Four Essays in Equity-Linked Life and Pension Insurance:

Financial Analysis of Surrender Guarantees, Pension Guarantee Funds and Pension Retirement Plans

Inaugural–Dissertation
zur Erlangung des Grades eines Doktors
der Wirtschafts– und Gesellschaftswissenschaften
durch die
Rechts– und Staatswissenschaftliche Fakultät
der
Rheinischen Friedlrich–Wilhelms–Universität Bonn

vorgelegt von

Filip Uzelac

aus Skopje (Mazedonien)

BONN 2014

Dekan:	Prof. Dr. Sandmann
Erstreferent:	Prof. Dr. Chen
Zweitreferent:	Prof. Dr. Sandmann
Tag der mündlichen Prüfung:	02.09.2014

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn http: // hss.ulb.uni–bonn.de/ diss_online elektronisch publiziert.

Acknowledgments

First and foremost I would like to thank my first supervisor Prof. Dr. An Chen for her continuous support during my entire PhD studies at the Bonn Graduate School of Economics. I have acquired so much useful and indispensable knowledge in mathematical finance and insurance from her. She brought me into the interesting research area of life and pension insurance and considerably contributed to two papers on which chapters 3 and 5 are based on. Moreover she taught me how to approach and develop research questions, how to write a paper and how to give a good research talk. Most importantly, she gave me much guidance at the beginning of my PhD studies, at the time where I barely had experience with writing a research paper, she always found time to discuss research ideas and also encouraged and motivated me to be persistent during my studies. I also appreciate very much that she gave me moral support and encouragement at those times where I lacked motivation and was a bit frustrated. All in all, I can say that I could not have writtent this dissertation without your support An! Thank you so much for all!

I also want to thank my second supervisor Prof. Dr. Klaus Sandmann. He gave me the opportunity to stay in the Department of Banking and Finance (now Institute of Financial Economics and Statitics) and supported me with productive discussions and valuable advice at any time during my PhD studies. I am also deeply grateful to my (co)author and mentor, Prof. Dr. Alexander Szimayer. I have learnt so much about financial mathematics in his intriguing lectures during my master studies. Moreover he quickend up my research interest in regime-switching models which resulted in a paper on which the first chapter is based on. Prof. Szimayer not only supported me during this collaboration, but he always gave me valuable advice and comments during my entire studies. I also would like to thank Prof. Dr. Stefan Ankirchner, whose lectures and topics courses have helped me to gain important technical knowledge and with whom I have also had interesting and helpful discussions.

I would like to give my gratitude to all the colleagues in the Department of Banking and Finance (Institute of Financial Economics) of the University of Bonn: Prof. Dr. Hendrik Hakenes, Prof. Dr. Rainer Haselmann, Prof. Dr. Jörg Budde, Christian Hilpert, Sebastian Ebert, Birgit Koos, Friedericke Schlegel, Franz Gödecke, Jing Li, Anne Ruston, Almut Lunkenheimer, Stefan Lehner, Cornelius Veith, Michael Hewer, Eva Schliephake and all student assistants at our department and I would also like to thank all my collegues from the Bonn Graduate School of Economics. I would like to particularly thank Franz Gödecke, with whom I spent the exchange year in Berkeley and a lot of time during the first two years at the BGSE and who was a very good discussant and who became a close friend, Christian Hilpert with whom I spent a lot of time on conferences and with whom I had a lot of helpful discussions, Sebastian Ebert for many good remarks and comments on my research, Chunli Cheng with with whom I have worked on a project chapter 4 is based on and Stefan Lehner, with whom I shared an office for almost 2 years and who has become a good friend of mine.

I sincerely thank Prof. Dr. Urs Schweizer and Dr. Silke Kinzig for the very good organizational support and the entire BGSE organisation for offering me outstanding research opportunities. I acknowledge the financial support from the German Research Council (DFG) through the Bonn Graduate School of Economics. I would like to further thank Prof. Dr. Allois Kneip for the organizational support of my studies at the UC Berkeley during my first year as a PhD student. In this case I also acknowledge the additional financial support from the Hausdorff Center for Mathematics (HCM) Bonn.

I would like to express my deepest gratitude to my whole family, especially to my girlfriend Linda, my parents Rajko and Ljubica and my brother Igor, for their unconditional love, faith in me and moral support during my entire PhD studies. Last, but not least I would like to thank all my friends for their entire support, encouragement and faith in me.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Concepts

This dissertation deals with three important types of insurance: equity-linked life, pension insurance and (pension) insurance guarantee funds.

Standard life insurance contracts, like a term or endowment insurance, provide either survival benefits or death benefits or both. In equity-linked products these benefits are linked to the performance of a portfolio, which is usually set up by the insurance company and which consists of one or several underlying assets. Accordingly, these products offer the policyholder the opportunity to participate in the financial markets. Another essential feature of equity-linked life insurance products is that they are usually equipped with certain guarantees, which make up a substantial fraction of the life insurer's liabilities. The two main guarantee types are the minimum guarantee, where the policyholder is for instance offered a guaranteed interest rate for his investment, and the surrender guarantee, which entitles the policyholder to receive a certain cash amount when terminating the insurance contract. In this dissertation both guarantee types are considered, where a particular emphasis is put on the surrender guarantee or surrender option.

A (private) pension insurance contract is a contract between the employee, the policyholder, and the employer, the insurer or often also called the sponsoring company. The two major types of private pension insurance are the Defined Benefit (DB) and the Defined Contribution (DC) pension plan. In a standard DC plan [1] the contributions are fixed. Contributions are usually paid by both the employee and the employer (monthly or quarterly) and usually constitute a constant percentage of the employee's salary. These contributions are paid to a (external) pension fund, which invests the contributions on behalf of the employee in financial assets. The pension payment is then determined as the market value of these backing assets. Accordingly, like in equity-linked life insurance products the benefits also crucially depend on the performance of a financial portfolio, thus on the earned

[1] We use the term standard in order to distinguish between those pension plans with hybrid features, which also exist in practice.

investment returns. However, the main difference is that the benefits further particularly depend on the evolution of the salary of the insured. In a standard DB pension plan the benefits are predefined, or more precisely the formula which specifies the benefits is known in advance. This formula mainly takes the years of service, age and the salary of the employee into account. As in DC plans a (external) pension fund is set up to manage the pension obligations. The main difference to the DC plan is that the contributions provided to the pension fund are variable [2] and that only the employer earns the surpluses of the pension assets. That is, unlike equity-linked life insurance products and DC pension plans the benefits of the employees in the DB plan are not directly linked to the performance of a financial portfolio. Althogh the pension payments differ, the DB and the DC plan have in common that the employee's salary is the main determinant of the benefits.

The main objective of an insurance regulator is to protect policyholders by ensuring that promised benefits are paid to the latter, particularly in the case when the insurer is in financial distress or insolvent. Similar to the banking industry, where depositors are additionally protected by a deposit insurance, insurance guarantee funds have been set up as a protection vehicle in many countries as well for the life and nonlife insurance field. [3] Insurance companies are often obliged to enter into these government-imposed protection schemes. From the policyholder's perspective insurance guarantee funds can be considered as a reinsurance. Specifically, the insurance companies (the employer in case of a pension insurance) pay premiums to the insurance guarantee funds, which invest the premiums under certain regulatory rules. If the insurance company is not able to pay the promised benefits either due to financial distress or insolvency the insurance guarantee fund steps in, takes the residual assets of the insurance company and provides the payments of the benefits up to certain limits.[4] In this thesis we consider a pension guarantee fund for private DB pension plans. The essential difference to other insurance guarantee funds is that the pension guarantee fund only pays benefits if both the pension fund and the sponsoring company are in financial distress.

1.2 Methodology

The term financial analysis encompasses in this dissertation the pricing of financial guarantees in insurance, the (constrained) expected utility optimization and the expected utility comparison.

In chapter 2 and chapter 3 we consider the corresponding insurance contracts as con-

[2] In private sector DB plans the contributions are mainly made by the employer and are often deficit contributions, which means that the employer provides payments to the pension fund if the market value of pension assets falls below the pension liabilities.

[3] For a detailled list of such insurance guarantee funds confer to Schmeisser and Wagner (2013).

[4] The benefit payments are usually capped, see for instance chapter 3

tingent claims and apply option pricing techniques in a specific stochastic financial market model to determine the market consistent value of the insurance contracts. For the equity-linked life insurance contract we particularly extract the market consistent values of the different guarantees, focussing on the surrender option value. An accurate pricing or market consistent valuation of such guarantees is of paramount importance for any life insurance company, not only for risk management, but also for regulatory purposes. For example in Europe the new regulatory regime Solvency II requires insurers to disclose the market values of their guarantees on the liability side of their balance sheets. In particular, studying surrender more closely is important since for instance European Union regulators have identified surrender risk as the main risk driver for life insurance companies after the interest rate risk. In fact, the recent Quantitative Impact Study (QIS 5) showed that surrender risk is the most important risk among life underwriting risks, see EIOPA (2011). For the pension guarantee fund insurance the market consistent value of the insurance contract is interpreted as the fair initial premium sponsoring companies should pay to the guarantee fund for providing the pension insurance. The premium is particularly risk-based since it takes both pension fund and sponsor risk into account. Consequently, such a premium does not give pension funds and sponsoring companies adverse incentives to introduce risk into the pension system. Another mechanism, which protects employees in DB plans, is that the pension guarantee fund can itself prematurely terminate underfunded DB plans. We study this mechanism in chapter 4 by applying a different methodology. We compute critical funding ratios under which the pension guarantee fund prematurely terminates the insured underfunded DB plan in a constrained expected utility optimization model. In chapter 5 we study the underlying DB contract more closely and compare it to the DC pension contract. As in the previous chapter, we assume that the policyholder has certain preferences and that he can choose between a DB and a DC pension plan. The policyholder faces a tradeoff between different types of risk, which are (more) present in one type of the pension plan than in the other. Finally, we compare the pension plans by comparing the expected utility the policyholder can achieve in either pension plan at the retirement date.

More specifically, in chapter 2 we closely study the valuation of a stylized equity-linked life insurance contract with a surrender guarantee. In order to obtain accurate market consistent values for such a contract one has to specify a model that adequately captures the dynamics of the financial portfolio and that models the surrender behavior of the policyholder in a realistic manner. The second requirement is not compatible with the standard assumption of a rational agent in the financial literature. In our context this is a policyholder who behaves monetary optimal in terms of only surrendering the contract when it is financially optimal to do so. A large body of the behavioral economics literature rejects this assumption. More importantly, the empirical literature in the life insurance field shows that policyholders surrender due to both exogenous and endogenous reasons. Exogenous reasons are mainly personal reasons, often driven by the own financial distress, while endogenous reasons are financial factors which make it monetarily optimal to surrender the contracts at appropriate moments. We model both exogenous and endogenous surrender in an intensity-based approach, where each type of surrender is modeled with a corresponding surrender intensity. This modeling framework is also referred to as the rational expectation framework, see e.g De Giovanni (2010). Moreover, an essential feature of (equity-linked) life insurance products is that they usually have very long maturities. One important consequence is that in the long run economic conditions typically change several times. These changing economic conditions affect both the dynamics of the financial portfolio, in terms of the market interest rate and its volatility, and also both the exogenous and endogenous surrender behavior of the policyholder. To capture this important empirically observable fact, we introduce a regime-switching model where the different regimes represent different economic states.

In chapter 3 we develop a risk-based premium calculation model for the insurance provided by the largest pension guarantee fund, which is managed by the US Pension Benefit Guarantee Corporation (PBGC). Although our analysis focuses on the american pension insurance mechanism, the modeling framework can be readily applied to any government-imposed pension guarantee fund. [5] More importantly, the qualitative results hold for any insurance guarantee fund in general. One crucial problem most insurance guarantee funds share is that their premium calculation is not risk-based. Specifically, most insurance guarantee funds charge either a flat or a volume-based premium, see chapter 3 and Schmeisser and Wagner (2013). Such a premium calculation practice is not reasonable from the economic perspective since it gives insurers the incentive to invest more riskily in order to increase the market value of equity. A consequence of this adverse incentive is that less risky insurers cross subsidize more risky ones. The most severe consequence of such an ineffective premium calculation is that the insurance guarantee fund's financial status might deteriorate and become so poor that it could not provide the required payments to policyholders if some companies are in financial distress. Our calculation model for a pension guarantee fund insurance takes account of the pension fund's and sponsoring companie's investment policy and assumes that these are correlated. Moreover, it also takes account of the fact that the pension guarantee fund has to provide payments only if both the pension fund and the plan sponsor are in financial distress, therefore we model this insurance as a residual or secondary guarantee. In addition, our model allows for a further realistic perspective that the pension fund can be terminated prematurely. Most importantly, this premature termination is triggered by the poor financial status of the plan sponsor, which is another realistic perspective since most pension fund terminations in practice are due to the sponsoring companie's underfunding. We study our pricing formula both theoretically and particularly empirically by comparing the risk-based premiums for the largest American DB plan sponsors.

In chapter 4 we study the other type of termination where an insured underfunded DB pension plan is closed by the pension guarantee fund. This type of termination is called

[5] Such pension guarante funds exist for instance in Canada, the U.K., Japan, Germany, Switzerland or Sweden

involuntary termination. In this chapter we do not primarily focus on the US pension mechanism, although the US and Canadian pension guarantee funds serve as good examples for our analysis. We propose a premature termination rule for a pension guarantee fund to manage its financial guarantee and to protect policyholders. To this end we determine an optimal termination ratio for an insured defined benefit (DB) pension plans in terms of a critical funding ratio under which the pension fund is prematurely closed by the pension guarantee fund. In our model the guarantee fund pursues a social welfare motive and acts in the interests of the pension beneficiaries by maximizing their expected utility. To better manage its financial guarantee and to better protect the policyholders the pension guarantee fund imposes two constraints for the insured DB plan: a shortfall probability constraint (SPC) where the guarantee fund defines a maximum one-year shortfall probability and an expected shortfall constraint (ESC) where the guarantee fund predefines the maximum one-year expected loss size of underfunded but not terminated DB plans. We then solve this constrained one-year expected utility maximization problem and study how the regulatory constraints and particularly the risk aversion of the beneficiaries affect the optimal intervention rule.

In chapter 5 we compare the DB and the DC pension plan from the policyholder's perspective in a continuous time expected utility framework. In this framework we take the essential tradeoff the policyholder faces when opting for one type of pension retirement plan into account, that is the tradeoff between salary, asset price (investment risk) and portability risk. As we described in chapter 1.1 the employee's salary is the crucial component in determining the benefits in either retirement plan, thus salary risk is present in both the DB and DC plan. Moreover, the policyholder only bears investment or asset price risk in a DC plan since in a DB plan the employee's benefits are not directly linked to the investment returns of the pension fund, thus only the employer participates in the surplus of the pension fund, but also bears the entire investment risk. On the other hand portability risk, the risk to lose parts of the benefits when changing the employer, is mainly present in the DB plan because the benefits of the DC plan are determined as the market value of the backing assets, which are usually transferable from one employer to another. In this chapter we govern these main risk factors in a model with stochastic wages, stochastic job moving and stochastic asset prices. We compare the DB and the DC plan by mainly computing the indifference job switching intensity, that is the intensity which makes the policyholder equally well off in expected utility terms in both pension plans. From this quantity we infer the average number of job moves after which a DC pension plan is preferred as an interesting statistic.

Chapter 2

Valuation of Equity-Linked Life Insurance Contracts with Surrender Guarantees in a Regime-Switching Rational Expectation Model[1]

2.1 Introduction

The surrender guarantee is a bonus right, which is included in most equity-linked life insurance contracts. It gives the policyholder the opportunity to receive a certain cash refund when walking away from the contract. The surrender decision of a policyholder can be of two types. He can surrender exogenously due to personal reasons or endogenously by observing that fluctuations in the financial environment make a surrender profitable from a monetary point of view. Exogenous surrender is mainly explained by the emergency fund hypothesis conjecturing that personal financial distress, especially unemployment forces the policyholder to terminate the contract. The interest rate hypothesis, which states that policyholders lapse their contract to exploit rising market opportunities in terms of gaining higher interest rates from alternative investments is the main hypothesis accounting for endogenous surrender.

The valuation approaches of equity-linked life insurance products with surrender guarantees differ depending on which type of surrender is incorporated. The purely exogenous surrender valuation approach by Bacinello (2003), where surrender rates are estimated from historical lapse data, is at odds with the interest rate hypothesis and the estimated funds needed to support the contract are too low on average. The most used approach is the purely financial approach by Grosen and Jorgensen (2000), Bacinello (2005) and others where the surrender option is modelled as an American style put option and its market value is obtained by solving an optimal stopping problem. This approach completely neglects ex-

[1]This chapter is based on Uzelac and Szimayer (2014)

ogenous surrender and it further relies on the unrealistic assumption that policyholders make optimal surrender decisions. As a consequence the purely financial approach considerably overestimates the funds needed to support the contract. A more realistic valuation approach is done in the rational expectation framework allowing for both exogenous and endogenous surrender. This was first applied by Albizzati and Geman (1994) and was more recently studied by De Giovanni (2010) and Li and Szimayer (2014) in an intensity-based approach. In particular, the empirical study of Kuo, Tsai and Chen (2003) supports the rational expectation model as it shows that both the emergency fund hypothesis and the interest rate hypothesis are significant, the first is statistically more significant while the second is economically more significant. Consequently, these findings suggest that an accurate pricing model of equity-linked products with surrender guarantees should model both types of surrender.

Our main contribution is that we refine the rational expectation framework of De Giovanni (2010) and Li and Szimayer (2014) by additionally including a regime-switching model. In our model both the dynamics of the reference fund as well as exogenous and endogenous surrender rates are linked to the corresponding economic state. The switching behavior of the economic states can be attributed to structural changes in the economic conditions, changes in business conditions, changes in political situations, the impact of economic news (financial or macroeconomic) and business cycles. The inclusion of a regime-switching model is of practical importance as most equity-linked contracts with surrender guarantees are relatively long-dated compared with financial products. There can be substantial fluctuation in economic variables, which affect as well the market value of the reference fund as both exogenous and endogenous surrender decisions of the policyholders, over a long period of time. In a numerical illustration, we consider 2 states and interpret them as the business cycles recession and expansion. The example is further supposed to show that our model can for instance incorporate both the emergency fund and the interest rate hypothesis. By imposing a higher exogenous surrender rate in the recession state we can incorporate the emergency fund hypothesis because personal financial distress is more likely to occur in that state. Combining a procyclical endogenous surrender intensity with a procyclical risk-free rate we can also incorporate the interest rate hypothesis since the hypothesis states that endogenous lapse should be higher in states with higher interest rates.

Regime-switching models have become popular in actuarial science in recent years and are particularly recommended by the American Academy of Actuaries and the Canadian Institute of Actuaries, see Hardy (2001). The use of regime-switching models for pricing and hedging long term guarantee products have been popularized by Hardy (2001), who successfully fitted the model to monthly data from the Standard and Poor's 500 and the Toronto Stock Exchange indices using a discrete time regime-switching lognormal model. Concerning the pricing of equity-linked life insurance products with surrender options, thus far only Siu (2005) employed a Markov regime-switching model in a Black-Scholes Merton economy, that is the volatility and the risk-free rate of the reference fund depend on the states of the economy, but are constant across states. However he also uses the less realistic purely financial approach and provides approximate solutions for the free boundary value problem. This chapter proposes the first valuation model for equity-linked products with surrender guarantees, which does not only use regime-switches to model the evolution of the market value of the reference fund, but most importantly, we also model the lapsing dynamics more realistically by allowing the exogenous and endogenous surrender intensities to change over time according to the evolution of the economy.

More specifically, we use a Markovian regime-switching model where the economic states change according to the evolution of a continuous-time observable Markov chain. The reference fund is modelled as a Markov modulated diffusion process. The market incompleteness resulting from the nontradability of the regime-switching risk is resolved by specifying a unique martingale measure with the well-known Esscher transform. To model the surrender action of a representative policyholder we follow the intensity-based approach of Li and Szimayer (2014). We also assume that the surrender intensity is bounded from below and from above. The lower bound is given by the exogenous surrender intensity and represents the rate of monetary suboptimal surrender. The upper bound represents the maximal surrender rate that is attributed when it is financially optimal to do so and is given by the sum of the endogenous and exogenous surrender intensity. This way of modeling the surrender intensity implies that the surrender intensity is a function of the contract value. Unlike the model of Li and Szimayer (2014) the surrender intensity is regime-dependent in our model, since both surrender intensities are also modulated by the continuous-time Markov chain. Moreover, the American style surrender model is included in our setup as a special case when the exogenous surrender intensity is set to zero and the endogenous surrender intensity to infinity in either regime. To find the value of the insurance contract in the corresponding state of the economy we establish a system of two coupled partial differential equations (PDEs). This PDE system is nonlinear since the surrender intensities need to be determined simultaneously with the contract values. The solution of this penalty problem is obtained numerically by combining the Crank-Nicolson scheme with the penalty scheme of Dai and You (2007) using a generalized Newton search algorithm proposed by Forsyth and Vetzal (2002).

The rest of the chapter is structured as follows. In the next section we motivate the use of a regime-switching model, especially for our application in the numerical analysis. In section 2.3, we describe the model and the contract under consideration. The valuation of the contract is carried out and theoretical comparative statics results are derived. In section 2.4 we explain the numerical methodology to compute the contract values. In section 2.5 we study the contract closely through numerical examples. As already mentioned, we assume in this experiment that the economic states are business cycles. We focus on the computation of the value of the surrender guarantee and also compare our regime-switching rational expectation model with the American style surrender model. Section 2.6 concludes.

2.2 Some Empirical and Economic Motivation

In this section we empirically motivate a regime-switching model for our numerical example where the two economic states are the business cycles recession and expansion. To this end, we outline that stock market volatility tends to evolve countercyclically, while the risk free rate tends to evolve procyclically. In addition, we motivate the emergency fund hypothesis and the interest rate hypothesis.

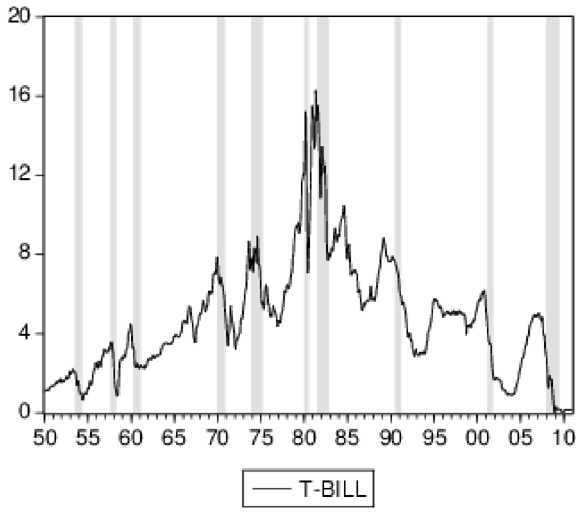

Figure 2.1: Monthly 3-month Treasury bill secondary market rates in % from 1950-2010, recession periods according to the NBER Business Cycle data are highlighted.

To highlight the cyclical behavior of the risk-free rate we use monthly 3-month US Treasury bill data from the US financial data base. In figure 1 we compare the interest rates on a 3-month US Treasury bill with the NBER Business Cycle data. The period from 1961 until 1975 outlines that the risk-free interest rate can behave countercyclically because the highest interest rates were observed in the two recessions in this period, which would confirm the results of King and Rebelo (1999). Nevertheless, in general we clearly observe that the interest rates substantially increase during an expansion and substantially decrease during a recession. More importantly, except the period mentioned above, the T-bill rates are on average significantly higher in the expansion period preceding a recession. We conclude that there is a tendency for a procyclical behavior and this tendency is especially clearly observed in the long period from 1976-2010. A procyclical evolvement of the risk free rate is also suggested by the majority of research studies, see for instance Blanchard and Watson (1986), Fama and French (1988) or more recently Ang and Bekaert (2002a) and Ang and Bekaert (2002b). A plausible argument for a procyclical risk-free rate is provided by Jouini and Napp (2011) arguing that during bad states of the world there is a pessimistic bias in the economy, whereas during good states of the world there is an optimistic bias.

To make some statements about the cyclical behavior of the volatility, we estimate the

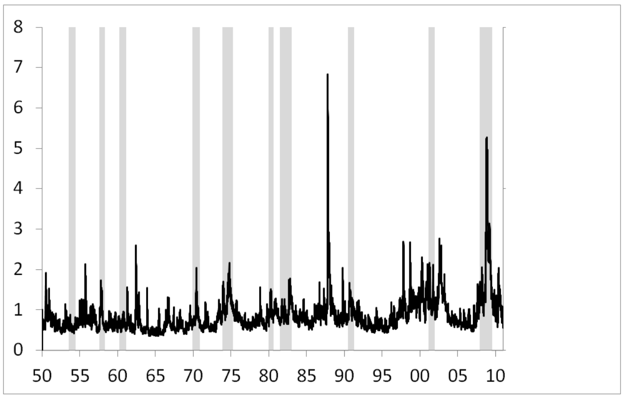

Figure 2.2: Estimated volatility of the S&P-500 daily log returns in % from 1950-2010 with a GARCH(1,1) model, recession periods according to the NBER Business Cycle data are highlighted.

volatility of the S&P 500 daily log returns from 1950 until 2010 with a GARCH(1,1) model, and then again compare the volatility estimates with Business Cycle data. The estimated volatility is depicted in figure 2. Having a closer look on the volatility and neglecting outliers like the stock market crash in 1987, we observe that the S&P 500 volatility has a clear tendency to evolve countercyclically. Volatility rises significantly in longer recession periods like those from 1973-1975 or the recent financial crises from 2007-2009. On the other hand, the volatility stays at a relatively low level during expansion periods like those from 1975-1980, 1983-1987 and especially the more recent expansion periods from 1991-1998 and 2003-2007. This observation seems to be fairly pronounced during the last 30 years. A countercyclical stock market volatility is empirically confirmed by many researchers, see for instance Schwert (1989) and Engle, Ghysels and Sohn (2008). One economic explanation of this stylized fact is that monetary tightening causes both a recession and increases stock market volatility.

We follow Dar and Dodds (1989)and Kuo et al. (2003) and consider the unemployment rate as the main driver of the emergency fund hypothesis. We collect a data set of voluntary termination rates in life insurance contracts from the American Life Insurance Council and acquire a sample of US unemployment rates from the Labour Force Statistics.[2] The evolution of both time series is shown in figure 3. The figure illustrates that the voluntary termination rates significantly depend on the unemployment rates because both time series have a tendency to move in the same direction. This observation, which motivates the emergency fund hypothesis, is very pronounced during the long time horizon from 1977-2009. More importantly, both time series are countercyclical as we observe that in general they are substantially higher in recession periods preceding an expansion period and the peaks are often reached in recession periods. Hence there is a clear tendency that the emergency fund hypothesis is strongly related to the state of the economy suggesting to model exogenous

[2] As in Kuo et al. (2003) there two typical limitations in the data set. First we cannot distinguish between pure lapse and surrender and second the lapse rates are obtained from different types of insurance products.

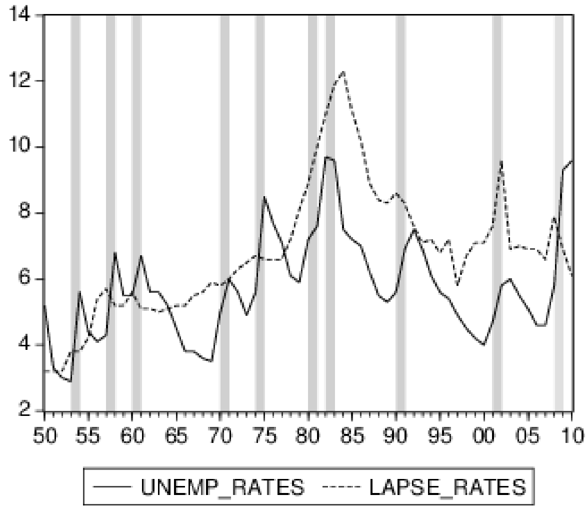

Figure 2.3: US annual unemployment rates in % (dashed), US annual voluntary termination rates in life insurance contracts in % (solid) from 1950-2010, recession periods according to the NBER Business Cycle data are highlighted.

surrender as a function of the macroeconomic process.

Kuo et al. (2003) show within a cointegration analysis that the interest rate hypothesis is statistically less significant than the emergency fund hypothesis. However, they point out that the interest rate is economically more significant. In an impulse response analysis they find shocks in the lapse rates respond substantially to shocks in the interest rate, whereas the response to shocks of the unemployment rate is small. This result and the result of procyclical risk-free rates gives rise to also incorporate the interest rate hypothesis into a valuation model for an insurance product with a surrender guarantee.

