), the dynamic programming approach seems to have some clear advantages. However, as a matter of fact, the theory examining and using the HJB approach has had a head start of about thirty years and, thus, its use has become more familiar.

The martingale approach certainly has its advantages within the theoretical discussion about more conceptional optimization aspects, most notably existence and uniqueness statements. Thus, it is widely used within the general theory of constrained optimization problems; see e.g. Cvitanić and Karatzas (1992) for constraints on the portfolio-proportion process, or e.g. Korn and Trautmann (1995) for constraints on the terminal wealth. Moreover, it allows the treatment of more general underlying processes than Itô-processes, namely semimartingales. These most general frameworks are treated by Kramkov and Schachermayer (1999), Mnif and Pham (2001), Karatzas and Žitković (2003) or Müller (2004). There is another advantage of the martingale approach in the area of viscosity solutions. The theory of viscosity solutions serves to study stochastic control problems without the strong regularity conditions required by the dynamic programming method (see the works of Zariphopoulou et al., e.g. Zariphopoulou (1994), Duffie, Fleming, Soner, and Zariphopoulou (1997) or Zariphopoulou (2001)).

The continuous-time models pioneered by Merton (1969) aim to optimize investment and consumption decisions during life time. Within this context there are several possibilities for enhancing the setting of the Merton problem. A useful enhancement for individuals who intend to optimize their investments and the consumption during life time, is it to consider a wage flow during their working time. Merton (1971) provided a setting with a (deterministic) income process and Duffie, Fleming, Soner, and Zariphopoulou (1997), Mnif and Pham (2001) or Karatzas and Žitković (2003) include stochastic income. Even closer to a life cycle model came Bodie, Merton, and Samuelson (1992), including a flexibility in varying the workers’ effort and including their choice of when to retire. In the same spirit, Bodie, Detemple, Otruba, and Walter (2004) examine a life-cycle-model with habit formation, stochastic opportunity set, stochastic wages and labor supply flexibility. They also discuss the effect of habits on post-retirement decisions. Beside the uncertainty
1.3 Literature Overview

About future income, there is the uncertainty about life time. Several studies including uncertain life time were presented; first Yaari (1965), followed by Hakansson (1969) and Merton (1971), who included the investment decision. Zilcha and Friedman (1985) especially focused on saving behavior in retirement. Many aspects of long-term investments and life-cycle planning are given in Campbell and Veceira (2002).

So far, we have considered the financial decision of an individual. Institutional investors, however, are subject to other criteria. Most notably they face liabilities. Among the first authors to address portfolio problems where fixed liabilities enter the optimization problem were Wise (1984a), Wise (1984b) and Wilkie (1985). Subsequently, the approach was developed by Sharpe and Tint (1990), Elton and Gruber (1992), Keel and Müller (1995), Browne (1999) and Denzler, Müller, and Scherer (2001). The studies by Browne (1999), in particular, contain the solution for an optimal portfolio strategy for outperforming a stochastic benchmark in a Merton problem. He, therefore, optimizes the ratio of the wealth process divided by the benchmark process. The benchmark process is assumed to be correlated to the wealth process but not hedgeable. If we substitute the benchmark process by a liability process, the ratio process becomes the funding ratio process. The optimization problem of Browne (1999) aims to maximize the probability that the Investor outperforms the benchmark by a predetermined percentage \( u \), before falling below it by another predetermined percentage \( l \). In a pension framework this would be the same as maximizing the probability that the funding ratio reaches a value \( u \), before falling below a value \( l \). Further, he solves the problem of minimizing the expected time until the benchmark is beaten by a predetermined percentage \( u \), as well as the related problem of maximizing the expected time before being beaten by the benchmark. Again, translated into a pension problem this would be similar to minimizing the expected time before the funding ratio reaches a value \( u \) or maximizing the expected time until the pension slips in a funding gap. We call these models “non-utility-based”, which is artificial from an economical point of view. As soon as we demand or restrict the movements of the funding ratio or other economical quantities, or even optimize expected values or higher moments, we implicitly assume a utility maximization containing the afforded performance of the quantity in the argument\(^{31}\). The result of Browne (1999) bears resemblance to a special case of a utility based framework with a CRRA utility-function, which is solved by Denzler, Müller, and Scherer (2001), who maximize the expected utility of the

\(^{31}\)Therefore, “non-utility-based” would more exactly be called “not explicitly utility-based”.
An extension for a DB plan including net contributions is treated in Zellweger (2003) and Filitti (2004).

In asset liability optimization models for DB plans the plan sponsor optimizes the funding ratio or the funding wealth considering several generally untradeable risks, such as inflation risk or salary risk (or contribution risk), depending on the kind of benefit that is defined at the time of retirement. The risk is fully borne by the plan sponsor. Therefore, this framework is closely related to the problems of hedging in incomplete markets as treated in Duffie, Fleming, Soner, and Zariphopoulou (1997) or Karatzas and Shreve (1998). More concrete studies can be found in, e.g., Haberman and Sung (1994) or Haberman (1997). The approach of Haberman et al., who consider the contribution rate risk using HJB-equations, is widely used. A further study including random benefits liabilities has very recently been presented by Josa-Fombellida and Rincón-Zapatero (2004).

Another approach was taken by Sundaresan and Zapatero (1997). They assume complete markets and calculate the pension valuation and an optimal asset liability policy. They further examine the employee’s retirement incentives using a partial equilibrium framework.

Models that come closer to reality by combining dynamic stochastic programming models with advances in computational capabilities lead to numerical results that are more popular in practice. Discussion concerning their technical issues and applications (scenario-based discrete-time optimizations) can be found, e.g. in Ziemba and Mulvey (1998) or Zenios and Ziemba (2003).

Public pension funds have a special role within the discussion of DB plans since the whole active population bears the risk of insolvency. Therefore, tax mechanisms should be included in economic models and the problem of optimal funding becomes a complex issue (see e.g. D’Acrý, Dulebohn, and Oh (1999)). It is obvious that in public pension funds the problem of moral hazard appears (see e.g. Chaney and Thakor (1985)). However, as stated above, moral hazard will also appear in private pension plans if the government is the lender-of-last-resort. In this case the plan sponsor will take more risk as shown in the empirical study of Niehaus (1990) who supports the hypothesis that moral hazard induced firms offering DB plans in the U.S. to increase promised pension after the

---

32 Among the first to calculate the portfolio strategy optimizing such a ratio process in an incomplete framework were Adler and Dumas (1983), who introduced a price index.

33 The theory of moral hazard is treated in almost every textbook in microeconomic theory. See e.g. Mas-Colell, Whinston, and Green (1995) or Laffont and Martimort (2002). Especially for insurance risk see Stiglitz (1983).
Over the risk sharing spectra, the 401k-plan is the form closest to an individual life-cycle-model. However, an optimal individual life-cycle planning requires the freedom of contribution decision. In contrast, DC plans generally contain a requirement of constant contribution rates. The employee can choose her preferred investment policy out of a variety offered by the plan sponsor, although the contribution rate is a predefined proportion of the salary. The employee has to choose the investment policy after considering her salary expectations and the funding target. A non-utility-based theoretical approach to this issue using a quadratic penalty function and assuming constant salary was done by Vigna and Haberman (2001). Actually, within a plan similar to the 401(k) plans there is no need for asset liability optimization by the pension sponsor. It is more reasonable that the optimization is done from the employee’s perspective, since she bears the whole financial risk. It would be problematic if the pension fund manager optimized the funding wealth based on his own risk preferences, since, we can realistically assume that the employee’s preferences do not coincide with those of the fund manager. If this was the case, it would lead to serious moral hazard problems. However, from the standpoint of financial theory, an asset liability optimization by the pension sponsor leads to interesting hedging results. The most recent study of this issue was done by Battocchio and Menoncin (2004). They presented an optimizing model in a stochastic framework, assuming that the price process and the salary process follow an Itô-process. Here, the fund manager optimizes the investment strategy while facing uncertainty about inflation and contributions.

The most common hybrid plan is a DC plan including a guarantee which is generally binding. Thus, the literature about hybrid plans is closely linked with individual asset (and liability) modeling for endowments with guarantees. This issue has become more and more common with private insurance policies containing guarantees, bonus provisions and surrender options. However, these policies are more complex to price than traditional insurance products and require financial pricing techniques to the valuation of insurance liabilities. The focus is shifting away from the traditionally static actuarial pricing to stochastic models; see Vanderhoof and Altman (1997) or Embrechts (2000). The pricing of the option embedded in some of the earlier insurance products with guarantees was focussed on the pioneering studies of Brennan and Schwartz (1976) and Boyle and Schwartz (1977). They analyzed unit-linked minimum guarantee policies. Later re-

\[34\] The “Sicherheitsfond” insures private pension benefits in case of insolvency. In public pension schemes the government stands in even before insolvency is attained.
search into this topic includes Delbaen (1990), applying the martingale theory (instead of the Black/Scholes formula), and Bacinello and Ortu (1993) who extended the model to the case where the minimum guarantees are endogenous, i.e. they are functions of the premium(s) paid. They conclude with numerical results. Grosen and Jørgensen (2000) separated the life insurance liability of participating policies with a minimum guarantee into a minimum guarantee (bond), a bonus option and a surrender option. They used numerical techniques to price both options in a unified and realistic framework. In the same spirit, Boyle and Hardy (1997) examined the pricing and reserving for certain guarantees that are associated with some insurance contracts.

Khorasanee (1995), who was among the first to contemplate a concrete hybrid plan, proposed a DC plan with an explicitly defined benefit formula. He discussed possible methods for distributing surpluses and simulated the behavior of the plan under a scenario of persistently unfavorable investment experience. Further, he suggested ways of making good investment shortfalls by reducing the level of future accruals of benefit. Later, they proposed a similar deterministic model (see Khorasanee (1997); Dufresne (1997)).

The impact on the funding ratio and the benefit for the employees considering a minimum yield guarantee combined with a participation guarantee in the surplus of the return on assets is simulated in Baumann, Delbaen, Embrecht, and Müller (2001).

In a theoretical setting, Boulier, Huang, and Taillard (2001) studied the optimal financial decision of a defined contribution plan in the presence of a minimum guarantee. Their argument used a three asset market (cash, bond and equity) which is complete and used a Vasiček specification of the term structure$^{35}$ and the contribution flow is a deterministic process. The guarantee is given and depends on the level of a stochastic interest rate when the employee retires. In particular, they optimized the pension wealth diminished by this guarantee using CRRA utility functions. The same optimization problem is studied by Deelstra, Grasselli, and Koehl (2003). They supplemented the study of Boulier et al. using a more general term structure which includes the CIR$^{36}$ and the Vasiček specification. Moreover, they provided a stochastic contribution flow and a generally random minimum guarantee. Most recently, Deelstra, Grasselli, and Koehl (2004) presented a two asset framework in complete markets to design an optimal guarantee for DC plans. Again, they used CRRA utility functions and focussed on the difference between the funding wealth and the guarantee optimized by the plan sponsor. In a second step the employee optimizes the minimal guarantee. Deelstra, Grasselli, and Koehl (2004) provide

$^{35}$see Vasicek (1977)

$^{36}$see Cox, Ingersoll, and Ross (1985)
1.4 Contents and Results of Thesis

a basis to calculate the optimal guarantee numerically.

For a general comparison of DB plans and DC plans see, e.g., Bodie, Marcus, and Merton (1988) or, to consider the Swiss case, see Baumann, Müller, and Keel (2003). Focusing on the tax issue, Black (1980) and Tepper (1981) showed that tax benefits can be obtained from DB plans that are not available to DC plans. For focussing on the saving implication; see Eugen, Gale, and Scholz (1996); Huppard and Skinner (1996) and Poterba, Venti, and Wise (1996).

In the next section we will provide a preview of the contents of the thesis. Further, we will give an abstract of our contribution to the topic and the results of our research.

1.4 Contents and Results of Thesis

The question as to whether an existing old-age insurance system is sustainable is closely related to the demographic development and levels of productivity. If the retirement age is too low or if we have a severe lack of productivity, then sooner or later every old-age insurance system will collapse. Although significant, these topics will not be covered in this thesis. Instead, we consider a running retirement arrangement with a risk component. The risk in question arises from a guarantee (or quasi-guarantee in chapter 5) which is additionally assumed to be determined by the government and which is assumed to be mandatory for every plan sponsor (only exception: section 5.2.3). In general the guarantee is assumed to be higher than the riskless yield (only exception: section 4.1). From this incomplete market setting evolves a non-hedgeable shortfall risk. For simplicity the plan sponsor is assumed to act as if that shortfall risk is borne by himself. We therefore suppose that the moral hazard problem either is under control or that the plan sponsor bears at least a part of the shortfall risk. The security primacy, arising from this shortfall risk and the direct mandate of the law to act “prudently”, leads us to suppose that the plan sponsor optimizes the funding ratio (only exception: again section 4.1).

The thesis’ primary focus will be on risks aspects evolving from the setting. In particular, we will consider the risk allocation between the plan sponsor and the employee and analyze the investment strategy of the plan sponsor. In order to study the main sources that affect the results the models are kept rather simple. Though the settings are idealized, they have the advantage that closed-form solutions are attainable enabling a thorough analysis of the major driving forces.
Chapter 1 Introduction

There are many differences to the related literature such as Boulier et al. or Deelstra et al.. Firstly, we include a mentionable simplification. To obtain closed form solutions we will not consider CIR term structure but only a deterministic riskless asset and risky assets following geometric Brownian motions including deterministic parameters. Additionally, there are some conceptional differences. As already stated, we optimize the funding ratio and not the surplus of pension wealth, diminished by the guarantee as done in Boulier et al. or Deelstra et al.. The studies of Boulier et al. or Deelstra et al. consider a guarantee which is an accumulated value at retirement and, therefore, they include past and future contributions. Unlike Boulier et al. or Deelstra et al., in our settings throughout the thesis the guarantee is expressed by an instantaneous yield. The plan sponsor has to guarantee a predetermined minimum yield at every point in time which is independent of future or past contributions. Thus the plan sponsor optimizes not for every employee individually considering a certain retirement date but optimizes in the long run ($t \to \infty$) considering the totalized contributions. The employee’s individual assets are determined at each point in time and independent of future contributions of both her own and the other employees. The relative guarantee stays constant. A technical advantage of our approach is the opportunity for including incomplete markets, i.e. allowing a minimum yield exceeding the riskless yield.

The thesis proceeds as follows.

Chapter 2 describes the general setting of the models of the subsequent chapters and provides some mathematical tools that are used throughout the thesis.

After briefly revisiting the well-known results of Merton (1969), chapter 3 explicitly models the optimization problem in the asset-liability context. Instead of using more realistic multi-stage stochastic programming models for dynamic asset-liability models, we restrict the presentation to a simpler model using the technique and results from stochastic control theory in order to get closed form solutions. We consider a plan sponsor who wants to optimize the long-term growth rate of the funding ratio. We will take the situation of a deterministic yield guarantee which is higher than the riskless yield. This framework corresponds to the Swiss situation. We will evaluate the optimal portfolio strategy for a pension and show that the strategy is independent of the yield guarantee. Finally, we will examine the impact of the yield guarantee on the funding ratio within the asset-liability setting. We will see that in a setting with a yield guarantee that exceeds the riskless yield a plan sponsor has to take substantial risk in order to survive.

In chapter 4 we consider a surplus in addition to the minimum yield. Here, the evalu-
ated surplus mechanism is mandatory for the plan sponsor\textsuperscript{37} and the liability becomes endogenous. We distinguish between the case with a low minimum yield and the case with a high minimum yield. The first setting contains a lower bound and an upper bound for the funding ratio as practiced in the UK pension law. Additionally, we include a minimum yield that is lower than the riskless yield similar the EU-directive for life insurance. We will see that it is possible that both plan sponsor and employee are better off if the minimum yield is very low.

In the second setting we include a surplus on the minimum yield that is credited for the employees in the optimization problem. Thus, we optimize both the funding ratio and the surplus in an adequate model. We will see that in the case of a CIES-utility function this surplus becomes a constant. If we additionally include a constraint on the funding ratio, then the surplus becomes stochastic and the plan sponsor transfers a part of the financial risk to the employee.

In chapter 5 we consider the individual risk preferences of the employees. We will study three mechanisms to evaluate an “optimal” risk sharing between the plan sponsor and the individual employee. Two settings are based on different contracts and the third assumes free choice of a pension fund. All settings provide a protection for the very risk averse employees and the possibility to participate in the returns, while bearing at least a part of the risk for the less risk averse employees. For each of these mechanisms, using CRRA-utilities, we will derive a closed form solution for the optimal portfolio strategy and the optimal participation. We will examine under what conditions the employees prefer the settings to either the yield guarantee plan or an autonomous investment policy, and under what conditions the plan sponsor would prefer the settings to the yield guarantee plan. As we will see, all three settings pass this test very well, in general, but since we still have a redistribution of risk, there are also shortcomings for some participants.

Additionally, we will show that the optimal portfolio in two of the settings is constant and, moreover, independent of any assumptions concerning the utility functions of the

\textsuperscript{37}This is a notable difference to the current Swiss system or the setting in chapter 3 where the yield guarantee is mandatory but where the form of the “surplus” is not transparent. Of course in practice, \textit{ex post}, the plan sponsor provides a surplus or contribution holidays in the case that the financial market has prospered. But since the plan sponsor is not bounded \textit{ex ante}, we treat the Swiss system as a \textit{fixed} yield guarantee setting and not as a \textit{minimum} yield setting. Note that, if the financial market prospers, intransparent additional payoffs to the settled surplus are also allowed to occur in a minimum yield setting.
plan sponsor and the employee.
Chapter 2

The Model Basics

In this chapter we provide the general setting for the models used throughout the thesis. As we intend to solve optimization problems within a stochastic and continuous framework of incomplete markets, we will give a short overview of the standard dynamic programming approach to stochastic control problems using the well-known Hamilton-Jacobi-Bellman equation. This approach will be presented for problems with finite and infinite time horizon and for the optimization of the growth rate of a state variable. Completing the preliminaries needed to solve the problems in the following chapters, we will present a theorem that allows us to treat continuous problems as static problems for CRRA utility functions.

2.1 Assumptions and Definitions

In this section we describe the general setting for the models used in the subsequent chapters. We will first provide a short survey of the model variables used before we define them properly.

2.1.1 Hybrid Pension Scheme

At time $t$, on the asset side of the pension fund balance is the accumulated wealth $W_t$. The liability is denoted by $L_t$. 
The wealth $W_{t+dt}$ will be composed of

<table>
<thead>
<tr>
<th>current wealth</th>
<th>$W_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ financial return on current wealth :</td>
<td>$+W_t d\Pi$</td>
</tr>
<tr>
<td>+ ${$ net contributions $}$ (contributions minus benefits)</td>
<td>$+dC_t$</td>
</tr>
<tr>
<td>− service costs</td>
<td>$-ds_t$</td>
</tr>
</tbody>
</table>

The liability $L_{t+dt}$ will be composed of

<table>
<thead>
<tr>
<th>the current value of the liability</th>
<th>$L_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ ${$ return (such as minimum yield) on retirement assets $}$</td>
<td>$+L_t d\rho_t$ if return is stochastic or $+L_t r_0 (t) , dt$ if return is deterministic</td>
</tr>
<tr>
<td>+ net contributions</td>
<td>$+dC_t$</td>
</tr>
<tr>
<td>+ ${$ additional changes to current retirement assets such as inflation adjustment $}$</td>
<td>$+dl_t$</td>
</tr>
</tbody>
</table>

The contributions $dC_t$ denote the sum of all individual contributions, i.e. $dC_t = \sum_i dc_i (t)$ and the liability $L_t$ denotes the sum of all individual retirement assets, i.e. $L_t = \sum_i X_i (t)$.

Note that in a DC plan usually the growth of the liability is basically determined by $r_0 dt$ or $d\rho_t$, respectively. In contrast $dl_t$ is rather small.

The individual accrued retirement assets $X_i (t + dt)$ will be composed of

<table>
<thead>
<tr>
<th>current value of individual retirement assets</th>
<th>$X_i (t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ return on retirement assets</td>
<td>$+X_i (t) d\rho_t$ if return is stochastic or $+X_i (t) r_0 (t) , dt$ if return is deterministic</td>
</tr>
<tr>
<td>+ individual net contributions</td>
<td>$+dc_i (t)$</td>
</tr>
<tr>
<td>+ ${$ additional changes to current retirement assets (e.g. inflation adjustment) $}$</td>
<td>$+X_i (t) \frac{dl_t}{dt}$</td>
</tr>
</tbody>
</table>
During the accumulation phase $dc_i$ will be positive and during the distribution phase $dc_i$ will be negative.

As stated in section 1.4, throughout the thesis the state variable of our problems will usually be the funding ratio. It is the ratio of pension assets to pension liabilities, denoted by $F_t = W_t/L_t$.

The technical characteristics of the variables will be explained in the next subsections.

### 2.1.2 Financial Market

Assumed is a risk-free asset. Its price is assumed to evolve according to

$$dS_0 = r_0 S_0 dt$$

where $r_t \triangleq \{r(t): t \in [0, T]\}$ is the instantaneous return of the riskless asset at time $t$ with $r(t) \geq 0$ and $\int_0^T r(t) dt < \infty$. Then we have $k$ risky assets, a stochastic net contribution, and a stochastic liability. The probabilistic setting is stated as follows: we have a standard Brownian motion $Z_t = (Z_1(t), Z_2(t), ..., Z_k(t), Z_\gamma(t), Z_\lambda(t))'$, $t \in [0, T]$ in $\mathbb{R}^{k+2}$ on a given probability space $(\Omega, \mathcal{F}, P)$, where $\{\mathcal{F}_t\}$ is the natural filtration generated by $\{Z_s; 0 \leq s \leq t\}$ completed by $P$-null sets\(^1\).

Henceforth, all statements containing stochastic variables are assumed to hold almost surely with respect to the probability measure $P$.

The prices of the $k$ risky assets $S_t$ satisfy the Itô-process

$$D_S^{-1} dS_t = \mu_t dt + \sigma_t dZ_t$$

where $D_S$ is a diagonal matrix with elements $S_i(t)$.

In general $\mu_t \triangleq \{\mu(t): t \in [0, T]\}$ is an $\mathbb{R}^k$-valued progressively measurable stochastic process of expected rates of return which is positive and integrable, that is satisfying $\int_0^T \|\mu_t\| dt < \infty$. Then $\sigma_t \triangleq \{\sigma(t): t \in [0, T]\}$ is an $\mathbb{R}^{k \times k}$-valued progressively measurable stochastic process of volatilities assumed to be square integrable, in the sense that $\sum_{i=1}^k \sum_{j=1}^k \int_0^T \sigma_{ij}^2 dt < \infty$. The matrix $\Sigma_t \triangleq \sigma_t \sigma_t'$ is the variance-covariance-matrix and is, therefore, symmetric and positive-semi-definite by definition and further it is assumed

\(^1\)For a more rigorous approach the reader is referred to the textbooks on mathematical finance, e.g. Karatzas and Shreve (1998).
to be regular. The process $\mu_t$ is assumed to fulfil $\mu_t \neq c\mathbf{1}$ for each $t \in [0,T]$. Thus it is $(\mu_t'\Sigma_t \mu_t) (1' \Sigma_t 1) - (1' \Sigma_t \mu_t)^2 > 0$ by the Cauchy-Schwarz inequality.

For sake of closed solutions, later, in the special cases, we will reduce $\mu$ and $\sigma$ to deterministic processes or even to constants.

### 2.1.3 Wealth

An investor holds a portfolio-proportion process of the $k$ risky assets $\pi_t \in \mathbb{R}^k$, $\pi_t \triangleq \{\pi(t) : t \in [0,T]\}$, which is a progressively measurable, self-financing vector control process with $\int_0^T \|\pi_t\|^2 \, dt < \infty$ to avoid ”doubling strategies”. The fraction invested in security $S_0(t)$ is given by $1 - \pi_t' \mathbf{1}$.

Further, we introduce the net contribution $C_t$ which is an integrable Itô-process and has finite variance in the sense that $-\infty < \int_0^T C_t \, dt < \infty$ and $\int_0^T \text{Var}[C_t] \, dt < \infty$. Then, we assume service costs denoted by $d_s$.

The value of the wealth $W_t^\pi$ then evolves as

$$dW_t^\pi = W_t^\pi \, d\Pi + dC_t - d_s$$

starting at $W_0$, where $d\Pi \triangleq (\pi_t' (\mu_t - r_t \mathbf{1}) + r_t) \, dt + \pi_t' \sigma_t \, dZ_t$. The service costs are defined as $d_s = \zeta_t W_t^\pi \, dt$, where $\zeta_t \triangleq \{\zeta(t) \geq 0 : t \in [0,T]\}$ is a positive stochastic process which is integrable and satisfies $\int_0^T \zeta_t \, dt < \infty$.

The portfolio-proportion process is called admissible if the stochastic differential equation (SDE) $dW_t^\pi = W_t^\pi \, d\Pi + dC_t - d_s$ with $W_0$ has a unique solution. The set of all admissible $\pi_t$ is denoted by $\mathcal{G}(\pi)$.

### 2.1.4 Liability

Firstly, similar to Browne (1999) but adding the net contribution and a yield $r_0(t)$, we define a liability process $L_t$ as follows

$$dL_t = dl_t + dC_t + L_t r_0(t) \, dt$$

(2.1)

where $dl_t \triangleq L_t (\mu_L(t) \, dt + b_t' \, dZ_t + \sigma_L(t) \, dZ_t^L)$. In general $\mu_L(t) \geq 0$ is a positive stochastic process satisfying $\int_0^T \mu_L(t) \, dt < \infty$, and $b_t \triangleq \{b(t) : t \in [0,T]\}$ is an $\mathbb{R}^k$-valued stochastic process of volatilities satisfying $\int_0^T b_t' b_t \, dt < \infty$ and $\int_0^T \sigma_L^2(t) \, dt < \infty$. The value $r_0(t) \geq 0$ is the deterministic instantaneous yield guarantee at time $t$ and $\int_0^T r_0(t) \, dt < \infty$. 

2.1 Assumptions and Definitions

$L_t$ is only partially correlated with the wealth process $W_t$, inasmuch as we choose $\sigma_L(t) \neq 0$ for complete generality. For $\sigma_L(t) = 0$ the liability can be fully tracked by assets.

In chapter 4 and chapter 5 we introduce the amount of pension benefits $\rho_t$ as a stochastic process controlled by $\xi_t \triangleq \{\xi(t) \in \mathcal{K} : t \in [0,T]\}$, where $\mathcal{K} \subseteq \mathbb{R}^k$ is a fixed subset. The control $\xi_t$, for example a bonus or a participation rate, is progressively measurable with $-\infty < \int_0^T \rho_t(\xi_t) \, dt < \infty$ and $\int_0^T \text{Var}[\rho_t(\xi_t)] \, dt < \infty$. The SDE of the liability becomes

$$dL_t = d\rho_t + dC_t + L_t \xi_t d\rho_t(\xi_t)$$

starting at $L_0$ where $d\rho_t(\xi_t)$ is the dynamics of the amount of benefits. The control variable $\xi_t$ is admissible if it satisfies $\xi_t \in \mathcal{K}$ and the SDE (2.2) has a unique solution. The set of all admissible $\xi_t$ is denoted by $\mathcal{Y}(\xi)$ if it is the only control variable. Otherwise, the set of all admissible $(\pi'_t, \xi_t)$ is denoted by $\mathcal{H}(\pi', \xi)$.

2.1.5 Funding Ratio

Since $W_l$ and $L_l$ are both Itô-processes, it follows directly that the ratio process $F_t \triangleq W_t / L_t$ is also an Itô-process. Applying Itô’s Lemma we get

$$\frac{dF_t}{F_t} = \frac{dW_t}{W_t} - \frac{dL_t}{L_t} - \frac{dW_t}{W_t} \cdot \frac{dL_t}{L_t} + \left(\frac{dL_t}{L_t}\right)^2. \tag{2.3}$$

For a pension fund $F_t$ would correspond to a generalized funding ratio. $F_t$ is an Itô-process. Therefore, we can write it in a simple form with a single Brownian motion $Z_t$.

The new diffusion term $\sigma$ is a stochastic process including all diffusion terms calculated in (2.3). The new drift term $\mu$ is also a stochastic process which includes all drift terms calculated in (2.3).

If the liability is not controlled we get

$$\frac{dF_t^{\pi}}{F_t^{\pi}} = \mu^{\pi}(t, F_t^{\pi}, L_t, C_t, S_t) \, dt + \sigma^{\pi}(t, F_t^{\pi}, L_t, C_t, S_t) \, dZ_t.$$

If it is (partly) controlled we get

$$\frac{dF_t^{\pi, \xi}}{F_t^{\pi, \xi}} = \mu^{\pi, \xi}(t, F_t^{\pi, \xi}, L_t^{\xi}, C_t, S_t) \, dt + \sigma^{\pi, \xi}(t, F_t^{\pi, \xi}, L_t^{\xi}, C_t, S_t) \, dZ_t. \tag{2.4}$$

In the following chapters we will evaluate the funding ratio more concretely.
2.1.6 Stopping Time

We will be interested in the first time when $W_t$ hits a certain percentage of $L_t$ or, in other words, the first time when $F_t$ hits a value denoted by $c$, which is constant. This leads us to

Definition 2.1 Stopping time

$$
\tau (F, c) \triangleq \inf \{ t > 0; F_t(\omega) = c \} \quad (2.5)
$$

2.1.7 Utility Functions

As soon as we try to describe economical coherence related to the behavior of individuals, we need assumptions about their attitude towards the economic quantities. Microeconomic theory of utility maximization is best suited as a basis for any proposition we make about economical mechanisms.

We use time-additive utility functions in their explicit form. To avoid preposterous solutions, we assume the economic individuals to be risk-averse. We formalize the utility function with $n + 1$ arguments, $t$ and $x_t = (x_1(t), x_2(t), \ldots, x_n(t))$ controlled by $\pi_t$ or controlled, respectively represented, by $(\pi_t, \xi_t)$, using the following definition Karatzas and Shreve (1998).

Definition 2.2 The utility function $U(t, x_t) : [0, T] \times \mathbb{R}^n \to (-\infty, \infty)$ is upper semicontinuous, nondecreasing, concave and progressively measurable for any $t \in [0, T]$. It satisfies the following properties for each $t \in [0, T]$

(i) dom $(U) \triangleq \{ x_t \in \mathbb{R}^n; U(t, x_t) > -\infty \} \subset [0, \infty)^n$ is not empty

(ii) $U'$ is continuous, positive and strictly decreasing in the interior of dom $(U)$, and

$$
\lim_{x_i \to \infty} \frac{\partial U(t, x)}{\partial x_i} = 0 \forall i
$$

(iii) Set $\bar{x}_t \triangleq \inf \{ x_t \in \mathbb{R}^n; U(t, x_t) > -\infty \}$ so that $\bar{x}_t \in [0, \infty)^n$.

Then

$$
\lim_{x_i \to \bar{x}^+_i} \frac{\partial U(t, x)}{\partial x_i} \in (0, \infty] \forall i
$$

(iv) Alternatively, $U \equiv 0$ is also called a utility function

Concavity of the utility function ensures risk-averse behavior. The first condition (i) ensures that we concentrate on a non-empty set. The second condition (ii) ensures the existence of an inverse map of the derivatives of $U(t, x_t)$. The third condition (iii) allows us utility functions which include subsistence levels of the arguments. The forth condition (iv) allows us to consider several problems at the same time.
2.1 Assumptions and Definitions

2.1.8 The Expected Utility Hypothesis

In order to discuss economic coherences related to the behavior of individuals, we have to assume that the actors always intend to maximize their “utility” \( U(t, x_t) \). Therefore we have rationality\(^2\) of individuals. But individuals are also assumed to maximize “utility” \( U(t, x_t) \) ex ante at time \( s < t, s, t \in [0, T] \). Since ex ante \( U(t, x_t) \) is a random variable (progressively measurable), we introduce an ex ante measure to compare possible outcomes of \( U(t, x_t) \). The corresponding measure in microeconomic theory is the expected utility.

**Definition 2.3** The expected utility is denoted by

\[
EU(t, x_t) = \int U(t, x_t) \, d\Phi(t, x_t),
\]

where \( \Phi(t, x_t) \) is the distribution function of \( x_t \) at time \( t \) on \( (\Omega, \mathcal{F}, P) \).

The theory of expected utility is rigorously introduced in almost every textbooks on microeconomic theory e.g. Mas-Colell, Whinston, and Green (1995). We simply state that \( \hat{x} \) is preferred to \( x \) if and only if \( EU(t, \hat{x}) > EU(t, x) \). Moreover, two expected utility functions \( EU_1(t, x) \) and \( EU_2(t, x) \) are considered equivalent if \( EU_1(t, x) = E[a + bU_2(t, x)] \) for constant parameters \( a \in \mathbb{R} \) and \( b > 0 \).

Although it is widely used, the expected utility hypothesis has drawn some criticism. Particularly, the independence axiom is empirically not confirmed (e.g. the Allais Paradox; see Mas-Colell, Whinston, and Green (1995)). Furthermore, several observed phenomena cannot easily be explained by the theory, most notably the equity premium puzzle. Kreps and Porteus (1978) presented an alternative to the expected utility hypothesis including preferences concerning the timing of the resolution of uncertainty. In the same spirit, Weil (1990) examined theoretical and empirical implications of these preferences for macroeconomics. Another approach is presented by Epstein and Wang (1994) who distinguish between risk and uncertainty where information is too imprecise to be summarized by probabilities including multiple priors.

\(^2\)We introduce here rationality in a very pragmatic sense, without axiomatical basics as in e.g. Mas-Colell, Whinston, and Green (1995). But our utility function is defined complete and monotone and, therefore, transitivity is satisfied.
2.1.9 Pareto Optimality

Since in chapter 3 there are several actors $i \in \{1, 2, ..., m\}$ with different utility functions $U_i(t, x)$, we need a criteria to compare different allocations of risk beyond the utility of only one actor. Following Mas-Colell, Whinston, and Green (1995), we use the criterion Pareto optimality:

**Definition 2.4** If we have $m$ participants, then a feasible allocation $(x_1, x_2, ..., x_m)$ is called Pareto optimal if there is no other feasible allocation $(\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_m)$ such that $\forall i \in \{1, 2, ..., m\}$; $EU_i(t, \tilde{x}_i) \geq EU_i(t, x_i)$ and $\exists i \in \{1, 2, ..., m\}; EU_i(t, \tilde{x}_i) > EU_i(t, x_i)$.

And, therefore, we define a change in $(x_1, x_2, ..., x_m)$ as a Pareto improvement if the following definition is satisfied.

**Definition 2.5** A change from a feasible allocation $(x_1, x_2, ..., x_m)$ to another feasible allocation $(\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_m)$ is called Pareto improvement if $\forall i \in \{1, 2, ..., m\}$; $EU_i(t, \tilde{x}_i) \geq EU_i(t, x_i)$ and $\exists i \in \{1, 2, ..., m\}; EU_i(t, \tilde{x}_i) > EU_i(t, x_i)$.

2.2 The Hamilton-Jacobi-Bellmann equation

The Hamilton-Jacobi-Bellmann equation (HJB equation) appears for the first time in Bellmann (1957) but is deduced in almost all textbooks on dynamic programming e.g. Øksendal (2003) or Björk (1998). We reduce the general problem here to the formulas that are required in the next chapters.

2.2.1 General HJB Equation

Given the utility function, an initial state $(t, W_0, L_0)$ and the expected utility hypothesis, we consider a finite-horizon setting within $[t, T]$ and solve the following problem

$$ J(t, F) = \sup_{(\pi', \xi) \in \mathcal{H}} E \left[ \int_t^T U_1\left(s, F^{\pi', \xi}_s, \xi_s\right) ds + U_2\left(T, F^T_T\right) \right]. \quad (2.6) $$

Here, $U_1 : [0, T] \times \mathbb{R}^2 \to (-\infty, \infty)$ is the instant utility function defined on a fixed domain $D \subseteq [0, T] \times \mathbb{R}^2$ and $U_2 : \mathbb{R} \to (-\infty, \infty)$ is called the terminal reward utility function
and is defined on the boundary $\partial D$ of $D$. The function $J(t,F)$ is called the *indirect utility function* or briefly the $J$-function. In this section, we deal with $(\pi_t^i, \zeta_t^i)$ as control variables. If only $\pi_t$ are controls, then the admissible set $\mathcal{H}$ can be replaced by $\mathcal{G}$.

For simplicity, we assume that $L_t$ and $W_t$ are such that the ratio process will be reduced to a univariate process, i.e.

$$
\frac{dE^{t\pi_t^\delta}}{F_t^{\pi_t^\delta}} = \mu^{\pi_t^\delta}(t,F_t^{\pi_t^\delta}) dt + \sigma^{\pi_t^\delta}(t,F_t^{\pi_t^\delta}) dZ_t. 
$$

**(Theorem 2.6) Hamilton-Jacobi-Bellman equation (necessary condition)**

We use $F_t^{\pi_t^\delta}$ as in (2.7). Suppose that $J(t,F) \in C^2$ (i.e. $J$ is twice continuously differentiable) is bounded, and that the optimal solution to (2.6) $(\pi_t^*, \zeta_t^*)$ exists. Then

1. $J(t,F)$ satisfies the HJB equation

$$
\frac{\partial J(t,F)}{\partial t} + \sup_{(\pi', \zeta') \in \mathcal{H}} \left\{ \begin{array}{l} U_1(t,F,\zeta) \\
+ \mu^{\pi_t^\delta}(t,F) F \frac{\partial J(t,F)}{\partial F} \\
+ \frac{1}{2} (\sigma^{\pi_t^\delta})^2(t,F) F^2 \frac{\partial^2 J(t,F)}{\partial F^2} \end{array} \right\} = 0,
\forall (t,F) \in D 
(2.8)
$$

$$
J(t,F) = U_2(t,F),
\forall (t,F) \in \partial D 
$$

2. The supremum in the HJB equation is attained for each $(t,F)$ by choosing $\pi = \pi_t^*$ and $\zeta = \zeta_t^*$.

If $(\pi_t^*, \zeta_t^*)$ exist, we call it an optimal control policy.

**(Theorem 2.7) Verification Theorem (sufficient condition)**

Suppose that we have two functions $\bar{J}(t,F)$, $\bar{\pi}(t,F)$ and $\bar{\zeta}(t,F)$, such that

(i) $\bar{J}(t,F)$ is sufficiently integrable and solves the HJB equation

$$
\frac{\partial \bar{J}(t,F)}{\partial t} + \sup_{(\pi', \zeta') \in \mathcal{H}} \left\{ \begin{array}{l} U_1(t,F,\zeta) \\
+ \mu^{\pi_t^\delta}(t,F) F \frac{\partial \bar{J}(t,F)}{\partial F} \\
+ \frac{1}{2} (\sigma^{\pi_t^\delta})^2(t,F) F^2 \frac{\partial^2 \bar{J}(t,F)}{\partial F^2} \end{array} \right\} = 0,
\forall (t,F) \in D 
$$

$$
\bar{J}(t,F) = U_2(t,F),
\forall (t,F) \in \partial D 
$$

(ii) $(\bar{\pi'}(t,F), \bar{\zeta}(t,F))$ is an admissible control law

(iii) For each fixed $(t,F)$ the supremum in the expression

$$
\sup_{(\pi', \zeta') \in \mathcal{H}} \left\{ \begin{array}{l} U_1(t,F,\zeta) \\
+ \mu^{\pi_t^\delta}(t,F) F \frac{\partial \bar{J}(t,F)}{\partial F} \\
+ \frac{1}{2} (\sigma^{\pi_t^\delta})^2(t,F) F^2 \frac{\partial^2 \bar{J}(t,F)}{\partial F^2} \end{array} \right\}
$$
is attained by $\pi = \tilde{\pi}(t, F)$ and $\xi = \tilde{\xi}(t, F)$

Then the following hold:
1. For all controls $(\pi', \xi)$ we have
   \[ J(t, F) \leq \tilde{J}(t, F) \]
2. There exists an optimal control policy $(\pi^*, \xi^*)$ and it is $(\pi^* t, \xi^* t) = (\tilde{\pi}'(t, F), \tilde{\xi}(t, F))$.


If we use the HJB-equation on the general form of the funding ratio (2.4), of course the HJB-equation would include the first-order and second-order derivatives of $J$ with respect to $L, C$ and $S_1, S_2, S_3, \ldots, S_k$ and all cross partial derivatives.

### 2.2.2 Optimal Terminal Value

If we set $U_1 \equiv 0$ then we have the following problem;

\[ J(t, F) = \sup_{(\pi', \xi) \in H} E U_2 \left( F_{\pi'} \right). \]

Then the corresponding HJB-equation turns out to be

\[
\frac{\partial J(t, F)}{\partial t} + \sup_{\pi \in \mathcal{G}} \left\{ \mu_{\pi} F_{\pi} \frac{\partial J(t, F)}{\partial F} + \frac{1}{2} \left( \sigma_{\pi} \right)^2 (t, F) F^2 \frac{\partial^2 J(t, F)}{\partial F^2} \right\} = 0 \quad \forall (t, F) \in (0, T) \times \mathbb{R},
\]

\[ J(T, F) = U_2(F) \quad \forall F \in \mathbb{R}. \]

### 2.2.3 Time-Homogeneous Setting

If we face an infinite-horizon setting, with $U_2 \equiv 0$ and $T \to \infty$ we have

\[
J(F) = \sup_{(\pi', \xi) \in \mathcal{H}} \left[ \int_0^\infty U_1 \left( s, F_{\pi'}(s, \xi) \right) ds \right].
\]

**Proposition 2.8** If $T \to \infty$ and $U_1 \left( s, F_{\pi'}(s, \xi) \right) = e^{-qt} u \left( F_{\pi'}(s, \xi) \right)$, where $q > 0$ is a constant, the corresponding HJB-equation turns out to be

\[
-\theta J(F) + \sup_{(\pi', \xi) \in \mathcal{H}} \left\{ u(F, \xi) + \mu_{\pi} (t, F) F_{\pi} \frac{\partial J(F)}{\partial F} + \frac{1}{2} \left( \sigma_{\pi} \right)^2 (t, F) F^2 \frac{\partial^2 J(F)}{\partial F^2} \right\} = 0
\]

\[ (2.10) \]
2.2 The Hamilton-Jacobi-Bellmann equation

Proof. The $J$-function satisfies

$$J(F) = E \left[ \int_0^\infty e^{-\varrho s} u \left( F_s^{\pi^*, \bar{y}_s}, \bar{y}_s \right) ds \right]$$

where $t > 0$.