2.3 Model

2.3.1 Regime-Switching, Financial and Insurance Risk

The starting point of our setup is an observable continuous-time two-state Markov chain $X = (X_t)_{t\geq 0}$ modelled under the real world measure $\mathbb{P}$ on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, that is used to describe the state of the economy. For the ease of exposition we limit ourselves to the two-state case and mention that the subsequent analysis can be readily extended to more states. The main argument for an observable Markov chain is that the market interest rates are observable and the volatility of the reference fund can be estimated with the quadratic variation of its logarithm, which is also observable, see Erlwein, Mamon and Siu (2008).

The state space of this stochastic process can be conveniently described by the canonical representation $\mathcal{S} = (e_1 \quad e_2)$, with the two unit vectors $e_1 = (1 \quad 0)'$ and $e_2 = (0 \quad 1)'$, with $'$

denoting the transpose, see Buffington and Elliot (2002). The two-state Markov chain generates the filtration $\mathbb{F}^X = (\mathcal{F}_t^X)_{t\geq 0}$.

The dynamics of our economy are described by the transition matrix $\mathbf{P}(t) = (P_{ij}(t))_{i,j=1,2}$ of X, with $P_{ij}(t) = \mathbb{P}(X_t = e_j | X_0 = e_i)$, $i, j = 1, 2$, $j \neq i$ and $t \geq 0$. The transition dynamics are governed by Kolmogorov's forward equation

$$\frac{\mathrm{d}\mathbf{P}}{\mathrm{d}t}(t) = \mathbf{P}(t)\, A\,, \quad \mathbf{P}(0) = I_2\,, \tag{2.1}$$

where I_2 is the identity matrix in dimension 2 and the generator A is given by

$$A = \begin{pmatrix} -\eta_1 & \eta_1 \\ \eta_2 & -\eta_2 \end{pmatrix}. \tag{2.2}$$

The interpretation is that η_i denotes the intensity with which the economy jumps from state i to state j, $i = 1, 2$ and $i \neq j$. In other words, during any time interval $\mathrm{d}t$, there is a time-invariant probability $\eta_i\, \mathrm{d}t$ that the process X changes from state i to state j. The expected time the economy stays in state i is given by η_i^{-1}. The generator A yields the following decomposition for the stochastic differential of X $\mathrm{d}X_t = A'\, X_t\, \mathrm{d}t + \mathrm{d}M_t$, where M is an $\mathbb{F}^X$-martingale, see Elliot, Aggoun and Moore (1994).

The financial market consists as usually of a risky non-dividend paying asset with a price process $S = (S_t)_{t\geq 0}$ and a riskless money market account with a price process $B = (B_t)_{t\geq 0}$. Under the real world measure $\mathbb{P}$, the stochastic processes are governed by the stochastic differential equations

$$\mathrm{d}B_t = r(t, X_t)B(t)\, \mathrm{d}t, \text{ for } 0 \leq t \leq T\,, B_0 = 1\,, \tag{2.3}$$

$$\mathrm{d}S_t = a(t, S_t, X_t)S(t)\, \mathrm{d}t + \sigma(t, S_t, X_t)S(t)\, \mathrm{d}W_t, \text{ for } 0 \leq t \leq T\,, S_0 \in \mathbb{R}^+\,, \tag{2.4}$$

where r denotes the possibly regime-dependent risk-free interest rate, a is the possibly regime-dependent local rate of return and σ is the possibly regime-dependent volatility.[3] The latter can be expressed more compactly by

$$r_t = r(t, X_t) = (r_1(t)\, r_2(t))\, X_t\,, \quad \text{and} \quad \sigma_t = \sigma(t, S_t, X_t) = (\sigma_1(t, S_t)\, \sigma_2(t, S_t))\, X_t\,,$$

for $0 \leq t \leq T$, where $r_i(t)$ and $\sigma_i(t, S_t)$ refer to the risk-free interest rate and volatility in regime i, for $i = 1, 2$. Note that the risk-free rate is deterministic within a regime. Moreover W refers to the standard Brownian motion under $\mathbb{P}$, that is independent of X and generates the filtration $\mathbb{F}^W = (\mathcal{F}_t^W)_{t\geq 0}$. The extended financial market filtration is then given by $\mathbb{F}^W \bigvee \mathbb{F}^X$, which we call $\mathbb{Z} = (\mathcal{Z}_t)_{t\geq 0}$.

[3]The model is formally written as a local volatility regime-switching model. If we make the simplifying assumptions that a,r and σ are constant within a regime, see section 5, then we have a regime-switching lognormal financial market model.

The financial market is free of arbitrage and incomplete, since the inclusion of the regime-switching risk leads to an additional source of risk that is not traded. Mathematically speaking, there exist infinitely many equivalent martingale measures. In general, there are two ways to resolve the market incompleteness in our model. We could either add a so called change of state security as an additional asset to complete the market, see Guo (2001) for details, or we have to select a specific martingale measure. We opt for a well-established approach in regime-switching models and select the Esscher martingale measure relying on the Esscher transform as the unique martingale measure. It can be shown that the martingale measure specified by the Esscher transform is the one that maximizes expected power utility, see Elliot, Chan and Siu (2005). Now, following Elliot et al. (2005) we briefly sketch the Esscher transform. First, define the regime-switching Esscher process $\theta = (\theta_t)_{t\geq 0}$ by $\theta_t = \theta(t, X_t, S_t) = (\theta_1(t, S_t)\, \theta_2(t, S_t))X_t$. Then, the regime switching Esscher transform $\mathbb{Q}_\theta \sim \mathbb{P}$ on $\mathcal{Z}_t$ is given by

$$\left.\frac{\mathrm{d}\mathbb{Q}_\theta}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{Z}_t} = \frac{\exp\left(\int_0^t \theta_u \mathrm{d}W_u\right)}{\mathbb{E}^P\left[\exp\left(\int_0^t \theta_u\, \mathrm{d}W_u\right)\Big|\, \mathcal{F}_t^X\right]}, \quad t \geq 0,$$

where $\mathbb{E}^P$ denotes the expectation taken under the real world measure. The Radon-Nikodym derivative of the Esscher transform can be expressed as

$$\left.\frac{\mathrm{d}\mathbb{Q}_\theta}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{Z}_t} = \exp\left(\int_0^t \theta_u\, \mathrm{d}W_u - \frac{1}{2}\int_0^t \theta_u^2\, \mathrm{d}u\right), \quad t \geq 0. \tag{2.5}$$

The risk-neutral regime-switching Esscher parameter $\tilde{\theta}(t, X_t, S_t)$ solves the martingale condition

$$S_t = \mathbb{E}^{\mathbb{Q}_{\tilde{\theta}}}\left[\exp\left(-\int_t^T r_u\, \mathrm{d}u\right) S_T \,\Big|\, \mathcal{Z}_t\right], \quad t \geq 0.$$

It can be shown that the above martingale condition implies that

$$\tilde{\theta}(t, X_t, S_t) = \frac{r(t, X_t) - a(t, X_t, S_t)}{\sigma(t, X_t, S_t)}, \quad t \geq 0. \tag{2.6}$$

Plugging $\tilde{\theta}$ in (2.5) we get the Radon-Nikodym derivative of $\mathbb{Q}_{\tilde{\theta}}$ with respect to $\mathbb{P}$. Then, by Girsanov's theorem one has $W^{\tilde{\theta}} = W - \int_0^{\cdot} \tilde{\theta}_t\, dt$ is a standard Brownian motion under the martingale measure $\mathbb{Q}_{\tilde{\theta}}$, which is called the Esscher martingale measure. Finally we end up with the usual risk-neutral dynamics of the reference fund S, i.e.

$$\mathrm{d}S_t = r(t, X_t)S_t\mathrm{d}t + \sigma(t, S_t, X_t)\, S_t\, \mathrm{d}W_t^{\tilde{\theta}}, \quad t \geq 0, \tag{2.7}$$

where $W^{\tilde{\theta}}$ is a standard Brownian motion under $\mathbb{Q}_{\tilde{\theta}}$. By construction of the change of measure X and $W^{\tilde{\theta}}$ are independent under $\mathbb{Q}_{\tilde{\theta}}$ and the probability law of X remains unchanged under the change of measure.

The insurance market is modelled by the jump process $H = (H_t)_{t\geq 0}$, with $H_t = 1_{\{\tau\leq t\}}$, for $0 \leq t \leq T$, where the random variable τ denotes the arrival of the death event of an individual aged y at $t = 0$. The jump process H generates the filtration $\mathbb{H} = (\mathcal{H}_t)_{t\geq 0}$ and has intensity μ, which is called the mortality intensity. We follow the rationale of Li and Szimayer (2014) and abstract from stochastic longevity since this plays a minor role for equity-linked products with a similar payoff structure at death and survival. In addition we make the assumption that, unlike the financial market, the insurance market is not modulated by the Markov chain X. We do not include mortality regimes in our framework since clear evidence that mortality risk is significantly linked to economic regimes has not been found yet in the literature. In fact, the mortality intensity is a regime-independent deterministic function of time and thus the mortality risk is unsystematic. This implies that we can work under the suitably extended Esscher martingale measure on the enlarged filtration $\mathbb{G} = \mathbb{Z} \bigvee \mathbb{H}$, where μ is the $(\mathbb{Q}_{\tilde{\theta}}, \mathbb{G})$-intensity of the jump process H, A is the $(\mathbb{Q}_{\tilde{\theta}}, \mathbb{G})$-generator of the Markov chain X, and $W^{\tilde{\theta}}$ is a $(\mathbb{Q}_{\tilde{\theta}}, \mathbb{G})$-standard Brownian motion, see Ch. 6 of Bielecki and Rutkowski (2004) for details.

2.3.2 Insurance Contract and Surrender Action

The contract we consider is a stylized equity-linked life insurance contract with a simple roll up minimum guarantee at death or survival and a surrender guarantee. The survival and the death benefit both entitle the policyholder to additionally participate in a profitable development of the risky asset. The surrender benefit is independent of the asset performance and depends on time only, as in Bernard and Lemieux (2008). We further assume that the policyholder pays a single premium P at the beginning of the contract with a maturity date T, which is a reasonable assumption since most equity indexed annuities contain a single premium payment, see Palmer (2006). When the policyholder survives time T and the contract is still active, the payment to him is

$$\Phi(S_T) = \alpha\, P(1+g)^T + \alpha\, P \,\max\left(\left(\frac{S_T}{S_0}\right) - (1+g)^T, 0\right), \tag{2.8}$$

where α denotes the percentage the initial premium is the provided with the minimum guaranteed rate g and the policyholder participates in the performance of the underlying asset. For an active contract the policyholder receives the death benefit

$$\Gamma(\tau, S_\tau) = \alpha\, P\,(1+g)^\tau + \alpha\, P \,\max\left(\left(\frac{S_\tau}{S_0}\right) - (1+g)^\tau, 0\right), \tag{2.9}$$

when the individual on whom the contract is written dies at time $\tau < T$. Specifically in case of death or survival the policyholder is offered a minimum guarantee given by the first term in (2.8) or (2.9) and he is offered a bonus option given by the second term. We also make use of the standard assumption in practice that death is treated as a natural event and thus it is reasonable to let the guaranteed rates be the same for either event. Furthermore, in practice the surrender benefit is often independent of the reference fund return and thus we make the assumption that the policyholder obtains the following amount $L(\lambda)$ at the surrender time λ, see Bernard and Lemieux (2008),

$$L(\lambda) = (1 - \beta_\lambda)P(1 + h)^\lambda, \quad 0 \le t \le T\,. \tag{2.10}$$

Here h refers to the minimum guaranteed rate at surrender and β_λ is the penalty charge. In Canada and the US regulation provides that h is not allowed to fall below g, see Bernard and Lemieux (2008). The penalty is further assumed to be constant over one calendar year and a decreasing function of time, see Palmer (2006).

The surrender action at a random time λ is described by the first arrival of a generalized Markov modulated Poisson process with stochastic intensity γ, which depends on the current state of the economy, the ratio of the surrender benefit L and the present value of the contract $V = (V_t)_{t\ge0}$. We refer to γ as the $(\mathbb{Q}_{\tilde{\theta}}, \mathbb{G})$ intensity. The new crucial assumption of our model is that exogenous and endogenous surrender depend on the current regime, that is the current state of the Markov chain. More formally, the arrival of an exogenous surrender is modelled with a Markov modulated Poisson process with intensity $\underline{\rho} X_t$, where the row vector $\underline{\rho}$ is given by $\underline{\rho} = (\underline{\rho_1} \quad \underline{\rho_2})$. Endogenous surrender is also modeled with a Markov modulated Poisson process with intensity $\rho^E\, X_t\, 1_{\{L(t)>V(t)\}}$, where $\rho^E = (\rho_1^E \quad \rho_2^E)$ is a again a row vector. Note that by definition the endogenous surrender intensity is 0 if it is not monetary optimal to surrender. The stopping time λ is the minimum of these two conditionally independent random times. Assuming that exogenous and endogenous surrender do not happen at the same time almost surely, the surrender intensity, that is the intensity of λ, is just the sum of these two intensities and it can be written more compactly as

$$\gamma_t = \begin{cases} \underline{\rho} X_t\,, & \text{if } \frac{L(t)}{V_t} < 1\,, \\ (\underline{\rho} + \rho^E) X_t\,, & \text{if } \frac{L(t)}{V_t} \ge 1\,, \end{cases} \tag{2.11}$$

The first case in (2.11) corresponds to the exogenous surrender bound and the second is the endogenous surrender bound. This formulation for the surrender intensity is inspired by Dai and You (2007) and can be traced back to Stanton (1995). In addition the European contract values and the American style contract values are included in this setup as limiting cases. That is, the European contract values in the two states of the economy are obtained by setting $\underline{\rho_i} = \rho_i^E = 0$, $i = 1, 2$, while the American style contract values are obtained by setting $\underline{\rho_i} = 0$ and sending $\rho_i^E \uparrow \infty$, $i = 1, 2$.

2.3.3 Contract Valuation

In this section we follow mainly Dai and You (2007) to derive a system of two coupled partial differential equations using the balance law of financial economics. A system of coupled partial differential equations is characterized by the property that the value function in state i satisfies a partial differential equation that depends on the value functions of all other states $j \neq i$. To find the contract value $V = (V_t)_{t\ge0}$ for an active contract, i.e. on the set $\{t < \tau, \lambda\}$, we have to solve this PDE system and to correctly identify the state of the economy. Note that if death or surrender occur at time t, $t \in]0, T[$, the contract value is trivially given by the death benefit, the surrender benefit respectively.

The balance law of financial economics is based on the no-arbitrage condition

$$r(t, X_t)\, V_t \,\mathrm{d}t = \mathbb{E}^{\mathbb{Q}_{\tilde{\theta}}}[\mathrm{d}V_t|\mathcal{G}_t]\,, \quad 0 \le t \le T, \tag{2.12}$$

on the set $\{t < \tau, \lambda\}$. On this set we can compute the following instantaneous probabilities under the assumption that the stopping times τ and λ are conditionally independent of each other:

(a) The conditional probability that death occurs over $(t, t + \mathrm{d}t)$ and surrender does not is $\mu(t)\mathrm{d}t(1 - \gamma_t \mathrm{d}t) = \mu(t)\mathrm{d}t$.

(b) The conditional probability that surrender occurs over $(t, t+\mathrm{d}t)$ while the death event does not is $\gamma_t \mathrm{d}t(1 - \mu(t)\mathrm{d}t) = \gamma_t \mathrm{d}t$.

(c) The conditional probability that both death and surrender occur over $(t, t + \mathrm{d}t)$ is 0.

Next, suppose that the contract value at time t is of the form $V_t = 1_{\{t<\lambda,\tau\}} v(t, S_t, X_t)$, on the set $\{t < \lambda, \tau\} \cap \{t \le T\}$. Then the surrender intensity γ specified in (2.11) is a function of the state variables t, v and x. Given that we are in state i, we can define $v_i(t, s) = v(t, s, e_i)$ for a suitably differentiable function $v_i : [0, T] \times \mathbb{R}^+ \to \mathbb{R}^+, (t, s) \to v_i(t, s)$, for $i = 1, 2$. Denoting the row vector of contract values for an active contract as $v(t, s) = (v_1(t, S_t)\, v_2(t, S_t))$, we can further write $v(t, S_t, X_t) = v(t, S_t)X_t$. Next, define γ_i depending on t and v by $\gamma_i : [0, T] \times \mathbb{R}^+ \to \mathbb{R}^+, (t, v) \to \gamma_i(t, v)$, for $i = 1, 2$. Denoting the corresponding row vector by $\gamma(t, v) = (\gamma_1(t, v)\, \gamma_2(t, v))$, we can write $\gamma_t = \gamma(t, V_t, X_t) = \gamma(t, v(t, S_t)X_t))X_t$. Now, we compute the differential of $v(t, S_t, X_t)$ by applying Ito's product rule, using the fact derived by Elliot et al. (1994) that X_t has the dynamics $\mathrm{d}X_t = A'X_t\mathrm{d}t + \mathrm{d}M_t$, where M_t is a two-dimensional martingale with respect to the filtration $\mathbb{G}$. We obtain that

$$\begin{aligned}\mathrm{d}v(t, S_t, X_t) &= \mathcal{L}v(t, S_t, X_t)\mathrm{d}t + v(t, S_t, X_t)\mathrm{d}X_t + \sigma(t, S_t, X_t)\, S_t\, \mathrm{d}W_t^{\tilde{\theta}} \\ &= \mathcal{L}v(t, S_t, X_t)\mathrm{d}t + v(t, S_t, X_t)A'X_t\, \mathrm{d}t + v(t, S_t, X_t)\mathrm{d}M_t + \sigma(t, S_t, X_t)\, S_t\, \mathrm{d}W_t^{\tilde{\theta}},\end{aligned} \tag{2.13}$$

where the regime-dependent differential operator $\mathcal{L}$ is defined as

$$\mathcal{L}f(t, s, x) = \frac{\partial f(t, s, x)}{\partial t} + r(t, x)\, s\, \frac{\partial f(t, s, x)}{\partial s} + \frac{1}{2}\, \sigma^2(t, s, x)\, s^2\, \frac{\partial^2 f(t, s, x)}{\partial s^2}.$$

The differential operator in state i $\mathcal{L}_i$ is then defined by

$$\mathcal{L}_i f(t, s) = \frac{\partial f(t, s)}{\partial t} + r_i(t)s\frac{\partial f(t, s)}{\partial s} + \frac{1}{2}\sigma_i^2(t, s)\, s^2 \frac{\partial^2 f(t, s)}{\partial s^2}\,, \quad i = 1, 2.$$

Next, we know that if either death or surrender occurs over $(t, t + \mathrm{d}t)$ the jump sizes are predictable. In particular, the change in the payment liability if death occurs and surrender does not is given by $\Gamma(t, S_t) - v(t, S_t, X_t)$ and by $L(t) - v(t, S_t, X_t)$ if surrender occurs and death does not. This implies that we can rewrite (2.12) as

$$\begin{aligned} r(t, X_t)v(t, S_t, X_t)\mathrm{d}t =& \mathbb{E}^{\mathbb{Q}_{\tilde{\theta}}}[\mathrm{d}v(t, S_t, X_t)|\mathcal{Z}_t] + (\Gamma(t, S_t) - v(t, S_t, X_t))\mu(t)\mathrm{d}t \\ &+ (L(t) - v(t, S_t, X_t))\gamma(t, V_t, X_t)\mathrm{d}t. \end{aligned} \tag{2.14}$$

Plugging (2.13) into (2.14) and using the fact that $W_t^{\tilde{\theta}}$ and M_t are $\mathbb{Q}^{\tilde{\theta}}$-martingales, we obtain $\mathrm{d}t \otimes \mathrm{d}\mathbb{Q}^{\tilde{\theta}}$-a.s.

$$\begin{aligned} 0 =& \mathcal{L}v(t, S_t) + \mu(t)\,\Gamma(t, S_t) + \gamma(t, V_t, X_t)L(t) \\ &- (r(t, X_t) + \mu(t) + \gamma(t, V_t, X_t))v(t, S_t, X_t) + v(t, S_t)\,A'\,X_t\,. \end{aligned} \tag{2.15}$$

Since this has to hold for any $t \in [0, T)$, $S_t \in \mathbb{R}^+$ and $X_t \in \mathcal{S}$ we have that

$$\begin{aligned} 0 =& \mathcal{L}_i v_i(t, s) + \mu(t)\Gamma(t, s) + \gamma_i(t, v_i(t, s))L(t) \\ &- (r_i(t) + \mu_t + \gamma_i(t, v_i(t, s)))v_i(t, s) + v(t, s)A'e_i\,, \end{aligned} \tag{2.16}$$

for $i = 1, 2$. Also, by no arbitrage we must have $v_i(T, s) = \Phi(s)$, for $s > 0$ and $i = 1, 2$. Carrying out the matrix multiplication, we derive a system of two coupled PDEs, which we summarize in proposition 2.3.1.

Proposition 2.3.1. *On the set $\{t \leq \tau \wedge \lambda \wedge T\}$ the contract value is given by*

$$V_t = 1_{\{\lambda>t,\tau>t\}}\left(1_{\{X_t=e_1\}}v_1(t, S_t) + 1_{\{X_t=e_2\}}v_2(t, S_t)\right) + 1_{\{\lambda>t,\tau=t\}}\,\Gamma(t, S_t) + 1_{\{\lambda=t\}}\,L(t),$$

where the price functions v_i, $i = 1, 2$, satisfy the following system of partial differential equations

$$\begin{aligned} 0 &= \mathcal{L}_1 v_1(t, s) + \mu(t)\Gamma(t, s) + \gamma_1(t, v_1(t, s))L(t) - (r_1(t) + \mu(t) + \gamma_1(t, v_1(t, s)))\,v_1(t, s) \\ &\quad +\eta_1\,(v_2(t, s) - v_1(t, s)), \qquad (2.17) \\ 0 &= \mathcal{L}_2 v_2(t, s) + \mu(t)\Gamma(t, s) + \gamma_2(t, v_2(t, s))L(t) - (r_2(t) + \mu(t) + \gamma_2(t, v_2(t, s)))\,v_2(t, s) \\ &\quad +\eta_2\,(v_1(t, s) - v_2(t, s)), \qquad (2.18) \end{aligned}$$

for $(t, s) \in [0, T) \times \mathbb{R}^+$ with terminal conditions $v_1(T, s) = v_2(T, s) = \Phi(s)$, for $s \in \mathbb{R}^+$.

Subsequently we theoretically investigate the contract values in the two regimes. In proposition 2.3.2 we show that some technical conditions can ensure that the contract value is always greater in one state than in the other. The proof is given in the appendix. In particular, the first two conditions that prominently depend on the delta and the convexity of the contract value, will in general not hold globally.

Proposition 2.3.2. *Assume the functions v_1 and v_2 in $C^{1,2}$ are solutions to the coupled PDE system in Proposition 2.3.1. Suppose that*

$$(r_2 - r_1)\,(s\,\frac{\partial v_1}{\partial s} - v_1) \geq 0\,, \quad (\sigma_2^2 - \sigma_1^2)\,\frac{\partial^2 v_1}{\partial s^2} \geq 0\,, \quad \textit{and} \quad \underline{\rho_2} + \rho_2^E \geq \underline{\rho_1} + \rho_1^E,$$

then $v_1 \leq v_2$ *on* $[0,T] \times \mathbb{R}^+$. *In particular, if* $r_1 = r_2$, $\sigma_1 = \sigma_2$ *and* $\underline{\rho_2} + \rho_2^E \geq \underline{\rho_1} + \rho_1^E$ *then* $v_1 \leq v_2$ *on* $[0,T] \times \mathbb{R}^+$, *and, if* $\underline{\rho_1} = \underline{\rho_2}$, ${\rho_1}^E = {\rho_2}^E$, $r_1 = r_2$ *and* $\sigma_1 = \sigma_2$ *then* $v_1 = v_2$ *on* $[0,T] \times \mathbb{R}^+$,

Finally we can give a very general comparative statics result for the endogenous surrender intensities ρ_i^E, $i = 1, 2$. Intuitively, a higher endogenous surrender intensity in state i increases the likelihood of monetary optimal surrender, which increases the contract value in either state of the economy. The following proposition states this fact precisely. The proof is given in the appendix.

Proposition 2.3.3. *Assume the functions* v_1 *and* v_2 *in* $C^{1,2}$ *are solutions to the coupled PDE system in Proposition 2.3.1. Further, assume that the functions* $\tilde{v}_1$ *and* $\tilde{v}_2$ *in* $C^{1,2}$ *are also solutions to the coupled PDE system that is identical except for the endogenous surrender parameters that are now given by* $\tilde{\rho_1^E}$, $\tilde{\rho_2^E}$, *respectively. Suppose that* $\rho_1^E \leq \tilde{\rho_1^E}$ *and* $\rho_2^E \leq \tilde{\rho_2^E}$. *Then we have* $\tilde{v}_1 \geq v_1$ *and* $\tilde{v}_2 \geq v_2$.

For the other regime-switching parameters we cannot give such general results, but we study their effects closely through numerical examples in section 4.

2.4 Numerical Methodology

To solve the PDE system (2.17) and (2.18) we need to determine reasonable boundary conditions as a first step. We approach this by specifying the surrender intensities at the boundary and using appropriate Neumann conditions for the partial derivatives. More formally our specifications are

$$\gamma_i(t, v_i) = (\underline{\rho_i} + \rho_i^E), \qquad \frac{\partial v_i(t,s)}{\partial s} = 0, \qquad s \downarrow 0,\ t \in [0,T),\ i = 1, 2, \tag{2.19}$$

$$\gamma_i(t, v_i) = \underline{\rho_i}, \qquad \frac{\partial v_i(t,s)}{\partial s} = \frac{\alpha\, P}{S_0}, \qquad s \uparrow \infty,\ t \in [0,T),\ i = 1, 2. \tag{2.20}$$

The intuitive idea behind our assumptions above is as follows. If the reference fund value is very small then for a realistic contract the surrender option will be almost surely in the money until it is exercised, accordingly we assume the surrender intensity coincides with the endogenous surrender bound. For a very small value of the reference fund changes in s have no effect on the contract value since the bonus option is deep out of money and the policyholder will almost surely receive the minimum guaranteed amount if he does not surrender. If the reference fund reaches a very large value then the surrender option will be deep out of the money, which implies that there are no incentives to surrender the contract endogenously throughout the entire lifetime of the contract. In that case changes in s will only affect the value of the bonus option and as this is deep in the money it will approximately move linearly with the reference fund. Next we can plug (2.19) and (2.20) separately in the PDE equation for state i in (2.16). Noting that by the above assumptions the second derivative with respect to s diminishes, the determination of the boundary conditions reduces to solving a system of two coupled linear ordinary differential equations (ODEs) for each boundary. Accordingly for $s \downarrow 0$ we have the coupled ODE system

$$\frac{\partial v_i}{\partial t} + \mu(t)\,\alpha\,P(1+g)^t + (\underline{\rho_i} + \rho_i^E)\,L(t) - (r_i(t) + \mu(t) + \underline{\rho_i} + \rho_i^E)v_i + \eta_i\,(v_j - v_i) = 0, \tag{2.21}$$

for $t \in [0,T), i \neq j = 1,2$ with the terminal condition $v_1(T) = v_2(T) = \alpha P(1+g)^T$. For the upper boundary we approximate $s \uparrow \infty$ with a sufficiently large maximum value we denote s_{max}, then we obtain the following coupled ODE system

$$\frac{\partial v_i}{\partial t} + r\,s_{max}\frac{\alpha P}{S_0} + \mu(t)\,\alpha\,P\frac{s_{max}}{S_0} + (\underline{\rho_i}\,L(t) - (r_i(t) + \mu(t) + \underline{\rho_i})v_i + \eta_i\,(v_j - v_i) = 0, \tag{2.22}$$

for $t \in [0,T), i \neq j = 1,2$ with the terminal condition $v_1(T) = v_2(T) = \alpha\,P\frac{s_{max}}{S_0}$. These ODE systems can be solved straightforwardly with a simple explicit scheme.

The next step is to solve the PDE system (2.17) and (2.18) subject to the terminal and boundary conditions. It is important to emphasize that we need to solve a nonlinear coupled PDE system since the surrender intensities are functions of the contract values and hence need to be determined simultaneously with the latter for each time step and each state of the economy. In this way, we need to solve a penalty problem, see Dai and You (2007). To do so we first show how the solution of the PDE system is approximated for fixed surrender intensities $\gamma_i(t, v_i)$, $i = 1,2$ with the well-known Crank-Nicolson scheme. Then we apply a generalized Newton iteration procedure, see Forsyth and Vetzal (2002), to simultaneously compute the contract values and the surrender intensities in each state of the economy.