Divided by $t$ we get

$$0 = E \left[ \frac{1}{t} \int_0^t e^{-\varrho s} u \left( F_s^{\pi^*, \bar{y}_s}, \bar{y}_s \right) ds \right]$$

$$+ \frac{E \left[ J(F_t^{\pi^*, \bar{y}_s}) \right]}{t} (e^{-\varrho t} - 1) + E \left[ \frac{J(F_t^{\pi^*, \bar{y}_s}) - J(F)}{t} \right].$$

Applying Itô’s Lemma and $t \to 0$ we obtain

$$0 = u(F, \bar{y}) - \varrho J(F) + \mu^{\pi^*, \bar{y}_s}(t, F) t \frac{\partial J(F)}{\partial F}$$

$$+ \frac{1}{2} \left( \sigma^{\pi^*, \bar{y}_s} \right)^2 (t, F) t^2 \frac{\partial^2 J(F)}{\partial F^2}.$$

This result will be used in Section 4.2.

2.2.4 Long-Term Growth Rate

Next we intend to maximize the long-term growth rate of expected utility of final wealth. Since a monotone transformation of the objective function has no impact on the optimiza-
tion results we may maximize
\[
\lim_{T \to \infty} \frac{1}{\beta T} \ln \left( \beta E \left[ U_2 \left( F_T^{\pi, \xi} \right) \right] \right)
\]
where \( E \left[ U_2 \left( F_T^{\pi, \xi} \right) \right] \neq 0 \) by assumption\(^3\). The artificial parameter \( \beta \) has the same sign as \( E \left[ U_2 \left( F_T^{\pi, \xi} \right) \right] \) in order that the logarithm is defined. Note that, if \( T \) is finite, maximizing \( \frac{1}{\beta T} \ln \left( \beta E \left[ U_2 \left( F_T^{\pi, \xi} \right) \right] \right) \) is equivalent to maximizing \( EU_2 \left( F_T^{\pi, \xi} \right) \). Set\(^4\)

\[
u_F \triangleq \sup_{(\pi', \xi') \in \mathcal{H}} \lim_{T \to \infty} \frac{1}{\beta T} \ln \left( \beta E \left[ U_2 \left( F_T^{\pi, \xi} \right) \right] \right). \tag{2.11}
\]

**Theorem 2.9** If a constant \( \hat{\nu}_F \) and a function \( J (F) \) exist such that:

(i) for all \( t > 0 \), \( J (F) \) satisfies the equation

\[ J (F) = \sup_{(\pi', \xi') \in \mathcal{H}} E \left[ e^{-\beta \hat{\nu}_F t} J \left( F_{T_1}^{\pi', \xi'} \right) \right] \] \tag{2.12}

(ii) the supremum in equation (2.12) is attained for each \( F \) by choosing \( \pi = \pi_t^* \) and \( \xi = \xi_t^* \)

(iii) there exist positive constants \( c_1 \) and \( c_2 \) such that \( c_1 \beta U_2 (F) \leq \beta J (F) \leq c_2 \beta U_2 (F) \), then

1) \( \hat{\nu}_F \) is the maximum long-term growth rate of expected utility of final wealth, i.e. \( \hat{\nu}_F = \nu_F \), and it is achieved by \( \pi = \pi_t^* \) and \( \xi = \xi_t^* \).
2) it is \( \hat{\nu}_F = \bar{\nu}_F \) where

\[ \bar{\nu}_F \triangleq \lim_{T \to \infty} \frac{1}{\beta T} \ln \left[ \beta \sup_{(\pi', \xi') \in \mathcal{H}} EU_2 \left( F_T^{\pi, \xi} \right) \right] \]

The proof of theorem 2.9 is given in appendix 2.A.1.

From theorem 2.9 we have \( J (F) = E \left[ e^{-\beta \hat{\nu}_F t} J \left( F_{T_1}^{\pi', \xi'} \right) \right] \) and hence

\[
0 = E \left[ e^{-\beta \hat{\nu}_F t} J \left( F_{T_1}^{\pi', \xi'} \right) \right] - J (F) = E \left[ J \left( F_{T_1}^{\pi', \xi'} \right) \right] \left( e^{-\beta \hat{\nu}_F t} - 1 \right) + E \left[ J \left( F_{T_1}^{\pi, \xi} \right) - J (F) \right].
\]

\(^3\)For justification of this growth rate see Taskar, Klass, and Assaf (1988); Grossman and Vila (1992); Grossman and Zhou (1993) or Cvitanić and Karatzas (1995).

\(^4\)The objective function is taken from Cvitanić and Karatzas (1995) or Grossman and Zhou (1993)
Dividing by $t$, applying Itô’s Lemma to the last term and $t \to 0$ we obtain the HJB-equation

$$0 = -\beta v_t (F) + \mu^{\pi^F} (t, F) F \frac{\partial J(F)}{\partial F} + \frac{1}{2} \left( \sigma^{\pi^F} (t, F) \right)^2 (t, F) F^2 \frac{\partial^2 J(F)}{\partial F^2}. \quad (2.13)$$

### 2.2.5 A Useful Theorem

Throughout the thesis we sometimes will face a simplification of the utility function but with constraints on the control $\xi_t$. All parameters $(\mu, \sigma, r_0, r, \zeta, \mu_L, \sigma_L)$ are assumed to be deterministic and $dC = 0$. The problem we will solve is

$$\sup_{(\pi', \xi') \in \mathcal{H}} EU_2 \left( F^{\pi'_F}_{T} \right)$$

subject to $\xi_t \in [a, b], t \in [0, T]$ where $a, b \in \mathbb{R}$. The utility function is assumed to exhibit a constant relative risk aversion (CRRA), i.e.

$$U_2 (x) = U (x) \triangleq \frac{1}{1 - R} x^{1 - R}, \quad R > 1.$$  

with $-\frac{U''(x)}{U'(x)} = R$ as the parameter of relative risk aversion. From Merton (1971) we know, that the above problem leads to constant portfolio weights. In Cvitanić and Karatzas (1992), Müller (2000) and Denzler, Müller, and Scherer (2001) it is shown that this result still holds under constant constraints on $\pi_t$. Since we will need constraints on $\xi_t$ we state the following theorem.

**Theorem 2.10** We set $U (x) \triangleq \frac{1}{1 - R} x^{1 - R}, \quad R > 1$ and assume that all model parameters $(\mu, \sigma, r_0, r, \zeta, \mu_L, \sigma_L)$ are deterministic. If there exists a $F^{\pi^F}_{T} \in \mathcal{H}$ solving

$$\sup_{(\pi', \xi') \in \mathcal{H}} EU_2 \left( F^{\pi'_F}_{T} \right)$$

subject to $\xi_t \in [a, b], t \in [0, T] \quad a, b \in \mathbb{R}$

with $\mu^{\pi^F} (t, F^{\pi^F}_t) = \mu^{\pi^F} (t)$ and $\sigma^{\pi^F} (t, F^{\pi^F}_t) = \sigma^{\pi^F} (t)$, then the solution is given by $(\pi_t', \xi_t) = (\pi^*_t, \xi^*_t)$ where $(\pi^*_t', \xi^*_t)$ solve

$$\max_{(\pi', \xi') \in \mathcal{H}} \left\{ \frac{2}{R} \mu^{\pi^F} (t) - (\sigma^{\pi^F} (t))^2 \right\}$$

subject to $\xi_t \in [a, b]$.
The proof of theorem 2.10 is given in appendix 2.A.2.

**Remark 2.11** Note that theorem 2.9 directly implies that the result of theorem 2.10 holds also for

$$\sup_{(\pi', \xi) \in \mathcal{H}} \liminf_{T \to \infty} \frac{1}{T} \ln \left( \beta E \left[ U \left( F_{\pi}^{\xi} \right) \right] \right)$$

s.t. \( \xi_t \in [a, b] \).
2.A Appendix

2.A.1 Proof of Theorem 2.9

Proof. for 1)

Case $\beta > 0$

For $(\pi', \xi) \in \mathcal{H}$, $c_{1} \beta U_{2} (F) \leq \beta J (F)$ obviously implies

$$
\sup_{(\pi', \xi) \in \mathcal{H}} \liminf_{T \to \infty} \frac{1}{\beta T} \ln \left( c_{1} E \left[ \beta U_{2} \left( F_{T}^{\pi', \xi} \right) e^{-\beta \hat{\upsilon} T} \right] \right)
$$

$$
\leq \sup_{(\pi', \xi) \in \mathcal{H}} \liminf_{T \to \infty} \frac{1}{\beta T} \ln \left( \beta J \left( F_{T}^{\pi', \xi} \right) e^{-\beta \hat{\upsilon} T} \right)
$$

because $e^{-\beta \hat{\upsilon} T}$ is constant.

Further, equation (2.12) implies

$$
\frac{1}{\beta T} \ln E \left[ \beta e^{-\beta \hat{\upsilon} T} J \left( F_{T}^{\pi', \xi} \right) \right] \leq \frac{1}{\beta T} \ln (\beta J (F))
$$

and thus for $T \to \infty$

$$
\sup_{(\pi', \xi) \in \mathcal{H}} \liminf_{T \to \infty} \frac{1}{\beta T} \ln \left( c_{1} E \left[ \beta U_{2} \left( F_{T}^{\pi', \xi} \right) e^{-\beta \hat{\upsilon} T} \right] \right)
$$

$$
\leq \sup_{(\pi', \xi) \in \mathcal{H}} \liminf_{T \to \infty} \frac{1}{\beta T} \ln \left( \beta J \left( F_{T}^{\pi', \xi} \right) e^{-\beta \hat{\upsilon} T} \right) \leq 0.
$$

Note that

$$
\sup_{(\pi', \xi) \in \mathcal{H}} \liminf_{T \to \infty} \frac{1}{\beta T} \ln \left( c_{1} E \left[ \beta U_{2} \left( F_{T}^{\pi', \xi} \right) e^{-\beta \hat{\upsilon} T} \right] \right)
$$

$$
= \sup_{(\pi', \xi) \in \mathcal{H}} \liminf_{T \to \infty} \frac{1}{\beta T} \ln E \left[ \beta U_{2} \left( F_{T}^{\pi', \xi} \right) \right] - \hat{\upsilon} \varnothing
$$

for any constant $\varnothing$. Therefore

$$
\hat{\upsilon} \varnothing \geq \sup_{(\pi', \xi) \in \mathcal{H}} \liminf_{T \to \infty} \frac{1}{\beta T} \ln E \left[ \beta U_{2} \left( F_{T}^{\pi', \xi} \right) \right]
$$

(2.14)

and especially with $(\pi^{*}', \xi^{*})$

$$
\hat{\upsilon} \varnothing \geq \liminf_{T \to \infty} \frac{1}{\beta T} \ln E \left[ \beta U_{2} \left( F_{T}^{\pi^{*}', \xi^{*}} \right) \right].
$$

(2.15)
For a particular strategy \((\pi^*_t, \xi^*_t)\), the relation \(\beta J(F) \leq c_2 \beta U_2(F)\) and equation (2.12) imply
\[
E \left[ c_2 \beta U_2 \left( F_T^{\pi^*_t, \xi^*_t} \right) e^{-\beta \hat{\upsilon}_F T} \right] \geq E \left[ \beta J \left( F_T^{\pi^*_t, \xi^*_t} \right) e^{-\beta \hat{\upsilon}_F T} \right] = \beta J(F) > 0
\]
and therefore
\[
\hat{\upsilon}_F \leq \liminf_{T \to \infty} \frac{1}{\beta T} \ln E \left[ \beta U_2 \left( F_T^{\pi^*_t, \xi^*_t} \right) \right]
\]
(2.16)
and thus
\[
\hat{\upsilon}_F \leq \sup_{(\pi^*_t, \xi^*_t) \in \mathcal{H}} \liminf_{T \to \infty} \frac{1}{\beta T} \ln E \left[ \beta U_2 \left( F_T^{\pi^*_t, \xi^*_t} \right) \right].
\]
(2.17)

From (2.14) and (2.17) we have
\[
\hat{\upsilon}_F = \upsilon_F = \sup_{(\pi^*_t, \xi^*_t) \in \mathcal{H}} \liminf_{T \to \infty} \frac{1}{\beta T} \ln E \left[ \beta U_2 \left( F_T^{\pi^*_t, \xi^*_t} \right) \right]
\]
and from (2.15) and (2.16) we have
\[
\hat{\upsilon}_F = \liminf_{T \to \infty} \frac{1}{\beta T} \ln E \left[ \beta U_2 \left( F_T^{\pi^*_t, \xi^*_t} \right) \right].
\]

**Case \(\beta < 0\)**

We use the same idea as above:
If \(\beta J(F) \leq c_2 \beta U_2(F)\) then
\[
\sup_{(\pi^*_t, \xi^*_t) \in \mathcal{H}} \liminf_{T \to \infty} \frac{1}{\beta T} \ln \left( c_2 E \left[ \beta U_2 \left( F_T^{\pi^*_t, \xi^*_t} \right) e^{-\beta \hat{\upsilon}_F T} \right] \right)
\]
\[
\leq \sup_{(\pi^*_t, \xi^*_t) \in \mathcal{H}} \liminf_{T \to \infty} \frac{1}{\beta T} \ln E \left[ \beta J \left( F_T^{\pi^*_t, \xi^*_t} \right) e^{-\beta \hat{\upsilon}_F T} \right].
\]

Note that
\[
\sup_{(\pi^*_t, \xi^*_t) \in \mathcal{H}} \liminf_{T \to \infty} \frac{1}{\beta T} \ln \left( c_2 E \left[ \beta U_2 \left( F_T^{\pi^*_t, \xi^*_t} \right) e^{-\beta \hat{\upsilon}_F T} \right] \right)
\]
\[
= \sup_{(\pi^*_t, \xi^*_t) \in \mathcal{H}} \liminf_{T \to \infty} \frac{1}{\beta T} \ln E \left[ \beta U_2 \left( F_T^{\pi^*_t, \xi^*_t} \right) \right] - \hat{\upsilon}_F
\]
for any constant \(\hat{\upsilon}_F\).

Thus, equation (2.12) implies
\[
\sup_{(\pi^*_t, \xi^*_t) \in \mathcal{H}} \liminf_{T \to \infty} \frac{1}{\beta T} \ln E \left[ \beta U_2 \left( F_T^{\pi^*_t, \xi^*_t} \right) \right] - \hat{\upsilon}_F
\]
\[
\leq \sup_{(\pi^*_t, \xi^*_t) \in \mathcal{H}} \liminf_{T \to \infty} \frac{1}{\beta T} \ln E \left[ \beta J \left( F_T^{\pi^*_t, \xi^*_t} \right) e^{-\beta \hat{\upsilon}_F T} \right] \leq 0
\]
and thus
\[ \vartheta_F \geq \sup_{(\pi', \xi) \in H} \liminf_{T \to \infty} \frac{1}{\beta T} \ln E \left[ \beta U_2 \left( \mathcal{F}_{\pi, \xi}^T \right) \right]. \]  
(2.18)

For a particular strategy \((\pi_1^*, \xi_1^*)\)
\[ \vartheta_F \geq \liminf_{T \to \infty} \frac{1}{\beta T} \ln E \left[ \beta U_2 \left( \mathcal{F}_{\pi_1^*, \xi_1^*}^T \right) \right]. \]  
(2.19)

Again \(\beta J(F) \geq c_1 \beta U_2(F)\) and equation (2.12) imply (same argumentation as above)
\[ \vartheta_F \leq \liminf_{T \to \infty} \frac{1}{\beta T} \ln E \left[ \beta U_2 \left( \mathcal{F}_{\pi_1^*, \xi_1^*}^T \right) \right]. \]  
(2.20)

and therefore
\[ \vartheta_F \leq \sup_{(\pi', \xi) \in H} \liminf_{T \to \infty} \frac{1}{\beta T} \ln E \left[ \beta U_2 \left( \mathcal{F}_{\pi', \xi}^T \right) \right]. \]  
(2.21)

Thus we have from (2.18) and (2.21) the equation
\[ \vartheta_F = \vartheta_F = \sup_{(\pi', \xi) \in H} \liminf_{T \to \infty} \frac{1}{\beta T} \ln E \left[ \beta U_2 \left( \mathcal{F}_{\pi', \xi}^T \right) \right] \]
and from (2.19) and (2.20) the equation
\[ \vartheta_F = \liminf_{T \to \infty} \frac{1}{\beta T} \ln E \left[ \beta U_2 \left( \mathcal{F}_{\pi_1^*, \xi_1^*}^T \right) \right]. \]

Thus the first part of theorem 2.9, i.e.1), is proven. The second part is very simple to prove:

Proof for 2)
If \(c_3 U_2(F) \leq J(F) \leq c_4 U_2(F)\) then, because \(e^{-\beta \vartheta T}\) is a constant,
\[ c_3 U_2 \left( \mathcal{F}_{\pi, \xi}^T \right) e^{-\beta \vartheta T} \leq J \left( \mathcal{F}_{\pi, \xi}^T \right) e^{-\beta \vartheta T} \leq c_4 U_2 \left( \mathcal{F}_{\pi, \xi}^T \right) e^{-\beta \vartheta T}. \]

Taking expectations and then the supremum over \((\pi', \xi) \in H\)
\[ \sup_{(\pi', \xi) \in H} E U_2 \left( \mathcal{F}_{\pi, \xi}^T \right) c_3 e^{-\beta \vartheta T} \leq J(F) \leq \sup_{(\pi', \xi) \in H} E U_2 \left( \mathcal{F}_{\pi, \xi}^T \right) c_4 e^{-\beta \vartheta T} \]
and multiplying by \(\beta\), taking logarithms, dividing by \(T\) and finally \(T \to \infty\) we get
\[ \vartheta_F = \liminf_{T \to \infty} \frac{1}{\beta T} \ln \left[ \beta \sup_{(\pi', \xi) \in H} E U_2 \left( \mathcal{F}_{\pi, \xi}^T \right) \right]. \]
2.A.2 Proof of Theorem 2.10

The proof of theorem 2.10 follows Müller (2000).

Proof. Obviously, we can decompose the expected utility as follows

\[
E \left[ \left( F_T^{\pi, \delta} \right)^{1-R} \right] = E_0^{1-R} E \left[ \left( \frac{F_t^{\pi, \delta}}{F_0} \right)^{1-R} \cdot \left( \frac{F_T^{\pi, \delta}}{F_t^{\pi, \delta}} \right)^{1-R} \right]
\]

where \( t \in [0, T] \). Using the law of conditional expectations we get

\[
E \left[ \left( F_T^{\pi, \delta} \right)^{1-R} \right] = E_0^{1-R} \left\{ \left( \frac{F_t^{\pi, \delta}}{F_0} \right)^{1-R} \cdot E \left[ \left( \frac{F_T^{\pi, \delta}}{F_t^{\pi, \delta}} \right)^{1-R} \left| \mathcal{F}_t \right. \right] \right\}.
\]

Since all model parameters are assumed to be deterministic, the probability law of the underlying processes

\[
\frac{S_i(t + h)}{S_i(t)}, i = 1, ..., k, h \in [0, T - t]
\]

does not depend on \( \mathcal{F}_t \). Therefore, the Bellman principle and the strict concavity of \( U(x) \equiv \frac{1}{1-R} x^{1-R}, R > 1 \) ensure that

\[
E \left[ \left( \frac{F_T^{\pi, \delta}}{F_t^{\pi, \delta}} \right)^{1-R} \left| \mathcal{F}_t \right. \right] = E \left[ \left( \frac{F_T^{\pi, \delta}}{F_t^{\pi, \delta}} \right)^{1-R} \right].
\]

Therefore, we have

\[
E \left[ \left( F_T^{\pi, \delta} \right)^{1-R} \right] = E_0^{1-R} \left\{ \left( \frac{F_t^{\pi, \delta}}{F_0} \right)^{1-R} \cdot E \left[ \left( \frac{F_T^{\pi, \delta}}{F_t^{\pi, \delta}} \right)^{1-R} \left| \mathcal{F}_t \right. \right] \right\}
\]

\[
= E_0^{1-R} \left( \frac{F_t^{\pi, \delta}}{F_0} \right)^{1-R} \cdot E \left[ \left( \frac{F_T^{\pi, \delta}}{F_t^{\pi, \delta}} \right)^{1-R} \left| \mathcal{F}_t \right. \right].
\]

Next we use the law of conditional expectations, once more, to get

\[
E \left[ \left( F_T^{\pi, \delta} \right)^{1-R} \right] = E_0^{1-R} E \left[ \left( \frac{F_t^{\pi, \delta}}{F_0} \right)^{1-R} \cdot E \left[ \left( \frac{F_T^{\pi, \delta}}{F_t^{\pi, \delta}} \right)^{1-R} \left| \mathcal{F}_t+dt \right. \right] \right].
\]

Arguing the same way as above and multiplying the equation by \( \frac{1}{1-R} \), finally, we have for the expected utility

\[
E \left[ \frac{F^\pi_T}{1-R} \right] = F^\pi_0 \cdot \frac{1}{1-R} E \left[ \frac{F^\pi_T}{F^\pi_0} \right] \cdot E \left[ \frac{F^\pi_T}{F^\pi_{t+dt}} \right] \cdot E \left[ \frac{F^\pi_T}{F^\pi_T} \right].
\]

Therefore, an investment strategy \( \pi_t^*, t \in [0, T] \) is optimal if and only if it solves for all \( t \in [0, T] \)

\[
\max_{(\pi', \xi) \in H} \frac{1}{1-R} E \left[ \frac{F^\pi_T}{F^\pi_{t+dt}} \right] \quad \text{s.t.} \quad \xi_t \in [a, b], t \in [0, T], a, b \in \mathbb{R}.
\]

Since

\[
\left( \frac{F^\pi_T}{F^\pi_{t+dt}} \right)^{1-R} = \exp \left\{ (1-R) \left[ \mu^\pi^\xi (t) - \frac{(\sigma^\pi^\xi (t))^2}{2} \right] dt + \sigma^\pi^\xi (t) dZ_t \right\}
\]

are lognormally distributed we get

\[
E \left[ \frac{F^\pi_T}{F^\pi_{t+dt}} \right] = \exp \left\{ (1-R) \left[ \mu^\pi^\xi (t) - \frac{R \left( \sigma^\pi^\xi (t) \right)^2}{2} \right] dt \right\}.
\]

Hence \( \pi_t^* \) maximizes S if and only if it solves

\[
\max_{(\pi', \xi)} \left\{ \frac{1}{R} \mu^\pi^\xi (t) - \frac{\left( \sigma^\pi^\xi (t) \right)^2}{2} \right\} \quad \text{s.t.} \quad \xi_t \in [a, b].
\]
Chapter 3

Asset Liability Optimization with Exogenous Liability

In this chapter we will consider a plan sponsor who wants to optimize the long-term growth rate of the funding ratio. We will take the situation of a deterministic yield guarantee which is mandatory for every pension fund - a setting which corresponds to the Swiss situation.\footnote{BVG Art. 15 Abs. 2} Further, as we will show, if the yield guarantee is higher than the riskless yield, then it causes a serious short-fall risk for the plan sponsor. The impact of the yield guarantee on the plan sponsor’s behavior is that he has to invest substantially into risky assets in order to avoid slipping into a funding gap. These issues will be examined within an incomplete markets setting.

But before we turn to this hybrid plan setting, we briefly revisit the well-known results of Merton (1969).

3.1 Merton’s Problem Revisited

As it is the standard portfolio optimization problem in modern portfolio theory, we will give a short review of the Merton problem (Merton (1969)). It will also serve as a standard with which we can compare the results of our more specific optimization problems.

3.1.1 Optimal Consumption and Bequest Motive

We assume a power utility function of the CRRA-type, \( U(t, x) \triangleq \frac{1}{1-R} e^{-\kappa t} x^{1-R}, q > 0, R > 1 \). Further, all model parameters are assumed to be constant. We face the problem

\[
J(t, W) = \sup_{(\pi', \xi) \in \mathcal{H}} E \left[ \int_t^T U(t, \xi_s) \, ds + \epsilon R U \left( T, W_{T}^{\pi', \xi} \right) \right]
\]
where \( \xi_t \) is the consumption, \( \epsilon \geq 0 \) a "bequest parameter" and

\[
dW_t^{\pi, r} = W_t^{\pi, r} (\pi'_t (\mu - r \mathbf{1}) + r) \, dt + \pi'_t \sigma \, dZ_t - \xi_t \, dt. \tag{3.1}
\]

This problem leads to the HJB-equation

\[
\frac{\partial J(t, W)}{\partial t} + \sup_{(\pi'_t, \xi_t) \in H} \left\{ e^{-\eta t} \frac{\xi_t}{1 - R} \left( W_t^{\pi, r} (\pi'_t (\mu - r \mathbf{1}) + r) \right) \right. \\
+ \left. \frac{1}{2} \pi'_t \Sigma \pi'_t W^2 \frac{\partial^2 I_t(W)}{\partial W^2} \right\} = 0 \tag{3.1}
\]

with \( J(T, W) = e^{R} e^{-\eta T \frac{W_1 - R}{1 - R}} \). For simpler notation henceforth we will write \( J_t \triangleq \frac{\partial I_t(W)}{\partial t} \), \( J_W \triangleq \frac{\partial^2 I_t(W)}{\partial W \partial W} \).

The optimal portfolio becomes

\[
\pi'_t = -\frac{J_W}{W J_W} \pi, \tag{3.2}
\]

where \( \pi \triangleq \Sigma^{-1} (\mu - r \mathbf{1}) \) is the growth-optimal portfolio.

The optimal consumption is

\[
\xi'_t = \left( J_W e^{\eta t} \right)^{-\frac{1}{\pi}}. \tag{3.3}
\]

Substituting (3.2) and (3.3) into (3.1) gives

\[
J_t + e^{-\eta t} \frac{R}{1 - R} J_W^{\frac{1}{\pi}} - \frac{1}{2} J_W^2 \pi' \pi W^2 + W R J_W = 0 \tag{3.4}
\]

with \( J(T, W) = e^{R} e^{-\eta T \frac{W_1 - R}{1 - R}} \).

To solve the partial differential equation 3.4 we use the method of undetermined coefficients, i.e. we try a value function \( J(t, W) \) in a parameterized form and solve for the parameters. The value function we try has the property that the state variable \( W \) disappears after substituting into (3.4). The variable \( t \) and the parameters can be found in a second step.

Thus, we try

\[
J(t, W) = h(t) e^{-\eta t} \frac{W_1 - R}{1 - R} \tag{3.5}
\]

and get

\[
h'(t) = \frac{R}{2} (kh(t) - (h(t))^{\frac{1}{\pi}}) \]
\[
h(T) = e^R \tag{3.6}
\]
where
\[ \kappa \triangleq \varrho + (R - 1) \left( \frac{\hat{\pi}' \hat{\pi} + r}{R} \right) \]
which is positive. The solution to (3.6) is
\[ h(t) = \left( 1 + (\kappa \varrho - 1) \exp(\kappa (t - T)) \right)^R. \]
Substituting into (3.5) the solutions are
\[ \pi^* = \frac{1}{R} \hat{\pi} \]
and
\[ \xi^*_t = \frac{\kappa}{1 + (\kappa \varrho - 1) \exp(\kappa (t - T))} W_t. \]
Due to the multiplicative structure of the model, i.e. power utility function and the assumption of stationary return distribution, the optimal portfolio weights are neither time-dependent nor wealth-dependent. The investor holds a constant proportion (reciprocal to the relative risk aversion) of the growth-optimal portfolio and invests \((1 - 1' \pi^*) \cdot 100\) per cent of his wealth into the riskless asset.

The optimal consumption is linear in wealth. If \(T \to \infty\) we get
\[ \xi^*_t = \kappa W_t. \]

### 3.1.2 Terminal Wealth
If we focus just on the expected utility of the terminal wealth only, the problem becomes simpler.

\[ J(t, W) = \sup_{\pi \in \mathcal{G}} E \left[ U(T, W_T^\pi) \right]. \]

We can use the HJB-equation (2.9).

\[ J_t + \sup_{\pi \in \mathcal{G}} \left\{ W (\pi'_t (\mu - r 1) + r) f_W + \frac{1}{2} \pi'_t \Sigma \pi_t W^2 f_{WW} \right\} = 0, \]
\[ J(T, W) = \frac{W_{1 - \varrho}}{1 - \kappa}. \]
The optimal portfolio becomes
\[ \pi^*_t = -\frac{J_t}{W_t J_{WW}} \hat{\pi}. \] (3.8)

Substituting the optimal portfolio into (3.7) we get
\[ J_t - \frac{J_t^2}{J_{WW}} \hat{\pi}' \Sigma \hat{\pi} + W r J_t = 0, \]
\[ J(T, W) = \frac{W_t^{1-R}}{1-R}. \]

To solve the remaining problem we try
\[ J(t, W) = h(t) \frac{W_t^{1-R}}{1-R} \]
and get a homogenous ordinary differential equation
\[ h'(t) + h(t) \left[ (1-R) r + \frac{1-R}{2R} \hat{\pi}' \Sigma \hat{\pi} \right] = 0, \]
\[ h(T) = 1, \]
for which of course a solution exists. Independent of the solution of \( h(t) \) the optimal portfolio becomes
\[ \pi^* = \frac{1}{R} \hat{\pi} \] (3.9)
which is the same as in the previous problem which included consumption.

Besides, using theorem 2.10 we can directly calculate
\[ \max_{\pi} \left\{ \frac{2}{R} \left( \pi' (\mu - r 1) + r \right) - \pi' \Sigma \pi \right\} \]
to get the optimal portfolio \( \pi^* = \frac{1}{R} \hat{\pi} \).

In the following section we examine a pension adequate problem.

### 3.2 Optimal Funding Ratio

In this section we develop the optimal portfolio choice for a DC plan sponsor who has to ensure a yield of \( r_0 (t) \). This yield \( r_0 \) is exogenously given by the government. The liability contains various components such that the plan sponsor faces the risk of slipping into a funding gap. As we assume that he aims to accomplish the postulated benefits and fears
3.2 Optimal Funding Ratio

A shortfall, the state variable is the funding ratio. Since a plan sponsor must be able to sustain his actions, he is assumed to optimize its long-term growth rate (2.11), i.e.

\[
\sup \liminf_{T \to \infty} \frac{1}{\beta T} \ln (\beta E [U_2 (F^T)]) .
\]

Hence the plan sponsor will intend to optimize the funding ratio instead of the funding wealth because of the liability risks.

As stated in section 1.4, our approach differs from Boulier et al. or Deelstra et al. relative to; the structure of the guarantee, the argument within the utility function and the assumption of the term structure. Additionally, since we act in incomplete markets, there is the opportunity of a guaranteed yield exceeding the riskless yield.\(^2\)

3.2.1 General Model

The value of the wealth \(W^\pi_t\) evolves as

\[
dW^\pi_t = W^\pi_t \left( (\pi^t (\mu_t - r_t 1) + r_t) dt + \pi^t \sigma_t dZ_t \right) + dC_t - \varsigma_t W^\pi_t dt \tag{3.10}
\]

and the liability process (2.1)

\[
dL_t = L_t r_0 (t) dt + dC_t,
\]

We slightly rewrite the liability process as

\[
dL_t = L_t \left[ \pi_L (t) \left( (\mu_t - r_t 1) dt + \sigma_t dZ_t \right) + \lambda_t \left( (\mu_\lambda (t) - r_t) dt + \sigma_\lambda (t) dZ_\lambda \right) + r_t dt \right] + dC_t. \tag{3.11}
\]

The components of the process correspond to the previous notation in the following way.

\[
\sigma_L (t) = \lambda_t \sigma_\lambda (t),
\]

and

\[
\mu_L (t) = \theta (\lambda_t (\mu_\lambda (t) - r_t) + r_t) + \pi_L (t) (\mu_t - r_t 1),
\]

and

\[
r_0 (t) = (1 - \theta) (\lambda_t (\mu_\lambda (t) - r_t) + r_t),
\]

\(^2\)In complete markets \(r_0 > r_R\) would imply an arbitrage opportunity since \(r_0\) could be traded.
where $\theta \in [0, 1]$ is a weight parameter.

The vector $\mathbf{\pi}_L(t) = \sigma_t^{-1} \mathbf{b}(t)$ is the risky and tradeable part of the liability.

The value $\lambda$ is the weight of the untradeable part of the liability

$$\frac{dL_\lambda^\lambda}{L_t^\lambda} = \mu_\lambda(t) \, dt + \sigma_\lambda(t) \, dZ_\lambda^\lambda.$$  

Lastly, $1 - \mathbf{\pi}_L \mathbf{1} - \lambda$ is the weight of the riskless asset in the liability. The setting with $\lambda = 1$ but $\mu_\lambda = \sigma_\lambda = 0$ and $\mathbf{\pi}_L = \mathbf{0}$ is the no liability case. If we ignore the net contribution at the moment (i.e. $dC_t = 0$), this notation leads us to the following interpretation:

For $\mathbf{\pi}_L = \mathbf{0}$ and $\sigma_\lambda \neq 0$, the liability risk cannot be hedged. For $\lambda = 0$ the liability can be fully tracked and, moreover, it is fully tradeable, i.e. markets are complete. For $\lambda \neq 0$ and $\sigma_\lambda = 0$ the liability can be tracked but as long as $\mu_\lambda \neq 0$ it is not tradeable.

Now let us fix the net contribution process. In order to reduce the funding ratio to a single variable process for technical simplicity, we assume the net contribution to be proportionally linked with the funding wealth

$$dC_t = W_t \left[ \pi_C(t)' \left( (\mu_t - r_1) \, dt + \sigma_t dZ_t \right) \right. + \gamma_t \left( \left( \mu_\gamma(t) - r_2 \right) \, dt + \sigma_\gamma(t) dZ_\gamma^t \right) + r_t \, dt \right].$$  

The vector $\mathbf{\pi}_C(t)$ is the tradeable part of the contributions.

For $\mathbf{\pi}_L = \mathbf{0}$, $\mathbf{\pi}_C = \mathbf{0}$ and $\lambda = \gamma = 1$ the net contribution and the liability are completely unhedgeable. For $\lambda = \gamma = 0$ the net contribution and the liability can be tracked and, moreover, both are fully replicable by assets.

As already stated in the previous chapter, the administrative costs are simply stated as a linear function of $W_t^2$ i.e. $ds_t = \zeta_t W_t^2 \, dt$. Note that we could easily add stochastic terms to $ds_t$, but firstly, we would not gain more insights, and secondly, in practice, these costs are rather rigid.

Taking everything into account, if and only if $\zeta = \lambda = \gamma = 0$, then markets are complete.

Now we focus on the funding ratio. Henceforth, we drop the $t$-subscripts of all parame-
3.2 Optimal Funding Ratio

The first order condition leads to the optimal portfolio

\[ \pi_t^* = - \frac{I_F}{F_t \Pi_{FF}} (\hat{\pi} - F_t \pi_C - \pi_L) + (F_t - 1) \pi_C + \pi_L. \]
Substituting into (3.12) the HJB-equation becomes

\[
0 = -\beta v_F \left( (1 - F) r - \zeta + \gamma (1 - F) \left( \mu_\gamma - r \right) - \lambda \left( \mu_\lambda - r \right) - F (1 - F) (\gamma \sigma_\gamma)^2 + (\lambda \sigma_\lambda)^2 \right) + F J_F \\
- \frac{F^2}{2} \left( \hat{\pi} - F \pi_C - \pi_L \right)' \Sigma (\hat{\pi} - F \pi_C - \pi_L) + \frac{F^2 I_{FF}}{2} (\lambda \sigma_\lambda)^2, 
\]

(3.14)

\[
J(T,F) = U_2(F).
\]

As usual, this general HJB-equation is not explicitly solvable.

### 3.2.2 Special Case: No Net Contributions

Suppose that \( dC_t = 0 \). This means that the contributions of active members equal benefits for retired members. Then the funding ratio evolves as

\[
\frac{dF_t}{F_t} = \left( (\pi_t - \pi_L)' \Sigma (\hat{\pi} - \pi_L) - \lambda (\mu_\lambda - r) + \lambda^2 \sigma_\lambda^2 - \zeta \right) dt \\
+ (\pi_t - \pi_L)' \sigma dZ_t - \lambda \sigma_\lambda dZ_t^\lambda
\]

and the HJB-equation becomes

\[
0 = -\beta v_F J + F J_F \left( r - \zeta - \lambda (\mu_\lambda - r) + (\lambda \sigma_\lambda)^2 \right) \\
- \frac{F^2}{2} \left( \hat{\pi} - \pi_L \right)' \Sigma (\hat{\pi} - \pi_L) + \frac{F^2 I_{FF}}{2} (\lambda \sigma_\lambda)^2,
\]

\[
J(T,F) = U_2(F)
\]

and the portfolio

\[
\pi_t^* = -\frac{F}{F_{1FF}} (\hat{\pi} - \pi_L) + \pi_L.
\]

(3.17)

Note that a tradeable part in the liability, while optimizing the funding ratio, results in a scale shift. The result is similar to optimizing the portfolio \( \bar{\pi}_t = \pi_t - \pi_L \) within an optimal wealth framework as in the Merton problem.

We will discuss this result within a special case.
3.2 Optimal Funding Ratio

The utility function is assumed to exhibit a constant relative risk aversion \( R > 1 \), i.e. \( U_2(x) = U(x) = \frac{1}{1-R} x^{1-R} \). It is further assumed that all model parameters depend only on time.

We intend to optimize the long term growth rate, i.e.

\[
\sup_{\pi \in \mathcal{G}} \liminf_{T \to \infty} \frac{1}{(1-R)T} \ln (1-R) E [U(F_T^\pi)] .
\]

Using the theorems 2.9 and 2.10 the optimization problem turns out to be

\[
\max_{\pi} \left\{ \frac{2}{R} (\pi - \pi_L)' \Sigma (\hat{\pi} - \pi_L) - (\pi - \pi_L)' \Sigma (\pi - \pi_L) \right\} .
\]

The portfolio becomes

\[ \pi^* = \frac{1}{R} (\hat{\pi} - \pi_L) + \pi_L. \quad (3.18) \]

The optimal portfolio is constant and consists of the growth-optimal and \( \pi_L \). The optimal portfolio can also be written as

\[ \pi^* = \frac{1}{R} \hat{\pi} + \frac{R-1}{R} \pi_L. \]

We can see that the weight of each portfolio depends on the relative risk aversion \( R \). A very risk averse plan sponsor (i.e. \( R \to \infty \)) will hedge the full risk of the tradeable part of the liability. He will only invest in \( \pi_L \), the liability hedge portfolio, in order to minimize the variance of the funding ratio. Note that for \( \pi_L = 0 \) we get the Merton portfolio 3.9.

The funding ratio evolves as

\[
\frac{dF_t}{F_t} = \left( \frac{1}{R} (\hat{\pi} - \pi_L)' \Sigma (\hat{\pi} - \pi_L) - \zeta \right) dt + \frac{1}{R} (\hat{\pi} - \pi_L) \sigma dZ_t
- \lambda \left( (\mu_\lambda - r) dt + \sigma_\lambda d\lambda Z_t^\lambda \right) + (\lambda \sigma_\lambda)^2 dt,
\]

or

\[
F_t = F_0 \exp \left[ \left( \frac{R-1}{R} (\hat{\pi} - \pi_L)' \Sigma (\hat{\pi} - \pi_L) - \zeta \right) t + \frac{1}{R} (\hat{\pi} - \pi_L) \sigma Z_t \right]
- \lambda \left( (\mu_\lambda - r) t + \sigma_\lambda Z_t^\lambda \right) + \frac{1}{2 \lambda} (\lambda \sigma_\lambda)^2 t .
\]

We now examine the optimal long-term growth rate of \( F_t \).

\[
\liminf_{T \to \infty} \frac{1}{(1-R)T} \ln (1-R) E [U_2(F_T)]
= \frac{1}{2R} (\hat{\pi} - \pi_L)' \Sigma (\hat{\pi} - \pi_L) - \zeta - \lambda (\mu_\lambda - r) + \frac{2-R}{2} (\lambda \sigma_\lambda)^2 .
\]
The untradeable part of the liability $L^\lambda_t$ has, therefore, two implications for the expected utility of the funding ratio. 

Firstly, the impact of the drift parameter $\mu_\lambda$. The higher the drift, the lower the expected utility. Of course, if the untradeable part of the liability $\lambda$ is big, then this effect will be even larger.

Secondly, the impact of the diffusion parameter $\sigma_\lambda$. If $R > 2$ then the effect of the unhedgeable variance $\sigma^2_\lambda$ on expected utility is negative. This is intuitive, since the inability to construct a perfect hedge portfolio has a negative effect on a risk-averse investor’s utility. But if $1 < R < 2$ the result is converse which is not intuitive. The reason for this effect is the optimization of a ratio process. Because of the Itô-Lemma the diffusion parameter $\sigma_\lambda$ of the liability affects positively the drift of the ratio process. If the plan sponsor is not risk averse enough, this effect is overwhelming.