For the reason of computational efficiency we perform the log-transformation $w = \ln(s)$ and set $u_i(t,w) = v_i(t,s)$. The transformed PDE system is then given by

$$\begin{aligned} 0 \;=\; & \frac{\partial u_i(t,w)}{\partial t} + (r_i(t) - \frac{1}{2}\sigma_i^2(t,w))\frac{\partial u_i(t,w)}{\partial w} + \frac{1}{2}\sigma_i^2(t,w)\frac{\partial^2 u_i(t,w)}{\partial w^2} + \mu(t)\Gamma(t,y) \\ & + \gamma_i(t, u_i(t,w))L(t) - (r_i(t) + \mu(t) + \gamma_i(t, u_i(t,w)) + \eta_i)u_i(t,w) + \eta_i u_j(t,w) \end{aligned} \tag{2.23}$$

$(t,w) \in [0,T) \times \mathbb{R}, i \neq j = 1,2$ and terminal conditions $u_1(T,w) = u_2(T,w) = \Phi(e^w)$.

To approximate (2.23) with we truncate the time domain $[0,T]$ in N equally spaced time intervals Δt and the log-price domain $[w_{min}, w_{max}]$ into M subintervals of length Δw, where w_{min} denotes the minimum log-price and w_{max} the maximum log-price the reference fund can attain. For each $i = 1,2$, $n = 0,..N$, $m = 0,...M$ let $u_i^{(n,m)}$ denote the discretized version of the contract value at node (n,m), where n is the time-point and m the log-price step. In the sequel this definition applies to all parameters depending on (n,m). Applying the Crank-Nicolson scheme to approximate the coupled PDE system (2.23) we have

$$\begin{aligned}
&- b_i u_i^{(n+1,m-1)} - \tilde{a}_i^{(n+1,m)} u_i^{(n+1,m)} - \tilde{b}_i u_i^{(n+1,m+1)} - \eta_j u_j^{(n+1,m)} \\
&= b_i u_i^{(n,m-1)} + a_i^{(n,m)} u_i^{(n,m)} + \tilde{b}_i u_i^{(n,m+1)} + \eta_j u_j^{(n,m)} + \quad c_i^{(n,m)} + c_i^{(n+1,m)},
\end{aligned} \tag{2.24}$$

where

$$\begin{aligned}
b_i &= -\left(r_i^{(n)} - \frac{1}{2}\sigma_i^{2(n,m)}\right)\frac{1}{2\Delta w} + \frac{1}{2}\frac{\sigma_i^{2(n,m)}}{(\Delta w)^2}, \\
\tilde{b}_i &= \left(r_i^{(n)} - \frac{1}{2}\sigma_i^{2(n,m)}\right)\frac{1}{2\Delta w} + \frac{1}{2}\frac{\sigma_i^{2(n,m)}}{(\Delta w)^2}, \\
a_i^{(n,m)} &= -\left(\frac{2}{\Delta t} + \frac{\sigma_i^{2(n,m)}}{(\Delta w)^2} + r_i^{(n)} + \mu^{(n)} + \gamma_i^{(n,m)} + \eta_i\right), \\
\tilde{a}_i^{(n+1,m)} &= -\left(-\frac{2}{\Delta t} + \frac{\sigma_i^{2(n+1,m)}}{(\Delta w)^2} + r_i^{(n+1)} + \mu^{(n+1)} + \gamma_i^{(n+1,m)} + \eta_i\right), \\
c_i^{(n,m)} &= \mu^{(n)}\Gamma^{(n,m)} + \gamma_i^{(n,m)} L^{(n)}, \\
c_i^{(n+1,m)} &= \mu^{(n+1)}\Gamma^{(n+1,m)} + \gamma_i^{(n+1,m)} L^{(n+1)}.
\end{aligned}$$

In order to obtain the contract value at time point n in each state we have to solve a system of $2(M-1)$ linear equations in $2(M-1)$ unknowns. We define

$$u^{(n)} = (u_1^{(n,M-1)}, u_2^{(n,M-1)}, u_1^{(n,M-2)}, u_2^{(n,M-2)}, ..., u_1^{(n,1)}, u_2^{(n,1)})', \tag{2.25}$$

as the vector of contract values, which is first ordered by the regime number and then in descending order of log-price steps. This enables us to write (2.24) compactly in matrix notation as

$$\tilde{D}^{(n+1)} u^{(n+1)} = D^{(n)} u^{(n)} + d^{(n,n+1)}, \tag{2.26}$$

where the $2(M-1) \times 2(M-1)$ matrix $D^{(n)}$ is given by

$$D^{(n)} = \begin{pmatrix}
a_1^{(n,M-1)} & \eta_1 & b_1 & 0 & 0 & 0 & 0 & \dots \\
\eta_2 & a_2^{(n,M-1)} & 0 & b_2 & 0 & 0 & 0 & \dots \\
\tilde{b}_1 & 0 & a_1^{(n,M-2)} & \eta_1 & b_1 & 0 & 0 & \dots \\
0 & \tilde{b}_2 & \eta_2 & a_2^{(n,M-2)} & 0 & b_2 & 0 & \dots \\
\vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
0 & 0 & \dots & \tilde{b}_2 & \eta_2 & a_2^{(n,2)} & 0 & b_2 \\
0 & 0 & 0 & \dots & \tilde{b}_1 & 0 & a_1^{(n,1)} & \eta_1 \\
0 & 0 & 0 & \dots & 0 & \tilde{b}_2 & \eta_2 & a_2^{(n,1)}
\end{pmatrix}.$$

$\tilde{D}^{(n+1)}$ is simply obtained by replacing each diagonal entry $a_i^{(n,m)}$ in $D^{(n)}$ by $\tilde{a}_i^{(n+1,m)}$ and multiplying each by -1, and $d^{(n,n+1)}$ is given by

$$d^{(n,n+1)} = \begin{pmatrix} c_1^{(n,M-1)} + c_1^{(n+1,M-1)} + \tilde{b_1}(u_1^{(n,M)} + u_1^{(n+1,M)}) \\ c_2^{(n,M-1)} + c_2^{(n+1,M-1)} + \tilde{b_2}(u_2^{(n,M)} + u_2^{(n+1,M)}) \\ c_1^{(n,M-2)} + c_1^{(n+1,M-2)} \\ c_2^{(n,M-2)} + c_2^{(n+1,M-2)} \\ \vdots \\ c_1^{(n,1)} + c_1^{(n+1,1)} + b_1(u_1^{(n,0)} + u_1^{(n+1,0)}) \\ c_2^{(n,1)} + c_2^{(n+1,1)} + b_2(u_2^{(n,0)} + u_2^{(n+1,0)}) \end{pmatrix}.$$

Note that $d^{(n,n+1)}$ depends via $c^{(n,m)}$ on $\gamma_i^{(n,m)}$, $i = 1, 2$, and that needs yet to be determined. We solve the matrix equation (2.26) for $u^{(n)}$ and destack this vector into the two vectors $u_i^{(n)} = (u_i^{(n,M-1)} ... u_i^{(n,1)})'$, $i = 1, 2$, to obtain the contract values in state i at time point n.

The algorithm to determine $\gamma^{(n,m)}$ and $u^{(n,m)}$ works backwards starting at $n = N - 1$. First we specify a starting value for the contract value in the corresponding state of the economy at time point n. A natural candidate for the starting value at time point n is the contract value in the corresponding state of the economy at the time point $n+1$. The next step is to specify the surrender intensities $\gamma_i^{(n,m)}$, $i = 1, 2$ according to (2.11). Then we solve the PDE system (2.23) with the Crank-Nicolson scheme as described above to determine the new contract values. For each time step this procedure is repeated until the maximum relative deviation of the contract values after the next iteration falls below a prespecified tolerance level.

Specifically, denote $u_i^{\tilde{k}}(n, m)$, the contract value in state i, $i = 1, 2$, at node (n, m) after the $\tilde{k}$-th iteration run and $\gamma_i^{\tilde{k}}(n, m)$ is the surrender intensity in state i at node (n, m) after the $\tilde{k}$-th iteration. Next define

$$\varepsilon_i^{\tilde{k}} = \max_m \frac{|u_i^{\tilde{k}+1}(n, m) - u_i^{\tilde{k}}(n, m)|}{\max(1, u_i^{\tilde{k}+1}(n, m))},$$

for $i = 1, 2$ and define $\varepsilon^{\tilde{k}} = \max(\varepsilon_1, \varepsilon_2)$. That is, ε is the maximum relative deviation the contract values can have in any of the two states after the $\tilde{k}+1$-th iteration run compared to the $\tilde{k}$-th iteration. Finally, let ξ be the error tolerance in the iteration. Using this notation the algorithm to determine simultaneously u_i and γ_i, $i = 1, 2$ at time $n = 0, 1, ...N - 1$ for each $m = 1, ...M - 1$ can be summarized as follows:

(a) take $u_i^0(n, m) = u_i(n + 1, m)$ as a starting value, compare $u_i^0(n, m)$ with $L(n)$, get $\gamma_i^0(n, m)$
compute $u_i^1(n, m)$ by solving the PDE system (2.23) with Crank-Nicolson
and compute ε^0,
for $\tilde{k} = 1, ...\tilde{k}^*$

(b) compare $u_i^{\tilde{k}}(n,m)$ with $L(n)$ and get $\gamma_i^{\tilde{k}}(n,m)$

(c) take each $\gamma_i^{\tilde{k}}(n,m)$ and solve PDE system (2.23) numerically to obtain $u_i^{\tilde{k}+1}(n,m)$

(d) compute $\varepsilon^{\tilde{k}}$

(e) repeat steps 2 until 4 until $\varepsilon^{\tilde{k}} < \xi$,

where $\tilde{k}^*$ is the last but one iteration.

2.5 Numerical Results

In this section we study the life insurance contract closely through numerical examples. We assume that the two economic states represent the business cycles recession and expansion. The recession state is denoted as state 1 and the expansion state as state 2, respectively. We focus on the surrender option values in the two economic states and compare those for the two surrender models, our regime switching rational expectation model and the American style surrender model.[4]

The benchmark parametrization is as follows. The Standard and Poor's 500 index is the underlying with $S_0 = \$1000$. The single premium is $P = \$100$ and the contract life time is $T = 10$ years. We assume that the percentage the minimum guarantee is provided with and the participation coefficient is $\alpha = 0.875$, and the minimum guaranteed rates are all equal and given by $g = h = 0.02$. The penalty rates are constant each calender year with $\beta_1 = 0.05$, $\beta_2 = 0.04$, $\beta_3 = 0.02$, $\beta_4 = 0.01$ and $\beta_t = 0$ for $t \geq 5$. Next the mortality intensity is assumed to follow the Makeham model $\mu(y) = A + Bc^y$ for an y-aged policyholder with $A = 5.0758 \times 10^{-4}$, $B = 3.9342 \times 10^{-5}$, $c = 1.1029$, see Li and Szimayer (2014). The representative policyholder is assumed to be 40-aged at the moment he enters into the contract. The switching intensities are estimated from the NBER business cycle data from 2000 until 2010. We obtain the estimates $\hat{\eta_1} = 0.8889$ and $\hat{\eta_2} = 0.2215$. This implies that the expected time the economy will stay in a expansion is about 4.5 years, while the economy is expected to stay about 1.125 years in a recession state. For the regime-switching parameters we make the following assumptions throughout this section: $\underline{\rho_1} \geq \underline{\rho_2}$, $\rho_1^E \leq \rho_2^E$, $r_1 \leq r_2$, $\sigma_1 \geq \sigma_2$. The first three assumptions are made such that both the emergency fund hypothesis and the interst rate hypothesis hold. The procyclcial risk-free rate and the countercyclical volatility are in line with empirical studies, see for instance Engle et al. (2008) for a countercyclical volatility and Ang and Bekaert (2002 a,b) for a procyclical risk-free rate. Note that we also assume for simplicity that the reference fund volatility and the risk-free rate are constant within a regime, i.e $\sigma_i(t,s) = \sigma_i$ and $r_i(t) = r_i$ for $i = 1, 2, (t,s) \in [0,T] \times \mathbb{R}^+$. Specifically, we set the exogenous surrender intensities as $\underline{\rho} = (0.05 \quad 0.02)$. The endogenous surrender intensities in the two states are $\rho^E = (0.25 \quad 0.28)$ such that the endogenous surrender bound is the same in either state

[4] European and American option values are understood here in the presence of mortality risk.

and equal to 0.3. This specification of the surrender intensities implies a monetary more optimal surrender behavior in the expansion state. The volatilities are estimated by constructing a recession subsample and a expansion subsample for the Standard and Poor's 500 daily log returns from 2000 until 2010. We have $\hat{\sigma_1} = 0.3397$ and $\hat{\sigma_2} = 0.1728$, which are the annualized historical volatilities of the corresponding subsamples. For the risk-free rate we assume $r_1 = 0.025$ and $r_2 = 0.04$. Finally the following values for the parameters in the numerical approximation are used: $\Delta t = 1/50$, $N = T/\Delta t$, $M = 250$, $w_{min} = 0$, $w_{max} = \ln(5000)$, $\Delta w = (w_{max} - w_{min})/M$ and $\xi = 0.0001$.

To better understand the contract intuitively we decompose it into its different components in table 2.1. The European contract value is given by the value of the minimum guarantee and the bonus option. The value of the surrender option is obtained as the difference between the contract value with surrender and the European contract value. Table 2.1 shows that the contract value in the recession state is about 2.31% higher than that in the expansion state. This is primarily the result of the higher minimum guarantee value due to the lower risk-free rate and a higher value for the bonus option due to the higher volatility in the recession state. Moreover we notice that the value of the surrender option is about 11 % higher in the recession state for both our regime-switching rational expectation model and the American style surrender model. This is the result of the countercyclical volatility and the procyclical risk-free rate effect, which will be studied more closely in table 2.5 and table 2.6 below. Note that in the regime switching rational expectation model the latter effects dominate the effect of monetary less optimal surrender behavior in the recession state. Most importantly we clearly observe that the American surrender option values in either regime are substantially, about 5 times larger than those in the regime switching rational expectation model.

Minimum Guarantee (1)	Bonus Option (2)	Surrender Option (3)	American Surrender Option (4)
74.738	29.652	1.951	9.366
73.724	28.461	1.756	8.473

European Contract Value (1)+(2)	Contract Value (1)+(2)+(3)	American Contract Value (1)+(2)+(4)
104.390	106.341	113.756
102.185	103.941	110.658

Table 2.1: Decomposition of the contract values (in \$) for the benchmark parameters in the two economic states, the upper value corresponds to the value in the recession state and the lower to that in the expansion state, respectively.

In table 2.2 we see how changes of the insurance parameters, the percentage the minimum guaranteed amount is provided with and the policyholder participates in the performance of the reference fund α, the minimum guaranteed interest rate g and the interest rate the surrender value grows with h, affect the value of the surrender option. A significant decrease in α makes the holding of the contract less attractive and hence it substantially increases the value of the surrender option in either state. On the other hand an increase of $g = h$ decreases the value of the surrender option in both states since the higher value of the minimum guarantee dominates the increase in the cash surrender value L.

	Surrender Option	American Surrender Option
$\alpha = 0.8$	7.097	14.053
	6.635	12.997
$g = h = 0.025$	1.604	8.996
	1.618	8.295

Table 2.2: Surrender Option values in the two economic states for different insurance parameters, the upper value corresponds to the value in the recession state and the lower to that in the expansion state, respectively.

Now we study more closely the effect of the regime-dependent parameters on the surrender option values. In tables 2.3 and 2.4 we see how changes in the regime-dependent surrender intensities affect the surrender option values in our rational expectation model, while the American style option remains unaffected by assumption. In table 2.3 we observe that increasing the endogenous surrender intensity ρ_i^E increases the surrender option value in either state. This is an immediate consequence of proposition 2.3.3. We see that the endogenous surrender intensity in the expansion state ρ_2^E has a substantial effect on the surrender option values, while the impact of the endogenous surrender intensity in the recession state ρ_1^E is low. It is interesting to observe that when the policyholder surrenders only exogenously in the recession state, i.e $\rho_1^E = 0$ and the endogenous surrender intensity in the expansion state is low, i.e $\rho_2^E \leq 0.1$, then the value of the surrender option can get negative. This at first glance surprising result is explained by the fact that the surrender option is exercised with a sufficient likelihood when it is out of the money, especially in the recession state, while it is exercised with a too low likelihood when it is in the money.

In table 2.4 we see that both exogenous surrender intensities $\underline{\rho_1}$ and $\underline{\rho_2}$ have a very strong impact on the values of the surrender option. In comparison to the previous table we clearly observe that the surrender option values vary more with shifts of the exogenous surrender intensities than with the endogenous surrender intensities. Moreover there is a clear symmetric regime-impact of the exogenous surrender intensities on the option values. More precisely, the exogenous surrender intensity in state 1 has a stronger impact on the option value in state 1 than in state 2 and vice versa. This fact is fairly pronounced for the exogenous surrender intensity in the recession state $\underline{\rho_1}$. We see that the surrender option

$\rho_1^E \setminus \rho_2^E$	0	0.1	0.2	0.5
0	-2.992 -2.625	-0.551 -0.472	0.814 0.742	2.514 2.143
0.2			1.330 1.373	2.719 2.535
0.5				2.816 2.658

Table 2.3: Surrender option values in the two economic states for various endogenous surrender intensities ρ_2^E in the first row and ρ_1^E in the first column, the upper value corresponds to the value in the recession state and the lower to that in the expansion state, respectively.

values in either state substantially decrease with increasing exogenous surrender intensities. This is not a general result but is explained by the fact that for this contract the surrender option is on average out of the money and hence a higher exogenous surrender intensity results in a more monetary suboptimal surrender behavior. This point further accounts for considerably negative surrender option values if the exogenous surrender intensities are high, that is $\underline{\rho_2} \geq 0.03$ and $\underline{\rho_1} \geq 0.1$.

$\underline{\rho_2} \setminus \underline{\rho_1}$	0	0.03	0.1	0.3
0	6.210 5.354	4.758 4.395	1.947 2.520	-2.939 -1.135
0.03		1.690 1.284	-0.636 -0.258	-4.538 -3.137
0.1			-4.700 -4.390	-7.157 -6.089

Table 2.4: Surrender option values in the two economic states for various values of the exogenous surrender intensities $\underline{\rho_1}$ in the first row and $\underline{\rho_2}$ in the first column, the upper value corresponds to the option value in the recession state and the lower to that in the expansion state, respectively.

Table 2.5 confirms that both volatilities σ_1 and σ_2 have a significant effect on the surrender option values, where we identify a stronger effect in the regime-switching rational expectation model. For the latter there is an asymmetric regime impact on the surrender option values, which is slight for shifts of σ_1 and moderate for shifts of σ_2. The asymmetric regime impact means that the volatility in state i has a stronger effect on the option value in state j than in state i. For the American option we identify a strong symmetric regime impact for both volatilities σ_1 and σ_2. Besides the opposing regime impacts, it is interesting to see that the overall effect of the volatility also completely differs in the two surrender models. While the surrender option values decrease with σ_i in the rational expectation model, they increase with σ_i in the American style surrender model. The first result can be explained by the effect that an increasing volatility increases the likelihood that the option goes out of the money and hence increases the likelihood of monetary suboptimal surrender, which here dominates the opposing effect that an increasing volatility can also make monetary optimal surrender more likely. For the American style surrender option only the effect that an increasing volatility increases the likelihood for the surrender option to be in the money plays a role since by assumption the policyholder behaves monetary optimal, hence the American style surrender option values are increasing with the volatilities.

$\sigma_2 \backslash \sigma_1$	0.10	0.20	0.30	0.40
0.10	4.463 (8.063) 4.085 (7.603)	3.891 (8.572) 3.415 (7.640)	3.139 (9.002) 2.648 (7.695)	2.344 (9.415) 1.838 (7.802)
0.20		2.227 (8.813) 2.183 (8.535)	1.728 (9.276) 1.682 (8.659)	1.283 (9.738) 1.145 (8.795)
0.30			0.476 (9.668) 0.694 (9.430)	0.135 (10.189) 0.324 (9.671)

Table 2.5: Surrender options values for the regime-switching rational expectation model and the American style surrender model in the two economic states for various volatilities σ_1 in the first row and σ_2 in the first column, the upper value corresponds to the value in the recession state and the lower to that in the expansion state, respectively, the values in brackets denote the American style surrender option values.

We learn from table 2.6 that the risk-free rate in the expansion regime has a stronger impact on the surrender option values in both states and for both surrender models than the risk-free rate in the recession state. Interestingly, for both surrender models we observe an asymmetric regime impact of the risk-free rate, which is slight for shifts of r_1 and very pronounced for shifts of r_2. For both surrender models the option values increase with the risk-free rates. This is intuitive and in line with the interest rate hypothesis because a higher risk-free rate increases the incentive to surrender the contract endogenously in order to exploit profitable alternative investments.

Comparing table 2.3 through table 2.6 we notice that for this contract the surren-

$r_1 \backslash r_2$	0.02	0.03	0.04	0.05
0.02	0.432	1.143	1.932	2.602
	(8.167)	(8.629)	(9.396)	(10.260)
	1.070	1.421	1.691	1.870
	(8.071)	(8.146)	(8.442)	(8.746)
0.03		1.177	1.934	2.622
		(8.665)	(9.402)	(10.284)
		1.571	1.826	1.968
		(8.218)	(8.514)	(8.805)
0.04			1.937	2.663
			(9.451)	(10.305)
			1.958	2.092
			(8.619)	(8.892)

Table 2.6: Surrender options values for the regime-switching rational expectation model and the American style surrender model in the two economic states for various risk-free rates r_2 in the first row and r_1 in the first column, the upper value corresponds to the value in the recession state and the lower to that in the expansion state, respectively, the values in brackets denote the American style surrender option values.

$\eta_2 \backslash \eta_1$	0.20	0.50	1.50
0.10	-0.444	1.935	3.182
	(9.815)	(9.817)	(9.253)
	0.894	2.009	2.883
	(8.126)	(8.342)	(8.455)
0.20	-1.498	0.898	2.664
	(9.511)	(9.545)	(9.185)
	-0.567	1.007	2.444
	(8.175)	(8.384)	(8.493)
0.50	-3.147	-1.011	1.443
	(9.167)	(9.223)	(9.081)
	-2.637	-0.827	1.363
	(8.465)	(8.573)	(8.609)

Table 2.7: Surrender options values for the regime-switching rational expectation model and the American style surrender model in the two economic states for various switching intensities η_1 in the first row and η_2 in the first column, the upper value corresponds to the value in the recession state and the lower to that in the expansion state, respectively, the values in brackets denote the American style surrender option values.

der option values are always greater in the recession state in the American style surrender model. This is the result of the symmetric regime impact of the countercyclical volatility and the asymmetric regime impact of the procyclical risk-free rate. From table 2.5 and 2.6 we know that both effects are relatively value increasing, hence the surrender option value is greater in the recession state. For the regime-switching rational expectation model this result does not hold since the effect of monetary less optimal surrender in the recession state and the slight asymmetric regime impact of the volatility, see also table 2.5, produce a countereffect to the asymmetric risk-free rate effect. If the latter effects dominate, which is for instance the case when the exogenous surrender intensity in the recession state is sufficiently large or the risk-free rates coincide, see table 2.4 and table 2.6, then the surrender option value is greater in the expansion state.

Finally we study the effect of the switching intensities η_1 and η_2 on the option values in both surrender models in table 2.7. The surrender option values in the two states are fairly sensitive with respect to changes in the switching intensities in the regime-switching rational expectation model, while they are clearly less sensitive for the American style surrender model. This can be explained by the property that the switching intensities have a substantial effect on the average surrender behavior in the regime-switching rational expectation model in the sense that they improve or deteriorate the average surrender behavior. In particular, the surrender option values in the two states are increasing in η_1 because the economy is expected to spend less time in the recession state, which on average results in a monetary better surrender behavior of the policyholder. The line of reasoning reverts for increases in η_2. On the other hand for the American style surrender model the degree of monetary optimal surrender is unaffected by the switching intensities since by assumption the policyholder behaves monetary optimal. This explains the lower sensitivity of the American option values. Unlike the regime-switching rational expectation model we do not observe a clear monotonic effect of the switching intensities on the surrender option values. However, for both surrender models higher values of η_i are leading to a more instable economy coupling the option values while lower values of η_i are corresponding to a more persistent economy decoupling the option values.

2.6 Conclusion

We propose a regime-switching rational expectation model, where both the market value of a reference fund and the surrender intensity of a policyholder change randomly over time according to the evolution of a continuous-time Markov Chain with a finite number of states. The main contribution of this chapter is that it extends the rational expectation model of De Giovanni (2010) and Li and Szimayer (2014) by allowing both exogenous and endogenous surrender to depend on the economic regime. Such economic regimes, can represent for example financial market regimes with high or low volatility, macroeconomic regimes with high or low interest rates or business cycles. More formally, we derive a coupled system of two partial differential equations whose solutions characterize the contract values in the two economic states and establishe comparative statistics. This PDE system is nonlinear since the surrender intensities are a function of the contract values and hence need to be determined simultaneously with the latter. The solution of this penalty problem is obtained numerically by combining the Crank-Nicolson scheme with a generalized Newton search algorithm.

We performe extensive numerical experiments, where the economic states are considered as business cycles and the model parameters are set such that the emergency fund and the interest rate hypothesis hold. Based on this experiment, we have the following main results. First, the state of the economy has a significant impact on the contract value and the surrender option value for both surrender models. The state impact is the more pronounced the more persistent the economy is in the two economic states. It further strongly depends on the difference between the regime-dependent paramter values and how pronounced the symmetric or asymmetric regime impacts are. Second, the surrender option value is greater in the recession state in the regime-switching American style surrender model, while in the regime-switching rational expectation model the surrender option value can also be greater in the expansion state. Third, the surrender option values are substantially lower in our regime-switching rational expectation model than those in the American style regime-switching surrender model in either state of the economy. We further found that the exogenous surrender intensities have the strongest impact on the surrender option values in our regime-switching rational expectation model. Finally, if the exogenous surrender intensities are sufficiently high, especially that in the recession state, the surrender guarantee can become negative in our model. In particular, the last three results underline that an incorporation of the emergency fund hypothesis into a valuation model is important in order to avoid a dramatic overpricing of the surrender guarantee.

The model presented in this chapter can be extended in several ways. A first possible extension could be to allow the regime-dependent surrender intensities to be additionally time-dependent. For example, one could assume that the surrender intensities are a decreasing function of time, that is policyholders are more likely to surrender when they are younger. A second extension that arises is to model the reference fund dynamics more realistically, e.g by including jumps when the state of the economy switches. From the theoretical perspective, it would be interesting to address the hedging problem in our setup. Though we obtain an adequate risk premium for the macroeconomic risk the latter is only justified by diversification arguments. A reasonable theoretical approach would be to follow the rationale of Guo (2001) and to include regime-switching bonds in order to first complete the extended financial market and then to derive hedging strategies for hedging the regime-switching risk.

2.7 Appendix

2.7.1 Appendix A: Proof of Proposition 2.3.2

Proof 2.7.1 (Proof of Proposition 2.3.2)**.** *The value functions v_1 and v_2 are the solutions of the coupled PDE system in proposition 2.3.1 with terminal condition $v_1(T,s) = v_2(T,s) = \Phi(s)$. Define the function z as their difference, i.e.: $z = v_2 - v_1$. It follows directly that $z(T,s) = v_2(T,s) - v_1(T,s) = \Phi(s) - \Phi(s) = 0$. To obtain the dynamics of z take the difference of the PDEs describing v_2 and v_1:*

$$\begin{aligned} 0 &= \mathcal{L}_2 v_2 + \mu\,\Psi + \gamma_2 L - (r_2 + \mu + \gamma_2)\,v_2 + \eta_2\,(v_1 - v_2) \\ &\quad - (\mathcal{L}_1 v_1 + \mu\,\Psi + \gamma_1 L - (r_1 + \mu + \gamma_1)\,v_1 + \eta_1\,(v_2 - v_1)) \\ &= \mathcal{L}_2 z - (r_2 + \mu + \gamma_1 + \eta_1 + \eta_2) z + (\gamma_2 - \gamma_1)(L - v_2) + (\mathcal{L}_2 - \mathcal{L}_1) v_1 + (r_1 - r_2) v_1 \\ &= \mathcal{L}_2 z - (r_2 + \mu + \gamma_1 + \eta_1 + \eta_2) z \\ &\quad +(\gamma_2 - \gamma_1)(L - v_2) + (r_2 - r_1)\left(s\frac{\partial v_1}{\partial s} - v_1\right) + (\sigma_2^2 - \sigma_1^2)\frac{\partial^2 v_1}{\partial s^2}\,. \end{aligned}$$

In fact, we want to show that $z \geq 0$ in turn implying $v_2 \geq v_1$. For doing so, the last line of the previous equation is of importance. Using, e.g., Feynman-Kac, the following sufficient condition for $z \geq 0$ can be obtained

$$\begin{aligned} 0 &\leq (\gamma_2(t, v_2(t,s)) - \gamma_1(t, v_1(t,s)))(L(t) - v_2(t,s)) \\ &\quad +(r_2(t) - r_1(t))\big(s\frac{\partial v_1}{\partial s}(t,s) - v_1(t,s)\big) + (\sigma_2^2(t,s) - \sigma_1^2(t,s))\frac{\partial^2 v_1}{\partial s^2}(t,s)\,. \end{aligned}$$

Now, we verify that all three summands are nonnegative. Addressing the first term consider the case $L \geq v_2$. Then by specification of the exercise behaviour in (2.11) *we see that $\gamma_2 = \underline{\rho_2} + \rho_2^E$. By assumption $\gamma_1 \leq \underline{\rho_1} + \rho_1^E \leq \underline{\rho_2} + \rho_2^E$, and thus the first summand is nonnegative, i.e.: $(\gamma_2 - \gamma_1)(L - v_2) \geq 0$. Consider the alternative case $L < v_2$, then $\gamma_2 = \underline{\rho_2}$. By assumption $\underline{\rho_2} \leq \underline{\rho_1}$ and accordingly $\gamma_2 \leq \gamma_1$. Also for this case we conclude that the first summand is nonnegative, i.e.: $(\gamma_2 - \gamma_1)(L - v_2) \geq 0$. The second and the third summand are nonnegative by assumption finishing the proof of the first assertion. The special case $r_1 = r_2$, $\sigma_1 = \sigma_2$ and $\underline{\rho_2} + \rho_2^E \geq \underline{\rho_1} + \rho_1^E$ follows immediately. For $\underline{\rho_1} = \underline{\rho_2}$, ${\rho_1}^E = {\rho_2}^E$, $r_1 = r_2$, and $\sigma_1 = \sigma_2$ we obtain by the special case just discussed that $v_1 \leq v_2$ and $v_2 \leq v_1$ (after switching indices) implying the identity, i.e. $v_1 = v_2$.*

2.7.2 Appendix B: Proof of Proposition 2.3.3

Proof 2.7.2 (Proof of Proposition 2.3.3)**.** *Define $\bar{\rho} = (\underline{\rho} + \rho^E)\,X_t$ and analogously $\tilde{\bar{\rho}}$. The specification of the surrender intensity γ in* (2.11) *indicates that it is the maximal point of the optimization problem*

$$\sup_{g \in G} \mathbb{E}^{\mathbb{Q}_{\tilde{\theta}}}\left[e^{-\int_0^T r(t,X_t)+\mu(t)+g_t \mathrm{d}t}\,\Phi(S_T) + \int_0^T e^{-\int_0^t r(u,X_u)+\mu(u)+g_u \mathrm{d}t}\,\big(\Gamma(t,S_t)\,\mu(t) + L(t)\,g_t\big)\mathrm{d}t\right],$$

where $G = \{g$ *is* $\mathbb{Z}$*-adapted* : $\underline{\rho}\, X_t \le g_t \le \bar{\rho}\, X_t\}$. *Further,* $v(t,s,x) = (v_1(t,s)\, v_2(t,s))x$ *as given in proposition 2.3.1 is the value function of the optimization problem. Now, define the set of admissible controls by* $\tilde{G} = \{g$ *is* $\mathbb{Z}$*-adapted* : $\tilde{\underline{\rho}}\, X_t \le g_t \le \tilde{\bar{\rho}}\, X_t\}$. *The solution of the corresponding optimization problem is* $\tilde{\gamma}$ *with value function* $\tilde{v}(t,s,x) = (\tilde{v}_1(t,s)\, \tilde{v}_2(t,s))x$. *By assumption* $\tilde{\bar{\rho}} \ge \bar{\rho}$ *pointwise and thus* $G \subseteq \tilde{G}$. *Consequently, we have that the value functions satisfy* $v(t,s,x) \le \tilde{v}(t,s,x)$,

Chapter 3

A Risk-Based Premium: What does it mean for DB Plan Sponsors?[1]

3.1 Introduction

In chapter 2 we have dealt with the market consistent valuation of a stylized equity-linked life insurance contract in a regime-switching model. Specifically, we were concerned with the pricing of some embedded options under mortality risk, that is a bonus option, which is a call option, and especially a surrender option which represents an American stlye put option. In the present chapter we also work in a contingent claim framework and derive the market consistent value for the insurance provided by the Pension Benefit Guarantee Corporation (PBGC). This market consistent value of the insurance is interpreted as the fair risk-based premium a sponsoring company should pay to the PBGC. To determine this risk-based premium we have to again price an embedded option, which in our application turns out to be a down-out put option with rebate payments. Although the corresponding insurance contract is also long dated and thus it would be realistic to incorporate mortality and economic risk into the model as in the previous chapter, we neglect these sources of risk and rather model the pension fund's and the plan sponsor's investment policy in a Black Scholes setup. Unlike chapter 2 we are now able to derive a closed-form solution for the corresponding embedded option. It is important to emphasize that although this chapter focuses on the PBGC insurance, the subsequent model holds for any pension guarantee fund, while the qualitative results carry over to any insurance guarantee fund in general.

The Pension Benefit Guaranty Corporation (PBGC) is a federal US corporation that was created by Congress in the 1974 Employee Retirement Income Security Act (ERISA) to provide pension insurance for participants in private defined benefit (DB) pension plans. By law any sponsoring firm of a qualified DB plan is required to get into an insurance contract with the PBGC. Broadly speaking, the PBGC can be considered as an insurance guarantee fund with the main difference that its clients are not solely insurance compa-

[1]This chapter is based on Chen and Uzelac (2014)

nies. The long-term financial viability of the PBGC has recently triggered hot discussion after the termination of several severely underfunded pension plans. Some prominent cases are Delphi, Lehman Brothers, Circuit City, GM and Chrysler, and some earlier cases like Anchor Glass Container Corporation Service Retirement Plan, the Pension Plan of Bethlehem Steel Corporation and Subsidiary Companies, and the Polaroid Pension Plan. These terminations have worsened the financial position of the PBGC dramatically. The PBGC reported a year-end deficit of $21.9 billion and $23 billion in its 2009 and 2010 Annual Report, although the PBGC still had a surplus of $9.7 billion in 2000. In fact, 2010 marked the eighth straight year that the PBGC had been on the Government Accountability Office's "high-risk" watch list. Experts warned that if the PBGC were forced to take over the pension plans of massive companies, its deficit could be even more substantial and the financial condition of the PBGC could be even worse.

There are a variety of drivers which have caused the severe deterioration of the PBGC's financial condition. One very important driver is the ineffective premium calculation. In 2010, about 70% of the PBGC premiums were flat [2] and 30% were variable rate premiums. [3] The flat part currently consists of a 42$ premium per participant while the variable rate premium is solely based on the pension funds' underfunding, i.e. the sponsoring companies pay 0.9% of the underfunding of their pension funds. It is important to emphasize that both the flat and variable premium fail to account for the overfunding of a pension fund, the credit risk of the firm, asset allocation risk in the pension fund, and the correlation between the assets of the firm and its pension fund. As already pointed out by Josh Gotbaum (director of the agency) in an interview in February 2011: "it is not fair to say to businesses that are financially sound and have plans in good shape that they should pay the same premiums as guys who are not." In the academic literature, Bodie (2006) and Wilcox (2006), Brown (2008) and Love, Smith and Wilcox (2009) point out a flat premium implies that the PBGC insurance is mispriced and mispriced pension insurance gives firms adverse incentives.[4] The economic rationale is that the charge of a flat premium leads to gains in market value of equity for more risky firms and thus gives sponsors the incentive to invest more riskily and to underfund their pension funds. A further consequence is that less risky sponsors and pension funds will subsidize more risky ones and therefore have an incentive to withdraw from the pension system,[5] see Stewart (2007).

A risk-based premium has been considered a possible solution to the adverse incentive and cross subsidization problem. President George W. Bush advocated a risk-based premium in his 2006 and 2007 government budget plan and in February 2011, President

[2] A flat premium charge is the practice in most of the existing insurance guarantee funds, see Schmeisser and Wagner (2013).

[3] Since 2007 there exists a third premium component called annual termination premium, but thus far this premium part is negligible, see pbgc.gov.

[4] Cummins (1988) applies the same argument to insurance guarantee funds in general.

[5] A pension fund could voluntarily drop out of PBGC when some specific criteria are fulfilled. See pbgc.gov for details.

Obama backed up the risked-based premium endorsed by Bush. The Obama's budget does not mention how to take into account the risks that different sponsors pose to their retirees and to the PBGC. This authority to adjust premiums is given to the PBGC board. Josh Gotbaum says "they would be based on the company and the plan. For example, company risk might be determined by the company's credit rating or the value of its debt securities. Plan risk could be based on the nature of its asset base". In the current stage, it is unclear for the PBGC how to enforce a risk-based premium.

In this chapter, we extend Chen's (2011) model to determine a risk-based premium for the insurance provided by the PBGC by modeling the conventional way of a premature termination, distress termination, instead of the less common involuntary termination. In an involuntary termination, when the pension fund's asset falls below or hits a pre-specified regulatory threshold before or at the maturity date T, the underfunded pension fund will be trusted by the PBGC. In a distress termination, the premature termination is triggered by the underfunding of the sponsoring company. The distress termination mechanism is motivated since 99% of all plan terminations until 1995 were distress terminations (see e.g. Kalra and Jain (1997)) and the more recent OECD paper by Blome, Fachinger, Franzen, Scheunstuhl and Yermo (2007) confirms that still the most part of terminations are of distress type. In addition, we consider a further realistic perspective by allowing for a capped PBGC insurance payoff.

We are able to obtain an analytical valuation formula for the premium. More importantly, this is the first study using recent data that empirically illustrates which sponsors could be charged a higher and which sponsors a lower premium. In particular, we acquired the relevant data of the 100 largest American DB sponsors. Our analysis shows that the premiums paid to the PBGC differ significantly according to the differences in the sponsor and pension fund risks. The empirical results nicely illustrate that our risk-based premium calculation does not give sponsors adverse incentives as an increase in pension fund or sponsor specific risks comes at the cost of paying a considerably higher premium to the PBGC. The new practice of the PBGC to also charge a variable rate premium, which is solely based on the underfunding of the pension funds, goes partly in the right direction to eliminate adverse incentives and cross subsidization as our results show that the funding ratio is the most significant driver of the risk-based premium. However, the variable rate premium does not take into account that overfunded pension funds should be charged a significantly lower premium than underfunded ones, a point which is an important implication of our model. Moreover, our results suggest that other financial risk factors need to be taken into account, in particular it is very important to also incorporate sponsor specific risks like the leverage of the sponsoring companies in order to obtain an adequate risk-based premium.

The remainder of the chapter is organized as follows. The next section briefly reviews the literature about insurance guarantee funds in general and about the PBGC insurance. In the sections 3.3 and 3.4 we first model the insurance guarantee provided by the PBGC under distress termination and then derive the valuation formula for the risk-based premium by using the contingent claim approach. In the following section some comparative statistics based on the analytical formula in the previous section are exhibited. In the sections 3.6.1 and 3.6.2 we use the real data to determine and compare the risk-based premiums for the 100 largest DB pension funds in the US. The section 3.7 discusses a more general distress termination procedure in which the premature termination of the pension fund is modeled as the event that the pension fund assets and the assets of the sponsoring company fall below the pension liabilities and the corporate debt. This model is used as a robustness check for our proposed distress termination model. Finally, we present our conclusions in section 3.8 and the detailed derivation of the risk-based premium calculation in section 3.9.

3.2 Literature Review

The academic literature about insurance guarantee funds in general is extensive. Cummins (1988) employs an option pricing approach to determine risk-based premia for the insurance guarantee fund under three different model assumptions. This model is extended by Duan and Yu (2005) to a multi-period setting taking risk-based capital allocations into account. Rymaszewski, Schmeisser and Wagner (2012) introduce the concept of utility-based premiums. Moreover, Han, Lai and Witt (1997) address the problem of a system with ex post charges [6] not being able to be organized in a truly risk-based way due to the fact that the insolvent company, which may have been the most at risk, is typically not charged at all. Yasui (2001) points out that ex ante levies have the advantage of enabling relatively quick handling of insolvency cases, as funds for policyholder compensation are always available, which is particularly important if large (insurance) companies go bankrupt.

Considering specifically the PBGC insurance, three strands of academic literature have been developed. A recent strand is built by Romaniuk (2011), who studies the investment problem of the PBGC. Another direction is to determine an optimal intervention policy for the PBGC in terms of finding critical funding ratios such that the PBGC prematurely terminates these underfunded DB pension plans. This was first studied by Kalra and Jain (1997) and is also studied in this dissertation in chapter 4. Since Sharpe (1976) the economically fair pricing problem of the PBGC insurance has also been widely analyzed in the literature. The realistic but in practice less common involuntary termination case is studied in a variety of papers. Most studies assume the term to maturity of the PBGC insurance is known and also ignore that the pension fund can be closed prematurely due to its underfunding, thus they model the PBGC insurance as a plain vanilla put option, see Treynor (1977), Chen, Ferris and Hsieh (1994) and others. The premature termination is first considered in Kalra and Jain (1997) and extended by Chen (2011) by modeling the PBGC insurance as a secondary guarantee. This chapter studies the fair pricing problem of the PBGC in the most relevant distress termination framework and can be seen as an

[6] In the context of the PBGC's premium practice the variable rate premium is an ex post charge, while the flat rate premium is as an ex ante charge.

extension to the influential papers of Marcus (1987) and mainly of Lewis and Pennacchi (1994). Marcus (1987) models the PBGC's liability as a contingent forward, which allows the PBGC to gain the surpluses from overfunded terminated plans. It is not very realistic because the PBGC's liability can even become negative in this model, but law does not allow the PBGC's liability to be negative. Lewis and Pennacchi (1994) extend the previous literature by considering the PBGC liability as a contingent put option.

Our modeling framework clearly differs from the above authors. That is, we consider the PBGC insurance from the perspective of a representative beneficiary who retires at a specific nonrandom time in the future. Accordingly the PBGC liability becomes a down and out put option in our model. In addition we incorporate the two realistic perspectives of a secondary guarantee and of the capped PBGC insurance payoff, which are not covered in Marcus (1987) or Lewis and Pennacchi (1994). Notwithstanding these differences with Lewis and Pennacchi (1994) the economic implications of our models are similar. First, in contrast to Marcus (1987), our models rule out a negative value for the PBGC liability. Second, our models suggest that sponsors with better funded pension funds should pay less premiums to the PBGC. Third, we also obtain the intuitive result that a sponsor with a higher firm net worth should pay less premiums to the PBGC.

3.3 Model Setup

The basic model setup is based on Chen (2011) and most of the notation is drawn from that paper. However, Chen (2011) models an involuntary distress termination while this chapter considers distress termination.

We consider the pension insurance for a single-employer's defined benefits pension plan. Let us assume that the pension plan is issued at time $t_0 = 0$ to a representative beneficiary and that the benefits are paid out as a lump-sum payment B_T at the beneficiary's retirement date T. In order to focus on the effect of the investment policy of the pension fund, of the plan sponsor, and their possible default on the insurance of the PBGC, we assume that the pension liability B_T is deterministic. In other words, we assume that the plan sponsor is obliged to pay a fixed pension benefit T years from now on. In chapter 5 we will model B_T stochastically, as a function of the beneficiary's salary and years of service, and also show how the usual life annuity payments can be converted into a lump-sum.

There is a risk free asset F and a traded risky asset A in the economy. The risk free asset evolves according to

$$\mathrm{d}F_t = rF_t\mathrm{d}t, \quad F_0 = 1 \tag{3.1}$$

for a deterministic risk free rate r. The traded risky asset A evolves according to

$$dA_t = rA_t dt + \sigma_A A_t dW_t^{Q_1}, \quad A_0 = a, \tag{3.2}$$

where σ_A is the constant volatility $\sigma_A > 0$ and W^{Q_1} is a standard Brownian motion under the risk neutral probability measure $\mathbb{Q}$. As the main focus of the paper is the determination of the PBGC insurance premium, the asset processes have been expressed immediately under the risk-neutral instead of the real-world probability measure, see also remark 3.3.1 for a justification. Pension funds typically follow a rebalancing strategy in which the actual asset allocation fluctuates closely around a given strategic asset allocation. To analyze rebalancing and to take account of the plan's investment portfolio, we assume that the pension fund trades only in the risk free asset F and the risky asset A in a self-financing way starting with initial wealth X_0. Using π to denote the fraction of wealth invested in the risky asset A and the remaining $(1-\pi)$ fraction invested in the risk free asset F, we can write down the following assets process of the pension fund under the risk neutral measure $\mathbb{Q}$:

$$dX_t = rX_t dt + \pi\, \sigma_A\, X_t\, dW_t^{Q_1}. \tag{3.3}$$

Compared to (3.2), the volatility of the pension fund's assets becomes $\pi\sigma_A$. For $\pi = 0$, the pension fund invests in the risk-free assets only; and for $\pi = 1$, the pension fund invests in the risky assets only.

We assume that the plan sponsor's market value of assets also follow Black-Scholes dynamics with a volatility $\sigma_c > 0$. Under the risk neutral probability measure $\mathbb{Q}$, the market value of assets evolves over time according to

$$dC_t = rC_t dt + \sigma_c C_t(\rho dW_t^{Q_1} + \sqrt{1-\rho^2} dW_t^{Q_2}), \quad C_0 = c, \tag{3.4}$$

where W^{Q_2} is again a standard Brownian motion under the risk-neutral probability measure Q, independent of W^{Q_1}, and C_0 is the initial value of the sponsoring company's assets. Note that the sponsoring corporation's and the pension fund's assets are correlated with a correlation coefficient $\rho \in (-1, 1)$.[7] For $\rho = 0$, these two assets are uncorrelated.

Remark 3.3.1. *We are going to use an option pricing approach to determine a risk-based premium for the PBGC insurance. It is important to emphasize that this approach crucially depends on the assumption that all cash flows in our model can be replicated. If we deviate from this implicit assumption the only alternative to derive a risk-based premium would be to rely on actuarial valuation techniques, which would require the specification of time and state preferences. However, as in our context it is not really clear whose preferences should be modeled we opt for the option pricing approach.*

[7] We exclude the perfect correlation cases: $\rho = 1$ and $\rho = -1$. In these extreme cases, the pension fund either invests directly in the stock of the sponsoring company or in a portfolio which is perfectly negatively correlated with the sponsor's assets. In both cases, the sponsor's and the pension fund's assets are fully driven the randomness W^{Q_1}. The valuation becomes simpler and differs from what we will present in the remaining texts.

3.4 The Distress Termination Framework

Distress termination is initiated by the plan sponsor by proving that it is unable to pay its liabilities and to remain in business.[8] In this simple distress termination model, we assume that the plan sponsor defaults if the threshold $\epsilon\phi C_0 e^{gt}$, $\epsilon \geq 1$ and $\epsilon\,\phi < 1$, is hit. That is, the sponsoring company initiates default if it is unable to pay its outstanding corporate debt $\phi\,C_0\,e^{gt}$ plus an additional buffer $(\epsilon - 1)\ C_0\,\phi\,e^{gt}$. The product $\phi\,C_0$ is the initial debt value and g the constant growth rate at which the debt level increases. The additional buffer $(\epsilon - 1)\,\phi\,C_0\,e^{gt}$ serves to partly or fully cover the possible underfunding of the pension fund. Note that including the regulatory parameter $\epsilon > 1$ can be interpreted economically. The additional buffer can be understood as the moral obligation of the plan sponsor to always be able to cover at least partly the deficits of the pension fund. From the economic perspective only a small value for ϵ is reasonable since a firm that performs well and has a sufficiently high net worth has no incentive to terminate its business. The inequality $\epsilon\phi < 1$ is a technical condition which makes sure that the sponsoring company is not yet defaulted at the contract-issuing time $t = 0$. As in Chen (2011) we formulate the termination event in a standard barrier option framework and the termination time τ is constructed as the first hitting time that the plan sponsor's assets fall below or cross the threshold $\epsilon\,\phi\,C_0\,e^{gt}$:[9]

$$\tau = \inf\{t | C_t \leq \ \epsilon\,C_0\,\phi\,e^{gt}\}. \tag{3.5}$$

If $\tau \leq T$, there is a permature/mature termination enforced by the regulator. If $\tau > T$, the pension plan is naturally closed at the maturity date T.

We need to impose the constraint $\epsilon \geq 1$ to incorporate the realistic perspective that the sponsoring company provides the primal support and the PBGC insurance is considered as a secondary guarantee. Specifically, the sponsoring company will not provide the financial support at all cost and its support depends on its own funding situation and the funding situation of the pension fund:

- when the sponsoring company defaults, but the pension fund is sufficiently funded, then the sponsoring company does not have to balance any deficits of the pension fund;
- when the sponsor defaults and the pension fund is underfunded, then the sponsoring company will fully cover the deficit if the buffer is sufficient or it will only partly cover it if the deficits exceed the buffer.

The residual deficits that the sponsor cannot cover will become a financial burden on the PBGC. This way of modeling distress termination is simplifying and it has two weaknesses.

[8] In practice, the plan sponsor has to meet one or more of four distress tests specified by ERISA. Our framework mostly corresponds to the business continuation distress test, which is the most commonly used one.

[9] Strictly speaking, the premature termination is not triggered by the sponsor's underfunding, but due to its poor financial condition.

First, we could have inefficient termination, i.e. the pension fund is terminated though it performs sufficiently well to cover the entire pension liabilities. It is important to note that this scenario is possible as long as we do not jointly incorporate the distress and involuntary termination case.[10] Second, there could arise a scenario where the pension fund performs very poorly, but the sponsoring company's assets do not hit the threshold such that neither the sponsor nor the PBGC covers the deficit in the pension liabilities. However these two scenarios occur rarely in practice. First, as Bodie, Light, Morck and Taggart Jr. (1987) show the DB plan's funding degree is considerably positively related to the sponsoring company's long run profitability, therefore it is unlikely that the the DB plan is very well funded while the plan sponsor is in financial distress. Second, if the sponsoring company is sufficiently solvent while the DB plan is fairly underfunded, then the sponsoring company is required to make considerable contributions to improve the pension fund's funding status. More importantly, this simple distress termination model has the advantage that we will obtain closed-form solutions for the PBGC premium.

If the plan sponsor initiates distress termination prematurely or at maturity ($\tau \leq T$), the sponsor support and the insurance payoff of the PBGC occur already at τ. If the pension plan is naturally terminated at maturity ($\tau > T$), both of the supports follow at time T. Based on the above assumptions, we discuss the possible deficits the plan sponsor and the PBGC need to cover for both cases: $\tau \leq T$; and $\tau > T$. In the former case, the assets of the pension fund will be examined right at $\tau \leq T$. If we observe $X_\tau < B_T e^{-r(T-\tau)}$, [11] i.e. the assets of the pension fund fall below the discounted promised pension payment, the pension fund is underfunded and the deficits of the fund are $(B_T e^{-r(T-\tau)} - X_\tau)$. In this scenario, the sponsoring company is obliged to provide a primal support to the underfunded pension fund. Whether the sponsor provides a partial or full support depends on the magnitude of the underfunding. The size of this primal support is given by

$$\begin{aligned} SP(\tau) =& (B_T e^{-r(T-\tau)} - X_\tau) 1_{\{X_\tau < B_T e^{-r(T-\tau)}\}} 1_{\{(B_T e^{-r(T-\tau)} - X_\tau) < (\epsilon-1)\phi C_0 e^{g\tau}\}} \\ &+ (\epsilon - 1)\phi C_0 e^{g\tau} 1_{\{X_\tau < B_T e^{-r(T-\tau)}\}} 1_{\{(B_T e^{-r(T-\tau)} - X_\tau) > (\epsilon-1)\phi C_0 e^{g\tau}\}}, \end{aligned} \tag{3.6}$$

where $1_{\{A\}}$ is the indicator which is 1 when event A occurs and 0 otherwise. The first term on the right-hand side of (3.6) corresponds to the case in which the buffer is sufficient to cover all the deficits of the pension fund. In this case, the covered deficit is the difference between the discounted liability $B_T e^{-r(T-\tau)}$ and the asset value of the pension fund X_τ. The second term corresponds to the case in which the buffer is insufficient to cover all the deficits of the pension fund. After paying back to its own debt holders, the corporate can provide what still remains, i.e. $(\epsilon - 1)\phi C_0 e^{g\tau}$, to the beneficiary.

For $\tau > T$, the sponsor guarantee at T is almost the same as that in Chen (2011) with

[10] By modeling the first hitting time as the minimum of the stopping time in Chen's (2011) model and our above specified stopping time we would eliminate inefficient termination.

[11] Since the payment follows at τ, we have adjusted the pension benefits to the discounted value accordingly.

the only difference that the value of C_T is unknown and we know that the sponsors assets have never hit the treshold $\epsilon\,\phi\,C_0\,e^{gT}$. That is, the sponsoring company needs to provide the guarantee

$$\begin{aligned} SP(T) =&(B_T - X_T)1_{\{C_T > \phi C_0 e^{gT} + (B_T - X_T)\}}1_{\{X_T < B_T\}} \\ &+ (C_T - \phi C_0 e^{gT})1_{\{\epsilon\phi C_0 e^{gT} < C_T < \phi C_0 e^{gT} + (B_T - X_T)\}}1_{\{X_T < B_T\}}. \end{aligned} \tag{3.7}$$

We can express the entire support provided by the plan sponsor in the following compact form:

$$SP = SP(\tau)1_{\{\tau \le T\}} + SP(T)1_{\{\tau > T\}}. \tag{3.8}$$

In addition to the sponsor support, the PBGC provides a secondary security to the pension plans, i.e. it covers the residual deficits that the sponsoring company is unable to cover. However, in practice the amount of residual deficits the PBGC covers is capped. The inclusion of a cap is important since if a pension fund is terminated that is highly underfunded the PBGC can only provide a not-too-high fraction of the retirement income. A good example is Delphi where the retirees typically received pension payments ranging from \$3000 to \$4000 a month, but a 55-year old retiree can receive at most \$2025 from the PBGC. Hence, we include an additional important perspective by allowing for a capped payoff $\bar{G}$, where we assume $\bar{G} < B_T$. We assume that $\bar{G}$ is the capped payoff binding for the retirement date T. If there is a premature termination at $\tau < T$, the capped amount is correspondingly adjusted to $\bar{G}e^{-r(T-\tau)}$. Denote the difference between the present value of the pension liabilities and the buffer by $BC(t) = B_T\,e^{-r(T-t)} - (\epsilon - 1)\,\phi\,C_0\,e^{gt}, t \in [0, T]$, then we can express the insurance of the PBGC at $\tau \le T$ as the minimum of the difference between the residual deficit $BC(\tau) - X_\tau$ and the capped amount $\bar{G}e^{-r(T-\tau)}$. Formally the insurance of the PBGC is then given by

$$\begin{aligned} G(\tau) =& \min(BC(\tau) - X_\tau, \bar{G}e^{-r(T-\tau)})\,1_{\{X_\tau < BC(\tau)\}} \\ =&(BC(\tau) - X_\tau)\,1_{\{\max(0,BC(\tau)-\bar{G}e^{-r(T-\tau)}) < X_\tau < \max(0,BC(\tau))\}} \\ &+\bar{G}e^{-r(T-\tau)}1_{\{X_\tau < \max(0,BC(\tau)-\bar{G}e^{-r(T-\tau)})\}}, \end{aligned} \tag{3.9}$$

where in the second step we just split the min function in the 2 possible cases.

If $\tau > T$, the insurance payoff of the PBGC is once again described by the minimum between the residual deficit and the capped amount. In this case, the residual deficit differs, since now C_T is unknown. We have

$$\begin{aligned} G(T) =&1_{\{X_T < B_T\}} \min\big(\bar{G}, \max(B_T - X_T - \Phi_c(T), 0)\big) \\ =& \max\big\{0, B_T - X_T - (C_T - \phi C_0 e^{gT})\big\}\,1_{\{\epsilon\phi C_0 e^{gT} < C_T < \phi C_0 e^{gT} + B_T - X_T\}} \\ &\cdot 1_{\{X_T < B_T\}}1_{\{B_T - X_T - (C_T - \phi C_0 e^{gT}) < \bar{G}\}} \\ &+ \bar{G}1_{\{\epsilon\phi C_0 e^{gT} < C_T < \phi C_0 e^{gT} + B_T - X_T\}}1_{\{X_T < B_T\}}1_{\{B_T - X_T - (C_T - \phi C_0 e^{gT}) > \bar{G}\}}. \end{aligned} \tag{3.10}$$

The PGBC balances the deficits of the pension fund only when the sponsoring company is unable to cover the entire deficits, i.e. when $\epsilon\phi C_0 e^{gT} < C_T < \phi C_0 e^{gT} + B_T - X_T$. The size of the PBGC's payoff depends on whether the capped amount $\bar{G}$ is binding. More compactly, the insurance of the PBGC can be expressed as follows:

$$G = G(\tau)1_{\{\tau \leq T\}} + G(T)1_{\{\tau > T\}}. \tag{3.11}$$

That is, the insurance payoff provided by the PBGC is a package of exotic put options.