After we have seen the impact of the untradeable part of the liability, for simpler notation, we undo the notational changes in (3.11), i.e. $r_0 = (1 - \theta) (\lambda (\mu_\lambda - r) + r)$ and $\sigma_L = \lambda \sigma_\lambda$ and $\mu_L = \theta (\lambda (\mu_\lambda - r) + r) + \pi'_L (\mu - r 1)$. However, to separate the tradeable and the untradeable part of $\mu_L$ we define $\bar{\mu}_L \triangleq \mu_L - \pi'_L (\mu - r 1)$ for the untradeable part of $\mu_L$. Thus, the liability becomes

$$dL_t = L_t \left( \pi'_L ((\mu - r 1) dt + \sigma dZ_t) + \bar{\mu}_L dt + r_0 dt + \sigma_L dZ^\lambda_t \right) + dC_t \quad (3.20)$$

Henceforth, throughout the thesis, we will keep this notation for the liability. Consequently, the funding ratio becomes

$$\frac{dF_t}{F_t} = \left( \frac{1}{R} (\hat{\pi} - \pi_L) \Sigma (\hat{\pi} - \pi_L) + r - \zeta - \bar{\mu}_L - r_0 + \frac{1}{2} \sigma^2_L \right) dt$$
$$+ \frac{1}{R} (\hat{\pi} - \pi_L) \sigma dZ_t + \sigma_L dZ^\lambda_t \quad (3.21)$$

or, solving the SDE,

$$F_t = F_0 \exp \left[ \left( \frac{R - 1}{R} (\hat{\pi} - \pi_L) \Sigma (\hat{\pi} - \pi_L) + r - \zeta - \bar{\mu}_L - r_0 + \frac{1}{2} \sigma^2_L \right) t \right]$$
$$+ \frac{1}{R} (\hat{\pi} - \pi_L) \sigma Z_t + \sigma_L Z^\lambda_t$$

### 3.3 The Impact of the Yield Guarantee

In this section we consider the concrete results from the special case of the previous section and examine the impact of different “high” yield guarantees. In particular, we are
interested in the following questions: If a pension has a funding gap, i.e. $F_0 < 1$, how long does it take until this gap vanishes, if the pension follows its strategy? If a pension is in the comfortable situation that $F_0 > 1$, what is the probability, that this situation can be maintained, if the pension follows its strategy?

Henceforth, we assume that all model parameters remain constant. Additionally, for the moment, let us suppose that the funding ratio follows the geometric Brownian motion

$$\frac{dF_t}{F_t} = \mu t + \sigma dZ_t$$

Then the term $\mu$ is the growth rate of the expected funding ratio $F_t$, i.e. $\frac{\partial \ln E F_t}{\partial t} = \mu$. It is quite intuitive that a plan sponsor facing $\mu < 0$ has a serious problem since the funding ratio is expected to decrease. Either, if the pension is in a funding gap, then we expect that by following the plan sponsor’s optimal strategy it will never leave this gap, or, if it has a debit balance, then with probability one it will slip into a funding gap. The crucial value is not the growth rate of the expected funding ratio, but the growth rate of the median of the funding ratio, i.e. $\frac{\partial \ln \text{Me}(F_t)}{\partial t} = \mu - \frac{1}{2} \sigma^2$. Even though the expected funding ratio might be growing, if the median is falling, then the pension is expected never to leave a funding gap. To show this, we will use the following proposition.

**Proposition 3.1** Suppose $\frac{dF_t}{F_t} = \mu t + \sigma dZ_t$ and $\mu, \sigma$ constant. If $\mu - \frac{1}{2} \sigma^2 > 0$ the probability $p_d$ that $F_t$ ever hits a constant $d < F_0$ is

$$p_d = \left( \frac{d}{F_0} \right)^{\frac{2\mu - \sigma^2}{\sigma^2}}$$

(3.22)

and the expected value of the stopping time $E \left[ \tau (F, h) \right]$, where $h > F_0$ becomes

$$E \left[ \tau (F, h) \right] = \ln \frac{h}{F_0} \frac{1}{\mu - \frac{1}{2} \sigma^2}$$

If $2\mu - \sigma^2 < 0$ the probability that $F_t$ ever hits a constant $d < F_0$ is

$$p_d = 1$$

and the expected value of the stopping time $E \left[ \tau (F, h) \right]$, where $h > F_0$, does not exist.

The proof of proposition 3.1 is given in appendix 3.A.1.
Now let us use this proposition to examine the special case of the previous section. We consider the case without net contributions. The funding ratio (3.21) in its simplest form is

$$\frac{dF_t}{F_t} = \mu dt + \sigma dZ_t$$

where

$$\mu = \frac{1}{R} (\hat{\pi} - \pi_L)' \Sigma (\hat{\pi} - \pi_L) + r - r_0 - \bar{\mu}_L - \zeta + \sigma^2 = \text{const}$$

and

$$\sigma^2 = \frac{1}{R^2} (\hat{\pi} - \pi_L)' \Sigma (\hat{\pi} - \pi_L) + \sigma^2_L = \text{const}$$

Thus, in this case we have for the growth rate of the median

$$\mu - \frac{\sigma^2}{2} = \frac{R - \frac{1}{R^2}}{2} (\hat{\pi} - \pi_L)' \Sigma (\hat{\pi} - \pi_L) + r - \bar{\mu}_L - \zeta - r_0 + \frac{\sigma^2_L}{2}$$

This term is positive, if

$$r_0 < \hat{r}_0 \triangleq r - \bar{\mu}_L - \zeta + \frac{\sigma^2_L}{2} + \frac{2R - 1}{2R^2} (\hat{\pi} - \pi_L)' \Sigma (\hat{\pi} - \pi_L)$$

Because \( \Sigma \) is positive-semi-definite, the last addend is always positive. Thus, if \( r_0 > r + \frac{\sigma^2_L}{2} - \bar{\mu}_L - \zeta \), a pension with a smaller risk aversion in its portfolio strategy can afford a higher minimum yield, than the converse. Or, in other words, if \( r_0 > r + \frac{\sigma^2_L}{2} - \bar{\mu}_L - \zeta \), only plan sponsors which do not invest too conservatively, i.e. \( R \) is low enough, will at least have a chance of surviving. In the case that \( r_0 \leq r + \frac{\sigma^2_L}{2} - \bar{\mu}_L - \zeta \), the plan sponsor is always save if he invests in the riskless asset, i.e. also a plan sponsor with \( R \to \infty \) can invest according to his preferences.

In the following subsections we will examine the effect of \( r_0 \) on an underfunded and an overfunded pension. Henceforth, we assume that \( \mu - \frac{\sigma^2}{2} > 0 \).

### 3.3.1 Underfunded Pension

From proposition 3.1 we have

$$E [\tau (F, h)] = \frac{\ln h - \ln F_0}{2R - 1} (\hat{\pi} - \pi_L)' \Sigma (\hat{\pi} - \pi_L) + r - r_0 - \zeta - \bar{\mu}_L + \frac{\sigma^2}{2}. \quad (3.23)$$

It is intuitively obvious, that \( E [\tau (F, h)] \) becomes smallest, if the pension takes as much risk as possible. Additionally, only if \( r_0 < r + \frac{\sigma^2_L}{2} - \bar{\mu}_L - \zeta \) the pension has the opportunity
3.3 The Impact of the Yield Guarantee

to invest very conservatively in order to reach \( h \). Otherwise, if \( r_0 \geq r + \frac{\sigma^2}{2} - \bar{\mu}_L - \varsigma \), does the plan sponsor have to take more risk if he intends to reach \( h \). However, let us turn the focus: If the government fixes the yield guarantee \( r_0 \), then what yield guarantee would be affordable for a representative pension with risk aversion \( R \) and current funding ratio \( F_0 < 1 \) if it is expected to reach a debit balance, i.e. \( h = 1 \), within \( E[\tau(F,1)] = T \) years? The answer can be found by solving (3.23) for \( r_0 \). We get

\[
 r_0 < \hat{r}_{0,T} \triangleq r - \bar{\mu}_L - \varsigma + \frac{\sigma^2}{2} \\
+ \frac{2R - 1}{2R^2} (\hat{\pi} - \pi_L)' \Sigma (\hat{\pi} - \pi_L) + \frac{\ln F_0}{T}.
\]

Thus, for increasing \( R \) the value \( \hat{r}_{0,T} \) becomes smaller. In Figure 3.1, different \( \hat{r}_{0,T} \) subject to \( R \) are shown for specified parameters. We can see that if the government sets \( r_0 = 2.5\% \) (dotted line) then an underfunded pension with \( F_0 = 0.9 \) and risk aversion \( R = 5 \), i.e. investing 20\% into the risky fund (dotted line), is expected to need more than five years to reach a debit balance. Thus, it takes a long time until a plan sponsor is safe even if he takes substantial risk. A conservative plan sponsor with risk aversion \( R = 15.3 \), or more, i.e. investing 6.54\% or less into the risky fund, cannot expect to reach a debit balance.

3.3.2 Overfunded Pension

Now we consider an overfunded pension, i.e. \( \bar{F}_0 > 1 \). The question we ask is: For the same representative pension with risk aversion \( R \) but with a current funding ratio \( \bar{F}_0 > 1 \), what yield guarantee would be affordable if the pension should never slip into a funding gap with a probability \( p \)?

From proposition 3.1 for \( F_0 > d \) we know that

\[
p_d = \left( \frac{d}{\bar{F}_0} \right)^{\frac{2\mu - \sigma^2}{\sigma^2}} \\
= \left( \frac{d}{\bar{F}_0} \right)^{\frac{2R - 1)(\bar{\pi} - \pi_L)' \Sigma (\bar{\pi} - \pi_L) + 2R^2(r_0 - \bar{\mu}_L - \varsigma) + \sigma^2 \bar{\mu}^2}{(\bar{\pi} - \pi_L)' \Sigma (\bar{\pi} - \pi_L) + \sigma^2 \bar{\mu}^2}.
\]
Figure 3.1: Upper boundary $\hat{r}_{0,T}$. Parameters: $F_0 = 0.9, \pi_L = 0, \bar{\mu}_L = \sigma_L = 0, r = 0.02, \zeta = 0.005, \mu = 0.1$ and $\sigma = 0.2$ for the risky fund in a two fund framework.
3.4 Conclusions

We set \( d = 1 \) and \( F_0 = \bar{F}_0 \) and solve for \( r_0 \)

\[
\begin{align*}
  r_0 & \leq \hat{r}_{0,p} \triangleq r + \frac{2R - 1}{2R^2} (\hat{\pi} - \pi_L)' \Sigma (\hat{\pi} - \pi_L) - \bar{\mu}_L - \varsigma + \frac{\sigma_L^2}{2} \\
  & \quad + \frac{\ln p_1}{2 \ln (\bar{F}_0) R^2} ((\hat{\pi} - \pi_L)' \Sigma (\hat{\pi} - \pi_L) + R^2 \sigma_L^2) \\
  & \quad - \bar{\mu}_L - \varsigma + \frac{\sigma_L^2}{2} \left( 1 + \frac{\ln p_1}{\ln (\bar{F}_0)} \right) 
\end{align*}
\]

In Figure 3.2, different \( \hat{r}_{0,p} \) subject to \( R \) are shown for \( \bar{F}_0 = 1.2 \) and the same parameters as in Figure 3.1. Again, we can see that the plan sponsor has to take substantial risk in order not to slip in a funding gap almost surely. However, if he takes too much risk, the probability of slipping into a funding gap will be considerable. If the government sets \( r_0 = 0.025 \), then an overfunded pension with \( \bar{F}_0 = 1.2 \) and risk aversion \( R = 5 \), i.e. investing 20% in the risky fund, will slip into a funding gap with a probability of 34%. A plan sponsor with about \( R = 13 \), i.e. 7.7% in the risky fund or \( R = 3 \), i.e. 33% in the risky fund, will slip into a funding gap with a probability of 50%.

It is obvious that \( \hat{r}_{0,p} \) increases if we consider the probability of hitting \( d \) within a finite time. This issue is treated in appendix 3.A.2.

3.4 Conclusions

In this chapter we assumed that \( r_0 > r_R (t) \) and that \( r_0 \) is given by the government. Thus, the plan sponsor is not able to hedge the yield \( r_0 \) and faces the risk of slipping into a funding gap. This circumstance caused us to take the funding ratio as state variable which leads to an interesting first result. If the relative risk aversion \( R \) is low enough, i.e. in the CRRA case \( R < 2 \), then the effect of the unhedgeable variance \( \sigma_L^2 \) on expected utility is positive. The reason for this result is that the diffusion parameter \( \sigma_L \) of the liability positively affects the drift of the ratio process. If the plan sponsor is not risk averse enough, this effect is preponderate.

The optimal portfolio of this asset-liability setting in the CRRA case without net contributions is constant and consists of a two-fund separation. The plan sponsor will invest in the liability hedge portfolio and the growth-optimal portfolio. The amount dedicated to each sub-portfolio depends on his relative risk aversion.
Figure 3.2: Upper boundary $\hat{r}_{0,p}$. Parameters: $F_0 = 1.2$, $\pi_L = 0$, $\beta_L = \sigma_L = 0$, $r = 0.02$, $\varsigma = 0.005$, $\mu = 0.1$ and $\sigma = 0.2$ for the risky fund in a two fund framework.
Finally, we have examined the impact of the yield guarantee on the funding ratio within the asset liability setting. We have seen that if $r_0$ is higher than the return of the riskless asset, then a plan sponsor of an underfunded pension has to take substantial risk in order to reach a debit balance. Moreover, we have to consider that it will take a long time until he reaches the debit balance. Also, a plan sponsor of an overfunded pension must take substantial risk in order not to slip into a funding gap with probability one, although taking too much risk will at the same time increase the probability of slipping into a funding gap. Generally we have seen, that, independent of the risk the plan sponsor takes, he faces a high probability of becoming underfunded at least once if the yield guarantee is higher than the return of the riskless asset.
3. Appendix

3.1 Proof of Proposition 3.1

Proof. We have

$$\frac{dF_t}{F_t} = \mu dt + \sigma dZ_t.$$ 

Now we use a 1-dimensional version of Dynkin’s formula\(^3\):

$$E[f(F_\tau)] = f(F_0) + E \left[ \int_0^\tau A f(F_t) \, ds \right] \quad (3.24)$$

with \(A f(F) \triangleq \mu F \frac{\partial f}{\partial F} + \frac{1}{2} \sigma^2 F^2 \frac{\partial^2 f}{\partial F^2}\), where it is assumed that \(f\) is twice differentiable and \(E[\tau] < \infty\). Further \(d < F_0\) is a lower bound and \(h > F_0\) an upper bound.

The probability of hitting \(h\) before \(d\),

$$P(\tau(F,d) < \tau(F,h)) = 1 - P(\tau(F,h) < \tau(F,d)),$$

can be calculated as follows:

Choose \(f(x) = \frac{\sigma^2}{\sigma^2 - 2\mu x} x^2 \frac{e^{-2\mu x}}{\sigma^2}\), then \(A f(x) = 0\) and thus

$$E[f(F_\tau)]
\begin{align*}
\quad & = P^x(\tau(F,d) < \tau(F,h)) f(d) + (1 - P(\tau(F,d) < \tau(F,h))) f(h) \\
\quad & = f(F_0)
\end{align*}$$

and, therefore,

$$P(\tau(F,d) < \tau(F,h)) = \frac{h^2 - 2\mu h}{h^2 - 2\mu d} - \frac{F_0^2 - 2\mu F_0}{F_0^2 - 2\mu d}.$$

The same procedure for \(P(\tau(F,h) < \tau(F,d))\) gives

$$P(\tau(F,h) < \tau(F,d)) = \frac{F_0^2 - 2\mu F_0}{h^2 - 2\mu d} - \frac{d^2 - 2\mu d}{h^2 - 2\mu d}.$$

It is \(\lim_{h \to \infty} P(\tau(F,d) < \tau(F,h)) = P(\tau(F,d) < \infty) = \left(\frac{d}{F_0}\right)^{2\mu - \frac{\sigma^2}{\mu}}\) if \(\mu > \frac{\sigma^2}{2}\) and if \(\mu - \frac{\sigma^2}{2} < 0\), then \(P(\tau(F,d) < \infty) = 1.\)

---

\(^3\)See Rogers and Williams (1994) or, for a more general version of this formula, see Dynkin (1965).
Using (3.24) once more with \( f(x) = \ln x \) gives \( Af(x) = \mu - \frac{1}{2}\sigma^2 \). Thus

\[
E \left[ \ln (F_\tau) \right] = \frac{\sigma^2 - d^2}{\sigma^2} \ln \frac{\sigma^2 - 2u}{\sigma^2 - d^2} \ln d + \frac{\sigma^2 - 2u}{\sigma^2 - d^2} \ln h \\
= \ln F_0 + E \left[ \int_0^\tau Af(F_s) \, ds \right] \\
= \ln F_0 + E \left[ \int_0^\tau \mu - \frac{1}{2}\sigma^2 \, ds \right] \\
= \ln F_0 + \left( \mu - \frac{1}{2}\sigma^2 \right) E(\tau_{(d,h)})
\]

where \( \tau_{(d,h)} \) denotes the stopping time when \( F_\tau \) hits either \( d \) or \( h \). Thus

\[
E(\tau(F,h)) = \lim_{d \to 0} E(\tau_{(d,h)}) = \ln \frac{\mu}{\mu - \frac{1}{2}\sigma^2}.
\]

\[ \Box \]

### 3.A.2 Distribution Function of \( \tau \)

In the finite time setting we consider the probability of hitting \( d \) before a time \( T \) can be calculated, as follows.

**Proposition 3.2** If \( dF_t = F_t(\mu dt + \sigma dZ_t) \) where \( \mu \) and \( \sigma \) are constant parameters and \( \mu > \frac{\sigma^2}{2} \), then the probability of hitting a constant \( d < F_0 \) before \( T \) is given by

\[
\Phi(T) \triangleq P(\tau(F,d) < T) \\
= \int_0^T (\ln F_0 - \ln d) \exp \left( -\frac{(\mu - \frac{\sigma^2}{2} t + \ln F_0 - \ln d)^2}{2t\sigma^2} \right) dt \\
\]

**Proof.** The SDE \( dF_t = F_t(\mu dt + \sigma dZ_t) \) can be rewritten as

\[
d \ln F_t = \tilde{\mu} dt + \sigma dZ_t
\]

where \( \tilde{\mu} \triangleq \mu - \frac{1}{2}\sigma^2 \)

We first calculate the Laplace transform using a martingale method, in order to get the
distribution function by calculating the inverse Laplace transform.

The Laplace transform is
\[ g(\lambda) \triangleq E(\exp(-\lambda \tau)) \]
where \( \tau \) is the short hand notation of \( \tau(F, h) \).

Suppose a stochastic process \( M_t \) with
\[ M_t = \exp \left( -\lambda t - \frac{\bar{\mu} + \sqrt{\bar{\mu}^2 + 2\sigma^2 \lambda}}{\sigma^2} (\ln F_t - \ln d) \right). \]

Note that \( M_{\tau} = \exp(-\lambda \tau) \). Using Itô’s Lemma we get
\[
\begin{align*}
\frac{dM_t}{M_t} &= -\lambda \exp \left( -\lambda t - \frac{\bar{\mu} + \sqrt{\bar{\mu}^2 + 2\sigma^2 \lambda}}{\sigma^2} (\ln F_t - \ln d) \right) dt \\
&\quad - \frac{\bar{\mu} + \sqrt{\bar{\mu}^2 + 2\sigma^2 \lambda}}{\sigma^2} \left( \frac{dX_t}{dF_t} \right) \\
&\quad + \frac{1}{2} \left( \frac{\bar{\mu} + \sqrt{\bar{\mu}^2 + 2\sigma^2 \lambda}}{\sigma^2} \right)^2 \left( \frac{dF_t}{F_t} \right)^2 \\
&= \exp \left( -\lambda t - \frac{\bar{\mu} + \sqrt{\bar{\mu}^2 + 2\sigma^2 \lambda}}{\sigma^2} (\ln F_t - \ln d) \right) \\
&\quad \cdot \left( -\frac{\bar{\mu} + \sqrt{\bar{\mu}^2 + 2\sigma^2 \lambda}}{\sigma^2} \left( \bar{\mu} dt + \sigma dZ_t \right) - \lambda dt \right) \\
&\quad + \exp \left( -\lambda t - \frac{\bar{\mu} + \sqrt{\bar{\mu}^2 + 2\sigma^2 \lambda}}{\sigma^2} (\ln F_t - \ln d) \right) \\
&\quad \cdot \left( \frac{\bar{\mu}^2 + \bar{\mu}\sqrt{\bar{\mu}^2 + 2\sigma^2 \lambda} + \sigma^2 \lambda}{\sigma^4} \right) \sigma^2 dt \\
&= -\exp \left( -\lambda t - \frac{\bar{\mu} + \sqrt{\bar{\mu}^2 + 2\sigma^2 \lambda}}{\sigma^2} (\ln F_t - \ln d) \right) \frac{\bar{\mu} + \sqrt{\bar{\mu}^2 + 2\sigma^2 \lambda}}{\sigma} dZ_t.
\end{align*}
\]
Thus \( \{M_{t \wedge \tau} \}_{t \geq 0} \) is a (bounded) martingale. It follows that \( E(M_\tau) = M_0 \) and as a consequence
\[ g(\lambda) = \exp \left( -\frac{\bar{\mu} + \sqrt{\bar{\mu}^2 + 2\sigma^2 \lambda}}{\sigma^2} (\ln F_0 - \ln d) \right). \]
To get the distribution of $\tau$ it suffices to find the inverse Laplace transform of $g(\lambda)$ i.e. $f(t)$ such that
\[
\int_0^\infty \exp(-\lambda t) f(t) \, dt = \exp \left( -\frac{\bar{\mu} + \sqrt{\bar{\mu}^2 + 2\sigma^2 \lambda}}{\sigma^2} (\ln F_0 - \ln d) \right).
\]
The distribution of $\tau$ is
\[
f(t) = \frac{(\ln F_0 - \ln d) \exp \left( -\frac{(\bar{\mu}t + \ln F_0 - \ln d)^2}{2\sigma^2 t^2} \right)}{\sqrt{2\pi\sigma^2 t^3}}.
\]
The probability of hitting $d$ before $T$ is
\[
P_{d,T} = \int_0^T \frac{(\ln F_0 - \ln d) \exp \left( -\frac{(\bar{\mu}t + \ln F_0 - \ln d)^2}{2\sigma^2 t^2} \right)}{\sqrt{2\pi\sigma^2 t^3}} \, dt
\]
\[= \Phi \left( -\frac{\ln \left( \frac{F_0}{d} \right) + \bar{\mu}t}{\sigma \sqrt{T}} \right) + \left( \frac{d}{F_0} \right)^{\frac{2\pi}{q}} \Phi \left( -\frac{\ln \left( \frac{F_0}{d} \right) - \bar{\mu}t}{\sigma \sqrt{T}} \right)
\]
where $\Phi(.)$ is the distribution function of the standard normal distribution.

Note that
\[
\lim_{T \to \infty} P(\tau < T) = \left( \frac{d}{F_0} \right)^{\frac{2\pi}{q}}
\]
which is of course equal to (3.22).

In Figure 3.3 $p_{d,T}$ is shown for different $r_0$. We can see that the lower the $r_0$ the faster converges $p_{d,T}$ to $p_{d,\infty} = p_d$ which is, of course, lower the lower the $r_0$.

In order to calculate the upper boundary $\hat{r}_{0,p}$ we can use proposition 3.2 and solve the following equation numerically for $\hat{r}_{0,p}$:

\[
\int_0^T \frac{\exp \left( -\left( \frac{\ln F_0 - \ln d}{\sigma^2} t + \ln F_0 - \ln d \right)^2 \right)}{\sqrt{2\pi \left( \frac{\ln F_0 - \ln d}{\sigma^2} \right)^2 t^3}} \, dt
\]
\[= \frac{p}{\ln F_0 - \ln d'}
\]
where $q \triangleq (\hat{\pi} - \pi_L)' \Sigma (\hat{\pi} - \pi_L)$. 
Figure 3.3: Distribution function of $\tau$. Parameters: $\bar{F}_0 = 1.2$, $\pi_L = 0$, $\mu_L = \sigma_L = 0$, $r = 0.02$, $\zeta = 0$, $\mu = 0.1$ and $\sigma = 0.2$ for the risky fund in a two fund framework.
Chapter 4

Minimum Yield Guarantee and Surplus

In chapter 3 the development of the retirement assets was predefined through an exogenously given fixed yield guarantee. In this chapter the yield guarantee becomes a lower boundary and there will be a surplus on the minimum yield which is linked to the funding ratio or to be optimized. Thus, we will endogenize the liability.

The first setting assumes a minimum yield that is lower than the riskless yield. We will state a lower and an upper boundary for the funding ratio. Since the minimum yield is lower than the riskless yield, a very conservative strategy, i.e. investing everything into the riskless asset, always increases the funding ratio and, thus, in a continuous-time framework avoids shortfall. The employees would always receive at least the minimum yield. If the funding ratio reaches the upper boundary, they additionally receive the full surplus such that the funding ratio does not exceed the upper boundary. Once the funding wealth decreases, the employees’ return falls to the minimum yield.

In the second setting the minimum yield is higher than the riskless yield. Thus, in this case as in the previous chapter, the state variable is the funding ratio. But, unlike the setting in chapter 3, the employees are assumed to influence the decisions of the plan sponsor. In a conjoint optimization problem the employees and the employers evaluate the optimal investment policy and the optimal surplus given a minimum return. In a first step we will calculate the optimal surplus without constraint, and in a second step we will include a constraint on the funding ratio.

Note that in both settings we still have an intra-generational and an inter-generational redistribution of risk. The intra-generational redistribution is based on the neutral treatment of the employees, and the inter-generational redistribution is based on the fact that the plan sponsor’s funding reserves increase in good times and decrease in bad times which affects the benefits of the employees.

1For comparison, we sometimes call the yield guarantee of chapter 3 “fixed” even though it is allowed to depend on time. In the context of the yield guarantee “fixed” means that the employee cannot expect an additional payout to the yield guarantee.
4.1 Low Minimum Yield Guarantee and Bounded Funding Ratio

In this setting we suppose a minimum yield that is lower than the riskless yield and assume that there is no further unhedgeable risk caused by the liability or the net contributions. Investing everything in the riskless asset would always make the funding ratio increase and, thus, the plan sponsor can avoid a shortfall. Such a plan can be profitable and is, of course, not as risky as if the minimum guarantee exceeds the riskless yield. This idea is supported by the EU-directive for life insurance (92/96/EWG) which states that the minimum yield is allowed to be at most 60% of the yield of government bonds.

Since the plan sponsor can avoid a shortfall, there is no incentive to optimize the funding ratio. We assume, that the plan sponsor benefits from the growth of wealth, e.g. by linking his wage to the funding wealth. In this way he optimizes the funding wealth and so far we have a retirement plan that is offered by most life insurance companies to individual clients.

However, in our approach, we assume further that there is an exogenously given minimum yield. Finally we add a boundary for the funding ratio as in the UK. Under UK pension law it is permissible for the funding ratio to vary between 90% and 105%. If the funding ratio exceeds the 105% limit the plan sponsor has at most 5 years to reduce the ratio to 100%. But if the funding ratio slips below 90% the plan sponsor has one year to raise the level to 90% and another 5 years to raise it to 100%. Of course such a lower boundary is only possible if the minimum yield does not exceed the riskless yield.

4.1.1 Model Setting

We assume that $F_0 = F_d > 1$. Thus, we assume the plan sponsor has reached the upper boundary before he implements this setting. This assumption has to be taken for technical reasons. Henceforth, we assume further that $dC_t = dl_t = 0$, and $r_0 \leq r - \varsigma$.

The wealth process evolves as

$$dW^\pi_t = W^\pi_t \left( (\pi'_t (\mu_t - r_t 1) + r_t) \, dt + \pi'_t \sigma_t dZ_t \right) - \varsigma_t W^\pi_t \, dt.$$

The plan sponsor optimizes the long-term growth rate of the finding wealth but has to consider a lower bound for the funding ratio. Thus his optimization problem is

$$\sup_{\pi \in \Pi} \inf_{T \to \infty} \frac{1}{\beta T} \ln \beta E \left[ U \left( W^\pi_T \right) \right]$$

subject to $F^\pi_t \geq F_d$. 

4.1 Low Minimum Yield Guarantee and Bounded Funding Ratio

where \( F_d \in [0, F_u) \). The liability is determined by

\[
L_t \triangleq \max \left[ L_0, \frac{W_\pi}{F_u} r_0(t-s) ; s \leq t \right].
\]

Thus, the liability process evolves at least as \( dL_t = L_t r_0 dt \). Moreover, the liability causes an upper boundary for the funding ratio. To show this, we examine the mechanism that is activated by the liability defined as above: The plan sponsor is assumed to start with \( F_0 = F_u \) when the framework is implemented. If the financial market prospers, then the funding wealth maximum increases as well as (the sum of) the retirement assets of the employees. If the return of the portfolio or the growth of the wealth decreases and falls below \( r_0, \) then the retirement assets grow with \( r_0 \), respectively. As the funding wealth decreases, the reserves are reduced. However, the plan sponsor will adjust the investment policy such that the funding ratio does not slip below \( F_d \). If the financial market prospers again, the plan sponsor still provides \( r_0 \) until the funding ratio has reached \( F_u \) again. Thus, through the constraint and the liability mechanism the plan sponsor is obliged to keep the funding ratio within the interval \([F_d, F_u]\).

Figure 4.1 depicts the mechanism is shown for a trajectory of the extremes, \( r_0 = 0 \) (gray) and \( r_0 = r \) (black). On the left side we have the development of the pension wealth \( W_\pi \), starting with 125 and the liability \( L_t \) starting with 100. Note that if we drop the individual contributions \( c_i \), which is necessary to analyze the effect of this setting, and assume that \( dC_t = dl_t = 0 \), then the development of the individual employee’s accrued retirement assets \( X_t \) will be equal to the development of the liability \( L_t \) as specified in section 2.1.1. Thus, in Figure 4.1 we can see the development of the individual retirement assets \( X_t \) starting with \( X_0 = 100 \) as well as the liability \( L_t \) starting with \( L_0 = 100 \). On the right side we have the corresponding funding ratio which is bounded by \( F_u = 1.25 \) and \( F_d = 1 \). In the trajectory \( r_0 = 0 \) we can clearly see that the plan sponsor is lowering the variance of the portfolio as the funding ratio approaches \( F_d = 1 \).

We assume that \( U(x) = \frac{1}{1-R} x^{1-R} \), where \( R > 1 \) and choose \( \beta = 1 - R \). Further, all parameters \( \mu, \sigma, r, \varsigma, r_0 \) are assumed to be constant.

Technically, this problem is exactly the same as the drawdown constraint problem in Grossman and Zhou (1993) who solve

\[
\sup_{\pi \in \Pi} \lim_{T \to \infty} \inf_{\tau \rightarrow -\infty} \ln \left( \frac{1}{1-R} E \left[ U(W_\pi^T) \right] \right) \quad \text{s.t.} \quad W_\pi^t \geq \eta \max \left[ W_0^\tau, (W_\pi^s)^{r_0(t-s)} ; s \leq t \right], \eta \in [0, 1).
\]
The same drawdown constraint problem is solved by Cvitanić and Karatzas (1995) who use a martingale approach for complete markets, i.e. \( r_0 = r \), and assume \( 0 < R < 1 \), but with the parameters allowed to depend on time.

### 4.1.2 Solution

Here we give only a short outline of the method used in Grossman and Zhou (1993). For a detailed argumentation the reader is referred to the original paper.

The \( J \)-function (2.12) can be changed to

\[
J (\bar{W}, F_u \bar{L}) = \sup_{\pi \in G} \mathbb{E} \left[ e^{-(1-R)dt} J (\bar{W}_{\pi t}, F_u \bar{L}_t) \right]
\]

where \( \bar{W}_t \triangleq W_t e^{-r_0 t}, \bar{L}_t \triangleq L_t e^{-r_0 t} \) and, therefore, \( \bar{\nu} \triangleq \nu - r_0 \). The new wealth process \( \bar{W}_t \) evolves as

\[
\frac{d\bar{W}_{\pi t}}{\bar{W}_{\pi t}} = (\pi'_t (\mu - r 1) + r - r_0 - \zeta) dt + \pi'_t \sigma dZ_t.
\]

Now the argument in Grossman and Zhou (1993) can be drafted as follows: If \( \bar{W}_t = \bar{L}_t \), then the pension needs to put all the money into the riskless asset in order to assure that
$W_{t+h} = L_{t+h}$ for a small $h > 0$. If $W_t = F_u L_t$, then obviously $F_u L_t$ is redundant for the optimization process. But if $W_t \in (F_d L_t, F_u L_t)$, then $dL_t = 0$. Therefore, if all parameters are assumed to be constant we can derive the HJB-equation as

$$0 = - (1 - R) \partial f (\tilde{W}, F_u \tilde{L}) + \left( \pi^\prime (\mu - r \mathbf{1}) + r - \zeta - r_0 \right) \tilde{W} \partial f (\tilde{W}, F_u \tilde{L})$$

$$+ \frac{1}{2} \pi^\prime \Sigma \pi \tilde{W}^2 \partial^2 f (\tilde{W}, F_u \tilde{L}) \quad (4.1)$$

The first order conditions yield

$$\pi^*_t = - \frac{f (\tilde{W})}{W \tilde{f}} \hat{\pi}. \quad (4.2)$$

Substituting (4.2) into (4.1) gives

$$0 = - (1 - R) \partial f (\tilde{W}, F_u \tilde{L}) + \left( r - \zeta - r_0 \right) \tilde{W} \frac{\partial f (\tilde{W}, F_u \tilde{L})}{\partial \tilde{W}}$$

$$\frac{1}{2} \hat{\pi}^\prime \Sigma \hat{\pi} \left( \frac{f^\prime (w)}{W f'' (w)} \right) \hat{\pi} \quad (4.1)$$

Next, we need the “homogeneity-proposition” 1.1 of Grossman and Zhou (1993):

**Proposition 4.1.** If $\tilde{W}$ exists such that $f (\tilde{W}, F_u \tilde{L})$ is finite for $\tilde{W} \geq F_d \tilde{L}$, then for $\tilde{W} \geq F_d \tilde{L}$, $f (\tilde{W}, F_u \tilde{L})$ is homogenous of degree $1 - R$ in $\tilde{W}$ and $F_u \tilde{L}$. That is, for all positive $k$, $f (k \tilde{W}, kF_u \tilde{L}) = k^{1 - R} f (\tilde{W}, F_u \tilde{L})$.

**Proof.** see Grossman and Zhou (1993)
Again, a substitution \( \phi(w) \triangleq \frac{f'(w)}{f''(w)} \) leads to

\[
0 = - (1 - R) \bar{v} f(w) + (r - \zeta - r_0) w f'(w) - \frac{1}{2} \hat{\pi}' \Sigma \hat{\pi} \phi(w)
\]

and after differentiation

\[
0 = \left( - (1 - R) \bar{v} + r - \zeta - r_0 - \frac{1}{2} \hat{\pi}' \Sigma \hat{\pi} \right) \phi(w) - \frac{1}{2} \hat{\pi}' \Sigma \hat{\pi} \phi'(w) \phi(w) + (r - \zeta - r_0) w. \tag{4.4}
\]

Note that \( \pi^*_t = \frac{\phi(w)}{w} \hat{\pi} \). Thus, once we have calculated \( \phi(w) \) we are through. The boundary condition to this first order differential equation (4.4) is \( \phi(F_d F_u) = 0 \) because at \( w = F_d F_u \) the weight of the risky assets should be zero. If we solve (4.4) for \( \phi(w) \), we get

\[
|\phi(w) - w g^-|^{1-h} |\phi(w) - w g^+|^{1+h} = \left( - \frac{F_d g^-}{F_u} \right)^{1-h} \left( \frac{F_d g^+}{F_u} \right)^{1+h}
\]

where

\[
h \triangleq - \frac{(1 - R) \bar{v} + \frac{1}{2} \hat{\pi}' \Sigma \hat{\pi} - r + \zeta + r_0}{\sqrt{((1 - R) \bar{v} + \frac{1}{2} \hat{\pi}' \Sigma \hat{\pi} - r + \zeta + r_0)^2 + 2 (r - \zeta - r_0) \hat{\pi}' \Sigma \hat{\pi}}} \]

and \( g^-, g^+ \) are the negative and positive solution to

\[
0 = \left( - (1 - R) \bar{v} + r - \zeta - r_0 - \frac{1}{2} \hat{\pi}' \Sigma \hat{\pi} \right) g - \frac{1}{2} \hat{\pi}' \Sigma \hat{\pi} g^2 + r - \zeta - r_0.
\]

Further, to calculate \( \bar{v} \) we use the heuristic argument that at \( w = 1 \) it is \( \frac{\partial J(W, F_u \bar{L})}{\partial (F_u \bar{L})} = 0 \). So \( f'(1) = (1 - R) f(1) \). Thus, this argument together with (4.3) gives

\[
\phi(1) = \frac{r - \zeta - r_0 - \bar{v}}{\hat{\pi}' \Sigma \hat{\pi}} = \frac{r - \zeta - \bar{v}}{\hat{\pi}' \Sigma \hat{\pi}}.
\]

Substituting into (4.5) gives

\[
\left| \frac{r - \zeta - r_0 - \bar{v}}{\hat{\pi}' \Sigma \hat{\pi}} - g^- \right|^{1-h} \left| \frac{r - \zeta - r_0 - \bar{v}}{\hat{\pi}' \Sigma \hat{\pi}} - g^+ \right|^{1+h} = \left( - \frac{F_d g^-}{F_u} \right)^{1-h} \left( \frac{F_d g^+}{F_u} \right)^{1+h}
\].
This is an algebraic equation for $\bar{\nu}$ which has to be solved numerically. Once we have calculated $\bar{\nu}$, we can evaluate $\phi(w)$ through (4.5) where we set $\tilde{\phi}(w) = \frac{\phi(w)}{w}$ and numerically solve for $\tilde{\phi}(w)$ in

$$w^2 |\tilde{\phi}(w) - g_-|^{1-h} |\tilde{\phi}(w) - g_+|^{1+h} = \left( \frac{g_-}{F_u} \right)^{1-h} \left( \frac{g_+}{F_u} \right)^{1+h}.$$

In the case where $r_0 = r - \zeta$ the term $h$ becomes one and, thus, we can achieve a closed form as was shown in Grossman and Zhou (1993). The optimal portfolio in this case becomes

$$\pi^*_t = \frac{1}{(1 - \frac{F_d}{F_u}) R + \frac{F_d}{F_u} \left( 1 - \frac{F_d}{F_u} \right)} \hat{\pi}.$$

Figure 4.2 on the left side shows a simulation of the development of the accrued retirement assets and of the funding ratio for $r_0 = r$ and $F_d = 1$. We can see that, starting at $F_u = 1.25$, 50% of the funding ratio trajectories will be above 1.2 in ten years. On the right side we have the plots of the case with $r_0 = 0$. Here 50% of the funding ratios decreased below 1.07. Since the minimum yield is very low, this result is not surprising. The plan sponsor does not need the same amount of reserves as in the other case. These reserves went to the employees. Thus, there is a substantial difference between the development of the two retirement assets. In the case with $r_0 = r$ we have a low standard deviation for the retirement assets since the plan sponsor has to invest rather conservatively to ensure the high minimum yield. On the right side we see that the development of the retirement assets is expected to increase about 50% faster. However, the 5%-quantile with $r_0 = 0$ is below the 5%-quantile with $r_0 = r$ on the left side. Therefore, very risk averse employees will still prefer the case with a high minimum yield. Note that in this example we assumed a rather moderate performance for the growth optimal portfolio with $\mu = 0.07$. In case that $\mu$ is higher, the arguments for a low minimum yield become even stronger.

Taking everything into account we can see that the less risk averse employees prefer a low minimum yield and the very risk averse employees prefer a high minimum yield, or put differently, the optimal minimum yield depends on the risk attitude of the employees.

Note that the quantiles of the funding ratio converge to a value within $(F_d, F_u)$. Since the most conservative investment strategy, i.e. investing all into the riskless asset, increases faster than $r_0 \leq r$, after bad times there will always be a rebound of the funding ratio.
Figure 4.2: Development of $X_t$ and $F_t$. Parameters: $R = 5, r = 0.02, \zeta = 0, \mu = 0.07$ and $\sigma = 0.2$ for the risky fund in a two fund framework, $F_u = 1.25$ and $F_d = 1$. Monte Carlo-simulation, 50'000 trajectories for ten years with daily data.
Thus, the quantiles do not converge to $F_d$. We simulate the expected values of $F$ in 100 years and got 1.09 for the case $r_0 = 0$ and 1.18 for the case $r_0 = r$.

Finally, a lower minimum yield is always preferred by the plan sponsor. If the minimum yield is low he would obviously be allowed to invest as conservatively as if it was high. But the opposite is not true because of the constraint. A high minimum yield forces the plan sponsor to be more careful than he would be with a lower minimum yield.

4.2 High Minimum Yield Guarantee and Surplus

In this section we provide a surplus model with a minimum yield that is assumed to be higher than the riskless yield. Thus, the plan sponsor bears a non-hedgeable shortfall risk and, therefore, is assumed to optimize the funding ratio as in chapter 3. However, since the plan sponsor optimizes the funding ratio, he won’t provide a surplus voluntarily. We have therefore to modify the setting of chapter 3.

There are the following three possibilities to include a surplus within a minimum return setting of which we will examine the last one in this section.

Firstly, as we have done in the previous section, it is possible to give an upper boundary for the funding ratio, i.e. give an upper limit for the funding reserves. But if the minimum return is higher than the riskless yield, an upper boundary for the reserves would be problematic since the plan sponsor is not able to protect himself against a shortfall risk. He will rely on the reserves.

Secondly, it is possible to stipulate a surplus exogenously. However, the shortfall risk for the plan sponsor is already substantial if the minimum return exceeds the riskless yield, as we have seen in chapter 3. Since a surplus would even increase this risk, the minimum yield has to be smaller than in the yield guarantee we have examined in chapter 3. But then the questions arise: How high should it be set? How to stipulate the surplus? Within a non-utility based simulation framework these questions are discussed in Baumann, Delbaen, Embrechts, and Müller (2001).