3.4.1 A Risk-Based Premium

Usually periodic (yearly) premiums are charged by the PBGC for providing the insurance. For simplicity, we assume that the PBGC receives an upfront premium for providing the security to the beneficiary. The upfront premium corresponds to the today's price of the insurance claim (3.11) on pension fund's and sponsoring company's assets. In our context, the risk-based premium of the PBGC insurance can be derived by computing the expected discounted payment under the risk-neutral measure $\mathbb{Q}$.

Proposition 3.4.1. *The risk-based premium paid by the plan sponsor to the PBGC is the expected discounted insurance payoff under the risk neutral probability measure:*

$$G_0 = \mathbb{E}^{\mathbb{Q}}\left[e^{-rT}G(T)1_{\{\tau > T\}}\right] + \mathbb{E}^{\mathbb{Q}}\left[e^{-r\tau}G(\tau)1_{\{\tau \leq T\}}\right], \tag{3.12}$$

where $\mathbb{E}^{\mathbb{Q}}$ denotes the expected value under the risk-neutral measure $\mathbb{Q}$. The closed-form solution is given in the appendix.

Proof 3.4.2. *A detailed derivation is provided in the appendix in section 3.9.*

3.5 Comparative Statistics

Before we move to empirically illustrate our risk-based premium calculation for the 100 biggest DB pension funds in the US, we exhibit some comparative statistics. The purpose of the analysis is to demonstrate the impact of several main parameters on the PBGC premium. The premium is expressed as the percentage of the promised pension payment B_T. By changing only one parameter each time, we can better understand what role each parameter plays in the PBGC premium. For this numerical calculation, we fix the relevant parameters as follows:

$$\begin{aligned} \epsilon =& 1.05,\ r = 0.05,\ \sigma_c = 0.25,\ \phi = 0.6,\ g = r,\ \pi = 0.6, \\ X_0 =& 100,\ B_T = 240,\ \sigma_A = 0.20,\ C_0 = 100,\ T = 15,\ \rho = 0.2, \bar{G} = 120. \end{aligned} \tag{3.13}$$

Figure 3.1 plots the PBGC premium as a function of the promised pension payment B_T for diverse correlation coefficient ρ. As we have fixed the initial asset value of the pension

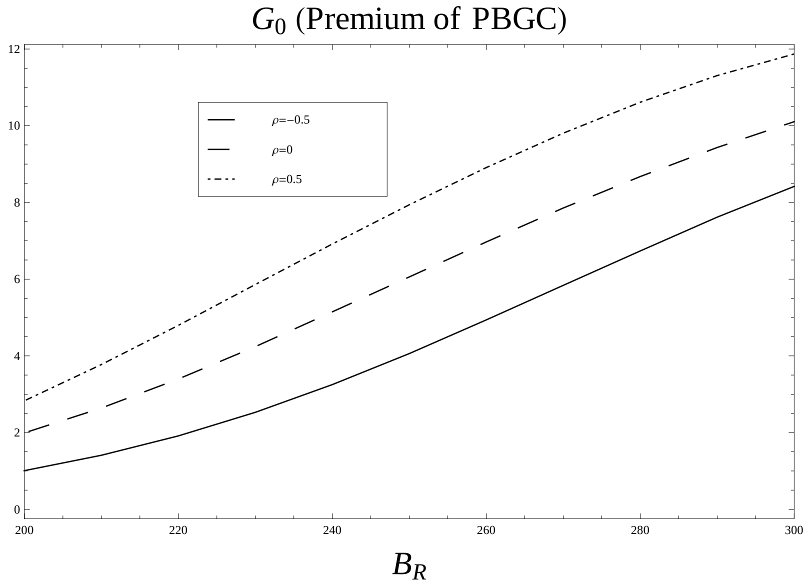

Figure 3.1: Fair premium G_0 as a function of B_T for different ρ levels.

fund to 100, an increase in B_T implies a deterioration of the funding situation of the pension fund, see the next section. As a result, it becomes more likely that the pension fund's terminal asset X_T is insufficient to provide the promised pension payment. Hence, the chance that the PBGC needs to cover the possible deficits of the pension fund rises, which leads to a higher PBGC premium. So underfunded pension funds are supposed to be charged with a higher risk-based PBGC insurance premium. Furthermore, it can be read from Figure 3.1 that the premium increases in the correlation ρ between the pension fund's and sponsor's assets. When the sponsor's and the pension fund's assets are strongly positively correlated, the likelihood is high that the sponsoring company is unable to provide full/partial guarantee when the pension fund is already at default. These results imply that the sponsoring company can free ride the PBGC much when the correlation coefficient is high.

How the volatility σ_c of the sponsor's asset influences the premium is exhibited in Figure 3.2. The volatility σ_c might show a non-monotone effect on the premium, depending on the regulatory parameter ϵ. However for a realistic scenario, say $\epsilon = 1.05$, we obtain the economically intuitive upward sloping curve for the premium as a function of σ_c, i.e. plan sponsors who invest in more risky investment portfolios should be charged with a higher premium for the PBGC insurance. A higher σ_c is more likely to cause the default of the sponsoring company. Hence, the premium part upon premature termination increases in σ_c. But it causes a simultaneous decrease of the premium part upon natural termination. For a low regulatory parameter, the former effect seems to dominate. For a higher and less realistic ϵ level (e.g. $\epsilon = 1.1$ or $\epsilon = 1.15$), an increase in σ_c could lead to a hump-shaped curve for the premium. The impact of the sponsor's assets volatility is most pronounced for the more realistic ϵ level, $\epsilon = 1.05$. In the same figure, we observe the negative relation between the premium and the regulatory parameter ϵ. This parameter influences the prob-

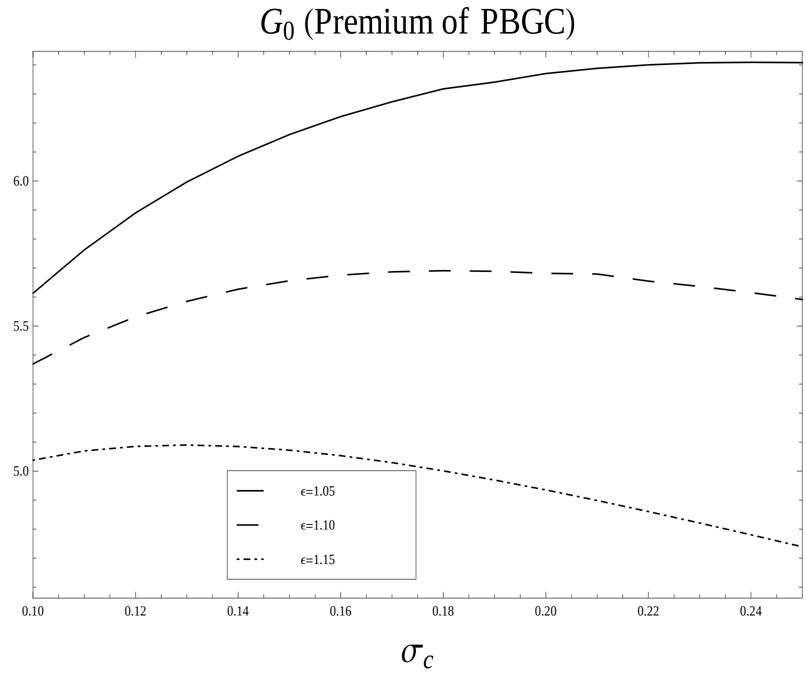

Figure 3.2: Fair premium G_0 as a function of the volatility of the sponsor's assets σ_c for different ϵ levels.

ability of premature termination and how high the buffer owned by the sponsoring company is to cover the deficits of the pension funds. The main effect is: the higher the ϵ level, the more deficits the sponsor company can balance. Hence, an increase in ϵ leads to a smaller insurance payoff of the PBGC, and consequently to a smaller PBGC premium.

In Figure 3.3, we observe the effect of ϕ on the PBGC premium for different σ values, where ϕ is the initial leverage ratio of the the plan sponsor and σ_A drives the volatility of the pension fund's assets. It is observed that the PBGC premium is (weakly) hump-shaped in the leverage ratio ϕ. The mainly increasing premium is primarily the result of an intuitive economic effect: the higher the leverage ratio, the more the sponsoring company needs to serve the outstanding liability and therefore the more likely it is that the sponsoring company will initiate distress termination before the retirement date, which increases the premature premium part. The (weak) decrease for higher leverage ratios is due to two effects. First there is a counter-effect that in our simple distress termination model a higher leverage ratio means that the sponsoring company uses a higher buffer in absolute terms to provide the primal guarantee. Second a higher ϕ also decreases the probability of natural termination, which might decrease the premium in charge of natural termination. In the same figure, we observe that the PBGC premium goes up in the volatility of the pension fund. The more risky the pension fund's portfolio, the more probably the pension fund becomes underfunded. Therefore, it is of high likelihood that the PBGC needs to balance the residual deficits of the pension fund. The PBGC premium rises.

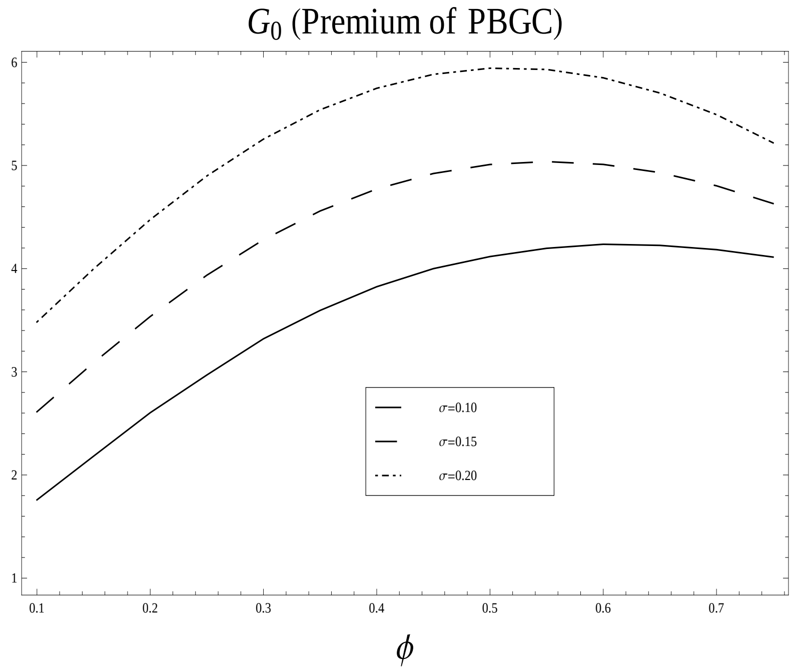

Figure 3.3: Fair premium G_0 as a function of ϕ for different σ levels.

Finally in Figure 3.4 two apparent effects on the PBGC premium are demonstrated: the increasing effect of the capped amount and the rising effect of the equity-holding (π) of the pension fund. When the PBGC promises a higher capped amount $\bar{G}$, it means a higher cost for the PBGC. A higher equity holding has a similar effect as holding a more risky portfolio for the pension fund. Therefore, the PBGC premium rises.

3.6 Empirical Example

In the following section we present an empirical example where we compute risk-based premiums for a representative subsample of the 100 largest US corporates and their defined benefit plans.The corresponding data set is obtained from P&I Investments. Representative means that we include sponsors from all industry sectors in this subsample. It is important to emphasize that the risk-based premiums we estimate can hardly be interpreted as real risk premiums since as well our theoretical model relies on some simplifying assumptions as does the estimation of the relevant parameters at hand. Nonetheless the major objective of this empirical exercise is to illustrate how premiums sponsoring companies had to pay would diverge if the premium calculation were risk-based.

3.6.1 Data and Estimation Methodology

Our data set contains the fair value of the pension plan assets, the benefit obligation, the funding ratio, the total corporate asset value and the equity value of the sponsor and the asset allocation of the pension fund for each sponsor.

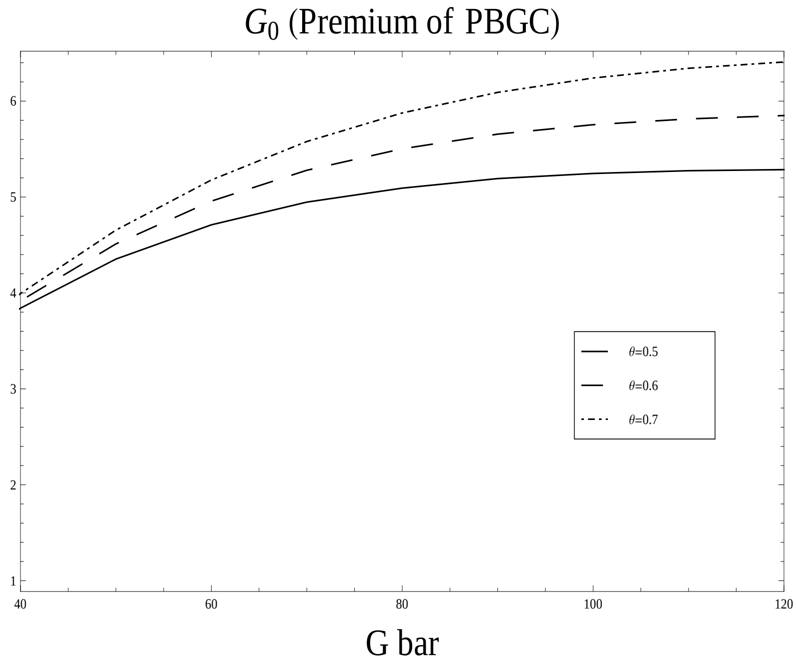

Figure 3.4: Fair premium G_0 as a function of $\bar{G}$ for different π levels.

First of all, we have to resolve the heterogeneity in our sample. This is necessary because the different firms have a different number of policyholders with different characteristics as age, income etc., but our model yields a theoretical premium for a single representative policyholder. This implies that we cannot readily infer the initial wealth of the pension fund X_0, the initial value of the firm C_0 and the present value of the pension obligations B_T from the data set. That is why we normalize the initial wealth level X_0 to 100 \$ and the initial asset value C_0 to 300 \$ for each sponsor in the sample.[12] Then, to obtain an adequate estimate for the lump sum payment to the beneficiary in T years B_T in our model we can use the observed funding ratio as the main input. First, we know that the funding ratio is defined as the ratio of the initial asset value and the initial outstanding accrued liability.[13] More precisely, let R_0 be the funding ratio and B_0 the initial pension liability, then $R_0 = \frac{X_0}{B_0} \Leftrightarrow B_0 = R_0^{-1} X_0$. In addition, we know that the initial outstanding liability in our model is simply given by the discounted pension liability B_T, that is $B_0 = e^{-rT} B_T$. Equating the latter two equations we estimate B_T as

$$\hat{B}_T^i = e^{rT} X_0 (R_0^i)^{-1},$$

where the superscript i denotes the pension liabilities of sponsor i. Put differently, our simplifications imply that the pension liabilities are a riskless asset under the risk-neutral probability measure with initial value $X_0 R_0^{-1}$.

[12]The sponsor's initial asset value is set higher since in practice sponsors assets usually take significantly higher values than those of pension funds.

[13]In chapter4 we model the funding ratio dynamically.

To estimate the fraction invested in the risky asset we use the asset allocation data in our sample. Our sample contains investment shares in equity, fixed income, alternatives, real estate, private equity, hedge funds, cash and other investments. As the investments within the different types of assets are not further characterized in the data set we classify the asset classes as risky or riskless by relying on long-term empirical evidence. Specifically, we classify investments in equity, private equity, hedge funds and real estate as investments in the risky asset, while we characterize cash and fixed income investments as riskless, since except hedge funds [14] the former have been significantly more volatile in the long run, see for instance Eychenne, Martinetti and Roncalli (2011). For the other assets and alternatives category in the sample, containing instruments like derivatives, commodities, balanced funds etc., in most cases we cannot infer from the data if these are rather equity or fixed income type investments. Nonetheless we cannot omit these categories because they are not negligible for a significant number of companies, i.e because their pension funds invest more than 10% in these categories. We decide to classify the half of them as risky and the other half as riskless as we think that this produces the smallest bias. The classification is summarized in table 1.

Risky Asset Class	**Riskless Asset Class**
Equity	Cash
Private equity	Fixed income
Hedge funds	Alternatives
Real Estate	Other investments
Alternatives	e.g Balanced Funds
Other investments	
e.g Derivatives	

Table 3.1: Classification of the assets into the risky and riskless asset class.

Then the fraction invested in the risky asset is estimated as

$$\hat{\pi}^i = \pi^i_E + \pi^i_{PE} + \pi^i_{HF} + \pi^i_{RE} + \frac{1}{2}(\pi^i_A + \pi^i_O),$$

where π^i_E denotes the observed percentage pension fund i invested in equity, analogously π^i_{PE} is the share invested in private equity, π^i_{HF} the share in hedge funds, π^i_{RE} in real estate, π^i_A the share in alternatives and π^i_O the share sponsor i invested in other assets.

To estimate the sponsor specific parameters as the leverage ϕ and the volatility of the sponsor's assets σ_c we generally need to estimate the market value of assets, which is defined as the sum of the market value of liabilities and the market value of equity. The market value of equity is easily determined because all sponsors in our sample are listed companies

[14]Hedge funds have on average a lower volatility than fixed income investments. However we consider them still as more risky, particularly because they have a substantially higher loss potential, see for instance Gaurav and Kat (2003).

and thus equity data is readily available. However it is hard to observe the market value of liabilities. That is why we follow the standard approach in the literature and consider the book value of liabilities, which is given in the balance sheet data of a company, as a proxy for the latter. Then we simply estimate the leverage as the ratio of the book value of liabilities and the approximated market value of assets, that is

$$\hat{\phi}^i = \frac{L_B^i}{L_B^i + E_M^i},$$

where L_B^i denotes the observed book value of liabilities of sponsor i and E_M^i the observed market value of equity, respectively.

For the estimation of the sponsor's asset volatility σ_c we use a historical data approach. The asset value is calculated as the sum of the market value of equity and the book value of liabilities. The book value of liabilities is available quarterly from the balance sheet data, so we construct a historical time series with quarterly data. For each data point and each sponsor we observe the market value of equity and infer the book value of liabilities. Then we compute the corresponding log returns for the estimated market value of assets. Eventually we estimate σ_c^i as the annualized volatility of these quarterly log return time series.

Finally we specify those parameters, which are not pension fund or sponsor specific. Parameters as the regulatory parameter ϵ, the maximum amount the PBGC provides $\bar{G}$ and the risk-free rate r are naturally not sponsor specific. The growth rate g of the sponsor liabilities is not sponsor specific here since it is a drift coefficient, which has to coincide with the risk-free rate under the risk-neutral probability measure. Other parameters like the correlation ρ of the sponsor's and pension fund assets and the volatility σ of the risky asset A in our model are sponsor and pension fund specific, but it is very hard to obtain adequate data to estimate them. Accordingly we keep these parameters constant across sponsors and pension funds.

We approximate σ by the S&P-500 volatility. Note that the estimated volatility of the pension fund i, which is given by $\hat{\pi}^i \hat{\sigma_A}$, still varies across pension funds because of the different share invested in risky assets.

We specify $\bar{G}$ such that the PBGC can provide a significant fraction c of the maximum pension liabilities in the sample, specifically we estimate

$$\hat{\bar{G}} = \max_i c\, \hat{B_T}^i .$$

This ensures that the PBGC covers a large part of the deficits in case of termination for any sponsor, which is in line with its legal obligation.

The other parameters are specified in the next section.

3.6.2 Estimation Results

First we have the following nonspecific parameter estimates and specifications:

$$r = g = 0.0413;\ \hat{\sigma} = 0.2022;\ \rho = 0.5;\ T = 15;$$
$$\epsilon = 1.05;\ \hat{\bar{G}} = 0.4 * 282.371 = 112.948.$$

The risk-free rate is approximated by the T-bill yield of a 20-year bond on 12/31/2010, σ_A is estimated as the annualized volatility of the S&P 500 daily log returns from 01/01/1996 until 12/31/2010. $T = 15$ is chosen because this is the average duration of pension liabilities. $\rho = 0.5$ is taken from Lewis and Pennacchi (1994). ϵ is set as in the comparative statics section. Note that this value satisfies the technical condition that no sponsor in our subsample defaults at $t = 0$. $\bar{G}$ is specified such as described in the previous section, where the maximum pension liabilities B_T are those of the sponsor Goodyear Tire & Rubber and the fraction of the maximum liabilities the PBGC can provide is set $c = 0.4$.

In table 3.2 below we present the sponsor and pension fund specific estimates for the representative subsample. The leverages are estimated with the corresponding market and balance sheet data on 12/31/2010 and the sponsors asset volatilities are estimated as described in the previous section for a period ranging from 01/01/2001 until 31/12/2010 [15] including 40 observation points for the approximated market value of assets. We normalize the premium sponsors paid to the PBGC by considering the percentage premium per pension liability, that is $\tilde{G} = \frac{G_0}{B_T}$. Due to the different levels of B_T, this ratio provides a better statistic for the premium instead of the absolute premium G_0.

First our estimation results illustrate that the percentage premiums per pension liability differ significantly across sponsors and their pension funds. The message of this important result is that if sponsor and pension fund risk are adequately taken into account then a mainly flat premium, which is still the current practice at the PBGC, is not justifiable at all.

More specifically, we see that the sponsor Bank of America pays the smallest percentage premium per liability though it has at the same time the highest leverage. Due to the hump shape of the premium as a function of leverage, the extremely high leverage dampens the magnitude of the premium slightly, so does the very low asset volatility. More importantly the very low premium is mainly justified because Bank of America has a very well funded pension fund, which ensures that even if it goes bankrupt in the near future it will be very likely that the corresponding pension liabilities can be met. The same justification also

[15] We cannot use a 15-year period for the historical estimation as some sponsors in our subsample were not listed prior to 2000.

applies to JP Morgan who pays the second smallest percentage premium per pension liability. On the other hand, Goodyear Tire&Rubber is clearly the sponsor carrying the most sponsor and pension fund risk because it has the worst funded pension fund, a very high leverage and a high investment share in risky assets. From the previous comparative statics section and these two examples we see that the funding ratio α, which accounts for the different sizes of the pension liabilities B_T, seems to be the most significant factor deciding the difference in the percentage premiums.

However our estimation results also confirm that other factors play a significant role in explaining differences in the premiums. To see this we can for instance compare sponsors having the same funding ratio as Ashland and Coca Cola. We observe that Ashland pays a percentage premium which is almost 3 times larger than that of Coca Cola. Here the substantial difference is clearly explained by the relatively high leverage of Ashland, which is about 2.7 times as large as that of Coca Cola and clearly dominates the effect of the more risky pension fund of Coca Cola due to the higher investment share in risky assets.

Another striking example illustrating the significant effect of the leverage is the comparison between the sponsors Wells Fargo and Coca Cola. Though the former has a considerably larger funding ratio, a substantially lower asset volatility and investment share in risky assets, it pays a 0.2% greater percentage premium per liability. This is purely explained by the very high leverage, which is more than 3 times larger than that of Coca Cola. Accordingly, we identify the leverage as the second major risk factor, in particular the major sponsor specific risk factor, that accounts for differences in the premiums per unit liability in our model.

Although their effect is smaller in magnitude, the investment share in risky assets and the sponsor's asset volatility together can produce a considerable effect on the PBGC premium. To illustrate this point we compare the sponsors Eli Lilly and Exxon Mobil. Eli Lilly has a moderately larger funding ratio and a slightly larger leverage than Exxon Mobil, but pays a about 2.6 times greater percentage premium per liability. As these two opposite effects nearly offset the significantly greater premium is mainly the result of the substantially greater asset volatility and the considerably greater investment share in risky assets.

3.7 Extension: A More General Distress Termination Model

A more realistic way to model distress termination is to model the default of the sponsoring company as the event that the pension fund assets and the assets of the sponsoring company fall below the pension liabilities and the corporate debt. More precisely, we could model the first hitting time as

Sponsor	$\hat{\tilde{G}}^i$	$\hat{\tilde{G}}_1^i$	$\hat{\phi}^i$	$\hat{\sigma_c}^i$	R_0^i	$\hat{B}_T^i$	$\hat{\pi}^i$
3M	1.648	1.570	0.186	0.174	0.940	197.660	0.678
Aetna	6.372	10.470	0.696	0.146	0.901	206.215	0.755
American Electric	8.704	11.250	0.680	0.159	0.803	231.382	0.693
Ashland	8.704	9.992	0.578	0.186	0.754	246.419	0.470
AT&T	5.996	8.370	0.475	0.229	0.883	210.419	0.527
Bank of America	0.040	7.330	0.938	0.086	1.123	165.450	0.632
Baxter International	5.451	5.370	0.266	0.213	0.784	236.990	0.630
Boeing	7.016	7.800	0.579	0.166	0.833	223.049	0.459
Caterpillar	5.928	6.210	0.472	0.140	0.826	224.939	0.736
Coca-Cola	3.120	3.010	0.215	0.178	0.754	246.419	0.697
Consolidated Edison	5.878	5.910	0.634	0.075	0.749	248.064	0.763
Dominion Resources	2.863	5.160	0.551	0.147	1.137	166.831	0.656
Dow Chemical	7.773	8.140	0.542	0.147	0.749	248.064	0.593
Eli Lilly	5.890	5.970	0.315	0.197	0.861	215.795	0.808
Exxon Mobil	2.287	2.240	0.288	0.144	0.822	226.034	0.609
FedEx	3.027	3.150	0.287	0.186	0.918	202.396	0.500
General Dynamics	7.972	8.210	0.418	0.180	0.677	274.446	0.722
Goodyear Tire & Rubber	13.529	13.960	0.836	0.139	0.658	282.371	0.667
Hewlett-Packard	5.261	7.150	0.468	0.227	0.865	214.800	0.392
Honey International	7.068	7.370	0.396	0.189	0.813	228.536	0.732
IBM	4.007	5.170	0.331	0.230	0.980	189.592	0.528
JP Morgan	0.099	7.580	0.922	0.078	1.301	142.813	0.814
United Technology	3.690	3.810	0.332	0.170	0.916	202.838	0.614
Walt-Disney	4.731	4.750	0.305	0.174	0.703	264.296	0.600
Wells-Fargo	3.307	8.430	0.875	0.060	0.932	199.356	0.641

Table 3.2: The premiums are expressed as the percentage ratio of the estimated promised benefit $\hat{B}_T^i$. The first column denotes the premium in the simple distress termination model and the second column the one in the extended distress termination model. $\hat{\phi}^i$ and $\hat{\sigma_c}^i$ are estimated the debt ratio and asset volatility of sponsor i. $\hat{\pi} i$ is the estimated equity holding of the pension fund i.

$$\tau = \inf\{t | C_t + X_t \leq \eta(C_0 \, \phi \, e^{gt} + B_T e^{-r(T-t)})\}, \tag{3.14}$$

where $\eta \in (0,1)$ is assumed.

Observe that unlike our simple distress termiantion model in this extended model the sponsoring company can at most provide a partial support if distress is initiated through the underfunding of the pension fund. The main advantage of this setup is that by construction of the hitting time the scenario where the pension fund performs very poorly but neither the sponsor nor the PBGC covers its deficits cannot arise. The price we have to pay for this more realistic setup is that we can no longer obtain analytic solutions for the PBGC premium. The problem is that we cannot derive the density and the distribution of the first hitting time in closed form, since the distribution of the minimum of the sum of two log-normal processes is unknown. Then the premium needs to be computed by Monte Carlo simulation.

We use the more realistic distress termination model as a robustness check for our simple distress termination model. The payoff functions, which we need for the simulation are given in the appendix in section 3.9.2.

The second column in 3.2 displays the percentage premiums per pension liability for the representative subsample. We observe that for most of the DB sponsors these premiums are very similar to the ones in our simple distress termination model and more importantly the extended distress termination model also confirms that the premiums differ significantly across sponsors. However there are some significant deviations.

Firstly, one can see that sponsors with a very high leverage, especially those from the banking industry, pay a substantially higher percentage premium in this extended distress termination model. For example, the sponsors Bank of America and JP Morgan do no longer have the smallest percentage premium, but they belong to the sponsors who would pay a higher percentage premium in this extended distress termination model. The reason for this deviation is that in our simple distress termination model the premium can be hump-shaped in the leverage, see figure 3.3. Accordingly a very high leverage might dampen the premium in this model, whereas the premiums increase in the leverage in the extended distress termination model. This is because here only the effect that a higher leverage increases the likelihood of premature termination is present. Secondly, we can observe that the sponsor's asset volatility has a stronger impact in the extended distress termination model, for instance sponsors with a fairly high asset volatility like IBM or Hewlett Packard pay a significantly higher percentage premium per pension liability. More generally, one can further notice that the percentage premiums are higher in the extended distress termination model, accordingly there is a tendency that the premiums in our simple distress termination model are downward biased.