The third possibility, which we examine in this section, is a setting where the employees are involved in the investment decision process. In a conjoint optimization problem, the employees and the employers evaluate the optimal investment policy and the optimal surplus, given a minimum return. The evaluated optimal surplus mechanism is mandatory for the plan sponsor. A premise for such a setting is that the employees have the
right to join the decision committee, as e.g. in Switzerland.

### 4.2.1 General Model

We assume that the employees and the plan sponsor optimize the funding ratio \( F_t^{\pi,\nu} \) as well as the surplus \( \nu_t \) at each point in time \( t \). Thus, we have a joint utility function \( U(F_t^{\pi,\nu}, \nu_t, t) \) the expected value of which is assumed to be maximized at each point in time \( t \). Additionally \( U(F_t^{\pi,\nu}, \nu_t, t) \) is assumed such that we have an autonomous problem, i.e. \( U(F_t^{\pi,\nu}, \nu_t, t) = e^{-\theta t} u(F_t^{\pi,\nu}, \nu_t) \). The problem becomes

\[
\max_{(\pi_t^\nu, \nu_t) \in \mathcal{H}} E \left[ \int_0^\infty e^{-\theta t} u(F_t^{\pi,\nu}, \nu_t) \, dt \right]
\]

s.t. \( \nu_t \geq 0 \)

similar to the case in section 2.2.3.

Henceforth, we drop the \( t \)-subscripts of all parameters in order to simplify the notation.

The funding wealth is equal to the one in the previous sections, i.e.

\[
dW_t^\pi = W_t^\pi \left( \pi_t (\mu - r I) + r \right) \, dt + W_t^\pi \pi_t \sigma \, dZ_t - W_t^\pi \xi dt + dC_t
\]

The liability contains the surplus \( \nu_t \) and is assumed to evolve as

\[
dl_t^\nu = L_t^\nu \left( r_0 + \nu_t \right) \, dt + dC_t + dl_t
\]

where

\[
\frac{dl_t}{L_t^\nu} = \pi_t^L ( (\mu - r I) \, dt + \sigma dZ_t ) + \mu_L \, dt + \sigma_L \, dZ_t^\lambda
\]

and

\[
dC_t = W_t \left( \pi_C^t \left( (\mu - r I) \, dt + \sigma dZ_t \right) + \mu_C \, dt \right).
\]

Thus, the funding ratio evolves as

\[
\frac{dF_t^{\pi,\nu}}{F_t^{\pi,\nu}} = \left[ (\pi_t^t + (1 - F_t^{\pi,\nu}) \pi_C^t - \pi_t^L) \Sigma \hat{n} + (1 - F_t^{\pi,\nu}) \pi_C^t \Sigma \pi_C - \pi_t^t \Sigma \left( \pi_C + \pi_L \right) - (1 - 2F_t^{\pi,\nu}) \pi_t^L \Sigma \pi_C + \pi_t^L \Sigma \pi_L - F_t^{\pi,\nu} (1 - F_t^{\pi,\nu}) \sigma_C^2 + \sigma_L^2 + (1 - F_t^{\pi,\nu}) \sigma_C \pi_C - \mu_L - r_0 + r - \xi - \nu_t \right] \, dt
\]

\[\quad + \left( \pi_t^t + (1 - F_t^{\pi,\nu}) \pi_C^t - \pi_t^L \right) \sigma \, dZ_t
\]

\[\quad + (1 - F_t^{\pi,\nu}) \sigma_C \, dZ_t^\gamma - \sigma_L \, dZ_t^\lambda.\]

\footnote{The parity of votes is settled in BVG Art. 51}
Using the HJB equation (2.10) we get

\[
0 = \sup_{(\pi_t, \nu_t) \in \mathcal{H}} \left\{ -qF + u(F, \nu) \right. \\
+ \frac{F J_F}{2} \left[ \begin{array}{l} 
\dot{\pi}_t' \Sigma \pi_t \\
- \pi_t' \Sigma (\pi_C + \pi_L) \\
-(1 - 2F) \pi_t' \Sigma \pi_C + \pi_t' \Sigma \pi_L \\
\end{array} \right] \\
\left. + (1 - F) \bar{\mu}_C - \bar{\mu}_L \\
-r_0 + r - \zeta - \nu \\
- F (1 - F) \sigma_C^2 + \sigma_L^2 \\
\right. \\
+ \frac{F^2 J_{FF}}{2} \left[ \begin{array}{l} 
\dot{\pi}_t' \Sigma \dot{\pi}_t \\
+ \sigma_C^2 (1 - F)^2 \sigma_C^2 \\
\end{array} \right] \right\}
\] (4.7)

where \( \bar{\pi}_t = (\pi_t + (1 - F_t) \pi_C - \pi_L) \) as in section 3.2.1.

The first order condition for the optimal portfolio strategy leads to the same optimal portfolio as in (3.13), i.e.

\[
\pi_t^* = - \frac{J_F}{F_J FF} \left( \bar{\pi} - F \pi_C - \pi_L \right) + (F - 1) \pi_C + \pi_L. 
\] (4.8)

The first order condition for the optimal surplus is

\[
h(v_t^*) \triangleq \frac{\partial u(F, v_t^*)}{\partial v_t} = J_FF 
\]

and, thus, the optimal surplus

\[
v_t^* = h^{-1}(J_FF) 
\] (4.9)

Substituting (4.8) and (4.9) into the HJB equation (4.7) we finally get

\[
0 = u \left( F, h^{-1}(J_FF) \right) - qF \\
+ \frac{F J_F}{2} \left[ \begin{array}{l} 
(1 - F) \bar{\mu}_C + \sigma_C^2 + F (1 - F) \sigma_C^2 \\
-r_0 + r - \zeta - h^{-1}(J_FF) \\
\end{array} \right] \\
+ \frac{1}{2} \frac{F^2 J_{FF}}{2} \left[ (1 - F)^2 \sigma_C^2 + \sigma_L^2 \right] \\
- \frac{F^2 J_{FF}}{2} \left( \bar{\pi} - F \pi_C - \pi_L \right) \Sigma \left( \bar{\pi} - F \pi_C - \pi_L \right) 
\]

for the HJB-equation of the general setting. As usual, it is not possible to solve this equation in general. In the next section we will provide a closed form solution for a special case.
4.2.2 Special Case: CIES-Utility

In this section we assume that $dC_t = 0$ and the joint utility function $U (F_t^{\pi, \nu}, \nu_t, t)$ has a constant intertemporal elasticity of substitution (CIES), i.e.

$$U (F_t^{\pi, \nu}, \nu_t, t) \triangleq e^{-\varrho t} \frac{(F_t^{\theta} v_1^{1-\theta})^{(1-\delta)}}{1-\delta}$$

where $\delta > 1$ and $\theta \in (0, 1)$ if

$$\varrho > \hat{\varrho} \triangleq \frac{(\delta - 1) \theta}{2 + 2(\delta - 1) \theta} \begin{bmatrix} (\hat{\pi} - \pi_L)' \Sigma (\hat{\pi} - \pi_L) \\ -2 (1 + (\delta - 1) \theta) \\ r - \bar{\mu}_L - \zeta - r_0 \\ + \frac{1}{2} (\theta (1 - \delta) + 1) \sigma_L^2 \end{bmatrix}.$$ 

Otherwise $\theta = 1$. As we will see afterwards, the differentiation of $\varrho$ ensures a surplus $\nu_t$ that is not negative.

The first order condition for $\nu_t^*$ becomes

$$(1 - \theta) F^{\theta(1-\delta)} (\nu^*)^{\theta(\delta-1)-\delta} = J_F F$$

and, therefore,

$$\nu_t^* = \left( \left(1 - \theta \right) \frac{F^{\theta(1-\delta)-1}}{J_F} \right)^{\frac{1}{\theta - (1-\delta)}}.$$ 

Substituting $\nu_t^*$ and $U (F_t^{\pi, \nu}, \nu_t, t)$ into the HJB equation (4.7) we get

$$0 = -\varrho J + \frac{1}{1-\delta} F^{\theta(1-\delta)} \left(1 - \theta \right) \frac{F^{\theta(1-\delta)-1}}{J_F} \left( \left(1 - \theta \right) \frac{F^{\theta(1-\delta)-1}}{J_F} \right)^{\frac{1}{\theta - (1-\delta)}}$$

$$- J_F F \left(1 - \theta \right) \frac{F^{\theta(1-\delta)-1}}{J_F} \left( \left(1 - \theta \right) \frac{F^{\theta(1-\delta)-1}}{J_F} \right)^{\frac{1}{\theta - (1-\delta)}}$$

$$+ J_F F \left( (r - \bar{\mu}_L - \zeta - r_0 + \sigma_L^2) \right)$$

$$- \frac{J_F^2}{2J_{FF}} (\hat{\pi} - \pi_L)' \Sigma (\hat{\pi} - \pi_L) + \frac{1}{2} J_{FF}^2 \sigma_L^2.$$ \hspace{1cm} (4.10)

We try

$$J (F) = \frac{c F^{\theta(1-\delta)}}{\theta (1 - \delta)}$$
where \( c > 0 \) is a parameter that has to be evaluated. The derivatives become

\[
J_F = c F^{\theta (1 - \delta) - 1}
\]

\[
J_{FF} = (\theta (1 - \delta) - 1) c F^{\theta (1 - \delta) - 2}
\]

and, therefore, the controls are

\[
v^* = \left( \frac{1 - \theta}{c} \right)^{\frac{1}{\theta (1 - \delta) - 1}},
\]

\[
\pi^* = \frac{1}{\theta (\delta - 1) + 1} (\hat{\pi} - \pi_L) + \pi_L
\]

which we substitute into (4.10) and get

\[
0 = -\theta c F^{\theta (1 - \delta)} \frac{1}{\theta (1 - \delta)} + \frac{1}{1 - \delta} F^{\theta (1 - \delta)} \left( 1 - \theta \right) \frac{c F^{\theta (1 - \delta) - 1}}{c F^{\theta (1 - \delta) - 1}} \frac{1}{\theta (\delta - 1) - 1}
\]

\[
-c F^{\theta (1 - \delta) - 1} F \left( 1 - \theta \right) \frac{c F^{\theta (1 - \delta) - 1}}{c F^{\theta (1 - \delta) - 1}} \frac{1}{\theta (\delta - 1) - 1}
\]

\[
+ c F^{\theta (1 - \delta)} (r - \bar{\pi}_L - \zeta - r_0 + \sigma^2_L)
\]

\[
- \frac{c F^{\theta (1 - \delta)}}{2 (\theta (1 - \delta) - 1)} (\hat{\pi} - \pi_L)' \Sigma (\hat{\pi} - \pi_L)
\]

\[
+ \frac{1}{2} (\theta (1 - \delta) - 1) c F^{\theta (1 - \delta)} \sigma^2_L
\]

Dividing by \( c F^{\theta (1 - \delta)} \) and rearranging we get

\[
0 = -\frac{\theta}{\theta (1 - \delta)} + \left( \frac{\delta + \theta (1 - \delta)}{(1 - \delta)} \left( \frac{1 - \theta}{c} \right)^{\frac{1}{\theta (1 - \delta) - 1}} \right)
\]

\[
+ r - \bar{\pi}_L - \zeta - r_0 - \frac{1}{2 (\theta (1 - \delta) - 1)} (\hat{\pi} - \pi_L)' \Sigma (\hat{\pi} - \pi_L) + \frac{1}{2} (\theta (1 - \delta) + 1) \sigma^2_L.
\]
Solving for \( c \) we get

\[
c = \frac{1}{1 - \theta} \left[ \frac{(1-\theta)\varrho}{\theta(\delta+\theta(1-\delta))} \left[ r - \bar{\mu}_L - \zeta - r_0 \right] - \frac{(1-\delta)(1-\theta)}{\delta + \theta(1-\delta)} \left( r - \bar{\mu}_L - \zeta - r_0 + \frac{1}{2} (\theta (1 - \delta) + 1) \sigma^2_L \right) \right]^{\delta + \theta(\delta - 1)}.
\]

Substituting \( c \) into (4.11) we get the solution of \( \nu^* \), i.e.

\[
\nu^* = \frac{(1 - \theta) \varrho}{\theta (\delta + \theta (1 - \delta))} \left( r - \bar{\mu}_L - \zeta - r_0 + \frac{1}{2} (\theta (1 - \delta) + 1) \sigma^2_L \right) \left[ \delta + \theta (1 - \delta) \right]^{\delta + \theta(\delta - 1)} (\hat{\pi} - \pi_L)^\prime \Sigma (\hat{\pi} - \pi_L)
\]

which is only positive if \( \varrho > \hat{\varrho} \). Otherwise, if \( \varrho \leq \hat{\varrho} \), then we have \( \nu_1 = 0 \) and \( \pi^* = \frac{1}{\hat{\varrho}} (\hat{\pi} - \pi_L) + \pi_L \) which is the result in section 3.2.2, i.e. (3.18). Thus, if either the time-preference parameter is not high enough or the minimum yield too high, then ceteris paribus the minimum yield becomes a yield guarantee as in chapter 3 because the surplus decreases to zero.

Note that in this setting the optimal surplus becomes a constant. Since \( \frac{\partial \nu^*}{\partial r_0} = \frac{(1-\delta)(1-\theta)}{\delta + \theta(1-\delta)} \right) < 0 \), the surplus is smaller for higher minimum yield. The individual employee’s accrued retirement assets evolves as

\[
\frac{dX_t}{X_t} = \left( r_0 + \nu^* (r_0) \right) dt + dc_t.
\]

Since \( \frac{\partial \nu^*}{\partial r_0} > -1 \), a higher minimum return is preferred by the employees within this framework. Additionally, because of \( -1 < \frac{\partial \nu^*}{\partial r_0} < 0 \), the shortfall probability for the plan sponsor increases as we can see in (4.6). We have \( r_{0,1} \gtrless r_{0,2} \Leftrightarrow X_{t,1} \gtrless X_{t,2} \Leftrightarrow F_{t,1} \gtrless F_{t,2} \), almost everywhere.

Since the surplus is constant, this setting is never a Pareto improvement to the yield guarantee setting in chapter 3. To show this we denote the yield guarantee of the setting in chapter 3 as \( \bar{r}_0 \) and the minimum yield of the current setting as \( r_0 \). Of course, if \( r_0 > \bar{r}_0 \) then the plan sponsor’s situation is worse than in the yield guarantee setting. If \( r_0 = \bar{r}_0 \), then in the case that \( \nu^* > 0 \) the plan sponsor’s situation gets worse again and if \( \nu^* = 0 \) nothing changes. In the interesting cases if \( r_0 < \bar{r}_0 \) the plan sponsor prefers the new
4.2 High Minimum Yield Guarantee and Surplus

setting if $\nu^* + r_0 < \bar{r}_0$. But obviously, in this case the employee prefers the yield guarantee setting. Conversely, if $\nu^* + r_0 > \bar{r}_0$ and, finally, both are unchanged if $\nu^* + r_0 = \bar{r}_0$. The reason for this result is that in this constant setting the plan sponsor still bears the whole risk. In the next section we consider this problem and include a constraint for the funding ratio.

4.2.3 Special Case: CIES-Utility and Constraint on Funding Ratio

Within this section we discuss the case where the funding ratio is exogenously bounded from below. Again, we assume that $dC = 0$. The liability is stated as follows

$$dL_t = L_t \left( (\mu - r) dt + \sigma dZ_t \right) + \bar{\mu}_L dt + r_0 dt - \nu_t dt .$$

Therefore, the liability is assumed to be trackable.

The funding ratio becomes

$$\frac{dF^\pi_t}{F^\pi_t} = \left[ \begin{array}{c} (\pi_t' - \pi_L') \Sigma \pi_t - \pi_L' \Sigma \pi_L + \pi_L' \Sigma \pi_L \\ + r - \bar{\mu}_L - r_0 - \varsigma - \nu_t \\ + (\pi_t' - \pi_L') \sigma dZ_t \end{array} \right] dt$$

We claim that the plan sponsor ensures a lower bound of the funding ratio. The constraint we assume is

$$F^\pi_t > \eta \Psi_t ,$$

where $\eta \in [0, 1)$ and $\Psi_t \triangleq F_{0e}(r - \bar{\mu}_L - r_0 - \varsigma)$. Note that the markets are incomplete. If $r < \bar{\mu}_L + r_0 + \varsigma$ the plan sponsor is forced to hold risky assets in order to reach a return that is equal to $\bar{\mu}_L + r_0 + \varsigma$. Thus, since he has to bear a downside risk, he is not able to ensure $F^\pi_t > a \in [0, F_0)$. With a probability one the constraint $F^\pi_t > a$ would be violated at least once in the future\(^3\). Moreover, a constraint containing only a (small) probability for violating $F^\pi_t > a$, i.e. $P(F^\pi_t \leq a) < b \in (0, 1)$, instead of $F^\pi_t > a$ would also be violated. Thus, it is not possible to demand a fixed lower bound. The best dynamics of a boundary that we can ensure is $F^\pi_t > \eta \Psi_t$, since the remaining risk can be hedged.

\(^3\)This follows directly from the fact that $P \left( \sup \limits_i Z_t = +\infty \wedge \inf \limits_i Z_t = -\infty \right) = 1$, which has been proven e.g. in Rogers and Williams (1994) lemma 3.6.
Now we create a new process $F_{t}^{\pi,v} = e^{F_{t}^{\pi,v}/\eta}$ for which we get
\[
\frac{dF_{t}^{\pi,v}}{F_{t}^{\pi,v}} = \left( (\pi'_{t} - \pi'_{L}) \Sigma \tilde{\pi} - \pi'_{L} \Sigma \pi_{L} + \pi'_{L} \Sigma \pi_{L} - v_{t} \right) dt \\
+ (\pi'_{t} - \pi'_{L}) \sigma dz_{t},
\]
\[F_{0} = 1.\]

As in the previous section, we consider the CIES utility function
\[
U(F_{t}^{\pi,v},v_{t},t) \triangleq e^{-\varrho t \left( F_{t}^{\pi,v,1-\theta} \right)^{1-\delta}}
\]
where again $\delta > 1$ and $\theta \in (0,1)$ if
\[
\varrho > \bar{\varrho} \triangleq \frac{(\delta - 1) \theta}{2 + 2(\delta - 1) \theta} \begin{bmatrix} (\tilde{\pi} - \pi_{L})' \Sigma (\tilde{\pi} - \pi_{L}) \\ -2(1 + (\delta - 1) \theta) \\ \cdot (r - \bar{\mu}_{L} - \zeta - r_{0}) \end{bmatrix}.
\]
otherwise $\theta = 1$.

Note that
\[
U(F_{t}^{\pi,v},v_{t},t) = e^{-\varrho t \left( F_{t}^{\pi,v,1-\theta} \right)^{1-\delta}} \\
= e^{-\varrho t \Psi_{t}^{\theta(1-\delta)}} \left( F_{t}^{\pi,v,1-\theta} \right)^{1-\delta} \left( 1 - \delta \right) \\
= F_{0}^{\theta(1-\delta)} e^{-\varrho t \left( F_{t}^{\pi,v,1-\theta} \right)^{1-\delta} / 1 - \delta}
\]
where $\bar{\varrho} = \varrho + \theta (\delta - 1) (r - \zeta - \bar{\mu}_{L} - r_{0})$.

Thus, the optimization problem becomes
\[
\max_{(\pi_{t},v_{t}) \in \mathcal{H}} E \left[ \int_{0}^{\infty} e^{-\varrho t} u(F_{t}^{\pi,v},v_{t}) dt \right] \\
\text{s.t.} \quad F_{t}^{\pi,v} > \eta \Psi_{t}
\]
\[
\iff \max_{(\pi_{t},v_{t}) \in \mathcal{H}} E \left[ \int_{0}^{\infty} e^{-\varrho t} u(F_{t}^{\pi,v},v_{t}) dt \right] \\
\text{s.t.} \quad F_{t}^{\pi,v} > \eta
\]

The HJB equation (2.10) fits for this problem. It turns out to be
\[
0 = \sup_{(\pi_{t},v_{t}) \in \mathcal{H}} \left\{ -\varrho I + u(F,v) + \left[ (\pi'_{t} - \pi'_{L}) \Sigma (\tilde{\pi} - \pi_{L}) \right] F_{t}^{\pi,v} \right\}
\]
\[
\left\{ -\bar{\mu}_{L} - r_{0} + r - \zeta - v_{t} \right\} F_{t}^{\pi,v} \right\}
\]
\[
\left\{ + \frac{\varrho}{2} ( (\pi'_{t} - \pi'_{L}) \Sigma (\pi_{t} - \pi_{L}) ) \right\}
\]
(4.12)
with boundary condition \( J(\eta) = 0 \).

The first order conditions are

\[
v^*_t = \left(1 - \theta \right) \frac{F^0(1 - \delta)^{-1}}{J_F} \left(1 - \theta \right) \frac{F^0(1 - \delta)^{-1}}{J_F}
\]

and

\[
\pi^* = \frac{J_F}{J_F F_1} (\hat{\pi} - \pi_L) + \pi_L.
\]

Substituting \( v^*_t \) and \( \pi^* \) into (4.12) we get

\[
0 = \frac{1}{1 - \delta} F^0(1 - \delta) \left(1 - \theta \right) \frac{F^0(1 - \delta)^{-1}}{J_F} \left(1 - \theta \right) \frac{F^0(1 - \delta)^{-1}}{J_F}
\]

\[-J_F F \left(1 - \theta \right) \frac{F^0(1 - \delta)^{-1}}{J_F} \left(1 - \theta \right) \frac{F^0(1 - \delta)^{-1}}{J_F}
\]

\[-\theta \bar{J} - \frac{J_F^2}{2 J_F F} (\hat{\pi} - \pi_L)' \Sigma (\hat{\pi} - \pi_L)
\]

\[
= \left(1 - \theta \right) \frac{1}{(1 - \delta)} \left( \frac{\delta + \theta (1 - \delta)}{(1 - \delta) (1 - \theta)} \right) F \left(1 - \delta \right)^{-1} \frac{J_F^2}{J_F^2 (1 - \delta)^{-1}}
\]

\[-\theta \bar{J} - \frac{J_F^2}{2 J_F F} (\hat{\pi} - \pi_L)' \Sigma (\hat{\pi} - \pi_L).
\]

As in the previous section, we try

\[
J(\bar{F}) = c (\bar{F} + g)^{\theta (1 - \delta)}
\]

where \( c \) is a parameter that has to be evaluated. The derivatives become

\[
J_F = c (\bar{F} + g)^{\theta (1 - \delta) - 1}
\]

\[
J_{FF} = (\theta (1 - \delta) - 1) c (\bar{F} + g)^{\theta (1 - \delta) - 2}.
\]

Substituting into the HJB-equation we get

\[
0 = \left(1 - \theta \right) \frac{1}{(1 - \delta)^{-1}} \left( \frac{\delta + \theta (1 - \delta)}{(1 - \delta) (1 - \theta)} \right) F \left(1 - \delta \right)^{-1} \frac{J_F^2}{J_F^2 (1 - \delta)^{-1}}
\]

\[-\theta \bar{J} \frac{c (\bar{F} + g)^{\theta (1 - \delta)}}{\theta (1 - \delta)} - \frac{c (\bar{F} + g)^{\theta (1 - \delta)}}{2 (\theta (1 - \delta) - 1)} (\hat{\pi} - \pi_L)' \Sigma (\hat{\pi} - \pi_L).
\]
This equation cannot be explicitly solved in general; we would have to apply numerical methods. However, for \( \theta = \frac{1}{2} \) we get a closed form solution. In this special case, equation (4.14) turns out to be

\[
0 = \left(2^{\frac{\delta}{1-\delta}} \left(\delta + \frac{1}{1-\delta}\right)\right) c^{\frac{1}{1-\delta}} - \bar{q} \frac{2}{1-\delta} + \frac{1}{1+\delta} (\hat{\pi} - \pi_L)' \Sigma (\hat{\pi} - \pi_L)
\]

where \( \bar{q} = q + \frac{1}{2} (\delta - 1) (r - \bar{\mu}_L - \zeta - r_0) \).

Solving for \( c \) we get

\[
c = \left[ -\bar{q}^{\frac{2}{1-\delta}} + \frac{1}{1-\delta} (\hat{\pi} - \pi_L)' \Sigma (\hat{\pi} - \pi_L) \right]^{\frac{1}{1-\delta}}.
\]

Substituting \( c \) into (4.13) and considering \( J(\eta) = 0 \), i.e. \( g = -\eta \), the solution for the optimal surplus becomes

\[
\nu_t^* = \left[ \frac{1}{2} \left( \frac{2 \bar{q}}{1-\delta} - \frac{1}{1-\delta} (\hat{\pi} - \pi_L)' \Sigma (\hat{\pi} - \pi_L) \right)^{\frac{1}{1-\delta}} \left( \frac{F_t}{F_t - \eta} \right)^{-\frac{1}{1-\delta}} \right]^{\frac{1}{1-\delta}}
\]

Unlike in the unconstrained case of section 4.2.2, in the constraint case the surplus depends on the funding ratio. A higher funding ratio causes a bigger surplus and vice versa. Thus, the employee participates in the financial risk. Therefore, unlike in the unconstrained case, the plan sponsor can transfer a part of the risk to the employee. Thus, a very risk averse plan sponsor is better off with the constraint setting than with the unconstrained setting, if \( r_0 < \bar{r}_0 \).

The optimal portfolio turns out to be

\[
\pi_t^* = \frac{2}{\delta + 1} \left(1 - \frac{\eta \Psi_t}{F_t}\right) (\hat{\pi} - \pi_L) + \pi_L.
\]

Because of the constraint, the plan sponsor takes less risk if the funding ratio decreases.
4.2 High Minimum Yield Guarantee and Surplus

The funding ratio becomes

\[
\frac{dF_t}{F_t} = \frac{2}{\delta + 1} \left( 1 - \frac{\eta F_0}{F_t} \right)
\cdot \left( -\bar{\nu} dt + \frac{\delta + 3}{2(1 + \delta)} (\bar{\phi} - \pi_L)' \Sigma (\bar{\phi} - \pi_L) dt + (\bar{\phi} - \pi_L)' \sigma dZ_t \right)
\]

or

\[
\frac{F_t}{F_0} = (1 - \eta) \exp \left[ \frac{2}{\delta + 1} \left[ -\bar{\nu} + r - \zeta - r_0 - \bar{\mu}_L \right] + \frac{\delta - 1}{(1 + \delta)} (\bar{\phi} - \pi_L)' \Sigma (\bar{\phi} - \pi_L) \right] + \eta \Psi_t.
\]

Compared with the exogenous liability case the funding ratio decreases more slowly. The reason is that the surplus decreases if \( r_0 \) increases.

The accrued retirement assets of the individual employee turns out to be

\[
\frac{X_t}{X_0} = \exp \left[ (r_0 + \vartheta_t) t \right] \tag{4.15}
\]

where

\[
\vartheta_t = \begin{bmatrix}
\frac{2 \bar{\nu}}{1 + \delta} + \frac{\delta - 1}{1 + \delta} (r - \bar{\mu}_L - \zeta - r_0) \\
+ \frac{\delta - 1}{(1 + \delta)} (\bar{\phi} - \pi_L)' \Sigma (\bar{\phi} - \pi_L)
\end{bmatrix} \left( 1 - \frac{\eta \Psi_t}{F_t} \right)
\]

The process \( X_t \) is denoted as a feedback formula of the funding ratio \( F_t \). Note that \( 0 < \left( 1 - \frac{\eta \Psi_t}{F_t} \right) < 1 \) and \( 0 < \frac{\delta - 1}{(1 + \delta)} < 1 \). Thus, \( X_t \) increases for increasing \( r_0 \) within this setting.

Compared with the yield guarantee setting of chapter 3 with exogenous liability, for the same \( r_0 \), the employees prefer of course the minimum yield. However, the employee might prefer this setting even with a lower minimum yield \( r_0 \) than in the yield guarantee case, i.e. \( \bar{r}_0 \) as we will show next.

The employee’s expected utility is \( EU_e (X_T) \) where \( T \) is the date of retirement. Thus, if we assume a CRRA utility for the employee, i.e. \( U_e (X) \triangleq \frac{X^{1 - R_e}}{1 - R_e} \), \( R_e > 1 \), then we have

\[
EU_e (X_T) = \frac{X_0^{1 - R_e}}{1 - R_e} E \left[ \exp \left[ (1 - R_e) (r_0 + \Omega_T) T \right] \right]. \tag{4.16}
\]

We compare this expected utility with the yield guarantee setting, i.e. with the employee’s utility of the assets at retirement date \( T \). This utility is

\[
EU_e (X_T)_{v=0, r_0=\bar{r}_0} = U_e (X_T^{yg})
= U_e (X_0 \exp [\bar{r}_0 T])
= \frac{X_0^{1 - R_e}}{1 - R_e} \exp \left[ (1 - R_e) \bar{r}_0 T \right].
\]
In particular, we are looking at the set \( \mathcal{E} \triangleq \{ r_0 \mid EU_e (X_T) \geq U_e (X_T^{yg}) \} \) and we define \( \tilde{r}_0 \triangleq \inf \mathcal{E} \). In order to get \( \tilde{r}_0 \) we need to substitute \( F_T \) into the feedback formula (4.15) and calculate the expected value. However, since we cannot calculate the expected value in a closed form we use a simple numerical method. Because we can calculate the quantiles of \( U_e (X_T) \), we can approximate the distribution function of \( U_e (X_T) \) with an arbitrary number of quantiles.\(^4\) In Figure 4.3 \( \tilde{r}_0 \) is plotted for different \( \bar{r}_0 \). We can see that the higher the yield guarantee \( \bar{r}_0 \) the smaller the difference between \( \tilde{r}_0 \) and \( \bar{r}_0 \). This is economically intuitive since a high \( \bar{r}_0 \) is hard to beat, and therefore the surplus will be smaller than in the case that \( \bar{r}_0 \) is rather low. If the employee is more risk averse, \( \tilde{r}_0 \) becomes higher since she demands more security. Of course, if \( R_e \to \infty \) then \( \tilde{r}_0 = \bar{r}_0 \).

\(^4\)Note that for very high \( R_e \) an approximation leads to a less exact comparison than for low \( R_e \). The reason is that \( \lim_{R_e \to \infty} U_e \left( X_T^{yg} \right) = \lim_{R_e \to \infty} EU_e (X_T) = 0 \).
4.3 Conclusions

In this chapter we allowed the liability to be endogenous.

The first setting contained a lower bound and an upper bound for the funding ratio as practiced in the UK pension law. Additionally, we included a minimum yield that is lower than the riskless yield as supported by the EU-directive for life insurance. We have seen in the low minimum yield setting that, depending on the employee’s preference structure, it is possible that both the employee and the plan sponsor are better off if the minimum yield is very low.

In the second setting we included a surplus on the minimum yield in the optimization problem. We have seen that in the case of a CIES-utility function this surplus becomes a constant. If we additionally include a constraint on the funding ratio, the surplus becomes stochastic. The plan sponsor transfers a part of the financial risk to the employee. Since the employee prefers this minimum yield setting to the fixed yield setting, even if the minimum yield is lower than the fixed yield, the minimum yield plan might be an improvement for both the employee and the plan sponsor if the latter benefits enough from the lower guarantee and the risk transfer.

However, in the first setting the minimum yield is lower than the riskless yield, and in the second setting the minimum yield should be lower than in the fixed yield plan since the shortfall risk is already substantial in the fixed yield setting, as we have shown in chapter 3. An additional surplus could hardly be afforded if the yield guarantee stays equal. But if the yield guarantee is lower, then the very risk averse employees would prefer the fixed yield plan.

In the next chapter we will focus on settings that display a distinct improvement in this respect. We will suggest plans which reduce the shortfall risk that the plan sponsor is obliged to bear and provide a high yield guarantee for the risk averse employees. This is only possible if we expand the model to include risk transfers, not only from the plan sponsor to a collective of employees, but also matched to the employee’s individual risk preferences.
Chapter 5

Individual Risk-Sharing Models

As we have already seen in the previous chapters, there is a trade off between the guaranteed growth of the employee’s retirement assets and the security of the funding ratio. It was shown that in a framework with a yield guarantee, as practised in Switzerland, the plan sponsor bears substantial risk. However, if the government lowers the guarantee, the employee’s situation gets worse.

In chapter 4 we have studied the situation where the employees have the possibility to influence the risk allocation of the plan sponsor and gain a surplus on the minimum yield. Consequently, the minimum yield has assumed to be lower than the yield guarantee in chapter 3. Including an endogenous surplus we lowered the inter-generational risk redistribution since the employee’s return depended on the current funding ratio. Compared to the yield guarantee in chapter 3, the first advantage of this setting is that the plan sponsor does not need not the same amount of reserves, since the guarantee is lower. In expectation that the employees are better off than in the fixed yield guarantee plan, and as the employees act collectively, we still have a substantial intra-generational risk redistribution. However, the minimum yield is lower than the yield guarantee in chapter 3, and therefore, the protection level for the employee has decreased. What is more the plan sponsor still bears substantial risk.

Keeping these issues in mind, in this chapter we introduce a risk-sharing model in three different settings, thereby focussing on improvements and more individuality.

The basics of the settings are as follows.

The plan sponsor intends to optimize the long-term growth rate of the funding ratio and the employee her assets at a specific retirement date. The participation is assumed to be linear to the difference between a predefined yield $r_0$ and the performance the plan sponsor has achieved. The linear factor, i.e. the participation rate $\alpha$, will be the crucial variable in this setting. Note that that difference could be negative and, therefore, in general $r_0$ is no longer a lower bound. The employee faces the probability that his accrued retirement assets decreases in $dt$ if $\alpha > 0$. 

83
To explain the basic mechanism we take a simple numerical example:

The current wealth of the pension is \( W_0 = 250. \)
The value of the liabilities is \( L_0 = 200. \)
Thus the funding ratio is \( F_0 = \frac{250}{200} = 1.25. \)
The service costs within the next period \( ds \) are 0.5\% of the current wealth.
We drop the net contributions within the next period \( dC = 0. \)
The value of her current retirement assets is \( X_0 = 1. \)
The contribution of the individual employee is \( dc_i \) is 0.5\% of her current retirement assets.
The yield \( r_0 \) is 3\%.
The return of the riskless asset \( r \) is 2\%.
Finally, the return on the risky assets within the next period \( d\Pi \) is 10\%.

\( d\Pi \) is the return on the risky assets, \( F_1 \) is the funding ratio at the end of the period, and \( X_1 \) is the retirement assets of an individual employee.

**Case I: no participation, i.e. \( \alpha = 0 \)**

<table>
<thead>
<tr>
<th>( d\Pi )</th>
<th>10%</th>
<th>-5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>( F_1 )</td>
<td>( \frac{250(1-0.5%+2%+10%)}{200(1+3%)} )</td>
<td>( \frac{250(1-0.5%+2%-5%)}{200(1+3%)} )</td>
</tr>
<tr>
<td>( X_1 )</td>
<td>( 1 + 3% + 0.5% = 1.035 )</td>
<td>( 1 + 3% + 0.5% = 1.035 )</td>
</tr>
</tbody>
</table>

**Case II: participation 60\%, i.e. \( \alpha = \frac{3}{5} \)**

<table>
<thead>
<tr>
<th>( d\Pi )</th>
<th>10%</th>
<th>-5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>( F_1 )</td>
<td>( \frac{250(1-0.5%+2%+10%)}{200(1+3%+\frac{3}{5}(2%+10%-3%))} )</td>
<td>( \frac{250(1-0.5%+2%-5%)}{200(1+3%+\frac{3}{5}(2%-5%-3%))} )</td>
</tr>
<tr>
<td>( X_1 )</td>
<td>( 1 + 3% + 0.5% + \frac{3}{5}(2% + 10% - 3%) )</td>
<td>( 1 + 3% + 0.5% + \frac{3}{5}(2% - 5% - 3%) )</td>
</tr>
<tr>
<td></td>
<td>( = 1.089 )</td>
<td>( = 0.999 )</td>
</tr>
</tbody>
</table>

If we compare the participation plan, with the yield guarantee plan it is obvious that the variance of the funding ratio decreases while the variance of the employee’s assets increases. The question arises what the “optimal” risk sharing is, i.e. the “optimal” \( \alpha. \)

We will study three mechanisms to evaluate an “optimal” participation rate \( \alpha. \)

1. The plan sponsor optimizes the portfolio and the participation rate \( \alpha \) and is obliged to offer the same contract \((\pi^*(r_0), \alpha^*(r_0))\) to every employee. The employee chooses
individually her optimal reference yield \( r^*_0 \) and the plan sponsor invests in \( \pi^* (r^*_0) \) and provides \( \alpha^* (r^*_0) \).

2. The plan sponsor optimizes the portfolio and the participation rate \( \alpha \). The yield \( r_0 \) is set exogenously and mandatory for every plan sponsor but the employee can choose the pension fund that provides her optimal \( \alpha \).

3. Again the yield \( r_0 \) is exogenously set and mandatory for every plan sponsor. The plan sponsor is obliged to offer to every employee the same contract \( \pi^* (\alpha) \) that contains the portfolio the plan sponsor would choose for every given \( \alpha \in [0, 1] \). The employee chooses \( \alpha \) and the plan sponsor invests in the portfolio.

It is assumed that, ex ante, the plan sponsor does not now anything about the risk preferences of the employees, for otherwise he might offer a “biased” contract in setting one and three.

Since the individual employee can optimize the risk/return profile in all three settings, these settings are very close to an individual life cycle model, provided that the individual employee is also allowed to optimize her contribution flow. Moreover, in the second and the third setting, employees with a high risk aversion have the same protection as in the plan with a yield guarantee \( r_0 \) since they can always choose a pension fund, or a contract, providing \( r_0 \) as yield guarantee.

Using CRRA utilities, we will provide for each of these mechanisms a closed form solution for the optimal portfolio strategy and the optimal participation rate.

Further, we will show that these schemes exhibit the following characteristics:

I. Most employees would prefer these settings to both the individual investment and the setting with a fixed yield \( r_0 \). The redistribution of risk in all three settings comes at the expense of the employees with very low risk aversion. The settings one and three are preferred to the autonomous investment (i.e. the Merton solution) by all employees except for those with very low risk aversion. In setting two employees with moderate risk aversion might also prefer to invest autonomously.

II. The change from the yield guarantee framework to setting one is at the expense of the very risk averse employees but only if the plan sponsor is also very risk averse, and at the expense of the plan sponsors with low risk aversion.

The change from the yield guarantee framework to setting two while keeping \( r_0 \)
equal is a Pareto improvement.
The change from the yield guarantee framework to setting three while keeping \( r_0 \)
equal is beneficial for the risk averse plan sponsors and all employees but at the
expense of the plan sponsor with low risk aversion.
The change from the yield guarantee framework to setting two or three with a cor-
responding raise of \( r_0 \) is beneficial for the risk averse employees but, in general, at
the expense of the plan sponsors with low risk aversion.

In the following section we define the model setting.

5.1 The Basics of the Settings

Throughout this chapter we will assume that \( r_{0,t} > r_t \). As in the previous chapters we
assume the wealth process following

\[
dW^π_t = W^π_t \left( (\pi'_t (\mu_t - r_1) + r_t) \, dt + \pi'_t \sigma_t dZ_t \right) + dC_t - \zeta_t W^π_t \, dt
\]

For simpler notation, henceforth we write \( \mu = \mu_t, \sigma = \sigma_t, r = r_t, \zeta = \zeta_t \).

The liability becomes

\[
dl_t = L_t (r_{0,t} dt + p_t (\pi_t)) + dl_t + dC_t
\]

It contains a new variable \( p_t (\pi_t) \) which expresses the participation proportion of the cur-
rent liability. In particular the participation is assumed to be proportional to the return of
the portfolio minus a yield \( r_{0,t} \). Thus, the participation \( p_t (\pi_t) \) is assumed to evolve as

\[
p_t (\pi_t) \triangleq \alpha_t \left((\pi'_t (\mu - r_1) + r) \, dt + \pi'_t \sigma dZ_t - r_{0,t} dt\right), \alpha_t \in [0, 1]
\]

The participation rate \( \alpha_t \) is a bounded control variable and we claim it to be admissible, i.e.
\( \alpha_t \in \mathcal{Y} \). The liability is given by

\[
dl^α_t = L^α_t (r_{0,t} dt + \alpha_t \left((\pi'_t (\mu - r_1) + r) \, dt + \pi'_t \sigma dZ_t - r_{0,t} dt\right)) + dl_t + dC_t.
\]

We set \( \bar{\mu}_t = \bar{\mu}_L(t), \sigma_L = \sigma_L(t), \pi_L = \pi_L(t) \) and

\[
dl_t = L_t \left((\pi'_t (\mu - r_1) \, dt + \sigma dZ_t) + \bar{\mu}_L dt + \sigma_L dZ^\lambda_t\right)
\]
5.1 The Basics of the Settings

as in the previous chapters.

Thus, the liability process $L_t$ evolves as

$$\frac{dL_t}{L_t} = r_0 dt + \alpha_t \left( (\pi'_L (\mu - r) + r) dt + \pi'_L \sigma dZ_t - r_0 dt \right) + \pi'_L (\mu - r) dt + \pi'_L \sigma dZ_t + \lambda dt + \sigma dZ_t.$$

The net contributions are also assumed to evolve as in the previous chapters, i.e.

$$dC_t = W_t \left( \pi'_C ((\mu - r) dt + \sigma dZ_t) + \mu_C dt + \sigma_C dZ_t \right)$$

where $\mu_C = \mu_C(t), \sigma_C = \sigma_C(t), \pi_C = \pi_C(t)$.

Henceforth, for simpler notation, we write $r_0 = r_{0,t}$.