3.8 Conclusion

In this chapter, we model the PBGC insurance taking account of distress termination, the most common type of termination in practice: the premature termination of the pension fund is caused by the underfunding of the sponsoring company. A risk-based premium is determined for the valuation of the PBGC insurance. Assuming both the pension fund's and the plan sponsor's assets follow Black-Scholes dynamics and are correlated, we obtain a analytic pricing formula for the risk-based premium. We extend the literature dealing with the fair pricing of the PBGC insurance in the distress termination framework by incorporating two realistic perspectives. First, the PBGC insurance is modeled as a secondary guarantee and second we allow for a capped insurance payoff.

This chapter also provides an important empirical contribution since this is the first study using recent data that empirically illustrates which sponsors could be charged a higher and which sponsors a lower premium. Specifically, using a data set for the 100 largest American DB sponsors, we show that the premiums paid to the PBGC should differ significantly according to the differences in the sponsor's and pension fund's risks. Most importantly, our results illustrate that our risk-based premium calculation does not give sponsors adverse incentives to introduce risk into the pension promises as an increase in pension fund or sponsor specific risks comes at the cost of paying a higher premium to the PBGC. The use of a variable rate premium which solely considers the underfunding (and ignores the overfunding) of the pension funds in the premium calculation is partly consistent with our results that the funding ratio is the most significant driver of the risk-based premium. An important implication of our model is that overfunded pension funds should be charged with a significantly lower premium than underfunded ones. Moreover our results suggest that sponsor specific risks play a very important role in the risk-based premium, where the leverage of the sponsoring company is the most pronounced sponsor specific risk factor in our model.

3.9 Appendix

3.9.1 Derivation of the fair premium PBGC receives

The risk-based premium of the PBGC insurance can be decomposed into two parts:

(a) $\mathbb{E}^{\mathbb{Q}}\left[e^{-rT}G(T)1_{\{\tau>T\}}\right]$

(b) $\mathbb{E}^{\mathbb{Q}}\left[e^{-r\tau}G(\tau)1_{\{\tau\leq T\}}\right]$.

In order to further calculate Part (a), we first rewrite

$$\begin{aligned}G_T =&\bar{G}1_{\{X_T<B_T\}}1_{\{\phi\epsilon C_0e^{gT}<C_T<\phi C_0e^{gT}+B_T-X_T-\bar{G}\}}\\ &+\max\left\{0,(B_T-X_T-(C_T-\phi C_0e^{gT}))1_{\{X_T<B_T\}}\right\}\\ &1_{\{\max\{\phi\epsilon C_0e^{gT},\phi C_0e^{gT}+B_T-X_T-\bar{G}\}<C_T<\phi C_0e^{gT}+B_T-X_T\}}.\end{aligned}$$

We proceed to determine Part (a) as follows

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{Q}}\left[e^{-rT}G(T)1_{\{\tau>T\}}\right]\\
=&\mathbb{E}^{\mathbb{Q}}\left[e^{-rT}G(T)\right]-\mathbb{E}^{\mathbb{Q}}\left[e^{-rT}G(T)1_{\{\tau\le T\}}\right]\\
=&\int_{-\infty}^{dx_1}\int_{dy_1}^{dy_2(x)} e^{-rT}\bar{G}\frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left\{-\frac{x^2+y^2-2\rho xy}{2(1-\rho^2)}\right\}dydx\\
&+\int_{-\infty}^{dx_1}\int_{dy_2(x)}^{dy_3(x)} e^{-rT}\max\left\{0, B_T-X_0e^{(r-\frac{1}{2}\pi^2\sigma^2)T+\pi\sigma\sqrt{T}x}-(C_0e^{(r-\frac{1}{2}\sigma_c^2)T+\sigma_c\sqrt{T}y}-\phi C_0e^{gT})\right\}\\
&\cdot\frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left\{-\frac{x^2+y^2-2\rho xy}{2(1-\rho^2)}\right\}dydx-\int_0^T\mathbb{E}^{\mathbb{Q}}[e^{-rT}G(T)|\tau=t]\mathbb{Q}(\tau\in\mathrm{d}t)
\end{aligned}
$$

where

$$
\begin{aligned}
dx_1=&\frac{\ln\frac{B_T}{X_0}-(r-\frac{1}{2}\sigma_A^2\pi^2)T}{\pi\sigma_A\sqrt{T}}\\
dy_1=&\frac{\ln\frac{\phi\epsilon C_0e^{gT}}{C_0}-(r-\frac{1}{2}\sigma_c^2)T}{\sigma_c\sqrt{T}}\\
dy_2(x)=&\frac{\ln\frac{\max\{\phi\epsilon C_0e^{gT},\phi C_0e^{gT}+B_T-X_0e^{(r-\frac{1}{2}\pi^2\sigma_A^2)T+\pi\sigma_A\sqrt{T}x}-\bar{G}\}}{C_0}-(r-\frac{1}{2}\sigma_c^2)T}{\sigma_c\sqrt{T}}\\
dy_3(x)=&\frac{\ln\frac{\phi C_0e^{gT}+B_T-X_0e^{(r-\frac{1}{2}\pi^2\sigma_A^2)T+\pi\sigma_A\sqrt{T}x}}{C_0}-(r-\frac{1}{2}\sigma_c^2)T}{\sigma_c\sqrt{T}}.
\end{aligned}
$$

Hereby we have used the fact that $\frac{\ln X_t-(r-\frac{1}{2}\sigma_A^2\pi^2)t}{\sigma_A\pi\sqrt{t}}$ and $\frac{\ln C_t-(r-\frac{1}{2}\sigma_c^2)t}{\sigma_c\sqrt{t}}$ follow the cumulative bivariate normal distribution with correlation coefficient ρ. Further note the difference between the real number π within the normal distribution function and the investment strategy.

To compute $\int_0^T\mathbb{E}^{\mathbb{Q}}[e^{-rT}G(T)|\tau=t]\mathbb{Q}(\tau\in\mathrm{d}t)$ we need to specify the stochastic processes (X_T, C_T) given $\tau=t$ and to use the density of the first hitting time τ. First, given τ, we have:

$$
\begin{aligned}
C_\tau=\epsilon\phi C_0e^{g\tau}=&C_0\exp\left\{\left(r-\frac{1}{2}\sigma_c^2\right)\tau+\sigma_cW_\tau^Q\right\}\\
\Rightarrow\ W_\tau^Q=&\frac{\ln(\epsilon\phi)-(r-g-\frac{1}{2}\sigma_c^2)\tau}{\sigma_c}.
\end{aligned}
$$

Next, we can write X_τ as

$$\begin{aligned} X_\tau =& X_0 \exp\left\{(r-\frac{1}{2}\sigma_A^2\,\pi^2)\tau+\sigma_A\,\pi\,W_\tau^{Q_1}\right\} \\ =& X_0(\epsilon\phi)^{\frac{\sigma_A\pi\rho}{\sigma_c}}\exp\left\{(r-\frac{1}{2}\sigma_A^2\pi^2)\tau-\frac{\sigma_A\pi\rho}{\sigma_c}(r-g-\frac{1}{2}\sigma_c^2)\tau+\sigma_A\pi\sqrt{1-\rho^2}\sqrt{\tau}z\right\}, \end{aligned}$$

where in the second line we have used that W_τ^Q is correlated with $W_\tau^{Q_1}$ with a correlation coefficient ρ and z is a standard normally distributed random variable under $\mathbb{Q}$ independent of W_τ^Q and $W_\tau^{Q_1}$. At time T, we have

$$\begin{aligned} C_T =& \epsilon\phi C_0 e^{g\tau}\exp\left\{\left(r-\frac{1}{2}\sigma_c^2\right)(T-\tau)+\sigma_c(W_T^Q-W_\tau^Q)\right\} \\ :=& \epsilon\phi C_0 e^{g\tau}\exp\left\{\left(r-\frac{1}{2}\sigma_c^2\right)(T-\tau)+\sigma_c\sqrt{T-\tau}y\right\} \\ :=& \kappa(\tau)\exp\{\sigma_c\sqrt{T-\tau}y\} \end{aligned}$$

with $k(\tau) := \epsilon\phi C_0 e^{g\tau}\exp\left\{\left(r-\frac{1}{2}\sigma_c^2\right)(T-\tau)\right\}$ and

$$\begin{aligned} X_T =& X_\tau\exp\left\{\left(r-\frac{1}{2}\pi^2\sigma_A^2\right)(T-\tau)+\pi\sigma_A(W_T^{Q_1}-W_\tau^{Q_1})\right\} \\ =:& X_\tau\exp\left\{\left(r-\frac{1}{2}\pi^2\sigma_A^2\right)(T-\tau)+\pi\sigma_A\sqrt{T-\tau}x\right\} \\ =:& h(\tau)\exp\left\{\sigma\theta\sqrt{1-\rho^2}\sqrt{\tau}z\right\}\exp\left\{\theta\sigma\sqrt{T-\tau}x\right\} \end{aligned}$$

with $h(\tau)=X_0(\epsilon\phi)^{\frac{\sigma\theta\rho}{\sigma_c}}\exp\left\{(r-\frac{1}{2}\sigma_A^2\pi^2)\tau-\frac{\sigma\theta\rho}{\sigma_c}(r-g-\frac{1}{2}\sigma_c^2)\tau\right\}\exp\left\{\left(r-\frac{1}{2}\pi^2\sigma_A^2\right)(T-\tau)\right\}$. x and y are bivariate normally distributed with a constant correlation coefficient ρ. Both x and y are independent of z, which implies that the joint density of (x,y,z), $f(x,y,z)=f(x,y)\,f(z)$.

Finally we compute the integral as

$$\begin{aligned} &\int_0^T \mathbb{E}^{\mathbb{Q}}[e^{-rT}G(T)|\tau=t]\mathbb{Q}(\tau\in \mathrm{d}t) \\ =&\int_0^T\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-z^2/2}\int_{-\infty}^{dx_5(z,t)}\int_{dy_5(t)}^{dy_6(x,z,t)}e^{-rT}\bar{G}\frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left\{-\frac{x^2+y^2-2\rho xy}{2(1-\rho^2)}\right\}\mathrm{d}y\mathrm{d}xf(t)\mathrm{d}z\mathrm{d}t \\ &+\int_0^T\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-z^2/2}\int_{-\infty}^{dx_5(z,t)}\int_{dy_6(x,z,t)}^{dy_7(x,z,t)}e^{-rT}\Big(\max\Big\{0,B_T-h(t)e^{\sigma_A\pi\sqrt{1-\rho^2}\sqrt{t}z}e^{\pi\sigma_A\sqrt{T-t}x} \\ &-\Big(\kappa(t)\exp\{\sigma_c\sqrt{T-t}y\}-\phi C_0e^{gT}\Big)\Big\}\Big)\frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left\{-\frac{x^2+y^2-2\rho xy}{2(1-\rho^2)}\right\}\mathrm{d}y\mathrm{d}xf(t)\mathrm{d}z\mathrm{d}t, \end{aligned}$$

where

$$dx_5(z,t) = \frac{\ln \frac{B_T}{h(t)\exp\left\{\sigma_A\pi\sqrt{1-\rho^2}\sqrt{t}z\right\}}}{\pi\sigma_A\sqrt{T-t}}$$

$$dy_5(t) = \frac{\ln \frac{\phi\epsilon C_0 e^{gT}}{\kappa(t)}}{\sigma_c\sqrt{T-t}}$$

$$dy_6(x,z,t) = \frac{\ln \frac{\max\{\phi\epsilon C_0 e^{gT}, \phi C_0 e^{gT} + B_T - \bar{G} - h(t)\exp\left\{\sigma_A\pi\sqrt{1-\rho^2}\sqrt{t}z\right\}\exp\left\{\pi\sigma_A\sqrt{T-t}x\right\}\}}{\kappa(t)}}{\sigma_c\sqrt{T-t}}$$

$$dy_7(x,z,t) = \frac{\ln \frac{\phi C_0 e^{gT} + B_T - h(t)\exp\left\{\sigma_A\pi\sqrt{1-\rho^2}\sqrt{t}z\right\}\exp\left\{\pi\sigma_A\sqrt{T-t}x\right\}}{\kappa(t)}}{\sigma_c\sqrt{T-t}}$$

The density of the first hitting time τ under $\mathbb{Q}$ is given by, see for instance Haug (2007)

$$f(t) = -\frac{\ln(\epsilon\phi)}{\sigma_c t^{\frac{3}{2}}} n\left(\frac{\ln(\epsilon\phi) - (r - \frac{1}{2}\sigma_c^2 - g)t}{\sigma_c\sqrt{t}}\right), \tag{3.15}$$

where $n(t) = \frac{1}{\sqrt{2\pi}}e^{-t^2/2}$.

In order to compute Term (b), define first $d_1(\tau)$ and $d_2(\tau)$ by letting $X_\tau = \max(0, BC(\tau))$ and $X_\tau = \max(0, BC(\tau) - \bar{G}e^{-r(T-\tau)})$: Hence

$$d_1(\tau) = \frac{\ln\left(\frac{\max(0,BC(\tau))}{X_0}\right) - \frac{\sigma\theta\rho}{\sigma_c}\left(\ln(\epsilon\phi) - (r - \frac{1}{2}\sigma_c^2 - g)\tau\right) - (r - \frac{1}{2}\sigma_A^2\pi^2)\tau}{\sigma_A\pi\sqrt{1-\rho^2}\sqrt{\tau}}$$

$$d_2(\tau) = \frac{\ln\left(\frac{\max(0,BC(\tau)-\bar{G}e^{-r(T-\tau)})}{X_0}\right) - \frac{\sigma_A\pi\rho}{\sigma_c}\left(\ln(\epsilon\phi) - (r - \frac{1}{2}\sigma_c^2 - g)\tau\right) - (r - \frac{1}{2}\sigma_A^2\pi^2)\tau}{\sigma_A\pi\sqrt{1-\rho^2}\sqrt{\tau}}.$$

Finally we obtain Term (b):

$$
\begin{aligned}
&\mathbb{E}^{\mathbb{Q}}\left[e^{-r\tau}G(\tau)1_{\{\tau\le T\}}\right]\\
=&\mathbb{E}^{\mathbb{Q}}\Bigg[e^{-r\tau}\Big(\left(BC(\tau)-X_\tau\right)1_{\{\max(0,BC(\tau)-\bar{G}e^{-r(T-\tau)})<X_\tau<\max(0,BC(\tau))\}}\\
&+\bar{G}e^{-r(T-\tau)}\,1_{\{X_\tau<\max(0,BC(\tau)-\bar{G}e^{-r(T-\tau)})\}}\Big)\,1_{\{\tau\le T\}}\Bigg]\\
=&\int_0^T e^{-rs}BC(s)\left(\Phi(d_1(s))-\Phi(d_2(s))\right)f(s)\mathrm{d}s\\
&-\int_0^T e^{-rs}X_0(\epsilon\phi)^{\frac{\rho\sigma_A\pi}{\sigma_c}}\exp\left\{\left(r-\frac{1}{2}\sigma_A^2\pi^2\right)s-\frac{\sigma_A\pi\rho}{\sigma_c}\left(r-\frac{1}{2}\sigma_c^2-g\right)s\right\}\\
&\cdot\left(\int_{d_2(s)}^{d_1(s)}\exp\{\sigma_A\pi\sqrt{1-\rho^2}\sqrt{s}z\}\frac{1}{\sqrt{2\pi}}e^{-z^2/2}\mathrm{d}z\right)f(s)\mathrm{d}s\\
&+\int_0^T e^{-rT}\bar{G}\Phi(d_2(s))f(s)\mathrm{d}s
\end{aligned}
$$

where $\Phi(s) := \int_{-\infty}^{s}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\mathrm{d}x$.

3.9.2 PBGC insurance payoff in the extended distress termination model

Define $U_t = C_t + X_t$, $t \in [0,T]$, as the total value of the sponsor's and pension fund's assets. If distress is initiated prematurely then the insurance of the PBGC is given by

$$
\begin{aligned}
G(\tau) =&(1-\eta)\left(B_T\,e^{-r(T-\tau)}+\phi\,C_0\,e^{g\,\tau}\right)1_{\{\phi\,C_0\,e^{g\,\tau}\le C_\tau\}}\,1_{\{(1-\eta)\,U_\tau\le\bar{G}\,e^{-r(T-\tau)}\}}\\
&+\bar{G}\,e^{-r(T-\tau)}\,1_{\{\phi\,C_0\,e^{g\,\tau}\le C_\tau\}}\,1_{\{(1-\eta)\,U_\tau>\bar{G}\,e^{-r(T-\tau)}\}}\\
&+(B_T\,e^{-r(T-\tau)}-X_\tau)\,1_{\{\phi\,C_0\,e^{g\,\tau}>C_\tau\}}\,1_{\{B_T\,e^{-r(T-\tau)}>X_\tau\}}\,1_{\{B_T\,e^{-r(T-\tau)}-X_\tau\le\bar{G}\,e^{-r(T-\tau)}\}}\\
&+\bar{G}\,e^{-r(T-\tau)}\,1_{\{\phi\,C_0\,e^{g\,\tau}>C_\tau\}}\,1_{\{B_T\,e^{-r(T-\tau)}>X_\tau\}}\,1_{\{B_T\,e^{-r(T-\tau)}-X_\tau>\bar{G}\,e^{-r(T-\tau)}\}}.
\end{aligned}
$$

In the first term we have used that at $t=\tau$, $U_\tau = \eta\,(B_T\,e^{-r(T-\tau)}+\phi\,C_0\,e^{g\,\tau})$. Further note that the triggering of distress termiantion without sponsor default, i.e $\phi\,C_0\,e^{g\,\tau} < C_\tau$, immediately implies that the pension fund must have defaulted, i.e the condition $B_T\,e^{-r(T-\tau)} > X_\tau$ is then always satisfied. The first two terms correspond to the case where the sponsor can provide a partial support, while the last two terms represent the case where both the sponsor and the pension fund default, therefore the PBGC carries the entire burden of the pension benefits.

Finally if $\tau > T$ the payoff of the PBGC insurance in the extended distress termination model can be written as

$$
\begin{aligned}
G(T) = & \max\left\{0, B_T - X_T - (C_T - \phi C_0 e^{gT})\right\} 1_{\{\phi C_0 e^{gT} < C_T\}} 1_{\{X_T < B_T\}} 1_{\{\max\{0, B_T - X_T - (C_T - \phi C_0 e^{gT})\} < \bar{G}\}} \\
& + \bar{G}\, 1_{\{\phi C_0 e^{gT} < C_T\}} 1_{\{X_T < B_T\}} 1_{\{\max\{0, B_T - X_T - (C_T - \phi C_0 e^{gT})\} > \bar{G}\}} \\
& + (B_T - X_T) 1_{\{\phi C_0 e^{gT} > C_T\}} 1_{\{X_T < B_T\}} 1_{\{B_T - X_T < \bar{G}\}} \\
& + \bar{G}\, 1_{\{\phi C_0 e^{gT} > C_T\}} 1_{\{X_T < B_T\}} 1_{\{B_T - X_T > \bar{G}\}},
\end{aligned}
$$

where the interpretation of the different components is the same as in the case of premature termination.

Chapter 4

An Optimal Termination Rule for a DB Pension Guarantee [1]

4.1 Introduction

In the previous chapter we have presented a risk-based premium calculation model for a pension guarantee fund and we have argued that such a risk-based premium calculation is an appealing approach to resolve some problems in the pension sector and thus to better protect employees in DB pension plans. In that chapter we have closely studied the distress termination mechanism, that is the termination of the pension fund is triggered by the sponsoring companie's poor financial status. This chapter is now devoted to the involuntary termination mechanism where an underfunded pension fund is terminated by the corresponding pension guarantee fund. The objective we pursue is to find an optimal involuntary termination, that is an optimal timing of intervention for the guarantee fund. Such an optmial intervention policy is a further protection mechanism for DB plan policyholders. It has the advantage that it is more applicable under current law in many countries.

Kalra and Jain (1997) argue that by law pension guarantee funds have no opportunity to control the investment riskiness of the pension plans and they are not allowed to adjust premiums according to their changing financial status. Therefore the only means of intervention a pension guarantee fund has to control its financial guarantee and to protect policyholders is to prematurely terminate an insured underfunded DB pension plan. [2] The pension guarantee fund can terminate and take over insured underfunded DB pension plans prematurely instead of waiting until the plans become severely underfunded and are then closed by their sponsoring companies, which would lead to even larger costs for the pension guarantee fund.

[1] This chapter is based on Cheng and Uzelac (2014)

[2] The PBGC states that initiating a premature termination helps to protect the interests of plan beneficiaries or of the PBGC insurance program, see pbgc.gov. However, there is no specified premature termination rule stated by PBGC to follow.

From the economic point of view two main arguments support an intervention policy. Firstly, as already suggested in the previous chapter, there is a moral hazard problem between the sponsors and the pension guarantee funds. Cooper and Ross (2003) find that introducing a pension guarantee fund creates a further incentive for the sponsors to underfund their DB pension plans and excessively increase pension fund investment risk while taking advantage of the pension guarantee fund. A possible premature termination can mitigate these adverse incentives of DB plan sponsors. Secondly, financially troubled sponsors contribute the minimum possible to DB pension plans, particularly by setting higher interest rates for discounting pension liabilities, see Bodie et al. (1987). As we mentioned in chapter 3, these authors also find that the DB pension plan's funding degree and the companie's long-run profitability are considerably positively correlated. This implies that financially troubled sponsoring companies are likely to not be able to meet their pension obligations. Accordingly, a premature termination of such insured DB pension plans can substantially reduce the liabilities of the pension guarantee funds.

The important question which needs to be addressed here is what the proper timing of intervention is. Or in other words how underfunded should pension funds be in order to be prematurely terminated? It is important to find the proper intervention timing since on the one hand, terminating insured underfunded DB pension funds too early takes the opportunity away from the funds to recover and the guarantee fund also loses potential future premiums; on the other hand, a too late intervention is likely to result in considerably larger liabilities for the pension guarantee fund. The latter is particularly the case since largely underfunded pension funds mainly belong to financially troubled companies which are likely to go bankrupt anyway, see Kalra and Jain (1997).

This chapter gives insights into the question above by first proposing a specified termination rule based on the funding status of the insured DB pension plan and finding an optimal termination ratio for the insured DB pension fund. In particular, we take the risk aversion of pension beneficiaries into account and we incorporate two regulatory constraints, a shortfall probability constraint (SPC) and an expected shortfall constraint, (ESC) into an expected utility maximization problem. The pension guarantee fund uses the regulatory constraints to control its current and ongoing liabilities. The SPC puts a restriction on the premature termination probability of the insured DB plan within a time horizon and reflects some current solvency regulations, see e.g. Solvency II. The solvency risk of the insured DB pension plan faced by the pension guarantee fund is controlled via adjusting the acceptable shortfall probability. Besides, the pension beneficiaries are protected under the SPC, which actually works as a security mechanism for protecting pension benefits, see Broeders and Chen (2013). We further include the ESC constraint into the expected utility maximization problem, which has the advantage that it can assess the size of the expected losses of underfunded but not terminated pension plans and therefore it can better identify DB pension funds with the highest cost of insolvencies, see Doff (2008). Although this constraint is not used in the current regulatory practice, it reflects for instance the law imposed on the PBGC-initiated termination, which says that the PBGC may terminate the pension plan in case the expected losses the PBGC incurs increase unreasonably if the pension plan is not closed.

More specifically, our specified termination rule states that the insured DB pension plan will be terminated and taken over by the pension guarantee fund once its funding ratio goes down a critical funding ratio. To obtain the optimal critical threshold, we set up a one-period model and maximize a power type utility function with the funding ratio as an independent variable to capture the interests of pension beneficiaries. Actually we assume the objective function of the pension guarantee fund is to maximize the beneficiaries expected utility. This is in so far reasonable as the fundamental goal of a pension guarantee fund is to protect pension beneficiaries. The use of the funding ratio as the argument in the utility function can on the one hand be motivated since this is a commonly used quantity in industry and particularly regulatory practice. More importantly, this quantity recognizes that what really matters in pension fund management is not the value of the assets on its own, but how the assets value compares to the liabilities value in each point in time, see Martellini and Milhau (2008).

We find that by considering the one-period expected utility maximization problem solely with the SPC, the specified termination rule is only applicable to more risk averse pension beneficiaries, but not to risk neutral and less risk averse ones. This result is to some extent in line with the passive behavior of the PBGC, as Kalra and Jain (1997) mention that until 1995 only 1% of the pension plan terminations were initiated by the PBGC. After adding the ESC into the maximization problem, the intervention policy is applied regardless of the risk aversion of the pension beneficiaries. Moreover, in the case where the two constraints are satisfied simultaneously, we obtain an optimal termination ratio which depends on the risk aversion of the pension beneficiaries. For instance, in the benchmark case the optimal termination ratio is 0.68 for risk neutral and less risk averse and 0.71 for more risk averse pension beneficiaries. In the case where the two constraints are inconsistent, we propose a suboptimal termination ratio, which does not depend on the risk preferences of the pension beneficiaries.

Related studies are Archarya and Dreyfus (1989) and particularly Kalra and Jain (1997). The former compute simultaneously premium policies and optimal dynamic termination policies for banks in terms of a threshold assets-to-deposits ratio by minimizing the insurer's net liabilities, below which an ailing bank should be closed. Kalra and Jain (1997) recommend that the PBGC follows an intervention policy where the PBGC insurance is considered as a down-and-out put option and the PBGC takes over a plan if the losses from terminating it are smaller than the losses from continuing it. Unlike our model the exercise boundary, which is also based on a critical funding ratio, is endogenous in their model.

The remainder of the chapter is organized as follows. Section 4.2 models the termination rule and describes the expected one-period utility and the regulatory constraints considered. Section 4.3 simplifies the utility maximization problem by deriving closed-form expressions for the one-period expected utility and the two regulatory constraints. In addition, it is shown that the constraints have monotonic properties. Section 4.4 provides a numerical analysis where we first numerically solve the utility maximisation problem solely with respect to the SPC and then add the ESC. Section 4.5 concludes the chapter and section 4.6 provides detailed calculations and proofs of the main results in the chapter.

4.2 Model Setup

As in the previous chapter we consider a DB pension plan for a single representative beneficiary, which is insured by a pension guarantee fund at time $t_0 = 0$. In this chapter we do not model the pension fund's assets and liabilities separately, but we directly model the funding ratio of the insured DB pension plan, which is the ratio of the plan's assets to its accrued liabilities. We denote the funding ratio at time t by R_t. If R_t takes a value less than 1 then pension fund is underfunded. We assume the funding ratio R_t follows a geometric Brownian motion (GBM) [3] under the market probability measure $\mathbb{P}$, i.e., it is governed by the stochastic differential equation (SDE)

$$\mathrm{d}R_t = \mu_R R_t \,\mathrm{d}t + \sigma_R R_t \,\mathrm{d}W_t \,,\; R_0 > 0 \,, \tag{4.1}$$

where W is a standard 1-dimensional Brownian motion under $\mathbb{P}$ and $\mu_R > 0$ and $\sigma_R > 0$ denote the drift and the volatility coefficients, respectively.

The pension guarantee fund uses a premature termination rule to intervene. Unlike chapter 3, we exclude distress termination and solely consider involuntary termination as the relevant premature termination mechanism. We further abstract from the financial status of the sponsoring company and also from contributions the plan sponsor makes to its DB pension fund. Accordingly, the premature termination solely depends on the funding ratio of the corresponding pension plan. We consider a specific premature termination rule which states that once the funding ratio of the insured DB pension plan R_t touches or falls bellow a predefined termination ratio η, the insured DB pension plan is closed and taken over by the pension guarantee fund. More precisely, the termination time τ is defined as the first hitting time the funding ratio reaches the predefined termination ratio

$$\tau := \inf\{t \,|\, R_t \leqslant \eta\} \,. \tag{4.2}$$

In our model, since we are not interested in the case where the premature termination can never happen, i.e., $\tau = +\infty$, the termination ratio is naturally required to be strictly larger than 0. In addition, we assume $\eta < R_0$. It is a technical condition which makes sure

[3] We mention that if we would model the pension assets and pension liabilities separately as geometric Brownian motions (GBMs), the dynamics of the funding ratio would still follow a GBM and our qualitative analysis and results would carry over.

that the pension plan is not terminated at the contract-issuing time $t_0 = 0$. Finally we require the termination ratio to be strictly smaller than 1, that is only underfunded pension plans can be prematurely terminated. This assumption is economically reasonable since employees involved in a fully funded DB pension plan can expect to receive full promised benefits on their retirement and therefore there is no reason to terminate such a pension fund. To summarize, the predefined termination ratio η is restricted to be chosen from the set $(0, \min\{R_0, 1\})$.