The funding ratio

$$F_t^{\pi,\alpha} = \frac{W_t}{L_t^\alpha}$$

is the state variable as in the previous chapters. In particular, the plan sponsor’s mechanisms to optimize the long-term growth rate are:

$$\sup_{\pi \in \mathcal{G}} \inf_{T \to \infty} \frac{1}{\beta T} \ln \beta E \left[ U_p \left( F_T^{\pi,\alpha} \right) \right]$$

used in section 5.4, and

$$\sup_{(\pi', \alpha) \in \mathcal{H}} \lim_{T \to \infty} \frac{1}{\beta T} \ln \beta E \left[ U_p \left( F_T^{\pi,\alpha} \right) \right]$$

s.t. $\alpha_t \in [0, 1]$ used in the sections 5.2 and 5.3. However, for technical reasons we claim that $\alpha_t \in [0, 1)$ at first.

The funding ratio evolves as

$$\frac{dF_t^{\pi,\alpha}}{F_t^{\pi,\alpha}} = \left( 1 - F_t^{\pi,\alpha} \right) \left( \pi'_C (\mu - r) dt + \pi'_C \sigma dZ_t + \mu_C dt + \sigma_C dZ_t^\lambda \right) - \pi'_L ((\mu - r) dt + \sigma dZ_t) - \mu_L dt - \sigma_L dZ_t^\lambda + \left( \frac{dl_t}{L_t^\alpha} + \frac{dC_t}{L_t^\alpha} \right)^2 - \frac{dC_t dl_t}{W_t^\alpha L_t^\alpha} - \frac{dC_t^2}{W_t^\alpha L_t^\alpha} + (1 - \alpha_t) \left( \pi'_L (\mu - r) + r - r_0 - \alpha_t \pi'_L \pi_t dt - \zeta dt \right) + (2 \alpha_t - 1) \pi'_L \pi dt + (2 \alpha_t - 1) F_t^{\pi,\alpha} \pi'_C \pi dt + (1 - \alpha_t) \pi'_L \sigma dZ_t.$$
The HJB equation (2.13) turns out to be

\[
0 = \sup_{(\pi^t, \alpha^t) \in \mathcal{H}} \left\{ J_F \left[ \begin{array}{c}
- \alpha_t (1 - \alpha_t) \pi^t_L \\
+ (2 \alpha_t - 1) \pi^t_L \\
+ ((2 \alpha_t - 1) F - \alpha_t) \pi^t_C \\
+ (1 - F) \bar{p}_C - \bar{p}_L \\
+ (1 - \alpha_t) (\gamma - \zeta + \psi \\
- \beta v_F^2 \left( \pi_i^t \Sigma \pi_t + (1 - F)^2 \sigma_2^2 + \sigma_L^2 \right) \\
- J_{FF} F \\
\end{array} \right] \right\}_{\pi^t, \alpha^t} (5.1)
\]

where now \( \pi^t_i = (1 - \alpha_t) \pi^t_i + (1 - F) \pi^t_C - \pi^t_L \) and \( \psi dt \equiv \left( \frac{dl_t}{l_t} + \frac{dC}{C_t} - \frac{dl_t}{l_t} - \frac{dC}{C_t} \right)^2 - \frac{dC}{C_t} dl_t - \frac{dC^2}{C_t^2} \).

Henceforth, to make the interpretation simpler, we assume that the utility function \( U_p \) is chosen such that \( J_F > 0 \) and \( J_{FF} < 0 \). This is plausible since the \( J \)-function can be interpreted as an intertemporal indirect utility function which is increasing and concave in \( F \). Similarly, to avoid more case differentiations in sections 5.2 and 5.3, we assume that \( J_{FF} F < -2J_F \), i.e. the indirect utility exhibits a risk aversion that is “high enough”. For the contract setting in section 5.4, we will also consider the possibility that \( J_{FF} F \geq -2J_F \).

The first order conditions of (5.1) are

\[
\pi^*_i = \frac{J_F}{2 \alpha_t^*[J_F - (1 - \alpha_t^*) J_{FF} F]} \left( \pi - \frac{1}{1 - \alpha_t^*} \pi_L + \frac{\alpha_t^* - F_i}{1 - \alpha_t^*} \pi_C \right) \\
+ \frac{1}{1 - \alpha_t^*} \pi_L - \frac{\alpha_t^* - F_i}{1 - \alpha_t^*} \pi_C . \tag{5.2}
\]

and

\[
\alpha^*_{ij} = \frac{J_F \left( \frac{\pi^*_i (\mu - r)}{\pi^*_t L / \pi_t} + \frac{r - r_0}{\pi^*_t L / \pi_t} - \frac{\pi^*_t \Sigma \pi^*_t}{\pi^*_t L / \pi_t} + 1 \right) + J_{FF} F_i}{2J_F + J_{FF} F_i} \\
- \frac{\pi^*_t \Sigma \pi^*_t}{\pi^*_t L / \pi_t} \left( 1 - F_i \right) \frac{\pi^*_t \Sigma \pi_C}{\pi^*_t L / \pi_t} \tag{5.3}
\]

Note that we needed to claim that \( \alpha_t \in [0, 1] \), otherwise \( \pi^*_t \) would not exist in this case.

Now we can substitute (5.2) into (5.3), then solve for \( \alpha^*_t \), and finally, substitute into (5.1) in order to calculate the \( J \)-function.

The value of an employee’s accrued retirement assets becomes

\[
\frac{dX_i}{X_i} = r_0 dt + \alpha_t \left( (\pi^*_t \mu - r_1) + r \right) dt + \pi^*_t \sigma dZ_t - r_0 dt + dc_i (t) \tag{5.4}
\]
where \( dc_i(t) \) are the individual net contributions at time \( t \). The employee’s aim is to optimize this value at her specific retirement date. The control variable in the first setting in section 5.2 is \( r_0 \in [r, r + h], h > 0 \) where \( r_0 \) is claimed to be admissible, i.e. \( r_0 \in Y \). In the settings of section 5.3 and 5.4 the control variable is \( \alpha_t \). Thus the employee’s three optimization mechanisms are

<table>
<thead>
<tr>
<th>Section 5.2</th>
<th>( \max_{r_0 \in Y} \mathbb{E} \left( X_T^{\pi_{r_0}} \right) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>s.t. ( r_0 \in [r, r + h], h &gt; 0 )</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Section 5.3</th>
<th>( \max_{\alpha_f \in Y} \mathbb{E} \left( X_T^{\pi_{\alpha_f}} \right) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>s.t. ( \alpha_f \in [0, 1] )</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Section 5.4</th>
<th>( \max_{\alpha_t \in Y} \mathbb{E} \left( X_T^{\pi_{\alpha_t}} \right) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>s.t. ( \alpha_t \in [0, 1] )</td>
<td></td>
</tr>
</tbody>
</table>

### 5.2 The Reference Yield Plan

#### 5.2.1 General Model

In this section the yield \( r_0 \) is called reference yield, since it is not a guarantee but the value with which the financial return is compared for evaluating the participation. In a first step we assume that the reference yield is exogenously given before we include it into the employee’s optimization mechanism in section 5.2.3.

The plan sponsor optimizes both the portfolio \( \pi_t \) and the participation rate \( \alpha_t \). Thus, from the plan sponsor’s perspective the problem in this section becomes

\[
\sup_{(\pi, \alpha) \in H} \liminf_{T \to \infty} \frac{1}{T} \ln \left( \beta E \left[ U_p (F_T^{\pi, \alpha}) \right] \right)
\]

s.t. \( \alpha_t \in [0, 1] \)

We will simplify our setting while assuming that the liability and the net contributions are not tradable at all, i.e. \( \pi_L = 0, \pi_C = 0 \). Thus, we state

\[
dl_t \triangleq L_t \left( \mu_L dt + \sigma_L dZ^L_t \right),
\]

\[
dC_t \triangleq W_t \left( \mu_C dt + \sigma_C dZ^C_t \right).
\]

Hence the liability process evolves as

\[
\frac{dL_t}{L_t} = r_0 dt + \alpha_t \left( \pi_t' (\mu - r \mathbf{1}) dt + (r - r_0) dt + \pi_t' \sigma dZ_t \right)
\]

\[
+ \mu_L dt + \sigma_L dZ^L_t + F_t \mu_C dt + F_t \sigma_C dZ^C_t.
\]
and the funding ratio becomes

\[
\frac{dF_t}{F_t} = \left[ (1 - \alpha_t) \left( \pi'_t (\mu - r1) + r - r_0 \right) - \zeta + (\alpha^2_t - \alpha_t) \pi'_t \Sigma \pi_t \right] dt
\]

\[
-\mu_L + (1 - F_t) \mu_C + \sigma^2_L + \sigma^2_C \pi'_t \pi_t - \mu L + (1 - F_t) \mu C + \sigma^2 L + \sigma^2 C \pi'_t \pi_t
\]

\[
+ (1 - \alpha_t) \pi'_t \sigma dZ_t - \sigma L dZ_t^\lambda + (1 - F_t) \sigma C dZ_t^\gamma .
\]

The first order conditions of (5.1) for the optimal portfolio (5.2) and the optimal participation rate (5.3) lead to a non-linear system of equations

\[
\pi^*_t = \frac{J_F}{2 \alpha^*_p, t J_F - (1 - \alpha^*_p, t) J_{FF} F_t} \hat{\pi} , \quad (5.5)
\]

\[
\alpha^*_p, t = \frac{J_F \left( \pi'_t (\mu - r1) \pi'_t \Sigma \pi_t + r - r_0 + 1 \right) + J_{FF} F_t}{2 J_F + J_{FF} F_t} . \quad (5.6)
\]

Substituting (5.5) into (5.6) we get

\[
0 = \zeta + \left( \frac{J_F}{2 \alpha^*_p, t J_F - (1 - \alpha^*_p, t) J_{FF} F_t} \right)^2 \pi^*_t , \quad (5.7)
\]

where \( \zeta \triangleq \frac{r - r_0}{\hat{\pi} \Sigma \hat{\pi}} \). Henceforth, we assume that \(-1 < \zeta < 0 \). The case \( \zeta \leq -1 \) leads to technical problems and would imply an utopian high \( r_0 \).

Solving (5.7) for \( \alpha^*_p, t \) we get

\[
\alpha^*_p, t = \frac{J_{FF} F_t \hat{\pi} \pm J_F \sqrt{-\zeta}}{2 J_F + J_{FF} F_t} \hat{\pi} \zeta \quad \alpha^*_p, t = \frac{J_{FF} F_t \hat{\pi} \pm J_F \sqrt{-\zeta}}{2 J_F + J_{FF} F_t} \hat{\pi} \zeta
\]

and substituting into (5.5) we have

\[
\pi^*_t = \pm \sqrt{-\zeta} \hat{\pi} .
\]

Thus we have two solutions. For practical reasons we decide to analyze the solution that, at least partly, leads to \( \alpha^*_p, t \in [0, 1] \) and always leads to \( \pi^*_t = c \hat{\pi} C > 0 \). So we consider only the solution

\[
\pi^*_t = -\sqrt{-\zeta} \hat{\pi} , \quad (5.8)
\]

\[
\alpha^*_p, t = \frac{1 - J_F + J_{FF} F_t}{2 J_F + J_{FF} F_t} . \quad (5.9)
\]
5.2 The Reference Yield Plan

Remember that we have assumed that $2J_F + J_{FF}I < 0$ for this section. Then, if $rac{1}{\sqrt{-\zeta}}J_F + J_{FF}F \geq 0$ we have the corner solution $(\pi_{t}', \alpha_{p.t}^*) = \left(\frac{1}{\sqrt{-\zeta}}\hat{\pi}', 0\right)$ which is equal to the solution of the yield guarantee plan in section 3.2.

If $\zeta \leq -\frac{1}{4}$ we have also a corner solution $(\pi_{t}', \alpha_{p,t}^*) = \left(\frac{1}{2}\hat{\pi}', 1\right)$. Note that if $\zeta = -\frac{1}{4}$ we have $(\pi_{t}', \alpha_{p,t}^*) = \left(\frac{1}{2}\hat{\pi}', 1\right)$. The investment strategy of this solution is not quite intuitive. A discussion of this issue is presented in section 5.4. At this point, we are interested in the inner solutions $\alpha_{p,t}^* \in (0, 1)$, which we attain if $\frac{1}{\sqrt{-\zeta}}J_F + J_{FF} < 0 \land \zeta > -\frac{1}{4}$. In this case we get $\pi_t^* = \sqrt{-\zeta}\hat{\pi}$. The solution for the portfolio is completely independent of any risk behavior by the plan sponsor. Every plan sponsor would invest the same way. Economically this is an interesting result, because if every plan sponsor fund holds the same portfolio weights, this leads to a very stable financial market. Moreover, the investment strategy is very transparent and the moral hazard problem in the investment strategy disappears. Also, the gambling-behavior in the presence of positive or negative net contributions also disappears in the investment strategy. Note that the optimal portfolio strategy is a rather general result, since, additional to the general utility function, we allow stochastic parameters, stochastic net contributions and an additional liability process $dl$.

Further, note that if we act in a two fund framework, then we get $\pi_t^* = \sqrt{-\zeta}\hat{\pi} = \sqrt{\frac{\pi_0 - \pi_0^2}{\hat{\pi}^2}}\hat{\pi} = \frac{\sqrt{\pi_0 - \pi_0^2}}{\hat{\pi}}$ for the weight of the risky fund. In this case even the expected return does not influence the optimal portfolio.

However, the preference structure appears in the participation rate $\alpha_{p,t}^*$. Hence, all the addressed problems are banned from the investment strategy but can appear in the optimal participation rate. Therefore, the problems may not have disappeared but are reduced to one single variable, which can be monitored by every single employee, since she is transparently affected by $\alpha_{p,t}^*$.

To calculate the $J$-function we can substitute (5.8) and (5.9) into (5.1) and we get

$$0 = \begin{bmatrix} -\beta v_F J_F + J_F \left((\frac{1}{2} - 1) \mu_C dt - \mu_L dt\right) \\ -\frac{E_f^2}{4J_F + 2J_{FF}^2} \left(1 - 2\sqrt{-\zeta}\right)^2 \hat{\pi} \Sigma \hat{\pi} \\ + \frac{1}{2}J_{FF}^2 \left((\frac{1}{2} - 1)^2 \sigma_C^2 + \sigma_L^2\right) \end{bmatrix}.$$ 

Now let us turn to a special case.

---

1What we need is that the funding ratio process is univariate and that liability and net contributions are completely untradeable.
5.2.2 Special Case: No Net Contributions and CRRA Utility

We assume CRRA-utility, i.e. \( U_p(z) = \frac{1}{1-R_p} z^{(1-R_p)} \) with relative risk aversion \( R_p > 2 \) for the plan sponsor, and \( U_e(z) = \frac{1}{1-R_e} z^{(1-R_e)} \) with relative risk aversion \( R_e > 1 \) for the employee, no net contributions, i.e. \( dC_t = 0 \) and \( r_0 < r + \frac{1}{2} \pi' \Sigma \pi \Leftrightarrow \zeta > -\frac{1}{4} \). The model parameters \( (\mu, \sigma, r, \mu_L, \sigma_L) \) are assumed to be deterministic. The optimization problem is

\[
\sup_{(\pi', \alpha)} \lim_{T \to \infty} \frac{1}{(1-R)T} \ln \left( (1 - R) E \left[ U_p(F_T^{\pi, \alpha}) \right] \right) \\
\text{s.t. } \alpha \in [0,1].
\]

The funding ratio becomes

\[
\frac{dF_t}{F_t} = \left[ (1 - \alpha_t) \left( \pi_t' (\mu - r1) + r - r_0 \right) - \mu_L - \zeta + (\alpha_t^2 - \alpha_t) \pi_t' \Sigma \pi_t + \sigma_L^2 \right] dt \\
+ (1 - \alpha_t) \pi_t' \sigma_t dZ_t - \sigma_L dZ_t^\lambda. \tag{5.10}
\]

Using the theorems 2.9 and 2.10 the optimization problem turns out to be

\[
\max_{\pi', \alpha} \left\{ \frac{2(1-\alpha_t)}{R} \left( \left( \pi' (\mu - r1) + r - r_0 \right) - \alpha_t \pi' \Sigma \pi \right) - (1 - \alpha_t) \pi' \Sigma \pi \right\} \\
\text{s.t. } \alpha \in [0,1]
\]

and the solutions are

\[
\left( \pi^*, \alpha^*_p \right) = \left( \sqrt{-\zeta} \pi', \frac{R_p - \frac{1}{\sqrt{-\zeta}}}{R_p - 2} \right). \tag{5.11}
\]

Thus if \( R_p < \frac{1}{\sqrt{-\zeta}} \) we have a corner solution \( \alpha^*_p = 0 \). Otherwise, if \( R_p > \frac{1}{\sqrt{-\zeta}} \) we get \( \alpha^*_p \in (0,1) \). So, if the plan sponsor is “more” risk averse, i.e. \( R_p > \frac{1}{\sqrt{-\zeta}} \), then he will share the risk with the employees. Otherwise he is risk tolerant enough to bear the whole risk of the constant\(^2\) portfolio \( \sqrt{-\zeta} \pi \) himself.

In a further step we intend to compare this plan with the framework in chapter three where we had a yield guarantee \( r_0 \). From the plan sponsor’s perspective it is obvious that this mechanism leads to an improvement since he could keep the status quo while choosing \( \alpha = 0 \). In order to obtain the conditions for which the employees would also prefer this model to the yield guarantee setting we calculate \( EU_e(X_t(\pi^*, \alpha^*)) \). Again

\(^2\)We call the portfolio “constant” even though in general it is allowed to depend on time since the parameters are deterministic.
ignoring the individual contributions \( dc_i = 0 \), since they do not influence the comparison, the individual accrued retirement assets (5.4) evolve as

\[
\frac{dX_t}{X_t} = r_0 dt + \alpha_p^* \left( (\pi^{*'} (\mu - r_1) + r) dt + \pi^{*'} \sigma dZ_t - r_0 dt \right)
\]

and, therefore,

\[
X_t = X_0 \exp \left[ (1 - \alpha_p^*) r_0 t + \alpha_p^* \left( (\pi^{*'} (\mu - r_1) + r) t - \frac{1}{2} \alpha_p^* \pi^{*'} \Sigma \pi^* t + \pi^{*'} \sigma^2 Z_t \right) \right]
\]

and for the expected utility

\[
EU_e (X_t) = X_0^{1 - R_e} \exp \left[ (1 - R_e) r_0 t + \alpha_p^* \left( (\pi^{*'} (\mu - r_1) + r - r_0) t - \frac{1}{2} \alpha_p^* R_e \pi^{*'} \Sigma \pi^* t \right) \right]. \tag{5.12}
\]

Substituting the optimal portfolio and the optimal participation (5.11) into the expected utility we have

\[
EU_e (X_t) = X_0^{1 - R_e} \exp \left[ (1 - R_e) r_0 t + \alpha_p^* \left( (\pi^{*'} (\mu - r_1) + r - r_0) t - \frac{1}{2} \alpha_p^* R_e \pi^{*'} \Sigma \pi^* t \right) \right] = X_0^{1 - R_e} \exp \left[ (1 - R_e) (r_0 + g \hat{\pi}^{*'} \Sigma \hat{\pi}) t \right]
\]

where

\[
g = \frac{2(R_p - 2)(R_p \sqrt{\varepsilon - 1} - 1 - \sqrt{\varepsilon - 1} - R_e(R_p \sqrt{\varepsilon - 1} - 1)^2)}{2(R_p - 2)^2}.
\]

Comparing with the yield guarantee plan \( U_r (X_0 e^{r_0 t}) \) we get

\[
EU_e (X_t) > U_r (X_0 e^{r_0 t}) \iff 1 < R_e < \hat{R}_e (R_p) \tag{5.13}
\]

where \( \hat{R}_e (R_p) \triangleq \frac{2(R_p - 2)(1 - \sqrt{\varepsilon - 1})}{R_p \sqrt{\varepsilon - 1} - 1} \). Figure 5.1 shows the function \( \hat{R}_e (R_p) \) for several \( r_0 \). Note that the curves start at \( R_p > \frac{1}{\sqrt{\varepsilon - 1}} \) (dotted lines) since for \( R_p \leq \frac{1}{\sqrt{\varepsilon - 1}} \) we have a corner solution \( \alpha_p^* = 0 \), which is the yield guarantee plan. We can see that for a moderate \( r_0 \) a wide range of employees prefer the participation plan, even though the participation rate is fixed by the plan sponsor.

However, so far this participation plan does not include the individual employee’s risk preference and there is no protection against a downside risk. As we will see in the next section, both disadvantages can be eliminated by a contract.
Figure 5.1: Reference yield plan or yield guarantee setting? The employee’s perspective. Parameters: $r = 0.02$, $\zeta = \bar{\mu}_L = 0$, $\mu = 0.1$ and $\sigma = 0.2$ for the risky fund in a two fund framework.
5.2 The Reference Yield Plan

5.2.3 Impact of the Reference Yield

For the sake of a closed form solution, we apply the special case of the previous section, i.e. $U_p(z) = \frac{z^{1-R_p}}{1-R_p}$, $R_p > 2$, $U_c(z) = \frac{z^{1-R_c}}{1-R_c}$, $R_c > 1$, $r_0 < r + \frac{1}{2} \hat{\pi}' \Sigma \hat{\pi}$ and $dC_t = 0$. Henceforth, we assume that $\mu_L = \sigma_L = 0$ in order to have a simpler notation. The parameters $\mu_L$, $\sigma_L$ do not influence the optimizing problems and therefore have the same effect on the optimal long-term growth rates for different reference yields. Thus, we don’t loose any insights if we set $\mu_L = \sigma_L = 0$.

To reduce the problem to two variables; we write the portfolio solution of (5.11) as $\pi^* = x^*(r_0) \hat{\pi}$. The general optimal long-term growth rate is

$$\Gamma_p \left( \alpha^*_p(r_0), x^*(r_0), r_0 \right) \triangleq \liminf_{T \to \infty} \frac{1}{(1-R) T} \ln \left( (1-R) E \left[ U \left( F_T^{\pi^*_t(r_0), \alpha^*_p(r_0)} \right) \right] \right)$$

(5.14)

$$= \frac{R_p \hat{\pi}' \Sigma \hat{\pi}}{2} \left( 1 - \alpha^*_p(r_0) \right) \begin{bmatrix} 2 R_p \left( x^*(r_0) + \frac{r_r_0}{\hat{\pi}' \Sigma \hat{\pi}} \right) \\ - \alpha^*_p(r_0) (x^*(r_0))^2 \\ - \left( 1 - \alpha^*_p(r_0) \right) (x^*(r_0))^2 \end{bmatrix} - \zeta.$$  

The comparative statics of (5.14) can be calculated by using the envelope theorem\(^3\)

$$\frac{d \Gamma_p \left( x^*(r_0), \alpha^*_p(r_0), r_0 \right)}{dr_0} = - \left( 1 - \alpha^*_p(r_0) \right) < 0$$

(5.15)

The plan sponsor prefers the lowest possible reference yield, which is not surprising, since the lower the reference yield the lower his risk ceteris paribus.

However, the question we are interested in is: How high can the reference yield $r_0$ be set if we want the plan sponsor to be at least as well off as in the case when he has to afford the yield guarantee $\bar{r}_0$? Thus we are looking for the upper boundary $\hat{r}_0$ below which the plan sponsor would still choose the participation plan instead of the yield guarantee setting providing $\bar{r}_0$. Note that the optimal growth rate in the yield guarantee setting is

$$\liminf_{T \to \infty} \frac{\ln E \left[ (1-R_p) U_p \left( F_T^{\pi^*_t(r_0), 0} \right) \right]}{(1-R_p) T} = \frac{\hat{\pi}' \Sigma \hat{\pi}}{2 R_p} + r - \bar{r}_0 - \zeta.$$  

The optimal solutions from the previous section are

$$\left( x^*(r_0), \alpha^*_p(r_0) \right) = \left( \sqrt{-\zeta}, \frac{R_p - \frac{1}{2}}{R_p - 2} \right)$$

(5.16)

\(^3\)For the envelope theorem see e.g. Léonard and Long (1992).
Substituting (5.16) into \( \Gamma_p \left( x^* (r_0) , a_p^* (r_0) , r_0 \right) \) we get

\[
\Gamma_p \left( a_p^* (r_0) , x^* (r_0) , r_0 \right) = \frac{\left( 1 - 2 \sqrt{\hat{\zeta}} \right)^2}{2 \left( R_p - 2 \right)} \bar{\Sigma} \hat{\Sigma} - \zeta.
\]

Having set \( \bar{\zeta} \triangleq \frac{\hat{\zeta} - r_p}{\bar{\Sigma} \hat{\Sigma}} \), the upper boundary can be found as follows

\[
\hat{r}_0 = \sup \left\{ r_0 \leq r + \frac{\hat{\pi}' \Sigma \hat{\pi}}{4} \left| \begin{array}{c}
\lim inf_{T \to \infty} \ln E \left[ \left( 1 - R_p \right) U_p \left( \bar{\pi}^* (0) \right) \right] \\
\geq \lim inf_{T \to \infty} \left( 1 - R_p \right) U_p \left( \bar{\pi}^* (\bar{r}_0) \right) \\
\geq \frac{\bar{\pi}' \Sigma \hat{\pi}}{2 R_p} + r - \hat{r}_0 - \zeta
\end{array} \right. \right\}
\]

\[
= \sup \left\{ r_0 \leq r + \frac{\hat{\pi}' \Sigma \hat{\pi}}{4} \left| \begin{array}{c}
\Gamma_p \left( a_p^* (r_0) , x^* (r_0) , r_0 \right) \\
\geq \frac{\bar{\pi}' \Sigma \hat{\pi}}{2 R_p} + r - \hat{r}_0 - \zeta
\end{array} \right. \right\}
\]

\[
= \sup \left\{ r_0 \leq r + \frac{\hat{\pi}' \Sigma \hat{\pi}}{4} \left| \begin{array}{c}
\left( 1 - 2 \sqrt{\hat{\zeta}} \right)^2 \\
\frac{1}{R_p} \geq \frac{1}{R_p} + 2 \hat{\zeta}
\end{array} \right. \right\}.
\]

Note that \( \frac{\left( 1 - 2 \sqrt{\hat{\zeta}} \right)^2}{R_p - 2} \geq 0 \) and, therefore, if \( \frac{1}{2 R_p} + \bar{\zeta} \leq 0 \) i.e. \( R_p \geq -\frac{1}{2 \hat{\zeta}} \), then the pension’s situation is always better or equal when choosing the participation system. Otherwise, if \( R_p < -\frac{1}{2 \hat{\zeta}} \) we can calculate \( \hat{r}_0 \) by solving

\[
\frac{\left( 1 - 2 \sqrt{\frac{r - h}{\bar{\pi}' \Sigma \hat{\pi}}} \right)^2}{R_p - 2} = \frac{1}{R_p} + 2 \hat{\zeta}
\]

for \( h \). This equation has two solutions

\[
h_{\pm} (R_p) = r + \frac{\hat{\pi}' \Sigma \hat{\pi}}{2 R_p} \left[ R_p - 1 + \frac{\left( R_p - 2 \right) R_p \hat{\zeta}}{R_p} \right] \pm 2 \sqrt{\frac{\left( R_p - 2 \right) R_p \hat{\zeta}}{R_p}} \left( 1 + 2 R_p \hat{\zeta} \right)
\]

As shown in section 5.2.2, for \( R_p = \frac{1}{\sqrt{-\hat{\zeta}}} \) we have \( a_p^* = 0 \). Thus, such a plan sponsor obtains no advantage from the participation plan and a higher reference yield would lead him into a worse situation either way. Therefore, we have the condition \( h_{\pm} \left( \frac{1}{\sqrt{-\hat{\zeta}}} \right) = \hat{r}_0 \) which leads to the solution for the upper boundary of reference yields

\[
\hat{r}_0 = h_{-} (R_p) = r + \frac{\hat{\pi}' \Sigma \hat{\pi}}{2 R_p} \left[ R_p - 1 + \frac{\left( R_p - 2 \right) R_p \hat{\zeta}}{R_p} \right] - 2 \sqrt{\frac{\left( R_p - 2 \right) R_p \hat{\zeta}}{R_p}} \left( 1 + 2 R_p \hat{\zeta} \right)
\]
5.2 The Reference Yield Plan

Figure 5.2 shows that the bigger the difference between the yield guarantee \( \bar{r}_0 \) and the return of the riskless asset, i.e. \( r = 0.02 \) (dashed line), the more the plan sponsors would accept a reference yield \( r_0 \) that clearly exceeds \( \bar{r}_0 \). In the case that the current yield guarantee \( \bar{r}_0 \) is e.g. 4\%, a plan sponsor with a relative risk aversion \( R_p > 4 = -\frac{1}{\zeta} \) (dash-dotted lines) would prefer any reference yield \( r_0 \) to the yield guarantee \( \bar{r}_0 \).

Hence, a change of the setting from a yield guarantee \( \bar{r}_0 \) to the participation model with reference yield \( r_0 > \bar{r}_0 \) is preferred by the more risk averse plan sponsors, but deprecated by the less risk averse plan sponsors. If the relative risk aversion is \( R_p \leq \frac{1}{\sqrt{\zeta}} \) (dotted lines) the plan sponsor will choose \( \alpha_p^* = 0 \), either way. Thus, such a plan sponsor would loose if \( r_0 > \bar{r}_0 \).

![Figure 5.2: Reference yield plan or yield guarantee setting? The plan sponsor’s perspective. Parameters: \( r = 0.02 \), \( \zeta = \bar{\mu}_L = 0 \), \( \mu = 0.1 \) and \( \sigma = 0.2 \) for the risky fund in a two fund framework.](image)

So far we have analyzed the impact of the reference yield on the plan sponsor. Now we take the employee’s perspective.
The expected utility of the accrued retirement assets we have from (5.12)

\[
EU_e \left( X_t^{\pi_t^o(r_0), a_p^o(r_0)} \right) = \frac{X_0^{1-R_s}}{1-R_e} \exp \left[ \frac{(1 - R_e) r_0 t}{1 - R_e} + \alpha_p^o (1 - R_e) \pi^{st} \left( \mu - r1 \right) + r - r_0 - \frac{a_p^o R_e}{2} \pi^{st} \pi^e \right] t.
\]

For comparative statics it suffices to focus on a monotone transformation of \( EU_e \left( X_t^{\pi_t^o(r_0), a_p^o(r_0)} \right) \) e.g. the growth rate

\[
\Gamma_e \left( a_p^o (r_0), x^* (r_0), r_0 \right) \triangleq \frac{1}{(1 - R_e)} \ln \left[ (1 - R_e) EU_e \left( X_t^{\pi_t^o(r_0), a_p^o(r_0)} \right) \right]
= r_0 + \frac{\pi^{st}}{\pi^e} \pi^{st} \left[ a_p^o (r_0) (x^* (r_0) + \zeta) - \frac{a_p^o}{2} x^* (r_0)^2 \right].
\]

(5.17)

Now we substitute the solutions (5.16) into (5.17). Taking the second derivative we get

\[
\frac{d^2 \Gamma_e \left( a_p^o (r_0), x^* (r_0), r_0 \right)}{dr_0^2} = \frac{2 - R_p (R_p + R_e - 1)}{4 \pi^{st} \pi^{st} \left( \sqrt{-\zeta} \right)^3 (R_p - 2)^2} < 0.
\]

Therefore \( \Gamma_e \left( a^o (r_0), a_p^o (r_0), r_0 \right) \) is concave in \( r_0 \) since \( R_p > \sqrt{2} \) by assumption and we have a global maximum of \( \Gamma_e \left( a_p^o (r_0), x^* (r_0), r_0 \right) \).

The employee's optimization problem is

\[
\max_{r_0 \in \mathcal{Y}} \left\{ \frac{2}{R_p} r_0 + \frac{2 a_p^o (r_0)}{R_p} (x^* (r_0) + \zeta) \pi^{st} \pi^{st} - a_p^o (r_0)^2 x^* (r_0) \pi^{st} \pi^{st} \right\}
\]

s.t. \( r_0 \in [r, r + \frac{1}{2} \pi^{st} \pi^{st}] \).

Using theorem 2.10 we get the equivalent mechanism

\[
\max_{r_0} \left\{ \frac{2 a_p^o (r_0)}{R_e} (x^* (r_0) + \zeta) \pi^{st} \pi^{st} - a_p^o (r_0)^2 x^* (r_0) \right\}
\]

s.t. \( r_0 \in [r, r + \frac{1}{2} \pi^{st} \pi^{st}] \)

or equivalently

\[
\max_{r_0} \left\{ a_p^o (r_0), x^* (r_0), r_0 \right\}
\]

s.t. \( r_0 \in [r, r + \frac{1}{2} \pi^{st} \pi^{st}] \)

which we attain by solving the first order condition of \( \frac{d \Gamma_e}{dr_0} = 0 \) for \( r_0 \). We get
5.2 The Reference Yield Plan

\[ r_0^* = r + \frac{\hat{\pi}' \Sigma \hat{\pi} (-2 + R_p (-1 + R_p + R_e))}{(-8 + R_p (4 + R_p R_e))^2} \]

Note that for \( R_p > 2 \) is \( \frac{\partial r^*(R_e)}{\partial R_e} < 0 \). This means that risk averse employees prefer a lower reference yield. The economic interpretation for this result is that a low reference yield implies a lower participation rate set by the plan sponsor which is preferred by risk averse employees.

If \( R_e > \frac{2(R_p-1)}{R_p} \), then \( r_0^* < r + \frac{\hat{\pi}' \Sigma \hat{\pi}}{4} \). Otherwise \( \frac{d\alpha}{dr_0} > 0 \) is in the domain, and the employee prefers the highest possible reference yield \( r_0 = r + \frac{\hat{\pi}' \Sigma \hat{\pi}}{4} \).

Hence, we can set the following contract: the plan sponsor offers for every individual employee \( (\alpha^*_p(r_0), x^*(r_0)) \) and the employee chooses \( r_0 \). If the employee is very risk averse, she has the opportunity to choose a very low \( r_0 \) so that she has to bear little risk, i.e. a small participation rate \( \alpha^*_p(r_0) \). If she chooses \( r_0 \) such that \( R_p < \frac{1}{\sqrt{\zeta}} \) i.e. \( r_0 < r + \frac{\hat{\pi}' \Sigma \hat{\pi}}{R_p} \), she will have a participation rate \( \alpha^*_p = 0 \). Thus, she has the possibility to choose a protection against downside risk. On the other hand, if she chooses \( r_0 \geq r + \frac{\hat{\pi}' \Sigma \hat{\pi}}{4} \) she will face \( \alpha^*_p = 1 \).

Figure 5.3: Reference yield setting: Contracts. Parameters: \( r = 0.02, \zeta = \bar{\mu} = 0, \mu = 0.1 \) and \( \sigma = 0.2 \) for the risky fund in a two fund framework.
In Figure 5.3 we can see these effects. The thin lines are offers of different plan sponsors \((R_p \in \{3, 5, 10, 20\})\) and the dashed lines are indifference curves of the employee with \(R_e = 5\) on the left hand side and \(R_e = 20\) on the right hand side. The bold line is the curve of the concluded contracts for different \(R_p\). For example, on the left side, where \(R_e = 5\), if \(R_p = 10\), then the contract would be set with \((r^*_0, \alpha_p^*(r^*_0)) = (0.31, 0.77)\). We can see that a more risk averse employee will choose a lower reference yield \(r_0\) in order to keep the participation rate low. A problem arises if both the plan sponsor and the employee are highly risk averse. The employee will have to choose a very low reference yield \(r_0\) in this case, as we see in the right plot, for \(R_p = 10\) and \(R_p = 20\) we get \(r^*_0 < \bar{r}_0 = 2.5\%\) (gray points).

Note that since \(x^* (r_0) = \sqrt{-\zeta}\) the lower the reference yield \(r_0\) chosen by the employee, the lower is the risk but also the expected return of the portfolio.

Next we will compare the participation setting with the yield guarantee plan. Thus, we will examine under what conditions the employee will prefer the participation fixed by the plan sponsor to the yield guarantee \(\bar{r}_0\). Thus we substitute \(r^*_0\) into the expected utility \(EU_e \left( X_t x^* (r_0), \alpha_p^* (r_0) \right)\) and compare \(EU_e \left( X_t x^* (\bar{r}_0), \alpha_p^*(\bar{r}_0) \right)\) with the utility of the fixed case \(U_r (e^{\bar{r}_0})\). In particular, we are looking for the conditions for

\[
EU_e \left( X_t x^* (\bar{r}_0), \alpha_p^*(\bar{r}_0) \right) > U_r (e^{\bar{r}_0}) \iff \Gamma_\epsilon \left( \alpha_p^* (r^*_0), x^* (r^*_0), r^*_0 \right) > \bar{r}_0.
\]

Solving this inequality for \(R_e\) we get.

\[
R_e < \hat{R}_e = \frac{7 + 16\zeta - R_p (2 + R_p + 8\zeta)}{2 + 2R_p^2\zeta}.
\]

If we compare this result in Figure 5.4 with the result (5.13) in Figure 5.1, we can now see that the range of \((R_p, R_e)\) that provides a preference for the participation plan is wider now. This is clear since in (5.13) we had \(r_0 = \bar{r}_0\), which is not optimal for the employee. Now we have \(r_0 = r^*_0\), which is optimal for the employee and, therefore, even an employee who is more risk averse prefers the participation fixed by the plan sponsor. But the very risk averse employee still rejects the participation setting if the plan sponsor is also very risk averse. In Figure 5.4 we can see e.g. for \(\bar{r}_0 = 2.5\%\) that, if the plan sponsor’s relative risk aversion exceeds 19, then the employee will not prefer the participation plan to the yield guarantee plan as long as her relative risk aversion exceeds 18.86. In Figure 5.1 this value was 11.87.

However, a notable advantage of this participation rate setting is that almost every employee would prefer it to the autonomous case, i.e. to the Merton solution (3.9). To see
5.2 The Reference Yield Plan

Figure 5.4: Reference yield plan with contract or yield guarantee setting? The employee’s perspective. Parameters: \( r = 0.02 \), \( \zeta = \beta_L = 0 \), \( \mu = 0.1 \) and \( \sigma = 0.2 \) for the risky fund in a two fund framework.
this, we will examine the conditions under which the employee would prefer the participation setting to the autonomous case. We consider the case where the employee would invest her money herself, i.e. $\alpha = 1$. She would invest as described in section 3.1.2. Her optimal portfolio would be $\pi_M = x_M \hat{\pi}$, where $x_M \triangleq \frac{1}{R_e}$.

Define the set $\mathcal{M}$ as

\[
\mathcal{M} \triangleq \left\{ r_0 \left| E U_e \left( X_t^{\pi'(r_0)_{\alpha p}(r_0)} \right) \geq E U_e \left( X_t^{\pi_M_{1}} \right) \right\} \right. \\
= \left\{ r_0 \left| \Gamma_e \left( \alpha_p^+ (r_0) , x^+ (r_0) , r_0 \right) \geq \Gamma_e (1, x_M, 0) \right\} \right. \\
= \left\{ r_0 \left| \Gamma_e \left( \alpha_p^+ (r_0) , x^+ (r_0) , r_0 \right) \geq \frac{\hat{\pi}' \Sigma \hat{\pi}}{2R_e} + r \right\} \\
\]

If $r_0^* \in \mathcal{M}$, then the employee prefers the participation solution to the autonomous solution. The inequality in $\mathcal{M}$ turns out to be quadratic and it can be written in the form

\[
\mathcal{M} = \left\{ r_0 \left| j_1 (R_p, R_e) - \sqrt{h (R_p, R_e)} \geq j_2 (R_p, R_e) \right\} \right. \\
\]

where $j_1 (R_p, R_e), j_2 (R_p, R_e), h (R_p, R_e) \in \mathbb{R}$.

Obviously, if $h (R_p, R_e) > 0$, then $r_0^* \in \mathcal{M}$ for sure, since $\Gamma_e$ is concave in $r_0$ and $r_0^*$ is its global maximum. Thus, in this case we have $\Gamma_e \left( \alpha_p^+ (r_0^*) , x^+ (r_0^*) , r_0^* \right) \geq \Gamma_e (1, x_M, 0)$.

But if $h (R_p, R_e) < 0$, then $\Gamma_e \left( \alpha_p^+ (r_0^*) , x^+ (r_0^*) , r_0^* \right) < \Gamma_e (1, x_M, 0)$ since $\Gamma_e \left( \alpha_p^+ (r_0^*) , x^+ (r_0^*) , r_0^* \right) < \frac{\hat{\pi}' \Sigma \hat{\pi}}{2R_e} + r \forall r_0$. Therefore, we just have to examine the sign of $h (R_p, R_e)$. Calculating $h (R_p, R_e)$ we get

\[
h (R_p, R_e) = 4 (R_p - 2)^2 R_e^3 ( -2 + R_p ( -1 + R_p + R_e) )^2 \\
\cdot (8 + 2R_p (R_e - 2) + R_e (2R_e - 7))
\]

And, therefore, for $R_p > 2$, we get

\[
R_e \geq \bar{R}_e \triangleq \frac{1}{4} \left( 7 - 2R_p + \sqrt{(2R_p - 3) (5 + 2R_p)} \right) \\
\iff h (R_p, R_e) \geq 0.
\]

Note that $\bar{R}_e (R_p)$ is strictly increasing and $\lim \limits_{R_p \to \infty} \bar{R}_e (R_p) = 2$. Therefore, if $R_e > 2$, this condition is always fulfilled.
Hence, if $R_e > 2$, then the employee will prefer the participation to the autonomous investment even though the participation rate and the portfolio strategy is chosen by the plan sponsor. The reason for this result is that the reference yield $r^*_0$ is optimal for the employee and therefore exhibits the employee’s risk preferences.