The pension guarantee fund pursues a social welfare motive, since its primary goal is that the employees get their promised pension benefits. [4] Accordingly, as the interests of the pension guarantee fund and the representative beneficiary are strongly connected, we assume in this chapter that the pension guarantee fund acts as an agent of the pension beneficiary by maximizing the beneficiary's expected utility. We use a power type utility function with the funding ratio of the insured DB pension plan as the argument to capture the interests of the representative beneficiary with a risk aversion parameter δ. The larger the funding ratio is, the more confidence the pension beneficiary has to get the promised pension benefits on his retirement date. So the beneficiary feels safer and his utility increases in the funding ratio of the insured DB plan. The line of reasoning reverts for a decreasing funding ratio.

In order to be consistent with the regulatory constraints that we will incorporate into the model, e.g. the SPC restricts the probability of the insured DB pension plan to be terminated within one period (i.e., one year), we set up a one-year utility maximization problem and calculate the one-year expected utility of the representative beneficiary. [5] If the premature termination is triggered within one year, we use the funding ratio at the premature termination time τ in the power utility function, i.e., R_τ. Since the funding ratio is assumed to be a continuous stochastic process, the latter exactly coincides with the predefined termination ratio η, i.e., the expected utility in this case is then given by $\mathbb{E}[U(R_\tau)] = \mathbb{E}\left[\frac{\eta^{(1-\delta)}}{(1-\delta)}\right]$ if $0 < \tau \leqslant 1$. Otherwise, we use the funding ratio at year one as the independent variable, which yields the expected utility $\mathbb{E}[U(R_1)] = \mathbb{E}\left[\frac{R_1^{(1-\delta)}}{(1-\delta)}\right]$ if $\tau > 1$. To sum up, the one-year expected utility [6] consisting of two parts conditional on whether the premature termination occurs within one year, can be written compactly as

$$\mathbb{E}[U(R_{\tau\wedge\{t=1\}})] = \mathbb{E}\left[\frac{1_{\{0<\tau\leqslant 1\}}\eta^{(1-\delta)}}{1-\delta}\right] + \mathbb{E}\left[\frac{1_{\{\tau>1\}}R_1^{(1-\delta)}}{1-\delta}\right], \ \delta \geqslant 0 \text{ and } \delta \neq 1\,, \tag{4.3}$$

where $\wedge$ denotes the minimum of τ and $t = 1$ and 1_X is an indicator which is 1 when event

[4] Salisbury's (1996) argument of viewing the pension guarantee fund as a social insurance program with intentional subsidies to the defined benefit system supports our point of view.

[5] The results remain qualitatively the same if we would consider a period which is longer than one year.

[6] Our model can also be understood in a multiperiod framework as a repeating one period model where the pension guarantee fund renews the insurance contract and sets up the termination ratio η annually.

X occurs and 0 otherwise.

As we mentioned in the introduction of this chapter, the premature termination is a way used by the pension guarantee fund to manage its financial guarantee, in terms of lessening potential insolvencies of ailing DB plans and consequently protecting the pension benefits in the long run. The financial guarantee is further protected by regulatory constraints, which put restrictions on the insolvency risk of the insured DB pension plan. In this chapter we consider two regulatory constraints.The first constraint is the shortfall probability constraint (SPC), which imposes a restriction on the one-year shortfall probability of the insured DB pension plan. The probability that the premature termination is triggered within one year is required to be less than a certain percentage, i.e. $\epsilon \in (0, 1]$, which is set up by the pension guarantee fund. Formally, the constraint is written as

$$P(0 < \tau \leqslant 1) \leqslant \epsilon,\ 0 < \epsilon \leqslant 1\,. \tag{4.4}$$

By fixing the maximum allowable shortfall probability the pension guarantee fund protects its current financial status by controlling the probability it has to step in and cover the pension deficits of an insured and underfunded pension fund within one year.

The SPC helps to reduce the number of current insolvencies the pension guarantee fund would have to manage. However, the SPC does not measure the size of the potential losses which can arise from underfunded but not terminated pension plans. We refer to these pension plans as ongoing pension plans. As the deficits of such ongoing pension plans can become large, the financial guarantee of the pension guarantee fund is exposed to a great risk in the future.

To mitigate this drawback of the SPC, we incorporate an expected shortfall type constraint (ESC) as the second constraint. This constraint puts a restriction on the size of expected deficits of the ongoing pension plan at year one. For simplicity, we assume that the insured DB pension plan's liabilities are constant within one year. With this assumption the maximum expected deficits can be written as the pension liabilities multiplied by a certain percentage $q > 0$, that is set up by the pension guarantee fund. Then the ESC has the following expression

$$\mathbb{E}\left[(1 - R_1)1_{\{\tau>1\}}1_{\{R_1\leqslant 1\}}\right] \leqslant q\,, \tag{4.5}$$

The same type of ESC is considered in Shi and Werker (2012).[7] By fixing the maximum tolerable expected deficits, the pension guarantee fund protects its future financial status.

[7]However, they considered the restriction on the induced expected shortfall subject to the VaR constraint.

4.3 Simplifying the Utility Maximization Problem

In this section we derive closed-form expressions for the one-year expected utility and the two regulatory constraints.

Lemma 4.3.1. *The one-year expected utility is calculated as follows:*

$$\begin{aligned}
&\mathbb{E}[U(R_{\tau\wedge\{t=1\}})] = \mathbb{E}\left[\frac{1_{\{0<\tau\leqslant 1\}}\eta^{(1-\delta)}}{1-\delta}\right] + \mathbb{E}\left[\frac{1_{\{\tau>1\}}R_1^{(1-\delta)}}{1-\delta}\right] \\
&= \frac{1}{1-\delta}\Bigg[\eta^{(1-\delta)}\big(\Phi(A-B)+\exp\{2AB\}\Phi(A+B)\big) \\
&+ R_0^{(1-\delta)}\exp\left\{(1-\delta)\big(\mu_R - \frac{1}{2}\delta\sigma_R^2\big)\right\}\Phi\big(-A+B+\sigma_R(1-\delta)\big) \\
&-R_0^{(1-\delta)}\exp\left\{2AB + 2A\sigma_R(1-\delta) + (1-\delta)\left(\mu_R - \frac{1}{2}\delta\sigma_R^2\right)\right\}\Phi\left(A+B+\sigma_R(1-\delta)\right)\Bigg] \\
&= u(\eta)\,, \qquad (4.6)
\end{aligned}$$

where $A = \frac{1}{\sigma_R}\ln\left(\frac{\eta}{R_0}\right)$, $B = \frac{\mu_R - \frac{1}{2}\sigma_R^2}{\sigma_R}$ *and* $\Phi(\cdot)$ *denotes the standard normal cdf. The one-year expected utility is a function of the termination ratio* $\eta \in (0, \min\{R_0, 1\})$*, denoted by* $u(\eta)$*.*

Appendix 4.6.1 provides a detailed derivation of $u(\eta)$.

Lemma 4.3.2. *The one-year shortfall probability is calculated as follows, see e.g. Haug (2007),*

$$\begin{aligned}
&P(0<\tau\leqslant 1) = \Phi(A-B) + \exp\{2AB\}\Phi(A+B) \\
&= \Phi\left(\frac{1}{\sigma_R}\left(\ln\left(\frac{\eta}{R_0}\right) - \mu_R + \frac{1}{2}\sigma_R^2\right)\right) + \left(\frac{\eta}{R_0}\right)^{\left(\frac{2\mu_R}{\sigma_R^2}-1\right)}\Phi\left(\frac{1}{\sigma_R}\left(\ln\left(\frac{\eta}{R_0}\right) + \mu_R - \frac{1}{2}\sigma_R^2\right)\right) \\
&= P(\eta) \qquad (4.7)
\end{aligned}$$

which is a function of the termination ratio $\eta \in (0, \min\{R_0, 1\})$*, denoted by* $P(\eta)$*.*

The one-year expected utility maximization problem with only the SPC can be written as follows:

$$\begin{aligned}
&\max_{\eta\in(0,\min\{R_0,1\})} \mathbb{E}\left[\frac{1_{\{0<\tau\leqslant 1\}}\eta^{(1-\delta)}}{1-\delta}\right] + \mathbb{E}\left[\frac{1_{\{\tau>1\}}R_1^{(1-\delta)}}{1-\delta}\right],\ \delta \geqslant 0 \text{ and } \delta \neq 1\,, \\
&\text{subject to } P(\eta) \leqslant \epsilon\,, \qquad (4.8)
\end{aligned}$$

where $0 < \epsilon \leqslant 1$ is set up by the pension guarantee fund.

Lemma 4.3.3. *The expected shortfall of the ongoing DB pension plan at year one is calculated as follows:*

$$\begin{aligned}\mathbb{E}\left[(1-R_1)1_{\{\tau>1\}}1_{\{R_1\leqslant 1\}}\right] =&\Phi\left(-\frac{1}{\sigma_R}\ln(R_0)-B\right)-\Phi\left(A-B\right)\\ &-\exp\left\{2AB\right\}\left[\Phi(A+B)-\Phi\left(2A+B+\frac{1}{\sigma_R}\ln(R_0)\right)\right]\\ &-R_0\exp\left\{\mu_R\right\}\Phi\left(-\frac{1}{\sigma_R}\ln(R_0)-B-\sigma_R\right)\\ &+R_0\exp\left\{\mu_R\right\}\Phi\left(A-B-\sigma_R\right)\\ &+R_0\exp\left\{\mu_R+2AB+2A\sigma_R\right\}\left[\Phi(A+B+\sigma_R)\right.\\ &\left.-\Phi\left(2A+B+\frac{1}{\sigma_R}\ln(R_0)+\sigma_R\right)\right]=S(\eta),\end{aligned} \tag{4.9}$$

which is a function of the termination ratio η on the interval $(0,\min\{R_0,1\})$, denoted by $S(\eta)$.

Appendix 4.6.2 provides a detailed derivation of $S(\eta)$.

Then the one-year expected utility maximization problem with the two regulatory constraints can be written as follows:

$$\begin{aligned}&\max_{\eta\in\min\{R_0,1\}}\mathbb{E}\left[\frac{1_{\{0<\tau\leqslant 1\}}\eta^{(1-\delta)}}{1-\delta}\right]+\mathbb{E}\left[\frac{1_{\{\tau>1\}}R_1^{(1-\delta)}}{1-\delta}\right],\ \delta\geqslant 0 \text{ and } \delta\neq 1\,,\\ &\text{subject to } P(\eta)\leqslant\epsilon \text{ and } S(\eta)\leqslant q\,,\end{aligned} \tag{4.10}$$

where $0<\epsilon\leqslant 1$ and $q>0$ are set up by the pension guarantee fund.

Next, we formally state the monotonic properties of the two regulatory constraints.

Proposition 4.3.4. *The one-year shortfall probability is a continuous and monotonically increasing function of the termination ratio η on the interval $(0,\min\{R_0,1\})$, i.e. $P(\eta)$, with $\lim_{\eta\to 0}P(\eta)=0$ and $\lim_{\eta\to\min\{R_0,1\}}P(\eta)=1$. Admissible termination ratios $\tilde{\eta}$ chosen from the set $(0,\min\{R_0,1\})$ which satisfy the SPC, i.e. $P(\tilde{\eta})\leqslant\epsilon$, $\epsilon\in(0,1]$, are:*

(1) in the case where $\epsilon=1$, $\tilde{\eta}\in(0,\min\{R_0,1\})$;

(2) in the case where $\epsilon\in(0,1)$, there exists a unique upper bound termination ratio $\bar{\eta}_\epsilon\in(0,\min\{R_0,1\})$ such that $P(\bar{\eta}_\epsilon)=\epsilon$ and $\tilde{\eta}\in(0,\bar{\eta}_\epsilon]$.

Proof 4.3.5. *A Proof is given in appendix 4.6.3.*

Proposition 4.3.6. *The expected shortfall of the underfunded ongoing DB pension plan at the end of year one is a continuous and monotonically decreasing function of the termination* *ratio* η *on the interval* $(0, \min\{R_0, 1\})$*, i.e.* $S(\eta)$*, with* $\lim_{\eta \to \min\{R_0,1\}} = 0$ *and* $\lim_{\eta \to 0} S(\eta) = \Phi\left(-\frac{1}{\sigma_R}\ln(R_0) - B\right) - R_0 \exp\{\mu_R\}\Phi\left(-\frac{1}{\sigma_R}\ln(R_0) - B - \sigma_R\right)$. *Admissible termination ratios* $\hat{\eta}$ *chosen from the set* $(0, \min\{R_0, 1\})$ *which satisfy the ESC, i.e.* $S(\hat{\eta}) \leqslant q$*,* $q > 0$*, are:*

(1) in the case where $q \geqslant \Phi\left(-\frac{1}{\sigma_R}\ln(R_0) - B\right) - R_0 \exp\{\mu_R\}\Phi\left(-\frac{1}{\sigma_R}\ln(R_0) - B - \sigma_R\right)$*,* $\hat{\eta} \in (0, \min\{R_0, 1\})$*;*

(2) in the case where $0 < q < \Phi\left(-\frac{1}{\sigma_R}\ln(R_0) - B\right) - R_0 \exp\{\mu_R\}\Phi\left(-\frac{1}{\sigma_R}\ln(R_0) - B - \sigma_R\right)$*, there exists a unique lower bound termination ratio* $\underline{\eta_\epsilon} \in (0, \min\{R_0, 1\})$ *such that* $S\left(\underline{\eta_\epsilon}\right) = q$ *and* $\hat{\eta} \in \left[\underline{\eta_\epsilon}, (0, \min\{R_0, 1\})\right)$*.*

Proof 4.3.7. *The Proof is given in appendix 4.6.4.*

It is important to learn from proposition 4.3.4 and proposition 4.3.6 that adjustments of the termination ratio η have an opposite impact on the two regulatory constraints. That is, setting a lower termination ratio decreases the shortfall probability, while at the same time this increases the expected shortfall. Technically these properties imply that the sets of admissible termination ratios for the two constraints may not overlap. From an economic point view this raises the question of finding a suboptimal solution in the case the two constraints are not satisfied simultaneously.

If for the given values of ϵ and q, the two sets of admissible termination ratios overlap, a termination ratio in the overlap region which maximizes the one-year expected utility is an optimal solution to the constrained utility maximization problem. In the case, where there is no intersection between the two sets, the pension guarantee fund faces a tradeoff between alleviating its current and future insolvencies. We introduce a suboptimal solution where we define the best suboptimal termination ratio as the one which satisfies the SPC but violates the ESC to the least extent. Accordingly, we give priority to the SPC since it has been applied in the regulatory practice for pension plans and alleviating current financial insolvencies of the pension guarantee fund is more urgent.

4.4 Numerical Analysis

The constrained optimization problems (4.8) and (4.10) cannot be solved in closed-form and therefore we solve them numerically. To do so we first compute the constraint set, which is available in closed-form, and then perform a numerical search algorithm to find the maximum of the one-year expected utility given the constraint set. For the numerical analysis at hand we fix the relevant parameters as follows:

$$\mu_R = 0.03, \sigma_R = 0.2, R_0 = 1.1, \epsilon = 0.025, q = 0.03. \quad (4.11)$$

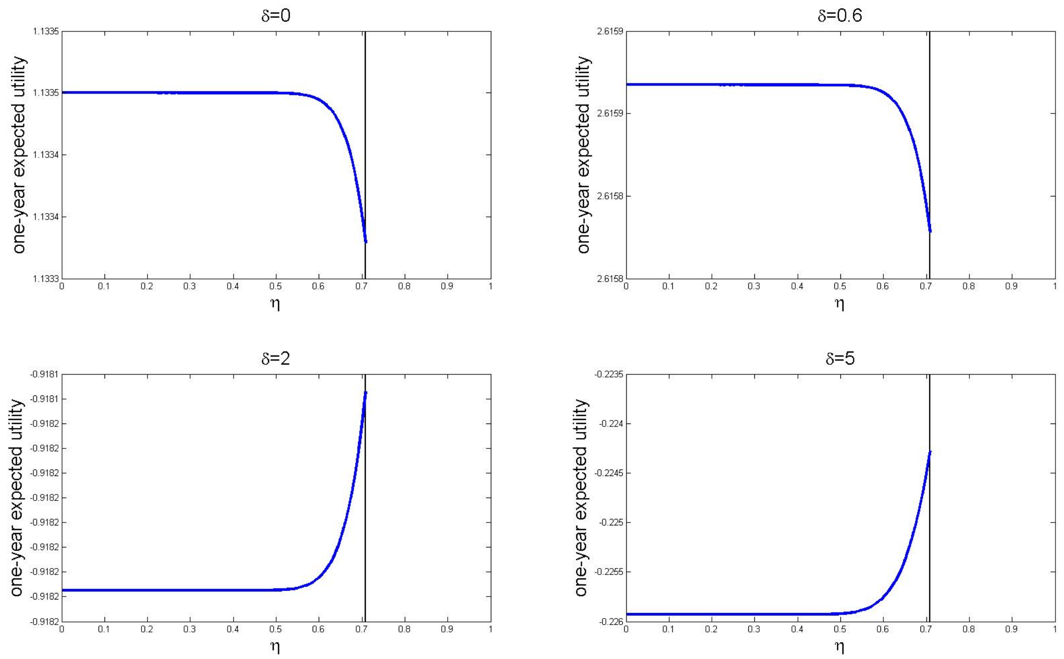

Figure 4.1: One-year expected utility as a function of the termination ratio $\eta \in (0, 0.71]$ for different risk aversion parameters δ.

In addition we choose the following risk aversion parameters:

$$\delta_1 = 0, \delta_2 = 0.6, \delta_3 = 2, \delta_4 = 5. \tag{4.12}$$

The parameters serve for illustration purposes.

4.4.1 Utility Maximization with only the SPC

We start with the solution of the one-year expected utility maximization problem when only the SPC is considered. With the given benchmark parameters the upper bound termination ratio $\bar{\eta}_\epsilon$ is equal to 0.71. Figure 4.1 plots the one-year expected utility for different risk aversion parameters over the constraint set, i.e. $\eta \in (0, 0.71]$. We see that for a risk neutral and less risk averse pension beneficiary, i.e. $\delta_1 = 0, \delta_2 = 0.6$, the utility value is not sensitive to changes of the termination ratio when the ratio is relatively small and as the termination ratio approaches $\bar{\eta}_\epsilon$, the one-year expected utility value decreases. In these cases, setting the termination ratio relatively small is beneficial for the pension beneficiary, which gives basically the same utility value to the beneficiary as if the termination ratio is set up extremely close to 0 such that the premature termination is hardly triggered. If we take termination costs into account, which exist in reality, e.g. administrative costs, setting the termination ratio extremely close to 0 with nearly no premature termination control from the pension guarantee fund is beneficial for both the pension beneficiary and the pension guarantee fund. Hence if only the SPC is considered, a premature termination rule is not

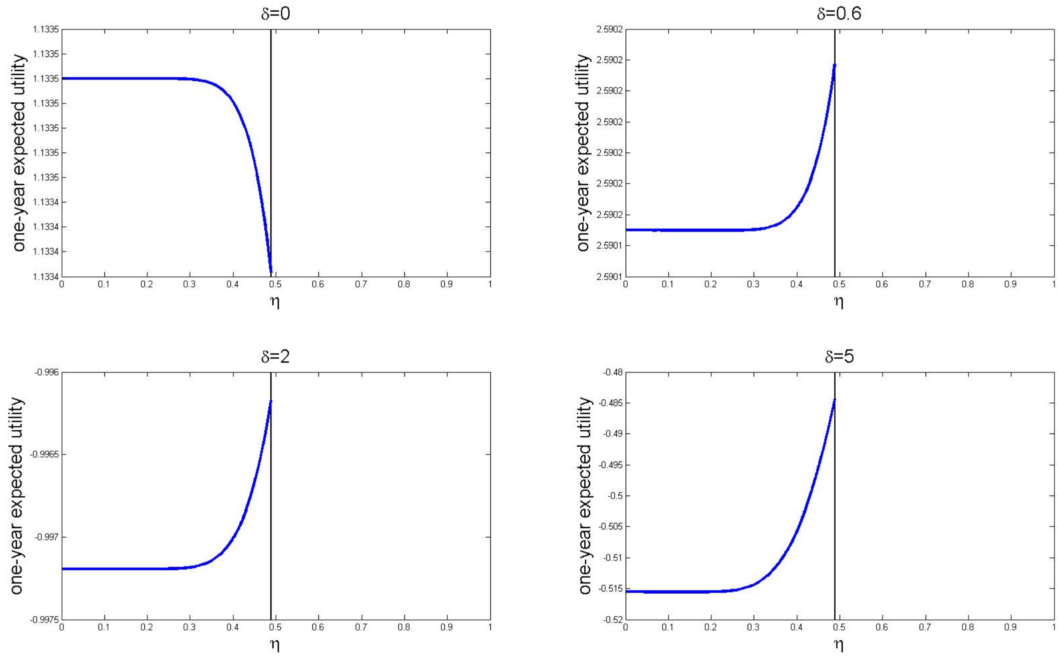

Figure 4.2: One-year expected utility as a function of the termination ratio $\eta \in (0, 0.49]$ for different risk aversion parameters δ when $\sigma_R = 0.3$.

applicable to risk neutral and less risk averse pension beneficiaries. Intuitively, since the pension benefits paid out by the guarantee fund are subject to a limit set by law, risk neutral and less risk averse pension beneficiaries prefer more the possibility that an underfunded DB pension fund can recover and therefore they will receive full promised pension benefits instead of being involved in a regulated DB pension plan with the possibility of losing benefits when the pension fund is prematurely terminated and the pension deficits exceed the legal guarantee limit. For the cases where the risk aversion parameters are relatively large, i.e., $\delta_3 = 2$, $\delta_4 = 5$, the one-year expected utility value remains nearly constant when the termination ratio is relatively small, which is the same as in the cases where $\delta = 0$ and $\delta = 0.6$, but it starts to increase as the termination ratio approaches the upper bound. In such cases the one-year expected utility is maximized at $\bar{\eta}_\epsilon$ where the SPC is binding. Intuitively, more risk averse beneficiaries prefer to have the premature termination control from the pension guarantee fund since they are concerned about potential benefit losses they might incur when an unregulated DB pension fund is terminated under distress termination with severe pension deficits which are much beyond the guarantee limit. The larger the termination ratio is set up, the safer the beneficiaries may feel since they are more likely to receive full promised pension benefits under the premature termination.

In figure 4.2, we show how the volatility of the funding ratio process affects the chosen optimal termination ratio. We increase σ_R from 0.2 to 0.35 and draw the one-year expected utility for the same chosen risk aversion parameters. First, we see that the constraint set narrows as the volatility increases, i.e. $\eta \in (0, 0.49]$. Moreover, we observe that in the plot of $\delta = 0.6$ the one-year expected utility now increases as η approaches the upper bound termination ratio, i.e. $\bar{\eta}_\epsilon = 0.49$. Therefore setting the termination ratio at 0.49 instead of extremely close to 0 maximizes the pension beneficiary's one-year expected utility. It is consistent with economic intuition. As volatility increases the probability that the financial status of the pension fund deteriorates and the beneficiary loses benefits increases. This now implies that even a less risk averse pension beneficiary looks for a premature termination control from the pension guarantee fund to benefit from protection against potential losses in their pension benefits. Accordingly, a larger volatility (uncertainty) of the DB pension plan's funding ratio can induce the premature termination rule used by the pension guarantee fund to be applicable in more cases.

4.4.2 Adding the ESC

In this section, we add the ESC and consider the one-year expected utility maximization problem (4.10). We calculate the lower bound termination ratio $\underline{\eta}_\epsilon$, which is equal to 0.68, so termination ratios which satisfy the ESC form the set $[0.68, 1)$. We have already found out the set of termination ratios which satisfy the SPC in section 4.1, i.e. $(0, 0.71]$. Figure 4.3 plots the one-year expected utility as a function of the termination ratio η over the interval $(0, 0.71]$ for different risk aversion parameters. The left and right vertical lines refer to the lower bound $\underline{\eta}_\epsilon$ and the upper bound termination ratio $\bar{\eta}_\epsilon$, respectively. The area in between refers to the set of termination ratios which satisfy both constraints. Accordingly, the constraint set turns out to be $[0.68, 0.71]$ and the optimal termination ratio for risk neutral and less risk averse pension beneficiaries changes to be the lower bound termination ratio $\underline{\eta}_\epsilon$. Although a risk neutral or less risk averse beneficiary would prefer to have no premature termination control on the insured DB pension plan, the pension guarantee fund sets up the lower bound termination ratio to better control its future financial guarantee. It is important to note that this result always holds when the ESC constraint is binding regardless of the SPC constraint. In other words, when the ESC constraint is satisfied a premature termination rule always exists. For more risk averse beneficiaries, the optimal termination ratio is still set at $\bar{\eta}_\epsilon$.

Figure 4.4 plots the one-year expected utility over the interval of termination ratios $(0, 1)$ for different risk aversion parameters when q is set to be 0.015. Now with $q = 0.015$, the lower bound termination ratio is calculated to be 0.8. The corresponding two sets of termination ratios, which satisfy the SPC and the ESC are $(0, 0.71]$ and $[0.80, 1)$, respectively. Hence, there is no termination ratio which satisfies the two regulatory constraints simultaneously and therefore we need to rely on a suboptimal solution. As introduced in section 4.3, a suboptimal termination ratio is defined to be the one which satisfies the SPC but violates the ESC to the least extent. By relying on this suboptimal solution, we say that protecting the guarantee fund's current financial guarantee has priority so that the pension guarantee fund pays more attention to the insolvency risk of the insured DB pension plan

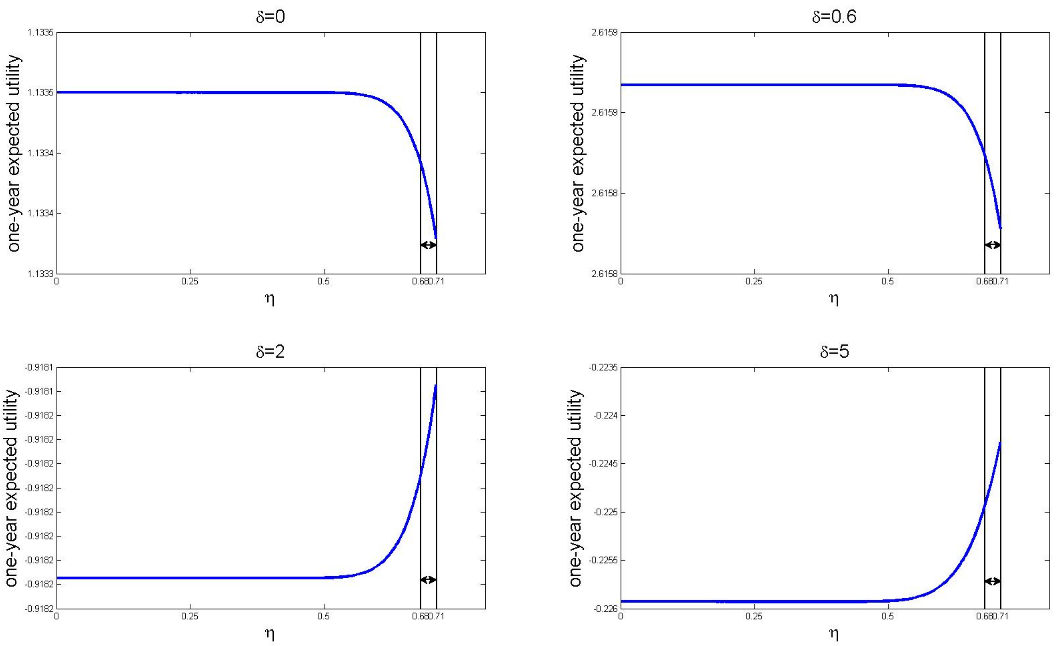

Figure 4.3: One-year expected utility as a function of the termination ratio $\eta \in (0, 0.71]$ for different risk aversion parameters δ when two admissible termination ratio sets overlap.