So far, we have examined the contract $\alpha^*_p (r^*_0)$ from the employee’s perspective. However, Figure 5.5 shows for a variety of parameters under what conditions the plan sponsor would still prefer the participation plan even if the employee can set the reference yield. The thick lines show the optimal $r^*_0$ for $R_e \in \{2, 3, 5, 20\}$ and the thin lines are the same as in Figure 5.2 showing $\hat{r}_0$ for different $\bar{r}_0$. Remember that $\hat{r}_0$ is independent of $R_e$ and $r^*_0$ is independent of $\hat{r}_0$. As we could also see in Figure 5.3, if $R_e = 5$ and $R_p = 10$ we have $r^*_0 = 0.31$ (gray dot). Now Figure 5.5 shows that in this case, the plan sponsor would prefer the participation plan to the yield guarantee plan with a yield of $\bar{r}_0 = 3\%$ as compared to a setting with $\bar{r}_0 = 2.5\%$. In Figure 5.5 we can also see, that the higher the $\hat{r}_0$ or the higher the $R_e$, the more the plan sponsors would prefer the participation framework. In case that $\bar{r}_0 = 2.5\%$ and $R_e = 5$, a plan sponsor with $R_p < 11.2$ would prefer to stay in the yield guarantee model.
Figure 5.5: Reference yield plan with contract or yield guarantee setting? The plan sponsor’s perspective. Parameters: $r = 0.02$, $\varsigma = \mu_L = 0$, $\mu = 0.1$ and $\sigma = 0.2$ for the risky fund in a two fund framework.
5.3 Free Pension Selection

In this section the employee is free to change her pension fund. If, for example, she is very risk averse, she can select a pension fund of a plan sponsor who is less risk averse and provides an $\alpha = 0$. Thus, a risk averse employee has the possibility to receive a yield guarantee $r_0$ which is assumed to be exogenously given and equal for every employee. Therefore, we call $r_0$ a “quasi-guarantee”.

However, our setting with free choice of pension fund implies that there is a sufficiently large supply of different pension funds. From a practical point of view this is crucial. The labour market argument fails, since the pension fund is no longer a component of the firm’s working conditions. Thus, in a setting with $r_0$ higher than the riskless yield and with the liability of the short-fall risk borne by the plan sponsor, few firms would offer such a pension plan voluntarily. Most probably it would have to be mandatory for the firms to offer such a pension plan (or at least only small firms are allowed to join a common pension fund) in order to provide a sufficient supply of funds.

Moreover, for our setting, the supply has to be broad enough to attain optimal results. Thus, we assume that for every $\alpha^*_f, t \in [0, 1]$, which is the optimal $\alpha_i$ for the employee, there exists a pension fund which offers $\alpha^*_f, t$, because $\alpha^*_f, t = \alpha^*_p, t$. Or put differently, there exist enough plan sponsors with different risk preferences.

5.3.1 General Model

The optimization problem for the employee is

$$\max_{\alpha_f, t \in \Psi} E U_e (X_T^f)$$

s.t. $\alpha_f, t \in [0, 1]$

As we have seen before, the portfolio, the plan sponsor invests in is $\pi = \sqrt{-\zeta} \hat{\pi}$, Thus, the employee’s accrued retirement assets evolve as

$$\frac{dX_t}{X_t} = r_0 dt + \alpha_f, t \left( \left( \sqrt{-\zeta} \hat{\pi}' \Sigma \hat{\pi} + r \right) dt + \sqrt{-\zeta} \hat{\pi}' \sigma dZ_t - r_0 dt \right)$$

and the HJB equation (2.9) turns out to be

$$0 = \max_{\alpha_f, t \in H} \left[ l_t + \frac{1}{2} \sigma^2 \frac{\partial^2}{\partial \alpha^2} \left( X^2 \zeta \alpha^2_f, t \hat{\pi}' \Sigma \hat{\pi} \right) - \frac{1}{2} \frac{\partial^2}{\partial \alpha^2} \left( X^2 \zeta \alpha^2_f, t \hat{\pi}' \Sigma \hat{\pi} \right) \right]$$
where \( I(t, X) \) is the employee’s indirect utility function and \( I_t = \frac{\partial I(t, X)}{\partial t}, I_X = \frac{\partial I(t, X)}{\partial X} > 0, I_{XX} = \frac{\partial^2 I(t, X)}{\partial X^2} < 0. \)

The optimal participation rate \( \alpha_{f,t}^* \) turns out to be
\[
\alpha_{f,t}^* = -\frac{I_X}{I_{XX}X} \left( \frac{1}{\sqrt{-\zeta}} - 1 \right)
\]

In order to get a \( \alpha_{f,t}^* \in (0, 1) \), we need that \( r_0 > r + \frac{\hat{\alpha} \Sigma \hat{\alpha}}{1 - \frac{I_{XX}}{I_X} \hat{\alpha}^2} \). If \( r_0 < r + \frac{\hat{\alpha} \Sigma \hat{\alpha}}{1 - \frac{I_{XX}}{I_X} \hat{\alpha}^2} \), we get a corner solution at \( \alpha_{f,t}^* = 1 \). This is the worst case for the employee, as she is restricted in her decisions but bears the whole risk.

Of course, the single employee does not receive the return and does not bear the risk of the portfolio \( \pi \), but receives the return and bears the risk for participation in this portfolio, i.e. \( \alpha \pi \). Let us call this portfolio the employee portfolio \( \pi_e \trianglerighteq \alpha \pi \). The optimal employee portfolio becomes
\[
\pi_{e,t}^* = \alpha_{f,t}^* \pi^* = -\frac{I_X}{I_{XX}X} \left( 1 - \sqrt{-\zeta} \right) \hat{\alpha}
\]
which would become \( -\frac{I_X}{I_{XX}X} \hat{\alpha} \) if \( r_0 = r \). In all the other cases, i.e. \( r_0 > r \), the portfolio for the employee would be less risky than in the autonomous Merton case (3.8). The reason is that the risk averse employee has the possibility to receive a higher risk-free return \( r_0 > r \) and with a less risky portfolio she gets the same expected utility as if \( r_0 = r \).

Thus, we can already state that, in case that \( r_0 > r + \frac{\hat{\alpha} \Sigma \hat{\alpha}}{1 - \frac{I_{XX}}{I_X} \hat{\alpha}^2} \), the employee prefers the participation mechanism which includes the free choice of a pension fund as to both the yield guarantee and the individual setting in the case that \( \alpha_{f,t}^* \in (0, 1) \).

The SDE of the accrued retirement assets becomes
\[
dX = Xr_0 dt - \frac{I_X}{I_{XX}} \left( 1 - \sqrt{-\zeta} \right) \hat{\alpha} \Sigma \hat{\alpha} - \frac{I_X}{I_{XX}} \left( \frac{1}{\sqrt{-\zeta}} - 1 \right) (r - r_0) dt - \frac{I_X}{I_{XX}} \left( 1 - \sqrt{-\zeta} \right) \hat{\alpha} \sigma dZ_t.
\]

### 5.3.2 Special Case: No Net Contributions and CRRA Utility

Again, let us turn to the special case with CRRA utility for the employee, i.e. \( U_e(z) = \frac{z(1 - R_e)}{(1 - R_e)} \) and assume no net contributions, i.e. \( dC_t = 0 \). As in the special case of the previous section, the model parameters are assumed to be deterministic. Using theorem 2.10 we
5.3 Free Pension Selection

get

$$\max_{\alpha_f, t \in Y} EU_c (X_T^\pi) \Longleftrightarrow \max_{\alpha_f} \left\{ \frac{2\alpha_f}{R_e} (x^* (r_0) + \zeta) \hat{\pi}' \Sigma \hat{\pi} \right\}$$

s.t. $\alpha_f, t \in [0, 1]$ \hspace{0.5cm} s.t. $\alpha_f \in [0, 1]$

(5.18)

where $x^* (r_0) = \sqrt{\zeta}$.

Figure 5.6: Optima for the plan sponsor and the employee for constant portfolio. Parameters: $R_e = 5, R_p = 5, r = 0.02, r_0 = 0.03, \zeta = \mu_L = 0, \mu = 0.1$ and $\sigma = 0.2$ for the risky fund in a two fund framework.

This problem is shown in Figure 5.6. The portfolio $x^* (r_0) = \sqrt{-\zeta}$ is constant and independent of any plan sponsor’s preferences. Suppose we have a plan sponsor with $R_p = 5$. He will set $(\alpha_p^*, x^*)$. However, $\alpha_p^*$ is not optimal for the employee. Thus, she will choose another pension fund providing $(\alpha_f^*, x^*)$, as is shown in Figure 5.7. The pension fund she will choose is managed by a plan sponsor with relative risk aversion $R_p = 7$.

Now we face the analytical solution. Solving the first order condition of (5.18) for $\alpha_f^*$ we get

$$\alpha_f^* = \frac{1}{R_e} \left( \frac{1}{\sqrt{-\zeta}} - 1 \right).$$
Figure 5.7: Free pension choice: The choice. Parameters: \( R_e = 5, \) \( R_p = 7, \) \( r = 0.02, \) \( r_0 = 0.03, \) \( \zeta = \bar{\mu}_L = 0, \) \( \mu = 0.1 \) and \( \sigma = 0.2 \) for the risky fund in a two fund framework.
5.3 Free Pension Selection

If \( r_0 < r + \frac{\hat{\alpha}^\pi \hat{\Sigma} \hat{\alpha}}{(1 + R_e)} \), we get the corner solution \( \alpha_f^* = 1 \). To attain an inner solution, i.e. \( \alpha_f^* \in (0, 1) \), the yield \( r_0 \) has to be high enough, i.e. \( r_0 > r + \frac{\hat{\alpha}^\pi \hat{\Sigma} \hat{\alpha}}{(1 + R_e)} \), or put differently, the relative risk aversion of the employee has to be big enough, i.e. \( R_e > \frac{1}{\sqrt{-\zeta}} - 1 \).

The employee portfolio becomes

\[
\pi_{e,f}^* = \alpha_f^* \pi^* = \frac{1}{R_e} \left( 1 - \sqrt{-\zeta} \right) \hat{\pi}.
\]

As we have already seen in the general case for \( r_0 > r \), the optimal portfolio for the employee is less risky than the one in the autonomous Merton case (3.9).

Next, we compare again \( EU_e(X_T^e) \) of this case with the utility from the yield guarantee plan, i.e. \( U_r(X_0 e^{rt}) \).

Substituting the solutions \( (\pi^*, \alpha_f^*) = \left( \sqrt{-\zeta} \hat{\pi}, \frac{1}{R_e} \left( \frac{1}{\sqrt{-\zeta}} - 1 \right) \right) \) into the SDE of the employee’s accrued retirement assets (5.4), ignoring the individual contributions and calculating the expected utility we get

\[
EU_e(X_t) = \frac{X_0^{1-R_e}}{1-R_e} \exp \left[ (1-R_e) r_0 t + \frac{R_e - 1}{2R_e} \left( 1 - \sqrt{-\zeta} \right)^2 \hat{\pi}^\pi \hat{\Sigma} \hat{\pi} t \right]
\]

Thus, not surprisingly in the case of free pension selection, the employees prefer the participation model. This is, of course, true for all kinds of risk averse preference structures, because, in the worst case, the employee can always choose a pension which offers the yield guarantee.

5.3.3 Impact of the Quasi-Guarantee

From the perspective of the plan sponsor the impact of the quasi-guarantee is of course the same as the impact of the reference yield in section 5.2.3 since he solves the same optimization problem. However, the perspective of the employees changes, since they are now free to choose a pension fund.

The optimal solution from the previous section for the employee is

\[
(x^* (r_0), \alpha_f^* (r_0)) = \left( \sqrt{-\zeta} \hat{\pi}, \frac{1}{R_e} \left( \frac{1}{\sqrt{-\zeta}} - 1 \right) \right).
\]
We have
\[ \frac{d^2 \Gamma_\epsilon \left( \alpha^*_f (r_0), x^*(r_0), r_0 \right)}{d r_0^2} = \frac{1}{2 \hat{\alpha}' \Sigma \hat{\alpha} R_e \sqrt{(-\zeta)^3}} \]
and, therefore, \( \Gamma_\epsilon \left( \alpha^*_f (r_0), x^*(r_0), r_0 \right) \) is convex. The first order condition for the minimum is
\[ \frac{d \Gamma_\epsilon \left( \alpha^*_f (h), x^*(h), h \right)}{d h} = 0 \Leftrightarrow h = r + \frac{\hat{\alpha}' \Sigma \hat{\alpha}}{(1 + 2R_e)^2}. \]
As we have seen previously, if \( r_0 < r + \frac{\hat{\alpha}' \Sigma \hat{\alpha}}{(1 + R_e)^2} \) we get a corner solution at \( \alpha^*_f = 1 \). In order to get a \( \alpha^*_f \in (0, 1) \), we need that \( r_0 > r + \frac{\hat{\alpha}' \Sigma \hat{\alpha}}{(1 + R_e)^2} > \frac{\hat{\alpha}' \Sigma \hat{\alpha}}{(1 + 2R_e)^2} \). Thus, the minimum of \( \Gamma_\epsilon \left( \alpha^*_f (h), x^*(h), h \right) \) is not in the domain for an inner solution and, therefore, not of interest. Thus, if \( \alpha^*_f \in (0, 1) \) we have \( \frac{d \Gamma_\epsilon}{d r_0} > 0 \). The employee, therefore, prefers a high quasi guarantee.

Finally, we consider the boundary for \( r_0 \) above which the employee is better off than in the yield guarantee plan, i.e. we compare with \( \alpha = 0 \) and \( r_0 = \bar{r}_0 \). The boundary can be calculated as follows;
\[ \bar{r}_0 = \inf \left\{ r_0 \left| EU_e \left( X^\pi (r_0, x^*(r_0)) \right) \geq U_e (\tilde{\pi}^{\text{guarantee}}) \right. \right\} \]
\[ = \inf \left\{ r_0 \left| \Gamma_\epsilon \left( \alpha^* (r_0), x^*(r_0), r_0 \right) \geq \bar{r}_0 \right. \right\} \]
\[ = \frac{1}{(1 + 2R_e)^2} \left[ \frac{1}{2 \sqrt{2}} \sqrt{\hat{\alpha}' \Sigma \hat{\alpha} R_e \left( \hat{\alpha}' \Sigma \hat{\alpha} - (\bar{r}_0 - r) (1 + 2R_e) \right)} + (1 + 2R_e) r \right]. \]

Next, we consider the case in which the employee would invest her money autonomously, i.e. \( \alpha = 1 \). Her optimal portfolio would be \( \pi_M = x_M \hat{\pi} \). We are looking for the lower boundary \( \bar{r}_{0, M} \) above which the employee would choose the free pension setting, i.e.
\[ \bar{r}_{0, M} \triangleq \inf \left\{ r_0 \left| EU_e \left( X^\pi (r_0, x^*(r_0)) \right) \geq EU_e \left( X^\pi_M (1) \right) \right. \right\} \]
\[ = \inf \left\{ r_0 \left| \Gamma_\epsilon \left( \alpha^*_f (r_0), x^*(r_0), r_0 \right) \geq \Gamma_\epsilon (1, x_M, 0) \right. \right\} \]
\[ = r + \frac{4 \hat{\alpha}' \Sigma \hat{\alpha}}{(1 + 2R_e)^2}. \quad (5.19) \]
Note that \( r + \frac{4 \hat{\alpha}' \Sigma \hat{\alpha}}{(1 + 2R_e)^2} > r + \frac{\hat{\alpha}' \Sigma \hat{\alpha}}{(1 + R_e)^2} \). Thus, \( \bar{r}_{0, M} \) is already binding if \( \alpha^*_f \in (0, 1) \), or put differently, if \( r_0 > \bar{r}_{0, M} \), then we will have \( \alpha^*_f < 1 \).

In Figure 5.8 the boundaries \( \bar{r}_{0, M} \) and \( \bar{r}_0 \) are presented for different \( \bar{r}_0 \).
5.3 Free Pension Selection

We can see that for $r_0 > \bar{r}_0$ and $r_0 > \bar{r}_0$ and $r_0 \leq \bar{r}_0$ the participation plan with free pension choice is a Pareto improvement compared to both the autonomous and the yield guarantee case. If $r_0 > \bar{r}_0$ then the employees’ situation is much better than in the yield guarantee case at the expense of the plan sponsors with low risk aversion.

From (5.19) we can calculate that employees with a risk aversion less than

$$\bar{R}_{e,f} \triangleq -\frac{1}{2} + \frac{1}{\sqrt{-\zeta}}$$

would prefer the autonomous setting. It is $\bar{R}_{e,f} > 2$ for $\zeta > -\frac{4}{25}$. Thus, if at least $\zeta > -\frac{4}{25}$ we have $\bar{R}_{e,f} > \bar{R}_e$, where $\bar{R}_e$ is the lower bound from the reference yield plan in section 5.2.3. We can see this result in Figure 5.8. If $r_0 = 0.025$, then plan sponsors with $R_e < 5.16$ would prefer to invest autonomously. The reason for this result is that, in case of an exogenously given $r_0$, no plan sponsor provides a $(x^*, \alpha^*_p)$ combination that is risky enough for the employee with that low risk aversion.

![Figure 5.8: Free pension choice vs. autonomous investment and yield guarantee setting: Boundaries for $r_0$. Parameters: $r = 0.02$, $\zeta = \mu_L = 0$, $\mu = 0.1$ and $\sigma = 0.2$ for the risky fund in a two fund framework.](image-url)
5.4 Contract Model

In this section we provide a setting within which the plan sponsor is obliged to offer a contract that contains the investment strategy he will set given different participation rates, i.e. $\pi^*_t(\alpha_t)$. In a second step, the employee will choose his optimal participation rate $\alpha^*$. Again, the reference yield $r_0$ is exogenously given. Since the employee is free to choose any participation rate $\alpha \in [0, 1]$, the reference yield is again a quasi-guarantee.

The advantage of this setting over to the free pension fund setting is that we can drop the strong assumption of a sufficiently broad supply of different pension funds. Therefore, in practice the contract model might be easier to implement. However, in the following sections we will see that there are other notable differences.

5.4.1 General Model

In the first step, the plan sponsor optimizes his portfolio given $\alpha$, i.e.

$$\sup_{\pi' \in \mathcal{H}} \liminf_{T \to \infty} \frac{1}{\beta T} \ln \left( \beta E \left[ U_p \left( F^{\pi, a}_T \right) \right] \right)$$

and offers his optimal set of contracts $\pi^*_t(a_t)$.

In the second step, the employee chooses the optimal participation rate $\alpha^*_t$ solving

$$\max_{\alpha_t \in [0,1]} EU_x \left( X^T, a \right)$$

$$\text{s.t. } \alpha_t \in [0,1].$$

We give the same simplification as in the previous sections, i.e.

$$dl_t \triangleq L_t \left( \mu_l dt + \sigma_l dZ^l_t \right),$$

$$dC_t \triangleq W_t \left( \mu_C dt + \sigma_C dZ^n_t \right).$$

The optimal set of portfolios $\pi^*_t(a)$ the pension will offer is then from (5.2)

$$\pi^*_t(a) = \frac{J_F}{2\alpha J_F - (1 - \alpha) J_F F_t} \tilde{\pi}$$

We can therefore write $\pi^*_t$ as a function $\pi^*: \mathbb{R} \to \mathbb{R}^k$ that assigns to each participation rate $\alpha_t$ an optimal portfolio strategy $\alpha_t \mapsto \pi^*_t(\alpha_t)$. This function $\pi^*_t(a)$ is the optimal set the pension will offer to the employee. Alternatively, we can set the contract as $x^*_t: \mathbb{R} \to \mathbb{R}$
that assigns to each participation rate $\alpha_t$ an optimal proportion of the growth optimal portfolio $\hat{\pi}$, i.e. $\alpha_t \mapsto x_t^* (\alpha_t)$.

We have already seen in the previous sections that $\lim_{\alpha \to 1} \pi^* (\alpha) = \frac{1}{2} \hat{\pi}$, which is not quite intuitive. Maybe one would think that for $\alpha \to 1$, the plan sponsor would aim for the growth-optimal portfolio and, therefore, take more risk since he is increasingly outsourcing the risk. But, technically, implicated by the aim to maximize the stochastic ratio process $F_T^\pi$, we get an “Itô term” into the drift.

Roughly speaking, the plan sponsor will be chary of increasing risk for a raising $\alpha$ because risk also appears with a negative sign in the drift of the funding ratio.

The explicit effect can be seen if we rewrite the portfolio strategy

$$\pi_t^* (\alpha) = \frac{J_F}{\alpha (2J_F + J_{FF}F) - J_{FF}F} \hat{\pi}$$

For $2J_F + J_{FF}F > 0$, while raising $\alpha$ the plan sponsor is confronted with a decreasing expected funding ratio which is not compensated for by the decreasing variance. Consequently he will further reduce the risk.

For $2J_F + J_{FF}F < 0$ and while raising $\alpha$, the decreasing expected funding ratio is indeed overcompensated for by the decreasing variance. So the plan sponsor takes more risk for a bigger $\alpha$. However, the compensation still leads to a moderate risk taking.

A plan sponsor with $2J_F + J_{FF}F = 0$ won’t change the portfolio strategy for different $\alpha$.

Henceforth, for simpler notation we declare $\pi^* (1) = \frac{1}{2} \hat{\pi}$.

In the general model the employee would solve the following problem

$$\max_{\alpha_t \in Y} \mathbb{E} U_e \left( X_T^{\pi^* (\alpha), \alpha} \right)$$

s.t. $\alpha_t \in [0, 1]$

with

$$\frac{dX_t}{X_t} = (1 - \alpha_t) r_0 dt + \alpha_t \left( \frac{J_F}{2\alpha_t J_F - (1 - \alpha_t) J_{FF}F} \hat{\pi}' \Sigma \hat{\pi} + r \right) dt$$

$$+ \frac{J_F}{2\alpha_t J_F - (1 - \alpha_t) J_{FF}F} \hat{\pi}' \sigma dZ_t$$

and, thus, the HJB-equation (2.9) becomes

$$0 = \max_{\alpha_t \in Y} \left\{ I_t + I_X X \left( (1 - \alpha_t) r_0 + \alpha_t \left( \frac{J_F}{2\alpha_t J_F - (1 - \alpha_t) J_{FF}F} \hat{\pi}' \Sigma \hat{\pi} + r \right) \right) \right\}$$

Now, let us turn to the well known special case.
5.4.2 Special Case: No Net Contributions and CRRA Utility

Again, we assume CRRA utility, i.e. \( U_p(z) = \frac{1}{1-R_p} z^{\left(1-R_p\right)} \) but this time with a more general relative risk aversion \( R_p > 1 \). The same as in the utility function of the plan sponsor, the employee is assumed to exhibit a constant relative risk aversion \( R_e > 1 \).

Further, we assume again that there are no net contributions, i.e. \( dC_t = 0 \). As opposed to the special cases in the previous sections we state no lower boundary for \( \zeta < 0 \).

The model parameters \((\mu, \sigma, r, \mu_L, \sigma_L)\) are assumed to be deterministic. The optimization problem of step one is

\[
\sup_{\pi \in G} \liminf_{T \to \infty} \frac{1}{(1-R_p) T} \ln \left( (1-R_p) E \left[ U_p(F_T^\pi) \right] \right)
\]

where \( \alpha_t \) is exogenously given and assumed to be deterministic for the moment. Then using theorem 2.10 and 2.9 the optimization problem turns out to be

\[
\max_{\pi} \left\{ \frac{2 (1-\alpha_t)}{R_p} \left( (\pi' (\mu - r1) + r - r_0) - \alpha_t \pi' \Sigma \pi \right) - (1-\alpha_t) 2 \pi' \Sigma \pi \right\}.
\]

Solving the first order condition of (5.20) for \( \pi^* \) we get the set of offers

\[
\pi^* (\alpha) = \frac{1}{2\alpha + (1-\alpha)R_p} \hat{\pi}.
\]

Figure 5.9 shows the offer sets for different \( R_p \). As we have already discussed in the previous section, the \( \lim_{\alpha \to 1} \pi^* (\alpha) = \frac{1}{2} \hat{\pi} \). Thus, the most risky employee portfolio for \( \pi_e \) is \( \frac{1}{2} \hat{\pi} \). Therefore, if \( R_e < 2 \) the employee would invest more risky in the autonomous case than he is able to in the contract case. Consequently the employees with \( R_e < 2 \) would prefer to invest autonomously.

A plan sponsor with \( R_p < 2 \) would further reduce the risk as \( \alpha \) increases. For this plan sponsor the decreasing expected funding ratio is not compensated for by the decreasing variance.

The optimization problem of step two is

\[
\max_{\alpha_t \in [0,1]} EU\left( X_T^{\pi^*(\alpha)} \right).
\]

Using once more theorem 2.10 and (5.21) the optimization problem turns out to be

\[
\max_{\alpha \in [0,1]} \left\{ \frac{2\hat{\alpha} \Sigma \hat{\pi} \left( \frac{1}{2\alpha + (1-\alpha)R_p} + \zeta \right)}{R_e} - \hat{\pi}' \Sigma \hat{\pi} \left( \frac{\alpha}{2\alpha + (1-\alpha)R_p} \right)^2 \right\}.
\]
Figure 5.9: Contract setting: Offers.

Since $\pi' \Sigma \hat{\pi} > 0$, this optimization problem is equivalent to

$$\max_{\alpha \in [0,1]} \Theta(\alpha) \quad \text{(P)}$$

where $\Theta(\alpha) \triangleq \frac{2a}{R_e} \left( \frac{1}{2a+(1-a)R_p} + \zeta \right) - \left( \frac{a}{2a+(1-a)R_p} \right)^2$.

Figure 5.10 shows this interaction problem for two different sets of relative risk aversions. The employee chooses the contract on the offer line in which his indifference curve is tangential to the offer line.

Before we face the general problem, we consider the simplification $r_0 = r \Leftrightarrow \zeta = 0$. In this case we have

$$\Theta(\alpha) = \frac{\alpha}{2a+(1-a)R_p} \left( \frac{2}{R_e} - \frac{1}{2a+(1-a)R_p} \right).$$

Solving the first order condition for $\alpha^*$ we get

$$\alpha^* = \frac{R_p}{R_p + R_e - 2}. \quad \text{(5.22)}$$

Thus, for $R_p > 1$ and $R_e \geq 2$ the constraint $\alpha^* \in [0,1]$ is not binding. However, for $R_e < 2$ the constraint is binding and it is $\alpha^* = 1$. 
Figure 5.10: Contract setting: Contract. Parameters: $r = 0.02$, $r_0 = 0.03$, $\zeta = \mu_L = 0$, $\mu = 0.1$ and $\sigma = 0.2$ for the risky fund in a two fund framework.
Firstly, let us turn to the case $R_e \geq 2$. The optimal portfolio strategy the plan sponsor chooses if the employee takes $\alpha^*$ becomes

$$\pi^* (\alpha^*) = \frac{1}{2\alpha^* + (1 - \alpha^*)} \pi = \frac{R_p + R_e - 2}{R_p R_e} \pi.$$

The employee portfolio becomes

$$\pi_e^* = \frac{R_p}{R_p + R_e - 2} \times \frac{R_p + R_e - 2}{R_p R_e} \pi = \frac{1}{R_e} \pi$$

which is the Merton portfolio (3.9). Thus the employee with $R_e \geq 2$ faces the same portfolio strategy she would choose autonomously.

For $R_e < 2$ and, therefore, $\alpha^* = 1$ the portfolio strategy becomes a corner solution

$$\pi^* (1) = \pi^* = \frac{1}{2} \pi < \frac{1}{R_e} \pi.$$

Now we turn to the start problem (P) when $r_0 > r \iff \zeta < 0$. The first order condition $\frac{\partial \Theta(\alpha)}{\partial \alpha} = 0$ turns out to be the cubic equation

$$-\zeta (R_p (1 - \alpha) + 2\alpha)^3 + R_p (\alpha (R_e + R_p - 2) - R_p) = 0.$$ 

Since this cubic equation with three parameters has one, two or three solutions and, additionally, since we claim that $\alpha \in [0, 1]$, it is not trivial to find the closed form solution to our optimization problem. There are many case differentiations. In the following proposition we provide the explicit solution to the problem (P).

**Proposition 5.1** *Explicit solution to the contract case:*

Define

$$y_0 (R_p) \triangleq -2 \left[ \frac{R_p + 2 (R_p - 2) R_p \zeta}{2 (-1 + (R_p - 2) \sqrt{R_p \zeta (1 + R_p \zeta)})} \right]$$

and

$$y_1 (R_p) \triangleq \frac{8 \zeta}{R_p} + 2$$

and

$$\tilde{\alpha}_3 \triangleq \sqrt{\frac{4 R_p (R_e + R_p - 2)}{3 (R_p - 2)^3 \zeta} \cos \left( \frac{\phi}{3} + \frac{4 \pi}{3} \right) + \frac{R_p}{R_p - 2}}$$
\[ \phi = \arccos \sqrt{\frac{-27 (R_p - 2) R_p R_p^2}{4 (R_p + R_p - 2)^3}} \] and finally
\[ \hat{\alpha} = \frac{3}{2} \sqrt{\frac{-R_p^2 R_p - \sqrt{Q} + \frac{3}{2} \sqrt{-R_p^2 R_p + \sqrt{Q} + \frac{R_p}{R_p - 2}}}{2 (R_p - 2)^4 \zeta}} \]

where \( Q = \frac{R_p^3 (4 (R_p + R_p - 2)^3 + 27 (R_p - 2) R_p R_p^2)}{108 (R_p - 2)^2 \zeta^3} \).

A: Set \(-1 < \zeta < 0\) and \(R_p > 2\) then
i) For \(R_p < -\frac{1}{\zeta}\) it is \(R_e > y_1 (R_p) \Leftrightarrow \alpha^* = \hat{\alpha}_3\) and \(R_e \leq y_1 (R_p) \Leftrightarrow \alpha^* = 1\)

ii) For \(\zeta < -\frac{1}{R_p} \wedge R_p < \frac{4 e}{e+1}\) or \(\zeta \geq -\frac{1}{R_p} \wedge R_p \geq -\frac{1}{\zeta}\) it is \(R_e > y_0 (R_p) \lor (R_p, R_e) = B \Leftrightarrow \alpha^* = 0, R_e = y_0 (R_p) \lor R_p > -\frac{1}{\zeta} \Leftrightarrow \alpha^*\) not unique, \(R_e \leq y_1 (R_p) \Leftrightarrow \alpha^* = 1\) and finally \(y_1 (R_p) < R_e < y_0 (R_p) \Leftrightarrow \alpha^* = \hat{\alpha}_3\)

iii) For \(\zeta < -\frac{1}{R_p} \wedge R_p \geq \frac{4 e}{e+1}\) it is \(R_e > 8 \zeta + 4 \Leftrightarrow \alpha^* = 0\) and \(R_e < 8 \zeta + 4 \Leftrightarrow \alpha^* = 1\) and \(R_e = 8 \zeta + 4 \Leftrightarrow \alpha^*\) not unique

B: Set \(-1 < \zeta < 0\) and \(R_p < 2\) then
i) For \(R_p < -\frac{1}{\zeta}\) it is \(R_e > y_1 (R_p) \Leftrightarrow \alpha^* = \hat{\alpha}\) and \(R_e \leq y_1 (R_p) \Leftrightarrow \alpha^* = 1\)

ii) \(R_p \geq -\frac{1}{\zeta} \Rightarrow \alpha^* = 0\)

C: Set \(-1 < \zeta < 0\) and \(R_p = 2\) then \(\alpha^* = \max \left[ \min \left[ \frac{2 (1+2 \zeta)}{R_e}, 1 \right], 0 \right] \)

D: Set \(\zeta \leq -1\) then \(\alpha^* = 0\)

In appendix 5.A.1 we provide the tedious proof of proposition 5.1.

**Remark 5.2** Directly from proposition 5.1 it follows that for the most usual case \(-\frac{1}{4} < \zeta < 0\) (i.e. \(r_0\) is not “too high”) we have
\(R_e \leq y_1 (R_p) \Leftrightarrow \alpha^* = 1,\)
\(R_p \geq -\frac{1}{\zeta} \wedge R_e > y_0 (R_p) \lor (R_p, R_e) = B \Leftrightarrow \alpha^* = 0,\)
\(R_p > -\frac{1}{\zeta} \wedge R_e = y_0 (R_p) \Leftrightarrow \alpha^*\) not unique,
and otherwise: \(\alpha^* = \hat{\alpha}_3\) for \(R_p > 2, \alpha^* = \hat{\alpha}\) for \(R_p < 2\) and \(\alpha^* = \max \left[ \min \left[ \frac{2 (1+2 \zeta)}{R_e}, 1 \right], 0 \right]\) for \(R_p = 2\).

Figure 5.11 shows two plots of \(\alpha^*\) for different \(r_0\). In the upper plot we have \(r_0 = 0.03\) and in the lower plot is \(r_0 = 0.035\). The latter case is once more presented in Figure 5.12 as a 3D-plot for a better view. The black thick lines in the plots of Figure 5.11 are \(y_0 (R_p)\) above which, in the white area, it is \(\alpha^* = 0\) and on which \(\alpha^*\) is not unique. On and below the white line, \(y_1 (R_p)\) it is \(\alpha^* = 1\) and finally the shades exhibit the inner solution.
\( \alpha^* \in (0, 1) \). Note that the 0.5% change of \( r_0 \) in Figure 5.11 increased the \( \alpha^* = 0 \)-area substantially. Thus, a small raise of \( r_0 \) causes a strongly increasing interest in the contract \( \alpha^* = 0 \).

Directly from Proposition 5.1 follow several implications which can be observed in Figure 5.11 and also in Figure 5.12.

If for \( \tilde{R}_c, \tilde{R}_p \) it is \( \alpha^* (\tilde{R}_c, \tilde{R}_p) = 0 \) then \( \alpha^* (\tilde{R}_c + a, \tilde{R}_p + b) = 0 \) \( \forall a \geq 0, b \geq 0 \). The economic interpretation of this fact is that the employee is not interested in a participation in \( (\tilde{R}_c, \tilde{R}_p) \). If she was even more risk averse, i.e. \( \tilde{R}_c + a \), she would of course not change her mind. However, if the plan sponsor lowers the investment risk, i.e. \( \tilde{R}_p + b \), the expected return will decrease which apparently is overwhelming the effect of the lower risk which, in turn, implies that the employee would rather participate.

If for \( \tilde{R}_c, \tilde{R}_p \) we have \( \alpha^* (\tilde{R}_c, \tilde{R}_p) = 1 \), then \( \alpha^* (\tilde{R}_c - a, \tilde{R}_p + b) = 1 \) \( \forall a \geq 0, b \geq 0 \) and \( \tilde{R}_c - a > 1 \). In case that \( \alpha^* (\tilde{R}_c, \tilde{R}_p) = 1 \) the less risk averse employee, of course, would still fully participate. However, if the plan sponsor lowers the investment risk i.e. \( \tilde{R}_p + b \), this time the effect of the lower risk is preponderate to the lower expected return.

Also interesting is the effect that if for \( \tilde{R}_c, \tilde{R}_p \) we have \( \alpha^* (\tilde{R}_c, \tilde{R}_p) \in (0, 1) \), then \( \alpha^* (\tilde{R}_c, \tilde{R}_p - b) \in (0, 1) \) \( \forall \tilde{R}_p - 1 > b > 0 \). Thus, an inner solution will never become a corner solution if \( R_p \) decreases. In such a situation neither the higher risk nor the higher expected return is preponderate to the extent that we get a corner solution.

It is intuitively obvious that in general the employee will choose a lower participation rate if she has a higher risk aversion. This issue is stated in the next corollary.

**Corollary 5.3** \( \alpha^* \in (0, 1) \) \( \Rightarrow \frac{\partial \alpha^*}{\partial \tilde{R}_c} < 0 \).

The proof of corollary 5.3 is given in appendix 5.A.2.

So far we have deduced the closed form solution of our optimization problem and discussed some implications. Next we will show under which assumptions this setting is preferred to the fixed yield guarantee \( r_0 \). Consequently we will compare with the case where \( \alpha = 0 \). Of course the employees are always better off with the participation plan than with the fixed yield plan since they are free to choose the case \( \alpha = 0 \).

But let us state the cases for which the plan sponsor prefers the participation plan. This is done in the next proposition.
Figure 5.11: Contract setting: Solutions. Parameters: \( r = 0.02, \zeta = \bar{\mu_L} = 0, \mu = 0.1 \) and \( \sigma = 0.2 \) for the risky fund in a two fund framework.
Figure 5.12: Contract setting: Solutions3D. Parameters: $r = 0.02, r_0 = 0.035, \zeta = \bar{\rho}_L = 0, \mu = 0.1$ and $\sigma = 0.2$ for the risky fund in a two fund framework.
Proposition 5.4 Set $R_p \neq 2$ and $\zeta > -\frac{1}{4}$ then

$R_p \leq \frac{1}{\sqrt{-\zeta}} \Rightarrow$ no Pareto improvement is possible and the plan sponsor would prefer the setting with $\alpha = 0$.

$\frac{1}{\sqrt{-\zeta}} < R_p < -\frac{1}{2\zeta} \Rightarrow$ Pareto improvement if $R_e > R_{e,\text{Pareto}} \triangleq -\frac{(R_p - 2)(1 + R_{e,\zeta})}{R_{e,\zeta}^2(1 + R_{e,\zeta})}$; no Pareto improvement otherwise.

$R_p \geq -\frac{1}{2\zeta} \Rightarrow$ we have always a Pareto improvement

The proof of proposition 5.4 is given in appendix 5.A.3.

Directly from proposition 5.4 it follows that, if the plan sponsor is less risk averse in the status quo, i.e. $R_p \leq \frac{1}{\sqrt{-\zeta}}$ (dashed lines), then the plan sponsor never wants to change the system. Otherwise if $R_p \geq -\frac{1}{2\zeta}$, which is equal to $r_0 \geq \frac{\kappa\bar{\mu}\bar{\pi}}{2R_p} + r$, than independent of $R_e$ the contract plan implies a Pareto improvement.

In Figure 5.13 the boundary $R_{e,\text{Pareto}}$ is plotted for different different $r_0$.

Figure 5.13: Contract setting: Pareto boundaries $R_{e,\text{Pareto}}$ Parameters: $r = 0.02$, $\zeta = \bar{\mu} = 0$, $\mu = 0.1$ and $\sigma = 0.2$ for the risky fund in a two fund framework.

We can see that, e.g. for $r_0 = 0.025$, if the plan sponsor has a relative risk aversion of less then $R_p = \frac{1}{\sqrt{-\zeta}} = 5.65$, then he will always prefer the yield guarantee setting. Another plan sponsor with $R_p = 8$ will prefer to stay in the yield guarantee plan with $r_0 = 0.025$
as long as the employee is less risk averse than \( R_{e, \text{Pareto}} = 5.65 \), (dotted lines) because the plan sponsor would like to bear more risk than the employee allows him. In Figure 5.14 we can see this situation more precisely. The outer thick line is the plan sponsor’s indifference curve which includes the situation of the yield guarantee plan \((0, x_{M, p})\), where \( x_{M, p} \) is the Merton portfolio (3.9) of the plan sponsor. The contract \((a^*, x^*(a^*))\) is on a lower utility level than the point \((0, x_{M, p})\). Thus, the plan sponsor prefers the yield guarantee plan.

\[
\begin{align*}
\text{Figure 5.14:} & \quad \text{Example for parameters where the contract setting is worse than the yield}\nonumber \\
& \quad \text{guarantee setting. Parameters: } R_e = 5, R_p = 8, r = 0.02, r_0 = 0.025, \zeta = \mu_L = 0, \mu = 0.1 \text{ and } \sigma = 0.2 \text{ for the risky fund in a two fund framework.}
\end{align*}
\]

Hence, we see under which conditions the contract setting is a Pareto improvement to the fixed yield guarantee. However, in general we get a contract that is not Pareto optimal because there is hidden information; the plan sponsor does not know the employee’s risk preferences ex ante. The contract implies a second best solution.\(^4\) To formalize this statement we give the following proposition.

\(^4\)Note that, as to \( a \) and \( x \), in the reference yield setting and in the free pension choice setting we were in a Pareto optimal situation. In particular, we were in the plan sponsor’s optimum. However, in the reference yield setting, as to the reference yield, we had a second best solution implied by a contract.
Proposition 5.5 $\alpha^* \in (0,1)$ is not Pareto optimal if $\alpha^* \neq \alpha_{\text{Po pt}} \triangleq \frac{R_p - 1}{R_p - 2}$ (which is the solution of the CRRA case in section 5.2) and it is $\alpha^* < \frac{R_p}{R_p + R_e - 2}$ (which is the solution of the case if $\zeta = 0$, i.e. (5.22))

The proof of proposition 5.5 is given in appendix 5.A.4.

In Figure 5.15 this issue is shown for a numerical example. Between the two indifference curves which intersect at the contract point (gray area) both the plan sponsor and the employee are better off than in the contract point $(\alpha^*, x^*(\alpha^*))$. The set of the Pareto optimal points, i.e. where no further Pareto improvement is possible, is called the Pareto set. Thus, the points on the Pareto set curve show the first best solutions of the general problem. The thick part of the Pareto set exhibits the first best solutions which are attainable from the contract point $(\alpha^*, x^*(\alpha^*))$. Let us assume that it is possible to renegotiate the contract. After the contract is settled, both the plan sponsor and the employee are informed about each other’s risk preferences. Thus, there are two possibilities for the solution after the renegotiation.

In the first case it is the plan sponsor who has the opportunity to offer an alternative contract. He will offer the contract in $P$. The employee will not refuse this offer and the new contract ameliorates the plan sponsor’s condition. In Figure 5.14 we see that in this case the plan sponsor would reach a better situation than in the yield guarantee setting if he could offer a new contract. Thus, in the case where the plan sponsor has the opportunity to renegotiate the contract the Pareto boundaries in Figure 5.13 would move to the left.

In the second case, if the employee is allowed to offer an alternative contract, he will choose $E$.

From proposition 5.5 we know that $\alpha^* < \frac{R_p}{R_p + R_e - 2}$. Thus, in the contract case $\alpha^*$ is always smaller than in the case where $\zeta = 0$. This is intuitively obvious since for $\zeta > 0$ a slightly lower $\alpha$ than in the case where $\zeta = 0$ leads to the same expected return but less risk.

In the next section we will study the impact of the quasi-guarantee on the contract plan.

5.4.3 Impact of the Quasi-Guarantee

The solution of the contract setting is

$$(x^* (r_0), \alpha^* (r_0)) = \left( \frac{1}{2\tilde{\alpha}_3 + (1 - \tilde{\alpha}_3) R_p}, \tilde{\alpha}_3 \right)$$
Figure 5.15: Contract setting: Pareto optimality. Parameters: $R_e = 3.5$, $R_p = 6$, $r = 0.02$, $r_0 = 0.03$, $\zeta = \bar{\mu}_L = 0$, $\mu = 0.1$ and $\sigma = 0.2$ for the risky fund in a two fund framework.
To avoid a lot of case differentiations, we assume that $R_p > 2$

From the employee’s perspective, the higher the quasi-guarantee $r_0$ the better. Using the envelope-theorem directly on the expected utility we can easily prove this point:

$$d\text{EU}_e \left( X_t^{\pi^*(r_0),x^*(r_0)} \right) dr_0 = q \left( 1 - \alpha^*(r_0) \right) t \exp \left[ g(r_0) t \right] > 0$$

where

$$g(r_0) \triangleq (1 - R_\alpha) (1 - \alpha^*(r_0)) r_0 + (1 - R_\alpha) \alpha^*(r_0) \left[ \pi^*(r_0)'(\mu - r1) + \frac{r}{2} \right]$$

and $q \triangleq X_0^{(1-R_c)} (1 - R_c)^2 > 0$.

As in the previous settings, we consider the lower boundary for $\bar{r}_0$ for which an employee with relative risk aversion $R_\alpha$ prefers the contract setting. Thus, we are looking for

$$\bar{r}_0 = \inf \left\{ r_0 < r + \frac{\hat{\pi}'\Sigma \hat{\pi}}{4} \left| \text{EU}_e \left( X_t^{\pi^*(r_0),x^*(r_0)} \right) \geq U_e (e^{\bar{r}_0}) \right. \right\}$$

$$\bar{r}_0 = \inf \left\{ r_0 < r + \frac{\hat{\pi}'\Sigma \hat{\pi}}{4} \left| \Gamma_e (\alpha^*(r_0), x^*(r_0), r_0) \geq \bar{r}_0 \right. \right\}.$$

The inequation $\Gamma_e (\alpha^*(r_0), x^*(r_0), r_0) \geq \bar{r}_0$ has to be solved numerically.

The more interesting issue is the plan sponsor’s view. The impact of the quasi-guarantee on the optimal long term growth rate can be calculated as in (5.15)

$$\frac{d\Gamma_e (\alpha^*(r_0), x^*(r_0), r_0)}{dr_0} = -1 + \alpha^*(r_0)$$

$$+ \left( r_0 - r - \frac{\hat{\pi}'\Sigma \hat{\pi}}{4} (R_p - \alpha^*(r_0) (R_p - 2))^2 \right) \frac{d\alpha^*}{dr_0}.$$

However, since $\alpha^*$ is not optimal from the plan sponsor’s view, we cannot apply the envelope-theorem. We state the following lemma:

**Lemma 5.6** If $R_p > 2$, then $\alpha^* \in (0, 1) \Rightarrow \frac{d\alpha^*}{dr_0} > 0$

**Proof.** $\alpha^* \in (0, 1) \Rightarrow \zeta > -\frac{4(R_\alpha + R_p - 2)^3}{27(R_p - 2) R_\alpha R_p^2}$ from lemma 5.9 in the appendix 5.A.1. Define

$$h (\zeta) \triangleq \arccos \sqrt{\frac{-27(R_p - 2) R_\alpha R_p^2 \zeta^3}{4(R_\alpha + R_p - 2)^3}}. \text{ We have } h' (\zeta) > 0 \text{ and } \frac{\partial \cos \left( \frac{3\theta + 12\pi}{4} \right)}{\partial \zeta} > 0. \text{ Thus, } \frac{d\alpha^*}{dr_0} > 0.$$

$\blacksquare$
Because of Lemma 5.6, we know that for \( R_p > 2 \) we have \( \alpha^* \in (0, 1) \Rightarrow \frac{d\alpha^*}{d\alpha} > 0 \) and, therefore, \( \frac{d\alpha^*}{dr_0} < 0 \). From lemma 5.3 we know further that \( \frac{d\alpha^*}{dR_e} < 0 \). So if \( R_e \) and \( R_p \) are high enough, then \( \frac{d\Gamma_p(\alpha^*(r_0), x^*(r_0), r_0)}{dr_0} < 0 \). But if \( R_e \) is “very low” and \( R_p \) not “too high”, then we might have an optimal quasi-guarantee from the plan sponsor’s perspective.

An economic interpretation of these results is the following:

If the relative risk aversion of the employee is high, she would choose a low \( \alpha^* \) from the beginning. Thus, in this case the plan sponsor needs to have a very low \( R_p \) if he still wants the employee to further reduce \( \alpha^* \). Since, additionally, in this case the plan sponsor’s risk increases greatly if \( r_0 \) is raised he would only like a higher \( r_0 \) if his \( R_p \) is below the assumed values. Thus, in this case, it is intuitive that \( \frac{d\Gamma_p(\alpha^*(r_0), x^*(r_0), r_0)}{dr_0} < 0 \).

However, if the relative risk aversion of both the plan sponsor and the employee are low, both would like to bear a substantial portion of the risk because in their view the additional return justifies the risk. Since the employee is allowed to choose, she will choose a big \( \alpha^* \) and only a substantially high \( r_0 \) could make this employee lower \( \alpha^* \). If the plan sponsor is willing to accept a higher \( r_0 \) because \( r_0 \) is “very low” and \( R_p \) not “too high”, then we might have an optimal quasi-guarantee from the plan sponsor’s perspective.

In Figure 5.16 some \( \Gamma_p \left( \alpha^*_p \left( r_0 \right), x^* \left( r_0 \right), r_0 \right) \) are plotted. In the upper plot there are some numerical examples for the case with an optimal \( r^*_0 \) and in the lower case there are some numerical examples for the case \( \frac{d\Gamma_p(\alpha^*_p(r_0), x^*(r_0), r_0)}{dr_0} < 0 \). Note that the relative risk aversion of the employee has to be rather small to receive an optimal \( r^*_0 \). In the upper plot we can see, that if \( R_e = 2.5 \), then the difference between \( r^*_0 \) and \( r \) is rather small. In case \( R_e = 1.5 \) and \( R_p = 5 \) we have for in the whole plotted domain \( \frac{d\Gamma_p(\alpha^*_p(r_0), x^*(r_0), r_0)}{dr_0} > 0 \). In case the relative risk aversion of the plan sponsor is high enough and since he does not bear too much risk anyway as \( R_e \) is low, e.g. the line with \( R_e = 2 \) and \( R_p = 10 \), the plan sponsor is nearly indifferent to the yield \( r_0 \). Besides, the jumps in the lower plot are caused by the crossing from \( \alpha^* \in (0, 1) \) to \( \alpha^* = 0 \) (compare with Figure 5.12).

Taking everything into account, if the plan sponsor and the employee are barely risk averse, then there could be an optimal \( r^*_0 \).

As we have seen in the previous section, compared with the yield guarantee plan, the less risk averse plan sponsor’s situation gets worse if \( r_0 = \bar{r}_0 \). The interesting point is that if, additionally, the \( r_0 \) increases and most of the employees also are barely risk averse, then this worsening could be compensated because the plan sponsor has a low risk aversion.
Figure 5.16: Contract setting: Possible \( r_0' \). Parameters: \( r = 0.02, \xi = \bar{\mu}_L = 0, \mu = 0.1 \) and \( \sigma = 0.2 \) for the risky fund in a two fund framework.
However, as Figure 5.16 illustrates, it is $R_e$ rather than $R_p$ which determines whether there is an optimal quasi-guarantee from the plan sponsor’s perspective. That value $R_e$ has to be very small. Otherwise the plan sponsor prefers a lower $r_0$ and the employee a higher $r_0$. Thus, if the employees’ risk aversions are in general moderate or high, then the Pareto frontier shown in Figure 5.13 will move to the right for increasing $r_0$. 
5.5 Conclusions

In the last chapter we presented three settings including a risk sharing between the employee and the plan sponsor. The first model included a contract \((\pi^* (r_0), \alpha^* (r_0))\), the second a free choice of pension fund and the third a contract \(\pi^* (\alpha)\). All settings provide a possibility for the employee to realize her individual risk preferences. In setting one the employee can choose her best reference yield, in setting two her best participation rate while choosing a pension fund and in setting three again her best participation rate but this time from a contract offered by “her” plan sponsor. This opportunity makes the settings attractive to almost all employees compared with the autonomous investment, i.e. the Merton solution. In settings one and three, if \(r_0 > r\), every employee would prefer these settings to the autonomous investment except those with very low risk aversion, i.e. \(R_p < 2\) in the CRRA case. In setting two, employees with moderate risk aversion, i.e. \(R_p < -\frac{1}{2} + \frac{1}{\sqrt{-\zeta}}\) in the CRRA case, also prefer to invest autonomously.

Summarizing, there is still an intra-generational redistribution of risk at the expense of the employees with (very) low risk aversion but, because of the risk-sharing, most employees benefit from the risk-sharing models compared to the autonomous investment. Additionally, there is an inter-generational redistribution of risk concerning those employees who choose \(\alpha^* < 1\).

Note that because none of these settings uses any assumptions about the individual contribution flow, in an even more individualized implementation the employee could simultaneously optimize her contribution flow. This framework would include all aspects of individual optimization in a life-cycle model. Moreover, all three frameworks provide a possibility for the employee to protect herself from a downside risk. Model one provides the least protection. There it is possible for the employee to choose a low reference yield, i.e. \(r_0 < r + \frac{\hat{\pi} \hat{\Sigma} \hat{\pi}^2}{R_p^2}\) in the CRRA case, such that the plan sponsor agrees to bear the whole risk and the reference yield becomes a yield guarantee. But if the plan sponsor’s risk aversion is high, this yield guarantee is low. Thus, compared with a yield guarantee plan, the situation of the more risk averse employees gets worse. However, the less risk averse employees would prefer the participation plan to the yield guarantee plan since they expect a higher return.

In settings two and three the reference yield is exogenously given. For all employees it is possible to choose a pension fund or a contract, respectively, such that the reference yield becomes a yield guarantee (at least) equal to the yield guarantee setting. Thus we
still face an inter-generational redistribution of risk. However, because of the additional opportunities, these two settings would be preferred to the yield guarantee framework by all employees.

From the plan sponsor’s perspective, the change from the yield guarantee plan to the first participation plan, i.e. the reference yield contract, is beneficial if either the yield guarantee is high or the employee has a high risk aversion. In the other cases the employee will choose a high reference yield $r_0^*$ in order to get a high participation rate $\alpha^*(r_0^*)$. Then the change would be at the expense of the less risk averse plan sponsors.

Keeping the reference yield equal, the change from the yield guarantee plan to setting two, i.e. the free pension choice setting, is an improvement for all kinds of plan sponsors. The reason for this is that the plan sponsor has the opportunity to choose the status quo, i.e. zero participation rate $\alpha$.

In setting three, i.e. the contract setting, the change comes at the expense of the plan sponsors with lower risk aversion. They would like to bear more risk than the employees allow. The more risk averse plan sponsors prefer the contract setting. However, as we have seen, if we allow to renegotiate the contract, then the situation becomes better for the participant who has the opportunity to offer a new contract. Thus, if the plan sponsor is allowed to offer an alternative contract, then he might prefer the contract setting to the yield guarantee plan even if his risk aversion is low.

If we change from the yield guarantee plan to the risk-sharing settings, what happens if we additionally raise the yield $r_0$? In setting two the plan sponsors who are less risk averse are worse off than before, while the more risk averse would still prefer the participation plans. In setting three the effect on the less risk averse plan sponsors is not clear. It depends on the risk attitude of the employees. If they are also less risk averse, a higher quasi-guarantee $r_0$ can even improve the plan sponsor’s situation. However, in general, as in the free pension setting, the more risk averse plan sponsors still benefit from the change while the less risk averse lose.

Taking everything into account, if we compare the three settings with both the yield guarantee plan and the autonomous investment, all three settings have clear advantages for most participants under the assumptions we made. The shortcomings can be summarized as follows: The reference yield plan has its comparative shortcoming for the more risk averse employees and in some circumstances, i.e. low $r_0$ or low $R_e$, also for the less risk averse plan sponsors. The free pension setting has its comparative shortcoming for the less risk averse employees and the contract plan for the less risk averse plan sponsors.
However, besides the direct implications concerning the economic well-being of the participants there are other effects of the participation plans. In the contract setting, the employee has to choose a single variable and the investment policy of the plan sponsor is specified through a contract. Thus, even if the structure of the contract is more complicated than in the examined CRRA case, it is possible to monitor the behavior of the plan sponsor. If he violates the contract, he is personally liable. Thus, implicitly, transparency is a part of the contract. In the reference yield plan, the monitoring is harder, since there are more components in the contract, but the argument is the same. Additionally, in this framework, the investment policy of the plan sponsor is constant and similar for all plan sponsors, i.e. \( \pi_t^* = \sqrt{-\zeta} \hat{\pi} \) (assuming univariate funding ratio process, untradeable net contributions, and untradeable liability). This simplifies the monitoring.

In the free pension setting, however, since \( r_0 \) is exogenously given, the investment policies of different plan sponsors are independent of any risk attitudes. Therefore, the investment strategy is exactly the same for every pension fund and transparency increases substantially. Additionally, due to the size of the pension funds, this setting would have a notable impact on the stability of financial markets.
5.A Appendix

5.A.1 Proof of Proposition 5.1

The very tedious and long proof of 5.1 is organized as follows:

Firstly, we will examine the general solutions of the cubic equation (5.23).

Secondly, we will state a lemma containing some conditions for which we get a corner solution for \( \alpha \).

Thirdly, we will state a technical lemma to eliminate some case differentiations.

Finally, we will use this tools to prove 5.1.

The cubic equation we have to examine is

\[
-\zeta \left( R_p (1 - \alpha) + 2\alpha \right)^3 + R_p \left( R_e + R_p - 2 \right) - R_p = 0. \tag{5.23}
\]

Remember that \( R_p > 1 \) and \( R_e > 1 \) by assumption. Henceforth, we will assume that \( R_p \neq 2 \) unless the opposite is explicitly written.

We separate the solutions of the cubic equation (5.23) into the well known three different cases:

Case I: If \( Q \triangleq \frac{R_p^3 (4(R_e + R_p - 2)^3 + 27(R_p - 2)R_p R_e^2 \zeta)}{108(R_p - 2)^3} > 0 \) which is true if and only if \( R_p > 2 \wedge \zeta < -\frac{4(R_e + R_p - 2)^3}{27(R_p - 2)R_p R_e^2} \) or \( R_p < 2 \) we have one real solution:

\[
\hat{\alpha} = \frac{3}{2} \left( \frac{-R_e^2 R_p}{R_p - 2} \right)^{\frac{3}{2}} - \sqrt{Q} + \frac{3}{2} \left( \frac{-R_p^2 R_e}{R_p - 2} \right)^{\frac{3}{2}} + \sqrt{Q} + \frac{R_p}{R_p - 2}
\]

Case II: If \( Q = 0 \) which is true if and only if \( \zeta = -\frac{4(R_e + R_p - 2)^3}{27(R_p - 2)R_p R_e^2} \), we get two real solutions to (5.23):

\[
\hat{\alpha} = R_p \left( \frac{4}{R_p - 2} - \frac{3}{(R_e + R_p - 2)} \right)
\]
\[
\hat{\alpha}_T = \frac{1}{2} R_p \left( \frac{-1}{R_p - 2} + \frac{3}{(R_e + R_p - 2)} \right)
\]

\[5^5\text{In the first two cases we simply use the Cardano formula and the third case is known as the casus irreducibilis. These solution methods for solving cubic equations can be found in almost every handbook of mathematics, e.g., Reinhardt and Soeder (1998).} \]
whereas \( \bar{\alpha}_T \) is also an interception point of \( \Theta(\alpha) \), and, therefore, only a solution to our maximization problem if \( \Theta(\bar{\alpha}_T) > \Theta(\bar{\alpha}) \) and \( \bar{\alpha}_T = 0 \) or \( \bar{\alpha}_T = 1 \).

Case III: If \( Q < 0 \) i.e. \( \zeta > -\frac{4(R_e + R_p - 2)^3}{27(R_p - 2)R_p R_e^2} \) we finally get three real solutions:

\[
\bar{\alpha}_i = \sqrt{-\frac{4R_p (R_e + R_p - 2)}{3 (R_p - 2)^3} \cos \left( \frac{\varphi}{3} + (i - 1) \frac{2\pi}{3} \right) + \frac{R_p}{R_p - 2} \quad i = 1, 2, 3}
\]

(5.24)

with \( \varphi = \arccos \sqrt{\frac{-27(R_p - 2)R_p R_e^2 \zeta}{4(R_e + R_p - 2)^3}} \).

We claim for optimal \( \alpha^* \) that \( \alpha^* \in [0, 1] \).

In the following lemma we evaluate some conditions for which we have \( \alpha^* \in (0, 1) \), \( \alpha^* = 1 \) and \( \alpha^* = 0 \).

**Lemma 5.7** Suppose \( \alpha^* \) is the solution to \( \arg \max_{\alpha \in [0,1]} EU(e|X_\pi^T) \). Define \( f_1 (R_e, R_p) \triangleq -\frac{4(R_e + R_p - 2)^3}{27(R_p - 2)R_p R_e^2} \) and \( f_0 (R_e, R_p) \triangleq -\frac{(R_e + 2R_p - 4)^2}{8(R_p - 2)R_p R_e} \) then we have for \( \alpha^* \) the following characteristics:

<table>
<thead>
<tr>
<th>Case I: one real solution ( \hat{\alpha} ) to (5.23)</th>
<th>(( R_p &gt; 2 ) &amp; ( \zeta &lt; f_1 (R_e, R_p) )) ( \lor R_p &lt; 2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>solution</td>
<td>if and only if</td>
</tr>
<tr>
<td>a) ( \alpha^* = \hat{\alpha} \in (0, 1) )</td>
<td>( -\frac{1}{R_p} &lt; \zeta &lt; \frac{R_p(R_p - 2)}{8} )</td>
</tr>
<tr>
<td>b) ( \alpha^* = 1 )</td>
<td>( \zeta &gt; \frac{R_e - 4}{8} \lor \zeta \geq \frac{R_p(R_e - 2)}{8} )</td>
</tr>
<tr>
<td>c) ( \alpha^* = 0 )</td>
<td>( \zeta \leq -\frac{1}{R_p} \land \zeta &lt; \frac{R_e - 4}{8} )</td>
</tr>
<tr>
<td>d) indifferent between ( \alpha = 0 ) and ( \alpha = 1 )</td>
<td>( \zeta &lt; -\frac{1}{R_p} \land \zeta = \frac{R_e - 4}{8} )</td>
</tr>
</tbody>
</table>
### 5.7 Appendix

<table>
<thead>
<tr>
<th>Case III: three real solutions</th>
<th>( R_p &gt; 2 \land \zeta &gt; f_1 (R_e, R_p) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( {\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3} = {\tilde{\alpha}_1, \tilde{\alpha}_2, \tilde{\alpha}_3} )</td>
<td>sufficient conditions</td>
</tr>
<tr>
<td>a) ( \alpha^* = \hat{\alpha}_2 \in (0, 1) ), ( \tilde{\alpha}_1 \notin (0, 1), \tilde{\alpha}_3 \notin (0, 1) )</td>
<td>(-\frac{1}{R_p} &lt; \zeta &lt; \frac{R_p(R_e-2)}{8})</td>
</tr>
<tr>
<td>b) ( \hat{\alpha}_1 \in [0, 1), \hat{\alpha}_2 \in (0, 1) ), ( \tilde{\alpha}_3 \notin (0, 1) )</td>
<td>(\zeta &gt; f_0 \left( R_e, R_p \right))  (\land \zeta &lt; \frac{R_p(R_e-2)}{8})  (\land \zeta &lt; -\frac{1}{R_p})</td>
</tr>
<tr>
<td>necessary conditions</td>
<td>sufficient conditions</td>
</tr>
<tr>
<td>c) ( \alpha^* = 1 ), ( \zeta \geq \frac{R_p(R_e-2)}{8} )  (\land \zeta &gt; \frac{R_e-4}{8})</td>
<td>(\zeta &lt; 0)</td>
</tr>
<tr>
<td>d) ( \alpha^* = 0 ), ( \zeta \leq -\frac{1}{R_p} )  (\land \zeta &lt; \frac{R_e-4}{8})</td>
<td>(\zeta &lt; f_0 \left( R_e, R_p \right))  (\lor R_e \geq 2 \left( R_p - 2 \right))  (\lor \zeta \geq \frac{R_p(R_e-2)}{8})</td>
</tr>
<tr>
<td>e) ( \alpha^* ) not unique</td>
<td>(\left( \zeta &lt; -\frac{1}{R_p} \right))  (\land \zeta = \frac{R_e-4}{8})  (\land \zeta \geq \frac{R_p(R_e-2)}{8})  (\lor \left( \zeta &lt; -\frac{1}{R_p} \right))  (\land \zeta &lt; \frac{R_e-4}{8})  (\land \zeta = f_0 \left( R_e, R_p \right))  (\land \zeta = \frac{R_p(R_e-2)}{3R_p-4} &lt; R_e)</td>
</tr>
</tbody>
</table>

### Case IV: degenerate cubic equation

<table>
<thead>
<tr>
<th>solution</th>
<th>if</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha^* = \max \left{ \min \left[ \frac{2(1+2\alpha)}{R_p}, 1 \right], 0 \right} )</td>
<td>( R_p = 2 )</td>
</tr>
</tbody>
</table>

Before we prove lemma 5.7, we state \( R_p \neq 2 \) and consider the following lemma.

#### Lemma 5.8

Define: \( \alpha_d \triangleq \frac{R_p}{R_p-2} \) and \( \alpha_W \triangleq -\frac{1}{2} R_p \left( \frac{1}{R_p-2} - \frac{3}{R_p+R_e-2} \right) \).

\( \alpha_W < 1 \) if and only if \( R_e > \frac{4(\alpha_p-2)}{3R_p-4} \). In case III, i.e. \( \zeta > f_1 \left( R_e, R_p \right) \) and \( R_p > 2 \), we have \( \hat{\alpha}_1 < \alpha_W < \hat{\alpha}_2 < < \hat{\alpha}_3 < \tilde{\alpha}_3 \) and \( \alpha_d > 1 \). Moreover, \( \hat{\alpha}_1 \) is a local minimum, \( \hat{\alpha}_2 \) and \( \hat{\alpha}_3 \) local maxima to (5.23) and, therefore, \( \alpha^* \in (0, 1) \Rightarrow \alpha^* = \hat{\alpha}_2 \).

**Proof.** \( \Theta(\alpha) \triangleq \frac{2\alpha}{R_p} \left( \frac{1}{2\alpha + (1-\alpha)R_p} + \zeta \right) - \left( \frac{\alpha}{2\alpha + (1-\alpha)R_p} \right)^2 \) has the following properties:
\( \Theta ( \alpha ) \) is continuous except at \( \alpha_d = \frac{R_v}{R_p-2} \). \( \lim_{\alpha \to \alpha_d \pm} \Theta ( \alpha ) = -\infty \). For \( \alpha \in [0, 1] \) \( \Theta ( \alpha ) \) is continuous. The first derivative of \( \Theta ( \alpha ) \) is continuous (except at \( \alpha_d \)).

\[ \lim_{\alpha \to \pm\infty} \Theta ( \alpha ) = \pm\infty \]

(3) \( \Theta'' ( \alpha ) = \frac{2R_p[2R_p(2R_p-R_v-4)-2(R_p-2)(R_p+R_v-2)\alpha]}{R_p(2R_p+2\alpha-R_v)\alpha} \) and, therefore, \( \Theta ( \alpha ) \) has only one inflection point at \( \alpha_W = -\frac{1}{2}R_p \left( \frac{3}{R_p-2} - \frac{3}{R_p+R_v-2} \right) \). We have \( \Theta' ( \alpha_W ) > 0 \iff \zeta > f_1 (R_v, R_p) \) and

\[ \alpha_W = -\frac{1}{2}R_p \left( \frac{3}{R_p-2} - \frac{3}{R_p+R_v-2} \right) < 1 \iff R_p > 4(R_p-2) \left( \sqrt{\frac{(2R_p+3R_v-12)^2}{16R_p^2}} - \frac{1}{R_p} \right) \]

From (1), (2) and (3) follows that for the only inflection point \( \alpha_W < \alpha_d \) and further \( \hat{\alpha}_3 > \alpha_d \). Because there are three real solutions there are two other solutions which have to be smaller than \( \alpha_d \), because \( \alpha_W < \alpha_d \). Thus, we have \( \hat{\alpha}_1 < \alpha_W < \hat{\alpha}_2 < \alpha_d \). Because of (1) and (2) \( \hat{\alpha}_3 \) is a maximum and because of (1), (2) and \( \hat{\alpha}_1 < \alpha_W < \hat{\alpha}_2 \) is \( \hat{\alpha}_1 \) a minimum and \( \hat{\alpha}_2 \) a maximum.

Now we will prove lemma 5.7:

**Proof. Proof of lemma 5.7.**

In addition to the characteristics that we already stated in the proof of lemma 5.8 we have the following properties for \( \Theta ( \alpha ) \):

(4) Zero points: If we multiply \( \Theta ( \alpha ) \) by \( (2\alpha + (1 - \alpha) R_p)^2 \) we have a third degree polynomial. Therefore, \( \Theta ( \alpha ) \) has maximally three zeros. The first zero point is \( \hat{\alpha}_1 = 0 \), the others are

\[ \hat{\alpha}_{2,3} = \frac{2R_p+R_v-4+4(R_p-2)R_p\zeta \pm \sqrt{(2R_p+R_v-4)^2 + 8(R_p-2)R_p\zeta}}{4\zeta (R_p-2)^2} \]

so in case that \( \zeta > f_0 (R_v, R_p) \land \zeta \neq -\frac{1}{R_p} \) there are three different zeros, with \( \zeta = f_0 (R_v, R_p) \land \zeta \neq -\frac{1}{R_p} \) or \( \zeta = -\frac{1}{R_p} \land \zeta > f_0 (R_v, R_p) \) we have two different zeros and otherwise one real zero.

(5) Corners: \( \Theta (0) = 0 \), \( \Theta (1) = \frac{1}{4} (1 + 2\zeta) - \frac{1}{4} \) and thus \( \Theta (1) \geq 0 \iff \zeta \geq \frac{R_v-4}{8} \)

\[ \Theta' (0) = \frac{2(1-R_p\zeta)}{R_p} \]

and thus \( \Theta' (0) \geq 0 \iff \zeta \geq -\frac{1}{R_p} \), \( \Theta' (1) = \frac{-R_p(R_v-2)+8\zeta}{4R_v} \) and thus \( \Theta' (1) \geq 0 \iff \zeta \geq \frac{R_v(R_v-2)}{8} \)

All the following implementations will apply the statements (1)-(5). As we will not explicitly mention it in every argument keep in mind that \( \Theta (0) = 0 \).

**Case IV:**

a) is straightforward. Inserting \( R_p = 2 \) in (5.23) and solve for \( \alpha \).

**Case I:**
a) \(-\frac{1}{R_p} < \zeta < \frac{R_p(R_p-2)}{8} \iff \Theta'(0) > 0 \land \Theta'(1) < 0 \) (see (5)). Because \(\Theta(\alpha)\) and \(\Theta'(\alpha)\) are continuous in \([0,1]\) (see (1)) it follows that \(\exists \hat{\alpha} \in (0,1) : \Theta'(\hat{\alpha}) = 0\) and thus \(-\frac{1}{R_p} < \zeta < \frac{R_p(R_p-2)}{8} \iff \hat{\alpha} \in (0,1) \) with \(\Theta'(\hat{\alpha}) > 0\).

b) \(\zeta > \frac{R_p-4}{8} \land \zeta > \frac{R_p(R_p-2)}{8} \iff \Theta'(1) < 0 \land \Theta'(1) > 0 \) (see (5)). Because \(\Theta(0) = 0\) it is clear that \(\Theta(1) > 0\) is a necessary condition for \(\alpha^* = 1\). If we had \(\Theta'(1) < 0\), then surely \(\exists \alpha \in (0,1) : \Theta'(\alpha) = 0\). So we also need \(\Theta'(1) > 0\). But these two conditions are not only necessary but also sufficient, because in case I there is only one solution \(\hat{\alpha} \in (0,1)\) possible for which we would have \(\Theta'(\hat{\alpha}) < 0\). If \(\hat{\alpha} \notin (0,1)\), then \(\Theta(\alpha)\) would be strictly increasing for \(\alpha \in (0,1)\).

c) \(\zeta \leq -\frac{1}{R_p} \land \zeta < \frac{R_p-4}{8} \iff \Theta'(0) < 0 \land \Theta(1) < 0 \) (see (5)). Because \(\Theta(0) = 0\) it is clear that \(\Theta(1) < 0\) is a necessary condition for \(\alpha^* = 0\). If we had \(\Theta'(0) > 0\), then \(\exists \alpha \in (0,1) : \Theta'(\alpha) > 0\). So we also need \(\Theta'(0) \leq 0\). But these two conditions are not only necessary but also sufficient, because in case I there is only one solution \(\hat{\alpha} \in (0,1)\) possible for which we would have \(\Theta'(\hat{\alpha}) < 0\). If \(\hat{\alpha} \notin (0,1)\), then \(\Theta(\alpha)\) would be strictly decreasing for \(\alpha \in (0,1)\).

d) \(\zeta < -\frac{1}{R_p} \land \zeta = \frac{R_p-4}{8} \iff \Theta'(0) < 0 \land \Theta(1) = 0 \) (5). The rest of the argument is similar to c).

Case III:

a) \(-\frac{1}{R_p} < \zeta < \frac{R_p(R_p-2)}{8} \iff \Theta'(0) > 0 \land \Theta'(1) < 0 \) (see (5)). So, because of the zero point theorem, it is clear that \(\exists \hat{\alpha} \in (0,1) : \Theta'(\hat{\alpha}) = 0 \land \Theta'(\hat{\alpha}) > 0\) and it is not possible, that there exists another \(\hat{\alpha} \in (0,1)\) because \(\Theta(\alpha)\) has only one inflection point (see (3)). Thus and because of lemma 5.8 we have \(-\frac{1}{R_p} < \zeta < \frac{R_p(R_p-2)}{8} \Rightarrow \alpha^* = \hat{\alpha}_2 \in (0,1)\), \(\hat{\alpha}_1 \notin (0,1)\), \(\hat{\alpha}_3 \notin (0,1)\). Moreover, because of lemma 5.8 we have \(\hat{\alpha}_1 < \alpha^* = \hat{\alpha}_2 < \hat{\alpha}_3\).

b) \((f_0(\hat{\alpha}, R_p)) < \zeta) \land \left(\frac{4(R_p-2)}{2R_p-4} < \hat{\alpha} < 2(R_p-2)\right) \land \left(\zeta < \frac{R_p(R_p-2)}{8}\right) \land \left(R_p > \frac{1}{\alpha}\right) \iff \Theta(\alpha)\) has three real zeros (see (4)) \land \(\alpha_W < 1\) (see (3)) \land \(\Theta'(0) > 0\) \land \(\Theta'(1) < 0\) \land \(\Theta'(0) < 0\) (see (5)). Because \(R_p > 2\) we have \(\alpha_d = \frac{R_p}{R_p-2} > 1\). Further, because of \(\lim_{\alpha \to \alpha_d} \Theta(\alpha) = -\infty\) (see (1)) while \(\alpha_d > 1\) and \(\Theta''(0) > 0\) it follows that \(\alpha_W > 0\). Thus \(\hat{\alpha}_2, 3 > 0\).

Because of \(f_0(\hat{\alpha}, R_p) < \zeta\) and \(R_p \neq \frac{1}{\alpha}\) we know that
\[\hat{\alpha}_{2,3} = \frac{2R_p + R_p - 4 + 4(R_p-2)R_p\zeta}{4\sqrt{(2 R_p + R_p - 4)^2 + 8(R_p-2)R_p R_p\zeta}}\] exist such that \(\hat{\alpha} < \frac{2R_p + R_p - 4 + 4(R_p-2)R_p\zeta}{4\sqrt{(2 R_p + R_p - 4)^2 + 8(R_p-2)R_p R_p\zeta}} < \hat{\alpha}_3\).

It is \(\frac{2R_p + R_p - 4 + 4(R_p-2)R_p\zeta}{4\sqrt{(2 R_p + R_p - 4)^2 + 8(R_p-2)R_p R_p\zeta}} < \alpha_d \iff R_p > -2(R_p-2)\) which is negative for \(R_p > 2\). Therefore we have \(\hat{\alpha}_2 < \alpha_d\). Now we have to show that \(\hat{\alpha}_2 < \alpha_d\) also. Because of lemma 5.8 this
is the case except if \( \hat{a}_2 = \hat{a}_2 \land \hat{a}_3 = \hat{a}_3 \). In that case we also had three zeros and lemma
5.8 would not be contradicted even though \( \hat{a}_3 > a_d \). Note that \( \hat{a}_2 = \hat{a}_2 \) implies \( \hat{a}_3 = \hat{a}_3 \)
because we need exactly three zeros. Thus, if we can show that either \( \hat{a}_2 \neq \hat{a}_2 \) or \( \hat{a}_3 \neq \hat{a}_3 \),
then \( \hat{a}_3 < a_d \). To see this imagine a positive shift of \( \Theta(\alpha) \) i.e. \( \Theta(\alpha) + h, h \geq 0 \). This
new function \( \Theta(\alpha) + h \) would also have maximally three zeros, because it has the same
structure as \( \Theta(\alpha) \) (see (4)). But if \( \hat{a}_2 = \hat{a}_2 \land \hat{a}_3 = \hat{a}_3 \) this new function \( \Theta(\alpha) + h \) would
have five zeros which is a contradiction.

Because \( \hat{a}_2 < \hat{a}_3 < a_d \) and \( a_W < 1 \) (see (3)) there is no opportunity for more than one zero
between 1 and \( a_d \). So we have \( \hat{a}_2 < 1 \).

Since \( \Theta'(0) < 0 \) and because of lemma 5.8 \( \hat{a}_3 > \alpha^* = \hat{a}_2 > \hat{a}_{1,2} \in (0, 1) \) and \( \hat{a}_3 \notin (0, 1) \).
c) \( \zeta > \frac{R_p(R_p - 2)}{8} \land \zeta > \frac{R_p - 4}{8} \implies \Theta'(1) > 0 \land \Theta(1) > 0 \) (see (5)) are obviously necessary
conditions for the corner solution \( \alpha^* = 1 \). Otherwise we would have \( \alpha^* < 1 \).

Sufficiency of \( \Theta'(1) \geq 0 \land \Theta(1) > 0 \land \zeta < 0 \):

Because of lemma 5.8 and \( \Theta'(1) \geq 0 \) we get for the only possible maximum \( 1 < \hat{a}_2 < a_d \).
Further, we have \( \hat{a}_1 < 1 \). So if \( \Theta(1) > 0 \) we have a corner solution at \( \alpha^* = 1 \).

d) Before we study the subcases i)-iii) note that \( \zeta = -\frac{1}{\lambda_p} \land \zeta < \frac{R_p - 4}{8} \implies \Theta'(0) \leq 0 \land \Theta(1) < 0 \) (see (5)) are obviously necessary
conditions for the corner solution \( \alpha^* = 0 \). Otherwise we would certainly have a value \( \Theta(\alpha) > 0 \) for \( \alpha \in (0, 1) \).

i) \( \zeta < \frac{1}{\lambda_p} \land \zeta < \frac{R_p - 4}{8} \) is equivalent to no zeros others than \( \hat{a}_1 = 0 \) (see (4)). If \( \Theta'(0) \leq 0 \land \Theta(1) < 0 \) and \( \Theta(\alpha) \) never crosses the \( \alpha \)-axes again, we have \( \alpha^* = 0 \).

ii) \( R_p \geq 2 \land \Theta'(0) \leq 0 \) (see (5)). \( R_p \geq 2 \) implies that \( a_d > 1 \). If \( \Theta(\alpha) \) is concave
with negative slope at \( \alpha = 0 \) there cannot be an inflection point \( a_W \in (0, a_d) \) because we
would need a second to reach \( \lim_{\hat{a}_2 \to a_d} \Theta(\alpha) = -\infty \) (see (1)). However, because of (see (3))
we only have one inflection point.

iii) \( \zeta > \frac{R_p(R_p - 2)}{8} \implies \Theta'(1) \geq 0 \) (see (5)) If \( \Theta'(1) \geq 0 \) and \( \Theta(1) < 0 \) we would need to
change the sign of the slope three times to get a \( \Theta(\alpha) > 0 \) for \( \alpha \in (0, 1) \) and \( \Theta'(0) \leq 0 \)
which is not possible with one inflection point (see (3)).

e) i) \( \zeta < -\frac{1}{\lambda_p} \land \zeta = \frac{R_p - 4}{8} \land \zeta = \frac{R_p(R_p - 2)}{8} \implies \Theta'(0) < 0 \land \Theta(1) = 0 \land \Theta'(1) \geq 0 \) thus, it
follows from lemma 5.8 and the fact that there is only one inflection point (see (3)) that \( \alpha^* 

is not unique. It is either 0 or 1.

ii) \( \zeta < -\frac{1}{\lambda_p} \land \zeta < \frac{R_p - 4}{8} \land \zeta = \frac{R_p(R_p - 2)}{8} \implies \Theta'(0) < 0 \land \Theta(1) < 0 \land \Theta(\alpha) \)
has exactly two different zeros. We have \( \Theta(\hat{a}_2) = 0 \) and it follows from lemma 5.8 and
the fact that there is only one inflection point (see (3)) that \( \alpha^* \) is not unique.

Before we face the proof of proposition 5.1 we state the following lemma.
Lemma 5.9 For $R_p \neq 2$ and $\zeta > -1$ we have $\zeta < f_1 (R_c, R_p) \Rightarrow \zeta < f_0 (R_c, R_p) \Rightarrow \alpha^* = 0$.

Proof. Again we use (1)-(5) from the proofs of the lemmata 5.8 and 5.7.

1. Firstly, we prove $\zeta < f_1 (R_c, R_p) \Rightarrow \zeta < f_0 (R_c, R_p)$:

The case with $R_p < 2$ is trivial since in this case $f_1 (R_c, R_p)$ and $f_0 (R_c, R_p)$ are positive and $\zeta$ is assumed to be negative.

The case with $R_p > 2$ is more complicated. We will proceed as follows. We will show, that $f_1 (R_c, R_p)$ and $f_0 (R_c, R_p)$ are tangent to each other and that there exist no other intersection. Thus, we first calculate the intersection of $f_1 (R_c, R_p)$ and $f_0 (R_c, R_p)$ and get $f_1 (R_c, R_p) = f_0 (R_c, R_p) \Leftrightarrow R_c = 2 (R_p - 2) \vee R_e = -\frac{8}{5} (R_p - 2)$

Next we show that $R_e = 2 (R_p - 2)$ are not only intersection points of $f_1 (R_c, R_p)$ and $f_0 (R_c, R_p)$ but also tangent points. To do this, we show that the gradients of $f_1 (R_c, R_p)$ and $f_0 (R_c, R_p)$ are colinear in $R_e = 2 (R_p - 2)$.

$$|\text{grad} f_1 (R_c, R_p), \text{grad} f_0 (R_c, R_p)| = \frac{(R_p + R_e - 2)^2}{54 (R_p - 2) R_p^3 R_e^4} \left(-2 (R_p - 2)^2 + \frac{1}{2} R_e^2\right)$$

Thus we get $|\text{grad} f_1 (R_c, R_p), \text{grad} f_0 (R_c, R_p)| = 0 \Rightarrow R_e = 2 (R_p - 2)$. Thus, $R_e = 2 (R_p - 2)$ are tangent points.

The gradients in the point $B^* \triangleq \left(-4 - \frac{2}{\zeta}, -\frac{4}{\zeta}\right)$ are

$$\text{grad} f_1 (B^*) = \text{grad} f_0 (B^*) = \left(\begin{array}{c}
\frac{2}{\zeta^2} \\
0
\end{array}\right)$$

Further $f_1 (R_c, R_p)$ and $f_0 (R_c, R_p)$ are continuous for $R_p > 2$ and have only on $R_e = 2 (R_p - 2)$ colinear gradients. This is a sufficient condition for the statement that the contour curves of the one function lie within the contour curves the other function and thus either $f_1 (R_c, R_p) \geq f_0 (R_c, R_p)$ or $f_1 (R_c, R_p) \leq f_0 (R_c, R_p)$. But which relation is true? It suffices to calculate one point within $\{(R_c, R_p) \in (1, \infty) \times (2, \infty) ; R_e \neq 2 (R_p - 2)\}$ and we get $f_1 (R_c, R_p) \leq f_0 (R_c, R_p) \forall (R_c, R_p) \in (1, \infty) \times (2, \infty)$.

2. Now we prove that $\zeta < f_0 (R_c, R_p) \Rightarrow \alpha^* = 0$

From lemma 5.7, case Ic) and case IIIId), we know $\zeta \leq -\frac{1}{R_p} \wedge \zeta < \frac{R_p - 4}{8} \wedge \zeta < f_0 (R_c, R_p) \Rightarrow \alpha^* = 0$.

Firstly, we prove that $\zeta < f_0 (R_c, R_p) \Rightarrow \zeta < -\frac{1}{R_p}$:
Define
\[ \mathcal{L} \triangleq \left\{ (R_e, R_p) \in \mathbb{R}^2; R_p \neq 2 \wedge -\frac{\partial f_0(R_e, R_p)}{\partial R_p} = 0 \right\} \]
\[ = \left\{ (R_e, R_p) \in \mathbb{R}^2; R_p \neq 2 \wedge R_e = 2 (R_p - 2) \right\} \quad (5.25) \]
where \( \frac{\partial f_0(R_e, R_p)}{\partial R_p} = \frac{dR_p}{dR_e} \) on the contour line \( f_2 \). Further, the intersection of \( \zeta \) and \( f_1(R_e, R_p) \) and \( f_0(R_e, R_p) \) is
\[ f_1(R_e, R_p) = f_0(R_e, R_p) = \zeta \Leftrightarrow (R_e, R_p)_1 = \left( \frac{16}{5} - \frac{1}{50}, \frac{1}{80} \right) \) and \((R_e, R_p)_2 = \left( -4 - \frac{2}{\zeta}, -\frac{1}{\zeta} \right) \). We assume that \( R_p > 1 \). Thus the point \((R_e, R_p)_1\) is irrelevant since \( \frac{1}{50} < 0 \).
Note that the point \( B^* \) lies in \( \mathcal{L} \). Additionally, twice implicitly differentiating, we find \( \frac{d^2 R_p}{dR_e^2}(B^*) > 0 \) (the contour curve is convex). Thus the point \((R_e, R_p)_2 = B^*\) is the point in \( f_0(R_e, R_p) = \zeta \) with smallest \( R_p \) and since the gradient in \( B^* \) is positive in \( R_p \)-direction it is \( \zeta < f_0(R_e, R_p) \Rightarrow \zeta < -\frac{1}{R_p} \Leftrightarrow \Theta'(0) < 0 \) by (5).

Finally, we prove that \( \zeta < f_0(R_e, R_p) \wedge \zeta < -\frac{1}{R_p} \Rightarrow \zeta < \frac{8}{R_p} - 4 \):
The relations \( \zeta < f_0(R_e, R_p) \) and \( \zeta < -\frac{1}{R_p} \) imply that \( \Theta(\alpha) \) has no zeros except \( \alpha_0 = 0 \) by (4), and respectively \( \Theta(\alpha) \) has negative slope in \( \alpha = 0 \) by (5). Therefore, and because of the continuity of \( \Theta(\alpha) \) it is strictly negative in \( \alpha \in (0, 1] \) and, therefore, also in \( \alpha = 1 \) which leads to \( \zeta < \frac{8}{R_p} - 4 \) by lemma 5.7, case IIId). Thus \( \zeta < f_0(R_e, R_p) \Rightarrow \zeta < -\frac{1}{R_p} \wedge \zeta < \frac{8}{R_p} - 4 \) which is sufficient for \( \alpha^* = 0 \).
Therefore, we have \( \zeta < f_0(R_e, R_p) \Rightarrow \alpha^* = 0 \).

Now we are ready for the proof of proposition 5.1.

**Proof.** Proof of proposition 5.1.

Again we use (1)-(5) from the proofs of the lemmata 5.8 and 5.7.

A: \( R_p > 2 \):
In a first step (a), we will show that \( R_e \geq 2 (R_p - 2) \wedge R_p \geq \frac{1}{\zeta} \Rightarrow \alpha^* = 0 \), then in a second step (b) we will derive the set union of
\[ \left\{ (R_e, R_p) \mid R_e \geq 2 (R_p - 2) \wedge R_p \geq -\frac{1}{\zeta} \right\} \cup \left\{ (R_e, R_p) \mid \zeta < f_0(R_e, R_p) \right\} \]
in order to unite the subcases of case IIIId) of lemma 5.7.

Next, we will prove iii), then ii), and finally i), before we turn to B and C.
At first, we have a look at the point \( B^* \). By substituting we find that
From lemma 5.7 (case IIId)) we have

\[
\zeta > f_0(R_e, R_p)
\]

and

\[
R_p < 0.8 (R_p - 2)
\]

Then by substituting \( B^s \) it can easily be found that for \((R_e, R_p) = B^s\) we have \( \Theta' (0) = \Theta'' (0) = 0 \), by (3) and (5), and thus \( a_{i_W} = 0 \). Because \( a_d > 1 \) and \( \lim_{a \to a_d} \Theta (a) = -\infty \) by (2) we have \((R_e, R_p) = B^s \Rightarrow \alpha^* = 0 \)

(a) Case \( R_p > -\frac{1}{\zeta} \wedge R_e > 2 (R_p - 2) \):

In the proof of lemma 5.9 we have seen that \( B^s \) lies in \( L \) and that it is the point on

\[
f_0(R_e, R_p) = \zeta
\]

with smallest \( R_p \). Thus for \( R_p > -\frac{1}{\zeta} \) it is \( \zeta < f_0(2(R_p - 2), R_p) \) and we have on \( R_e = 2(R_p - 2) \) strictly \( \zeta < f_0(R_e, R_p) \) and, therefore, \( \alpha^* = 0 \) from lemma 5.9.

Case \( R_p \geq -\frac{1}{\zeta} \wedge R_e > 2 (R_p - 2) \):

From lemma 5.7 we have \( R_e \geq 2(R_p - 2) \wedge R_p \geq -\frac{1}{\zeta} \wedge R_e > 8\zeta + 4 \Rightarrow \alpha^* = 0 \), see case IIIId).

Therefore, and since \(-4 - \frac{2}{\zeta} + \epsilon > 8\zeta + 4 \iff 0 < 4\zeta^2 + 4\zeta + 1 \) is always true, for \( R_p \geq -\frac{1}{\zeta} \) and \( R_e > 2(R_p - 2) \) we have always \( R_e > 8\zeta + 4 \) and, thus,

\[
R_e > 2(R_p - 2) \wedge R_p \geq -\frac{1}{\zeta} \Rightarrow \alpha^* = 0 \tag{5.27}
\]

(b) From lemma 5.9 we know that \( \zeta < f_0(R_e, R_p) \Rightarrow \alpha^* = 0 \).

Again we suppose \( R_p \geq -\frac{1}{\zeta} \). It is easy to calculate that we have

\[
\zeta < f_0(R_e, R_p)
\]

\[
\iff y_0(R_p) < R_e < y_0''(R_p) \triangleq -2 \left[ \frac{R_p + 2(R_p - 2)R_p\zeta}{2(1 - (R_p - 2)\sqrt{R_p\zeta + (1 + R_p\zeta))}} \right]
\]

Because of (5.25) we have \( R_e = y_0''(R_p) \Rightarrow R_e = 2(R_p - 2) \) in \( B^s \) and otherwise \( y_0''(R_p) < 2(R_p - 2) < y_0''(R_p) \) and, thus, \( R_e \geq y_0''(R_p) \Rightarrow R_e > 2(R_p - 2) \) i.e. the upper solution \( y_0''(R_p) \) is irrelevant because of (5.27). Hence, we have

\[
R_p \geq -\frac{1}{\zeta} \wedge R_e > y_0(R_p) \Rightarrow \alpha^* = 0 \tag{5.28}
\]

iii) \( \zeta < -\frac{1}{8} \) and \( R_p \geq -\frac{1}{\zeta} \)

From lemma 5.7 (case IIIId)) we have \( \zeta \leq -\frac{1}{R_p} \wedge \zeta < \frac{\zeta}{8} \wedge \zeta \geq \frac{R_p(R_p - 2)}{8} \Rightarrow \alpha^* = 0 \)
\[\zeta \geq \frac{R_p (R_e - 2)}{8}\] is only possible if \(R_e < 2\); Therefore, and because of \(\zeta < \frac{R_e - 4}{8}\) it is necessary that \(\zeta < -\frac{1}{4}\) in order to get \(\zeta < \frac{R_e - 4}{8} \land \zeta \geq \frac{R_p (R_e - 2)}{8}\).

It is \(R_e = 8\zeta + 4 \land R_e = y_1 (R_p) \Rightarrow R_p = \hat{R}_p = \frac{4\zeta}{4\zeta + 1}\) and, therefore, easy to see that taking everything into account we have

\[\zeta < \frac{R_e - 4}{8} \land \zeta \geq \frac{R_p(R_e - 2)}{8} \Rightarrow R_e > 8\zeta + 4 \land R_e \leq y_1 (R_p) \Leftrightarrow R_p > \hat{R}_p\] (5.29)

The last equivalency follows from the strict monotony of \(y_1 (R_p)\).

It is easy to show that \(y_0'\) is continuos for \(R_p > -\frac{1}{4}\). For \(y_0' (\hat{R}_p) = 0\) and \(\zeta < -\frac{1}{4}\) the only possible solution is \(\hat{R}_p = \hat{R}_p\). Thus, \(y_0 (R_p)\) is monotone for \(R_p > \hat{R}_p\).

Further, note that \(y_0 (\hat{R}_p) = 8\zeta + 4 < 2\) and \(\lim_{R_p \rightarrow +\infty} y_0 (R_p) = -\frac{1}{\zeta} < 2 = \lim_{R_p \rightarrow +\infty} y_1 (R_p)\).

Thus, it is \(y_0 (R_p) < 2\) for \(R_p > \hat{R}_p\) because \(y_0 (R_p)\) is monotone and there exists at least one point \(\hat{R}_p > \hat{R}_p\) for which \(y_0 (\hat{R}_p) < y_1 (\hat{R}_p)\) holds. Let us examine whether this holds in general.

Since \(y_0 (R_p)\) and \(y_1 (R_p)\) are monotone for \(R_p > \hat{R}_p\) and \(y_0 (\hat{R}_p) = y_1 (\hat{R}_p) = 8\zeta + 4 < 2\) and \(\lim_{R_p \rightarrow +\infty} y_0 (R_p) = -\frac{1}{\zeta} < 2\) and \(\lim_{R_p \rightarrow +\infty} y_1 (R_p) = 2\), respectively, we know that \(y_0 (R_p)\) and \(y_1 (R_p)\) move within \((-\infty, 2]\) for \(R_p > \hat{R}_p\).

Next, we are looking for another intersection point for \(R_p > \hat{R}_p\) besides \((R_e, R_p) = (8\zeta + 4, \hat{R}_p)\). We find that

\[
\begin{align*}
\left| \frac{\text{grad} R_p (R_e - 2)}{8}, \text{grad} f_0 (R_e, R_p) \right| &= 0 \Rightarrow R_{p,1} = 2 - \frac{R_e \zeta}{2} < 2, \\
R_{p,2,3} &= \frac{-16 + R_e (2 + 3 \zeta + \sqrt{(R_e - 2) (4 (R_e - 2))})}{4 (R_e - 2)} \notin \mathbb{R} \text{ for } 1 < R_e < 2.
\end{align*}
\]

Thus for \(1 < R_e < 2\) the contour curves of the functions \(f_0 (R_e, R_p)\) and \(\frac{R_p (R_e - 2)}{8}\) have no common slope. Therefore, the contour curves have at most one intersection point, namely \((R_e, R_p) = (8\zeta + 4, \hat{R}_p)\).

Hence for \(R_p > \hat{R}_p\) we have always \(y_0 (R_p) < y_1 (R_p)\).

Summarizing (5.28) and \(R_p \geq -\frac{1}{\zeta} \land R_e > 8\zeta + 4 \land R_e \leq y_1 (R_p) \Rightarrow \alpha^* = 0\) from lemma 5.7 (case IIIId) and (5.29) and finally \(R_p > \hat{R}_p \Rightarrow y_0 (R_p) < y_1 (R_p)\) we have

\[R_p \geq \hat{R}_p \geq -\frac{1}{\zeta} : R_e > 8\zeta + 4 \Rightarrow \alpha^* = 0\] (5.30)

and for \(R_p \geq \hat{R}_p\) we have \(y_1 (R_p) \geq 8\zeta + 4\) and because \(R_e < y_1 (R_p) \Leftrightarrow \zeta \geq \frac{R_p (R_e - 2)}{8}\) and \(R_e < 8\zeta + 4 \Leftrightarrow \zeta > \frac{R_e - 2}{8}\) with lemma 5.7 (case IIIc) we get

\[R_p \geq \hat{R}_p \geq -\frac{1}{\zeta} : R_e < 8\zeta + 4 \Rightarrow \alpha^* = 1.\] (5.31)
Again, because of lemma 5.7 (case IIIe)) and the same argumentation for the special case we get that
\[ R_p \geq \hat{R}_p \geq - \frac{1}{\xi} : R_e = 8\zeta + 4 \Rightarrow \alpha^* \text{ is not unique.} \] (5.32)

Since the left sides of the implications (5.30), (5.31) and (5.32), and, thus, the sets \( \{ R_e \mid R_e > 8\zeta + 4 \} \), \( \{ R_e \mid R_e = 8\zeta + 4 \} \) and \( \{ R_e \mid R_e < 8\zeta + 4 \} \) for \( R_p \geq \hat{R}_p \geq - \frac{1}{\xi} \) are disjoint and cover the whole domain \( R_e > 1 \) the equivalency in (5.30), (5.31) and (5.32) holds and, therefore, is iii) proven.

ii) For \( - \frac{1}{\xi} \leq R_p < \frac{4\zeta}{4\zeta + 1} \) the function \( y_0 (R_p) \) is monotone. Moreover, since the slope in \( y_0 \left( - \frac{1}{\xi} \right) \) is orthographic and \( y_0 \left( - \frac{1}{\xi} \right) = - 4 - \frac{2}{\zeta} \), and for \( R_p > - \frac{1}{\xi} \) we have \( y_0 (R_p) < 2 (R_p - 2) \) the function \( y_0 (R_p) \) is decreasing in \( R_p \). Thus, if \( R_p \geq - \frac{1}{\xi} \) and, additionally, \( y_0 \left( - \frac{1}{\xi} \right) \leq 1 \Leftrightarrow \zeta \leq - \frac{2}{\xi} \), then we have \( \alpha^* = 0 \).

Now we consider the case \( - \frac{2}{\xi} < \zeta < - \frac{1}{4} \):

Note that \( y_0 \left( - \frac{1}{\xi} \right) = - 4 - \frac{2}{\zeta} > y_0 (\hat{R}_p) = 8\zeta + 4 \Leftrightarrow 1 + 4\zeta^2 + 4\zeta > 0 \) is for \( \zeta \neq - \frac{1}{\xi} \) always true. Therefore, assuming \( - \frac{1}{\xi} \leq R_p < \hat{R}_p \) leads to \( y_0 (R_p) > 8\zeta + 4 \).

Additionally, we have \( y_1 (\hat{R}_p) = 8\zeta + 4 \) and \( y_1 (R_p) \) is strictly increasing and therefore, \( y_1 (R_p) < 8\zeta + 4 \) for \( R_p < \hat{R}_p \). Thus, assuming \( - \frac{1}{\xi} \leq R_p < \hat{R}_p \), the relation \( y_1 (R_p) < 8\zeta + 4 < y_0 (R_p) \) holds.

For \( - \frac{1}{\xi} \leq R_p < \hat{R}_p, R_e \leq y_1 (R_p) \Rightarrow R_e < 8\zeta + 4 \) and, thus, equivalently \( \zeta \geq \frac{R_p (R_p - 2)}{8} \Rightarrow \zeta > \frac{8 - 4}{8} \Leftrightarrow \alpha^* = 1 \) which leads to
\[ R_e \leq y_1 (R_p) \Leftrightarrow \alpha^* = 1. \] (5.33)

If \( R_e = y_0 (R_p) \), then \( \Theta (\alpha) \) has two zeros \( \hat{\alpha}_1 = 0 \) by (4) and \( \hat{\alpha}_2 = g (R_e) \triangleq R_p \left( \frac{1}{R_p - 2} + \frac{4}{R_p + 2R_p - 4} \right) \) and \( \Theta' (\hat{\alpha}_2) = \Theta (\hat{\alpha}_2) = 0 \). Further, we can calculate \( g' (R_e) < 0, g^{-1} (1) = \frac{2(R_p - 2)}{R_p - 1} \) and \( g^{-1} (0) = 2 (R_p - 2) \). Thus, if \( \frac{2(R_p - 2)}{R_p - 1} < R_e < 2 (R_p - 2) \) we have \( \hat{\alpha}_2 \in (0, 1) \). Since for \( R_p < \hat{R}_p, R_e = y_0 (R_p) > \frac{2(R_p - 2)}{R_p - 1} \) and for \( R_p > - \frac{1}{\xi}, 2 (R_p - 2) > R_e = y_0 (R_p) \), it follows that for \( R_e = y_0 (R_p) \) the solution is not unique.

\[ R_p > - \frac{1}{\xi} \wedge R_e = y_0 (R_p) \Rightarrow \text{solution not unique} \] (5.34)

For the special case where \( R_p = - \frac{1}{\xi} \wedge R_e = y_0 (R_p) \) it is easy to see that
\[ (R_e, R_p) = B^* \Rightarrow \alpha^* = 0 \] (5.35)
From lemma 5.7 (case IIIb)) and \( R_p > -\frac{1}{\zeta} \) we know that

\[
f_0 (R_e, R_p) < \zeta \land \frac{4(R_p - 2)}{3R_p - 4} < R_e < 2 (R_p - 2) \land \zeta < \frac{R_p(R_e - 2)}{8} \Rightarrow \alpha^* \in (0, 1) \land \alpha^* > \hat{\alpha}_1 \in [0, 1). \]

Remember \( \zeta < \frac{R_e(R_e - 2)}{8} \iff R_e > y_1 (R_p) \). Since \( \zeta > -\frac{R_e^2}{4(3R_p - 4)} \Rightarrow y_1 (R_p) > \frac{4(R_e - 2)}{3R_p - 4} \) and \(-\frac{R_e^2}{4(3R_p - 4)} < -\frac{4}{9} < -\frac{2}{3} \) for \( \zeta > -\frac{2}{3} \), we have \( y_1 (R_p) > \frac{4(R_e - 2)}{3R_p - 4} \) and, therefore, the condition \( \frac{4(R_e - 2)}{3R_p - 4} < R_e \) is redundant. Finally, note that \( f_0 (R_e, R_p) < \zeta \land R_e < 2 (R_p - 2) \iff R_e < y_0 (R_p) \) from (b).

Thus, for \( R_p \geq -\frac{1}{\zeta} \) we have \( f_0 (R_e, R_p) < \zeta \land R_e < 2 (R_p - 2) \land \zeta < \frac{R_p(R_e - 2)}{8} \iff y_1 (R_p) < R_e < y_0 (R_p) \) and it is

\[
y_1 (R_p) < R_e < y_0 (R_p) \Rightarrow \alpha^* \in (0, 1) \quad (5.36)
\]

The lacking implication is (5.28).

Since the left sides of the implications (5.33), (5.36), (5.34) together with (5.35) and (5.28), and, thus, the sets \{ \( R_e \mid R_e \leq y_1 (R_p) \) \}, \{ \( R_e \mid y_1 (R_p) < R_e < y_0 (R_p) \) \}, \{ \( R_e \mid R_e = y_0 (R_p) \) \} and \{ \( R_e \mid R_e > y_0 (R_p) \) \} for \(-\frac{2}{3} < \zeta < -\frac{1}{4} \) and \( R_p \geq -\frac{1}{\zeta} \) are disjoint and cover the whole domain \( R_e > 1 \) so equivalency in (5.36), (5.34) holds and it is \( R_e > y_0 (R_p) \lor (R_e, R_p) = \left(-4 - \frac{2}{\zeta}, -\frac{1}{\zeta}\right) \iff \alpha^* = 0 \).

Finally we consider the case \( \zeta \geq -\frac{1}{4} \): \( R_e > 8\zeta + 4 \land R_e \leq y_1 (R_p) \) is not possible if we have \( R_p < \hat{R}_p \). Thus, it is always \( y_1 (R_p) < 8\zeta + 4 < y_0 (R_p) \). Thus, the proof for this case is exactly the same as the proof for the case \(-\frac{2}{3} < \zeta < -\frac{1}{4} \) unless we do not have to care about the boundary \( \hat{R}_p \).

Hence, we have ii) almost proven. The remaining issue is the exact value of \( \alpha^* \in (0, 1) \).

For \( R_p > 2 \) and \( \zeta < 0 \) we have from lemma 5.9 \( \alpha^* \in (0, 1) \Rightarrow \zeta > f_1 (R_e, R_p) \). Therefore, we know from lemma 5.7 that there are three real solutions to the cubic equation (5.23). As already stated in (5.24) these solutions are

\[
\hat{\alpha}_i = \sqrt[3]{-\frac{64R_p^3 (R_e + R_p - 2)^3}{27 (R_p - 2)^9 \zeta^3} f (i) + \frac{R_p}{R_p - 2}}
\]

where \( f (i) \triangleq \cos \left( \frac{\varphi}{3} + (i - 1) \frac{2\pi}{3} \right), \quad i = 1, 2, 3 \) and \( \varphi = \arccos \left( \frac{-27(R_p - 2)R_p R_p^2 \zeta}{4R_e + R_p - 2} \right) \). The solutions discussed above \{\( \hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3 \)\} are somehow linked with the set \{\( \hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3 \)\}. We know from lemma 5.8 that if there exists a solution \( \alpha^* \in (0, 1) \), then \( \hat{\alpha}_1 < \alpha^* = \hat{\alpha}_2 < \hat{\alpha}_3 \) and \( \hat{\alpha}_3 > \alpha_d = \frac{R_p}{R_p - 2} > 1 \). The solution \( \hat{\alpha}_3 \) implies \( f (i) > 0 \), whereas \( \hat{\alpha}_1 \) and \( \alpha^* \) imply...
f(\(i\)) < 0. From lemma 5.9 we have, further, \(\alpha^* \in (0, 1) \Rightarrow \zeta > f_1(R_e, R_p)\) and, therefore, \(0 < \frac{-27(R_p-2)\eta_\zeta^2}{4(R_e+R_p-2)} < 1\) which means \(0 < \varphi < \frac{\pi}{2}\). Thus, \(\sqrt{3}\frac{1}{2} < f(1) < 1, -\sqrt{\frac{3}{2}} < f(2) < -\frac{1}{2}\) and \(-\frac{1}{2} < f(3) < 0\) and, therefore, \(\hat{\alpha}_2 < \hat{\alpha}_3 < \hat{\alpha}_1\). Thus, together with lemma 5.8 we have \((\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3) = (\hat{\alpha}_2, \hat{\alpha}_3, \hat{\alpha}_1)\).

Hence ii) is proven.

Let us turn to the last case within A.

i) \(R_p < -\frac{1}{\zeta}\)

From lemma 5.7 we know that \(\zeta \leq -\frac{1}{R_p} \iff R_p \geq -\frac{1}{\zeta}\) is a necessary condition for \(\alpha^* = 0\). Thus, in case i) only \(\alpha^* \in (0, 1]\) remains. From the same lemma 5.7, case Ib), we have \(\zeta \leq -\frac{1}{R_p} \wedge \zeta < \frac{R_p - 4}{8} \iff \alpha^* = 1\).

Further the relation \(y_1 \left(-\frac{1}{\zeta}\right) = -8\zeta^2 + 2 \leq 8\zeta + 4 \iff 1 + 4\zeta^2 + 4\zeta \geq 0\) holds for \(\zeta < 0\). Because \(y_1(R_p)\) is strictly increasing and, therefore, \(R_e \leq y_1(R_p) < 8\zeta + 4 \iff \alpha^* = 1, i)\) proven.

B: \(R_p < 2\)

i) Directly from lemma 5.7, case Ia), we have \(-\frac{1}{R_p} < \zeta < \frac{R_p(R_p - 2)}{8} \iff R_p < -\frac{1}{\zeta} \wedge R_e > y_1(R_p) \iff \alpha^* = \hat{\alpha} = (0, 1)\), and \(\zeta > \frac{R_p - 4}{8} \wedge \zeta \geq \frac{R_p(R_p - 2)}{8} \iff R_e < 8\zeta + 4 \wedge R_e \leq y_1(R_p) \iff \alpha^* = 1\), by lemma 5.7, case Ib).

Since for \(R_p < 2\) and \(\zeta > -\frac{1}{\zeta}\) we have \(y_1(R_p) < 8\zeta + 4\) it is \(R_e \leq y_1(R_p) \iff \alpha^* = 1\)

ii) For \(\zeta \leq -\frac{1}{\zeta}\) the relation \(R_e \leq y_1(R_p) = \frac{8\zeta}{R_p} + 2 \leq 1\) is not possible. However, the relation \(\zeta \leq -\frac{1}{R_p} \wedge \zeta \leq \frac{R_p - 4}{8} \iff R_p \geq -\frac{1}{\zeta} \wedge R_e > 8\zeta + 4 \iff \alpha^* = 0\) (by lemma 5.7, case Ic)) is only possible, if \(\zeta < -\frac{1}{\zeta}\) and, therefore, \(R_e > 8\zeta + 4\) holds. Thus, for \(R_p < 2\) we have: \(R_p \geq -\frac{1}{\zeta} \iff \alpha^* = 0\).

C: see lemma 5.7, case IV.

D: Note that because of lemma 5.7 (case 1c) for \(R_p < 2\) we have \(\zeta \leq -1 \Rightarrow \alpha^* = 0\). For \(R_p = 2\), see C. Note that the point \((R_e, R_p) = B^s\) is the point in \(f_0(R_e, R_p) = \zeta\) with smallest \(R_p\) (see proof of lemma 5.9) or, put differently, the domain of \(y_0(R_p)\) is \([-\frac{1}{\zeta}, \infty)\). Additionally, \(y_0 \left(-\frac{1}{\zeta}\right) = -4 - 2\zeta < 1\) and \(\lim_{R_p \to +\infty} y_0(R_p) = -\frac{1}{\zeta} < 1\) and \(y_0(\hat{R}_p) = 8\zeta + 4 < 1\) and, therefore, \(y_0(R_p) < 1\) because of monotonicity for \(R_p > \hat{R}_p\) and for \(R_p < \hat{R}_p\). Since \(R_p \geq -\frac{1}{\zeta}\) holds for \(\zeta \leq -1\), and because of (5.28) we have \(\alpha^* = 0\).
5.A.2 Proof of Corollary 5.3

Proof. Case \( R_p \leq \frac{1}{\xi} \):

On (5.23) we can calculate with implicit differentiation

\[
\frac{d\alpha^*}{dR_c} = -\frac{R_p\alpha^*}{R_p(R_p + R_c - 2) - 3(R_p - 2)(2\alpha^* + (1 - \alpha^*) R_p)\alpha^*\xi}.
\]

If the denominator \( f(R_p, \alpha^*) \triangleq R_p(R_p + R_c - 2) - 3(R_p - 2)(2\alpha^* + (1 - \alpha^*) R_p)\alpha^*\xi \) is strictly positive, then \( \frac{d\alpha^*}{dR_c} < 0 \).

We have \( \frac{\partial f(R_p, \alpha^*)}{\partial \alpha^*} = -6(R_p - 2)(2\alpha^* + (1 - \alpha^*) R_p)\alpha^*\xi > 0 \) for \( \alpha^* \in (0, 1) \).

Further we have \( f(R_p, 0) = R_p^2 (1 + R_p) \geq 0 \) for \( R_p \leq \frac{1}{\xi} \). Thus, \( f(R_p, \alpha^*) > 0 \) for \( \alpha^* \in (0, 1) \) and, therefore, \( \frac{d\alpha^*}{dR_c} < 0 \) for \( R_p \leq \frac{1}{\xi} \).

Case \( R_p > \frac{1}{\xi} \):

In this case the evaluation of the denominator’s sign turns out to be delicate. We choose a more indirect way.

For \( R_p > \frac{1}{\xi} \) we know from lemma 5.7, case IIIb) that if \( \alpha^* \in (0, 1) \), we have at most one other, inferior solution \( \hat{\alpha}_1 \) to (5.23) such that \( \alpha^* > \hat{\alpha}_1 \in (0, 1) \). From (5.23) we have \( R_c(\hat{\alpha}, R_p) \triangleq \left( \frac{(R_p(1-\hat{\alpha})+2\alpha)}{\alpha R_p} \right)^3 + \frac{R_p}{\alpha} + 2 - R_p \) which is unique and continuous for every \( \hat{\alpha} \in (0, 1) \). In the previous case we showed that for \( \hat{R}_p = \frac{1}{\xi} \) and two optimal solutions \( \alpha^*_{1,2} = \hat{\alpha}_{1,2} \in (0, 1) \), \( \alpha^*_{1,2} > \alpha^*_1 \Rightarrow R_c(\alpha^*_{1,2}, \hat{R}_p) < R_c(\alpha^*_1, \hat{R}_p) \). Because \( R_c(\hat{\alpha}, R_p) \) is continuous there exist some \( \hat{R}_{p,1} > \hat{R}_p \) for which the implication \( \alpha^*_{2} = \hat{\alpha}_2 > \alpha^*_1 = \hat{\alpha}_1 \Rightarrow R_c(\hat{\alpha}_2, \hat{R}_{p,1}) < R_c(\hat{\alpha}_1, \hat{R}_{p,1}) \) still holds and \( \hat{\alpha}_{1,2} = \alpha^*_{1,2} \), i.e. both solutions to (5.23) with different \( R_c \) are still optimal even if \( R_p > \frac{1}{\xi} \). But because \( \frac{\partial^2 R_c(\hat{\alpha}, R_p)}{\partial \alpha^*_2} < 0 \) it is possible at a certain point that \( R_p = R_{p,21} > R_{p,1} \) and we find that \( R_c(\hat{\alpha}_{2}, \hat{R}_{p,21}) = R_c(\hat{\alpha}_{1}, \hat{R}_{p,21}) \). Then we have for the same parameters, i.e. the same equation, two solutions \( \hat{\alpha}_{1,2} \) to (5.23), where \( \hat{\alpha}_2 > \hat{\alpha}_1 \).

In this case \( \hat{\alpha}_1 \) is the inferior solution to (5.23) if \( \alpha^*_2 = \alpha^*_3 \) is optimal. From lemma 5.7, case IIIb), we know that, if \( R_p > \frac{1}{\xi} \), then there exists for every \( R_{p,ij} \) at most one such pair \( \hat{\alpha}_i = \alpha^*_i > \hat{\alpha}_j \). Because of \( \frac{\partial^2 R_c(\hat{\alpha}, R_p)}{\partial \alpha^*_2} < 0 \) and the fact that \( \hat{\alpha}_3 \) is continuous at least in \( [\hat{\alpha}_1, 1] \), we will find an \( \alpha^* > \hat{\alpha}_1 \) to every \( R_p \geq \hat{R}_{p,21} \) with \( R_c(\alpha^*, R_p) = R_c(\hat{\alpha}_1, R_p) \) as long as \( \hat{\alpha}_1 \) is a solution to (5.23). So, \( \hat{\alpha}_1 \) is no longer optimal for any \( R_p \geq \hat{R}_{p,21} \).

So, if we have \( \hat{\alpha}_i > \hat{\alpha}_j \) and \( R_c(\hat{\alpha}_i, R_p) \geq R_c(\hat{\alpha}_j, R_p) \) then \( \hat{\alpha}_j \) can not be optimal. But as long as \( \alpha^*_i > \alpha^*_j \) are both optimal solutions, we have \( R_c(\alpha^*_i, R_p) < R_c(\alpha^*_j, R_p) \). }
5. A.3 Proof of Proposition 5.4

**Proof.** From theorem 2.10 and 2.9 we know that the optimal long-term growth rate is attained if we find max \( \Xi_p(\alpha, x) \) where

\[
\Xi_p(\alpha, x) \triangleq \frac{2}{R_p} \left[ (1 - \alpha)x \hat{\pi}' \Sigma \hat{\pi} + \alpha^2 (1 - \alpha)^2 \Sigma \hat{\pi} + (1 - \alpha) (r - r_0) \right] - (1 - \alpha)^2 x^2 \hat{\pi}' \Sigma \hat{\pi} = \frac{2 \hat{\pi}' \Sigma \hat{\pi}}{R_p} (x - ax^2 + \zeta) - \hat{\pi}' \Sigma \hat{\pi} ((1 - \alpha) x)^2.
\]

To have the conditions for a Pareto improvement we will evaluate the intersection \( \{ (\alpha, x) \mid x = \frac{1}{2a + (1 - a)R_p} \} \cap \{ (\alpha, x) \mid \Xi_p^0(\alpha, x) > \Xi_p(0, \frac{1}{R_p}) \} \).

In the point where \( \alpha = 0 \) the term \( \Xi_p(0, \frac{1}{R_p}) \) becomes

\[
\Xi_p(0, \frac{1}{R_p}) = \frac{\hat{\pi}' \Sigma \hat{\pi} (1 + 2R_p \zeta)}{R_p^2}
\]

The indifference curve through \( \alpha = 0 \) is

\[
\Xi_p^0(\alpha, x) = \Xi_p(0, \frac{1}{R_p})
\]

\[
\frac{2(1 - \alpha)}{R_p} (x - ax^2 + \zeta) - (1 - \alpha)^2 x^2 = \frac{1 + 2R_p \zeta}{R_p^2}
\]

\[
\frac{2(1 - \alpha)}{R_p} (x - ax^2 + \zeta) - (1 - \alpha)^2 x^2 - \frac{1 + 2R_p \zeta}{R_p^2} = 0
\]

The intersection points of \( \frac{2(1 - \alpha)}{R_p} (x - ax^2 + \zeta) - (1 - \alpha)^2 x^2 - \frac{1 + 2R_p \zeta}{R_p^2} = 0 \) and the contract \( x = \frac{1}{2a + (1 - a)R_p} \) are

\[
\alpha_{0,1} = 0, \alpha_{0,2} = \frac{1}{R_p} \frac{R_p^2 \zeta}{(R_p - 2)}
\]

Case \( R_p > 2 \):

Note that \( R_p \geq -\frac{1}{2\zeta} \iff \alpha_{0,2} \geq 1, R_p \leq \frac{1}{\sqrt{-\zeta}} \iff \alpha_{0,2} \leq 0 \) and \( \frac{1}{\sqrt{-\zeta}} < R_p < -\frac{1}{2\zeta} \iff 0 < \alpha_{0,2} < 1 \).

The gradient of \( \text{grad} \Xi_p^0 \) in the point \( (0, \frac{1}{R_p}) \) is \( \text{grad} \Xi_p^0(0, \frac{1}{R_p}) = \left( -\frac{2(1 + R_p^2 \zeta)}{R_p^3} \right) \)

Thus, we have \( R_p \geq -\frac{1}{\sqrt{-\zeta}} \iff -\frac{2(1 + R_p^2 \zeta)}{R_p^3} \geq 0 \) and, therefore, starting from \( \alpha_{0,1} = 0 \) the
utility $\Xi$ increases if $\alpha$ moves into the direction of $\alpha_{0.2}$. Thus, between $\alpha_{0.1}$ and $\alpha_{0.2}$ we have a Pareto improvement.

Finally, we have to examine the characteristics of $\alpha_{0.2}$ in order to give the conditions for a Pareto improvement.

The subcase $R_p \geq -\frac{1}{\sqrt{\zeta}}$ leads always to a Pareto improvement since $\alpha_{0.2} \geq 1$ and in the subcase $R_p \leq -\frac{1}{\sqrt{\zeta}}$ we have no Pareto improvement since $\alpha_{0.2} \leq 0$ and the gradient in $\alpha_{0.1} = 0$ for the indifference curve through $\alpha_{0.1} = 0$, and $\alpha_{0.2} \leq 0$ is negative in $R_p$-direction. Thus, there is no point on the contract that provides a higher utility for the plan sponsor than at $\alpha_{0.1} = 0$.

In the subcase $-\frac{1}{\sqrt{\zeta}} < R_p < -\frac{1}{\sqrt{\zeta}}$ we have a Pareto improvement if $\alpha_{0.1} < \alpha^* < \alpha_{0.2} = \frac{1+R_p^2\zeta}{R_p^2(1-R_p^2\zeta)}$ since the gradient at $\alpha_{0.1} = 0$ of the indifference curve through $\alpha_{0.1} = 0$ and $\alpha_{0.2} \leq 0$ is positive in the $R_p$-direction.

Finally we will use the cubic equation (5.23) to evaluate the value $\bar{R}_e$ for which $\alpha^* < \frac{1+R_p^2\zeta}{R_p^2(1-R_p^2\zeta)}$.

We find that $\alpha^* = \frac{1+R_p^2\zeta}{R_p^2(1-R_p^2\zeta)} \Leftrightarrow \bar{R}_e = \frac{(R_p-2)(1+R_p^2\zeta)}{R_p^2(1-R_p^2\zeta)}$.

Because of corollary 5.3 we have $\frac{d\alpha^*}{d\bar{R}_e} < 0$ and, therefore, $\bar{R}_e > \frac{(R_p-2)(1+R_p^2\zeta)}{R_p^2(1-R_p^2\zeta)} \Leftrightarrow \alpha^* < \frac{1+R_p^2\zeta}{R_p^2(1-R_p^2\zeta)}$.

Case $R_p < 2$:

Because $\zeta > -\frac{1}{4}$ we have $R_p < 2 \Rightarrow R_p < -\frac{1}{\sqrt{\zeta}}$ and $R_p < 2 \Rightarrow R_p < -\frac{1}{\sqrt{\zeta}}$. Thus, $\alpha_{0.2} > 1$ and the gradient of the indifference curve through $\left(0, \frac{1}{R_p}\right)$ i.e. $\text{grad} \Xi_p^0 \left(0, \frac{1}{R_p}\right)$ is negative in the $R_p$-direction. Thus there is no point on the contract that provides a higher utility for the plan sponsor than the point $\left(0, \frac{1}{R_p}\right)$.

5.A.4 Proof of Proposition 5.5

Proof. From theorem 2.10 we know that the optimal terminal wealth is attained if we find $\max \Xi (a, x)$ where $\Xi (a, x) \triangleq \frac{2a}{K_e} (x+\zeta) - a^2 x^2$. The gradient of $\Xi (a, x)$ is

$$\text{grad} \Xi (a, x) = 2 \left( \frac{-ax^2 + \frac{x+\zeta}{K_e}}{\frac{a}{K_e} - x a^2} \right)$$
And the gradient in the contract

\[ \text{grad} \Xi \left( \alpha, \frac{1}{2\alpha + (1 - \alpha) R_p} \right) = 2 \left( -\frac{\alpha}{(2\alpha + (1 - \alpha) R_p)} + \frac{1 + (2\alpha + (1 - \alpha) R_p) \zeta}{R_c (2\alpha + (1 - \alpha) R_p)} \right) \frac{\zeta}{\alpha} \]

If \( \alpha < \frac{R_p}{R_p + R_c - 2} \) then \( \alpha \left( \frac{1}{R_c} - \frac{\alpha}{2\alpha + (1 - \alpha) R_p} \right) \) is strictly positive. Since, because of theorem 2.10, the plan sponsor optimizes the “static utility function” \( \Xi_p (\alpha, x) \) we have \( \frac{\partial \Xi_p (\alpha, x)}{\partial x} = 0 \) on the contract line. Thus, the gradient of \( \Xi_p (\alpha, x) \) is horizontal on the contract line. Therefore, the indifference curves \( \Xi_p (\alpha^*, x) \) and \( \Xi (\alpha^*, x) \) intersect in \( \alpha^* \). Thus, we have \( x_{p,\text{opt}} > x^* \) and, therefore, no Pareto optimum for \( \alpha < \frac{R_p}{R_p + R_c - 2} \).

It remains to show that \( \alpha^* < \frac{R_p}{R_p + R_c - 2} \):
\( \Theta (\alpha) \) is continuous in \( \alpha \in (0, 1) \). Further we have

\[ \Theta' (\alpha) = \frac{2}{R_c} \left( -R_p \frac{\alpha (R_c + R_p - 2) - R_p}{R_p (1 - \alpha) + 2\alpha} \right) \]

and at the point \( \alpha = \frac{R_p}{R_p + R_c - 2} \)

\[ \Theta' \left( \frac{R_p}{R_p + R_c - 2} \right) = \frac{2\zeta}{R_c} < 0 \]

Thus, \( \Theta (\alpha) \) has negative slope in \( \alpha = \frac{R_p}{R_p + R_c - 2} \). Further it is

\[ \Theta'' (\alpha) = \frac{2R_p \left( R_p (2 (R_p - 2) - R_c) - 2\alpha (R_p - 2) (R_p + R_c - 2) \right)}{R_c (R_p (1 - \alpha) + 2\alpha)^4} \]

Thus, \( \Theta (\alpha) \) has only one inflection point and it is for \( \alpha = \frac{R_p}{R_p + R_c - 2} \)

\[ \Theta'' \left( \frac{R_p}{R_p + R_c - 2} \right) = -\frac{2 (R_c + R_p - 2)^4}{R_p R_c^4} < 0 \]

Therefore, \( \Theta (\alpha) \) is concave at \( \alpha = \frac{R_p}{R_p + R_c - 2} \). Since for \( \alpha > \frac{R_p}{R_p + R_c - 2} \) there is only one possibility for an inflection point, and because \( \Theta' \left( \frac{R_p}{R_p + R_c - 2} \right) < 0 \), and \( \Theta'' \left( \frac{R_p}{R_p + R_c - 2} \right) < 0 \),

no solution \( \alpha^* < 1 \) can appear for \( \alpha > \frac{R_p}{R_p + R_c - 2} \). The only possible inner solution has to be smaller than \( \frac{R_p}{R_p + R_c - 2} \). Thus, either we have \( \alpha^* = 0, \alpha^* = 1 \) or \( 0 < \alpha^* < \frac{R_p}{R_p + R_c - 2} \).
Bibliography


Bibliography


Curriculum Vitae

Born on 21th July 1975 in St. Gallen as son of Irène and Hansruedi Baumann; one sister.

1999 – 2004   Doctoral student, Department of Economics,  
              University of St. Gallen (HSG)

1997 – 2004   Research and teaching assistant, Mathematics and Statistics,  
              University of St. Gallen (HSG)

2002   Research semester, Department of Statistics,  
       London School of Economics and Political Science (LSE)

1995 – 1999   Study in Economics, University of St. Gallen (HSG)

1995   Matura Typus C, Kreuzlingen