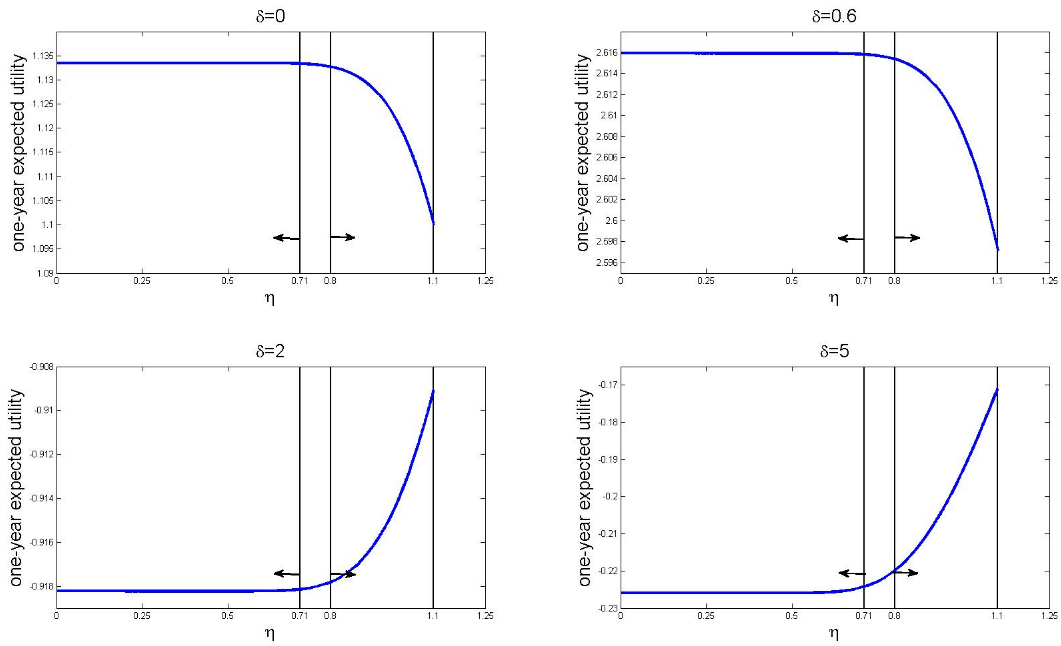

Figure 4.4: One-year expected utility as a function of the termination ratio $\eta \in (0, 1)$ for different risk aversion parameters δ when two admissible termination ratio sets do not overlap.

it faces currently. Moreover, violating the ESC to the least extent encourages the pension guarantee fund to search for new funding or adjust its investment strategy in order to cover unexpected pension deficits after one year. Therefore in this case where two admissible termination ratio sets do not overlap, the suboptimal termination ratio is $\bar{\eta}_\epsilon = 0.71$, which does not depend on the risk aversion of the pension beneficiaries any more.

To summarize, the main result we learn from figure 4.3 and figure 4.4 is that by adding the ESC a premature termination rule is applicable independent of the preferences of the pension beneficiaries, thus even to risk neutral and less risk averse pension beneficiaries.

4.4.3 Utility Losses

Finally we briefly discuss the question what type of beneficiary is (most) affected by imposing the additional regulation on the expected shortfall. To do so we compute the annualized lose rate in utility, which is brought by the ESC constraint. Formally we compute this loss rate l_{t_0+1} as

$$l_{t_0+1} = \ln\left(\frac{u\left(\eta^*_{\text{SPC\&ESC}}\right)}{u\left(\eta^*_{\text{SPC}}\right)}\right), \tag{4.13}$$

where $u\left(\eta^*_{\text{SPC\&ESC}}\right)$ is the maximum value of the expected utility if both the SPC and the ESC constraints are considered, while $u\left(\eta^*_{\text{SPC}}\right)$ denotes the value of the expected utility evaluated at the optimal termination ratio under the SPC only[8].

δ	l_{t_0+1} (SPC and ESC overlap)	l_{t_0+1} (SPC and ESC do not overlap)
0	-0.48	-1.07
0.6	-0.16	-0.34
2	0	0
5	0	0

Table 4.1: Annualized loss rates in basis points (bp) when the ESC constraint is added for different risk aversion parameters. The second column gives the annualized loss rate if both constraints are satisfied, i.e., $q = 0.03$, while the third column gives the annualized loss rate for the case where both constraints are not satisfied, i.e., $q = 0.015$.

Table 4.1 shows that risk neutral and less risk averse beneficiaries suffer a loss in utility from the additional regulation, while more risk averse beneficiaries are not affected by the additional regulation. More specifically, we notice that the utility loss is largest for the risk neutral beneficiaries and that it is larger for a suboptimal regulation where both constraints are not satisfied simultaneously. The economic intuition behind this result is that the more regulation a risk neutral and less risk averse beneficiary faces, the more disutility he obtains.

[8] Since we have discussed in subsection 4.1 that setting the termination ratio extremely close to 0 is optimal for risk neutral and less risk averse beneficiaries, the corresponding utility is given by the limit as the termination ratio η approaches 0, i.e., $\lim_{\eta \to 0} u(\eta)$.

In particular, these beneficiaries would prefer a fully unregulated DB plan since this has the highest likelihood to pay out full pension benefits. More risk averse beneficiaries on the other hand are more concerned about potential pension losses than the full pension payment and therefore they are not harmed by additional regulation.

4.5 Conclusion

In this chapter we study when a pension guarantee fund should prematurely close an underfunded DB pension plan in a one-year expected utility maximization model. We assume that the pension guarantee fund (perfectly) acts in the interests of the pension beneficiaries and maximizes a power type utility function of the beneficiaries, whose argument is the funding ratio of the insured DB pension plan. In addition, we incorporate two regulatory constraints into the maximization problem, the SPC, which reflects the current regulatory practice, and the ESC, which better assesses the expected losses of ongoing DB pension plans. The SPC and the ESC restrict current and ongoing-concern liabilities of the pension guarantee fund, respectively. We find that the power type utility maximization problem with the SPC solely cannot account for an intervention policy for risk neutral and less risk averse pension beneficiaries. The inclusion of the ESC induces the premature termination rule to be applicable to any pension beneficiary independent of his risk preferences. More specifically, in our benchmark case where the two constraints overlap, we obtain an optimal termination ratio of 0.68 for risk neutral and less risk averse beneficiaries and 0.71 for more risk averse ones. In the end, in a utility loss analysis we show that adding the additional ESC constraint worsens risk neutral and less risk averse beneficiaries' utility, whereas more risk averse beneficiaries are not harmed by this additional regulation.

As a possible extension of this chapter one could also consider a utility function which has the loss aversion property, see Siegmann (2011) and the next chapter. A second extension could be to provide a model which distinguishes between the short-term regulatory practice, usually performed on an annual basis, and the long-term pension obligation horizon T and derives an optimal dynamic intervention policy. More specifically, it would be interesting to consider $m = T$ regulatory annual tests of the shortfall probability and the expected shortfall and to evaluate the expected utility at the long-term pension obligation horizon time T. A similar modeling framework is considered in Shi and Werker (2012) for long-term investors.

4.6 Appendix

4.6.1 Derivation of Lemma 4.3.1

The one-year expected utility can be decomposed into two parts:

(a) $\mathbb{E}\left[\frac{1_{\{0<\tau\leqslant 1\}}\eta^{(1-\delta)}}{1-\delta}\right]$

(b) $\mathbb{E}\left[\frac{1_{\{\tau>1\}}R_1^{(1-\delta)}}{1-\delta}\right]$

Part (a) can be calculated as follows:

$$\begin{aligned}\mathbb{E}\left[\frac{1_{\{0<\tau\leqslant 1\}}\eta^{(1-\delta)}}{1-\delta}\right] &= \frac{\eta^{(1-\delta)}}{1-\delta}P(0<\tau\leqslant 1)\\ &= \frac{\eta^{(1-\delta)}}{1-\delta}\left(\Phi(A-B)+\exp\{2AB\}\Phi(A+B)\right),\end{aligned}$$

where $A=\frac{1}{\sigma_R}\ln\left(\frac{\eta}{R_0}\right)$, $B=\frac{\mu_R-\frac{1}{2}\sigma_R^2}{\sigma_R}$ and $\Phi(\cdot)$ is the cdf of a standard normal distribution. The second equality uses the result of the one-year shortfall probability , i.e. $P(0<\tau\leqslant 1)=(\Phi(A-B)+\exp\{2AB\}\Phi(A+B))$, for $\eta\in(0,R_0)$. (See Lemma 4.3.2)
Part (b) can be calculated as follows:

$$\begin{aligned}\mathbb{E}\left[\frac{1_{\{\tau>1\}}R_1^{(1-\delta)}}{1-\delta}\right] &= \frac{1}{1-\delta}\mathbb{E}\left[\left(1-1_{\{0<\tau\leqslant 1\}}\right)R_1^{(1-\delta)}\right]\\ &= \frac{1}{1-\delta}\left[\mathbb{E}\left(R_1^{(1-\delta)}\right)-\mathbb{E}\left(1_{\{0<\tau\leqslant 1\}}R_1^{(1-\delta)}\right)\right].\end{aligned}$$

In the first equality, we use $1_{\{\tau>1\}}=1-1_{\{0<\tau\leqslant 1\}}$. Now the right hand side of the above equation consists of two parts. The first one can be directly calculated by using the fact that $\ln(R_1)$ follows a normal distribution. So we have

$$\mathbb{E}\left(R_1^{(1-\delta)}\right)=R_0^{(1-\delta)}\exp\left\{(1-\delta)(\mu_R-\frac{1}{2}\delta\sigma_R^2)\right\}.$$

The second one can be split into two parts by using $1_{\{0<\tau\leqslant 1\}}=1_{\{\inf_{t\in[0,1]}R_t\leqslant\eta\}}=1_{\{\inf_{t\in[0,1]}R_t\leqslant\eta,R_1\leqslant\eta\}}+1_{\{\inf_{t\in[0,1]}R_t\leqslant\eta,R_1>\eta\}}$:

$$\begin{aligned}\mathbb{E}\left[1_{\{0<\tau\leqslant 1\}}R_1^{(1-\delta)}\right] =&\mathbb{E}\left[1_{\left\{\inf_{t\in[0,1]}R_t\leqslant\eta,R_1\leqslant\eta\right\}}R_1^{(1-\delta)}\right]+\mathbb{E}\left[1_{\left\{\inf_{t\in[0,1]}R_t\leqslant\eta,R_1>\eta\right\}}R_1^{(1-\delta)}\right]\\ =&\mathbb{E}\left[1_{\{R_1\leqslant\eta\}}R_1^{(1-\delta)}\right]+\mathbb{E}\left[1_{\left\{\inf_{t\in[0,1]}R_t\leqslant\eta,R_1>\eta\right\}}R_1^{(1-\delta)}\right]\\ =&R_0^{(1-\delta)}\exp\left\{(1-\delta)(\mu_R-\frac{1}{2}\delta\sigma_R^2)\right\}\Phi\left(A-B-(1-\delta)\sigma_R\right)\\ &+\mathbb{E}\left[1_{\left\{\inf_{t\in[0,1]}R_t\leqslant\eta,R_1>\eta\right\}}R_1^{(1-\delta)}\right].\end{aligned}$$

In the last equality, we use the fact that $\ln(R_1)$ follows a normal distribution.
The second part $\mathbb{E}\left[1_{\left\{\inf_{t\in[0,1]} R_t\leqslant\eta, R_1>\eta\right\}} R_1^{(1-\delta)}\right]$ can be calculated by using the Girsanov theorem and the reflection principle:

$$\begin{aligned}
\mathbb{E}\left[1_{\left\{\inf_{t\in[0,1]} R_t\leqslant\eta, R_1>\eta\right\}} R_1^{(1-\delta)}\right] =&\mathbb{E}\left[1_{\left\{\inf_{t\in[0,1]} Bt+W_t\leq A, B+W_1>A\right\}}\right.\\
&\left. R_0^{(1-\delta)}\exp\left\{(1-\delta)(\mu_R-\frac{1}{2}\sigma_R^2)+(1-\delta)\sigma_R W_1\right\}\right]\\
=&\mathbb{E}^{\tilde{\mathbb{P}}}\left[\exp\{BW_1^{\tilde{P}}-\frac{1}{2}B^2\}1_{\left\{\inf_{t\in[0,1]} W_t^{\tilde{P}}\leq A, W_1^{\tilde{P}}>A\right\}}\right.\\
&\left. R_0^{(1-\delta)}\exp\left\{(1-\delta)\sigma_R W_1^{\tilde{P}}\right\}\right]\\
=&R_0^{(1-\delta)}\exp\left\{2AB-\frac{1}{2}B^2+2\sigma_R A(1-\delta)\right\}\mathbb{E}^{\tilde{\mathbb{P}}}\left[1_{\left\{W_1^{\tilde{P}}\leq A\right\}}\right.\\
&\left.\exp\left\{-BW_1^{\tilde{P}}-\sigma_R(1-\delta)W_1^{\tilde{P}}\right\}\right]\\
=&R_0^{(1-\delta)}\exp\left\{2AB+2A\sigma_R(1-\delta)+(1-\delta)\left(\mu_R-\frac{1}{2}\delta\sigma_R^2\right)\right\}\\
&\Phi\left(A+B+\sigma_R(1-\delta)\right)\ ,
\end{aligned}$$

where $W_t^{\tilde{P}}$ is a standard Brownian motion under the probability measure $\tilde{\mathbb{P}}$.
Finally, summing up all the parts gives the one-year expected utility function as a function of the termination ratio $\eta\in(0,\min\{R_0,1\})$, denoted as $u(\eta)$, as follows:

$$\begin{aligned}
u(\eta) =&\mathbb{E}\left[\frac{1_{\{\tau\leqslant 1\}}\eta^{(1-\delta)}}{1-\delta}\right]+\mathbb{E}\left[\frac{1_{\{\tau>1\}}R_1^{(1-\delta)}}{1-\delta}\right]\\
=&\frac{1}{1-\delta}\left[\eta^{(1-\delta)}\Big(\Phi(A-B)+\exp\{2AB\}\Phi(A+B)\Big)\right.\\
&+R_0^{(1-\delta)}\exp\left\{(1-\delta)\left(\mu_R-\frac{1}{2}\delta\sigma_R^2\right)\right\}\Phi\left(-A+B+(1-\delta)\sigma_R\right)\\
&\left.-R_0^{(1-\delta)}\exp\left\{2AB+2A\sigma_R(1-\delta)+(1-\delta)\left(\mu_R-\frac{1}{2}\delta\sigma_R^2\right)\right\}\Phi\left(A+B+\sigma_R(1-\delta)\right)\right],
\end{aligned}$$

$\delta\leqslant 0$ and $\delta\neq 1$.

4.6.2 Derivation of Lemma 4.3.3

The expected shortfall of the underfunded ongoing pension plan at year one can be calculated as follows:

$$\begin{aligned}\mathbb{E}\left[(1-R_1)1_{\{\tau>1\}}1_{\{R_1\leqslant 1\}}\right] &= \mathbb{E}\left[1_{\{\tau>1\}}1_{\{R_1\leqslant 1\}}\right] - \mathbb{E}\left[R_1 1_{\{\tau>1\}}1_{\{R_1\leqslant 1\}}\right] \\ &= \mathbb{E}\left[1_{\{R_1\leqslant 1\}}\right] - \mathbb{E}\left[1_{\{\inf_{t\in[0,1]} R_t\leqslant \eta\}}1_{\{R_1\leqslant 1\}}\right] - \mathbb{E}\left[R_1 1_{\{\tau>1\}}1_{\{R_1\leqslant 1\}}\right].\end{aligned}$$

In the second equality, we use $1_{\{\tau>1\}} = 1 - 1_{\{0<\tau\leqslant 1\}}$.
The first term $\mathbb{E}\left[1_{\{R_1\leqslant 1\}}\right]$ can be easily calculated by using the fact that $\ln(R_1)$ follows a normal distribution

$$\mathbb{E}\left[1_{\{R_1\leqslant 1\}}\right] = \Phi\left(-\frac{1}{\sigma_R}\ln(R_0) - B\right).$$

The second term $\mathbb{E}\left[1_{\{\inf_{t\in[0,1]} R_t\leqslant \eta\}}1_{\{R_1\leqslant 1\}}\right]$ can be decomposed into two parts by using $1_{\{\inf_{t\in[0,1]} R_t\leqslant \eta\}} = 1_{\{R_1\leqslant\eta\}} + 1_{\{\inf_{t\in[0,1]} R_t\leqslant \eta, R_1>\eta\}}$ as follows:

(a) $\mathbb{E}\left[1_{\{R_1\leqslant\eta\}}1_{\{R_1\leqslant 1\}}\right]$,

(b) $\mathbb{E}\left[1_{\{\inf_{t\in[0,1]} R_t\leqslant \eta, R_1>\eta\}}1_{\{R_1\leqslant 1\}}\right]$.

Since the termination ratio η is chosen from the set $(0, \min\{1, R_0\})$, it is strictly smaller than 1, so we have

$$\mathbb{E}\left[1_{\{R_1\leqslant\eta\}}1_{\{R_1\leqslant 1\}}\right] = \mathbb{E}[1_{\{R_1\leqslant\eta\}}] = \Phi(A-B),$$

where the second equality uses the fact that $\ln(R_1)$ follows a normal distribution.
Part (b) can be calculated by using the Girsanov theorem and the reflection principle as follows:

$$\begin{aligned}\mathbb{E}\left[1_{\{\inf_{t\in[0,1]} R_t\leqslant \eta, R_1>\eta\}}1_{\{R_1\leqslant 1\}}\right] =&\mathbb{E}\left[1_{\left\{\inf_{t\in[0,1]} W_t+Bt\leqslant A, W_1+B>A\right\}}1_{\left\{W_1+B\leqslant -\frac{1}{\sigma_R}\ln R_0\right\}}\right] \\ =&\mathbb{E}^{\tilde{\mathbb{P}}}\left[\exp\left\{BW_1^{\tilde{P}} - \frac{1}{2}B^2\right\}1_{\left\{\inf_{t\in[0,1]} W_t^{\tilde{P}}\leqslant A, W_1^{\tilde{P}}>A\right\}}1_{\left\{W_1^{\tilde{P}}\leqslant -\frac{1}{\sigma_R}\ln R_0\right\}}\right] \\ =&\mathbb{E}^{\tilde{\mathbb{P}}}\left[\exp\left\{2AB - BW_1^{\tilde{P}} - \frac{1}{2}B^2\right\}1_{\left\{W_1^{\tilde{P}}\leqslant A\right\}}1_{\left\{W_1^{\tilde{P}}\geqslant 2A+\frac{1}{\sigma_R}\ln(R_0)\right\}}\right] \\ =&\exp\{2AB\}\left[\Phi(A+B) - \Phi\left(2A+B+\frac{1}{\sigma_R}\ln(R_0)\right)\right].\end{aligned}$$

The third term $\mathbb{E}\left[R_1 1_{\{\tau>1\}}1_{\{R_1\leqslant 1\}}\right]$ can be split into two parts by using $1_{\{\tau>1\}} = 1-1_{\{0<\tau\leqslant 1\}}$

$$\mathbb{E}\left[R_1 1_{\{\tau>1\}}1_{\{R_1\leqslant 1\}}\right] = \mathbb{E}\left[R_1 1_{\{R_1\leqslant 1\}}\right] - \mathbb{E}\left[R_1 1_{\{\inf_{t\in[0,1]} R_t\leqslant \eta\}}1_{\{R_1\leqslant 1\}}\right].$$

$\mathbb{E}\left[R_1 1_{\{R_1\leqslant 1\}}\right]$ is easily calculated by using the fact that $\ln(R_1)$ follows a normal distribution

$$\mathbb{E}\left[R_1 1_{\{R_1\leqslant 1\}}\right] = R_0 \exp\{\mu_R\}\Phi\left(-\frac{1}{\sigma_R}\ln(R_0) - B - \sigma_R\right).$$

$\mathbb{E}\left[R_1 1_{\{\inf_{t\in[0,1]} R_t\leqslant\eta\}} 1_{\{R_1\leqslant 1\}}\right]$ can be further decomposed into two parts by using $1_{\{\inf_{t\in[0,1]} R_t\leqslant\eta\}} = 1_{\{R_1\leqslant\eta\}} + 1_{\{\inf_{t\in[0,1]} R_t\leqslant\eta, R_1>\eta\}}$ as follows:

(c) $\mathbb{E}\left[R_1 1_{\{R_1\leqslant\eta\}} 1_{\{R_1\leqslant 1\}}\right]$,

(d) $\mathbb{E}\left[R_1 1_{\{\inf_{t\in[0,1]} R_t\leqslant\eta, R_1>\eta\}}] 1_{\{R_1\leqslant 1\}}\right]$.

Since η chosen from the set $(0, \min\{R_0, 1\})$ is strictly smaller than 1 and $\ln(R_1)$ follows a normal distribution, part (c) can be easily calculated

$$\mathbb{E}\left[R_1 1_{\{R_1\leqslant\eta\}} 1_{\{R_1\leqslant 1\}}\right] = \mathbb{E}\left[R_1 1_{\{R_1\leqslant\eta\}}\right] = R_0 \exp\{\mu_R\}\Phi\left(A - B - \sigma_R\right).$$

Part (d) can be calculated by using the Girsanov theorem and the reflection principle as follows:

$$\begin{aligned}
\mathbb{E}\left[R_1 1_{\{\inf_{t\in[0,1]} R_t\leqslant\eta, R_1>\eta\}}] 1_{\{R_1\leqslant 1\}}\right] =& \mathbb{E}\Bigg[R_0 \exp\left\{\left(\mu_R - \frac{1}{2}\sigma_R^2\right) + \sigma_R W_1\right\} 1_{\left\{\inf_{t\in[0,1]} W_t + Bt\leqslant A, W_1+B>A\right\}} \\
& 1_{\left\{W_1+B\leqslant -\frac{1}{\sigma_R}\ln(R_0)\right\}}\Bigg] \\
=& \mathbb{E}^{\tilde{\mathbb{P}}}\Bigg[R_0 \exp\left\{\mu_R - B\sigma_R - \frac{1}{2}\sigma_R^2 - \frac{1}{2}B^2\right\} \exp\left\{(B+\sigma_R)W_1^{\tilde{P}}\right\} \\
& 1_{\left\{\inf_{t\in[0,1]} W_t^{\tilde{P}}\leqslant A, W_1^{\tilde{P}}>A\right\}} 1_{\left\{W_1^{\tilde{P}}\leqslant -\frac{1}{\sigma_R}\ln(R_0)\right\}}\Bigg] \\
=& \mathbb{E}^{\tilde{\mathbb{P}}}\Bigg[R_0 \exp\left\{\mu_R - B\sigma_R - \frac{1}{2}\sigma_R^2 - \frac{1}{2}B^2 + 2A(B+\sigma_R)\right\} \\
& \exp\left\{-(B+\sigma_R)W_1^{\tilde{P}}\right\} 1_{\left\{W_1^{\tilde{P}}\leqslant A\right\}} 1_{\left\{W_1^{\tilde{P}}\geqslant 2A+\frac{1}{\sigma_R}\ln R_0\right\}}\Bigg] \\
=& R_0 \exp\left\{\mu_R + 2AB + 2A\sigma_R\right\}\Bigg[\Phi(A+B+\sigma_R) \\
& - \Phi\left(2A + B + \frac{1}{\sigma_R}\ln(R_0) + \sigma_R\right)\Bigg].
\end{aligned}$$

Finally, summing up all the parts yields the closed-form expression for the ESC with for a certain value $q > 0$ as a function of the termination ratio $\eta \in (0, \min\{R_0, 1\})$, denoted by

$S(\eta)$:

$$
\begin{aligned}
\mathbb{E}\left[(1-R_1)1_{\{\tau>1\}}1_{\{R_1\leqslant 1\}}\right] =&\Phi\left(-\frac{1}{\sigma_R}\ln(R_0)-B\right)-\Phi\left(A-B\right)\\
&-\exp\left\{2AB\right\}\left[\Phi(A+B)-\Phi\left(2A+B+\frac{1}{\sigma_R}\ln(R_0)\right)\right]\\
&-R_0\exp\left\{\mu_R\right\}\Phi\left(-\frac{1}{\sigma_R}\ln(R_0)-B-\sigma_R\right)\\
&+R_0\exp\left\{\mu_R\right\}\Phi\left(A-B-\sigma_R\right)\\
&+R_0\exp\left\{\mu_R+2AB+2A\sigma_R\right\}\Bigg[\Phi(A+B+\sigma_R)\\
&-\Phi\left(2A+B+\frac{1}{\sigma_R}\ln(R_0)+\sigma_R\right)\Bigg]=S(\eta)\leqslant q.
\end{aligned}
$$

4.6.3 Proof of Proposition 4.3.4

We have the following one-year shortfall probability as a function of the termination ratio $\eta\in(0,\min\{R_0,1\})$

$$
\begin{aligned}
P(\eta)&=\Phi\left(A-B\right)+\exp\left\{2AB\right\}\Phi\left(A+B\right)\\
&=\Phi\left(\frac{1}{\sigma_R}\left(\ln\left(\frac{\eta}{R_0}\right)-\mu_R+\frac{1}{2}\sigma_R^2\right)\right)+\left(\frac{\eta}{R_0}\right)^{\left(\frac{2\mu_R}{\sigma_R^2}-1\right)}\Phi\left(\frac{1}{\sigma_R}\left(\ln\left(\frac{\eta}{R_0}\right)+\mu_R-\frac{1}{2}\sigma_R^2\right)\right).
\end{aligned}
$$

Since $\Phi(\cdot)$ is a cdf $P(\eta)$ is trivially continuous. The first derivative of $P(\eta)$ with respect to η is given as follows:

$$
\frac{\partial P(\eta)}{\partial\eta}=\frac{2}{\eta\sigma_R}\exp\left\{2AB\right\}\left[n(A+B)+(A+B)\Phi(A+B)\right]-\frac{2}{\eta\sigma_R}\exp\left\{2AB\right\}A\Phi(A+B),\tag{4.14}
$$

where $n(\cdot)$ is the pdf of a standard normal distribution.
$n(A+B)+(A+B)\Phi(A+B)$ is a function of the termination ratio $\eta\in(0,\min\{R_0,1\})$ which has the following properties

$$
\frac{\partial(n(A+B)+(A+B)\Phi(A+B))}{\partial\eta}=\frac{1}{\eta\sigma_R}\Phi(A+B)>0\quad\text{and}
$$

$$
\lim_{\eta\to 0}\left(n(A+B)+(A+B)\Phi(A+B)\right)=0\,.
$$

Thus, $n(A+B)+(A+B)\Phi(A+B)$ is positive for $\eta\in(0,\min\{R_0,1\})$, hence the first term in (4.14) is positive. Since $A < 0$ for the termination ratio $\eta \in (0, \min\{R_0, 1\})$, the second term in (4.14) is negative. As a result we have $\frac{\partial P(\eta)}{\partial \eta} > 0$.

Next it is easy to see that as the termination ratio η approaches 0, the function has a limit which is 0, i.e., $\lim_{\eta \to 0} P(\eta) = 0$, and as the termination ratio η approaches $\min\{R_0, 1\}$ we have

$$\lim_{\eta \to \min\{R_0,1\}} P(\eta) = \Phi\left(-\frac{\mu_R - \frac{1}{2}\sigma_R^2}{\sigma_R}\right) + \Phi\left(\frac{\mu_R - \frac{1}{2}\sigma_R^2}{\sigma_R}\right) = 1\,.$$

In the end, it is trivial to see that when the acceptable shortfall probability ϵ is 1, any termination ratio from the set $(0, \min\{R_0, 1\})$ satisfies the SPC. The more interesting case is when $\epsilon \in (0, 1)$. In this case the upper bound termination ratio is the solution of the equation

$$P(\bar{\eta}_\epsilon) - \epsilon = 0.$$

Bolzano's theorem proves that there exist a solution $\bar{\eta}_\epsilon$. The above proved monotonicity property of $P(\eta)$ implies that $\bar{\eta}_\epsilon$ is the unique solution. Since $P(\eta)$ is monotonically increasing, only termination ratios from the set $(0, \bar{\eta}_\epsilon]$ satisfy the SPC.

4.6.4 Proof of Proposition 4.3.6

We have the following expected deficits of the ongoing pension plan at year one

$$\begin{aligned}
S(\eta) &= \mathbb{E}\left[(1 - R_1)1_{\{\tau > 1\}}1_{\{R_1 \leqslant 1\}}\right] \\
&= \Phi\left(-\frac{1}{\sigma_R}\ln(R_0) - B\right) - \Phi\left(A - B\right) \\
&- \exp\left\{2AB\right\}\left[\Phi(A + B) - \Phi\left(2A + B + \frac{1}{\sigma_R}\ln(R_0)\right)\right] \\
&- R_0 \exp\left\{\mu_R\right\}\Phi\left(-\frac{1}{\sigma_R}\ln(R_0) - B - \sigma_R\right) \\
&+ R_0 \exp\left\{\mu_R\right\}\Phi\left(A - B - \sigma_R\right) \\
&+ R_0 \exp\left\{\mu_R + 2AB + 2A\sigma_R\right\}\left[\Phi(A + B + \sigma_R)\right. \\
&\left.- \Phi\left(2A + B + \frac{1}{\sigma_R}\ln(R_0) + \sigma_R\right)\right],
\end{aligned}$$

which is a continuous function of the termination ratio $\eta \in (0, \min\{R_0, 1\})$.
The expected deficit function $S(\eta)$ can also be calculated as follows:

$$\begin{aligned}
S(\eta) &= \mathbb{E}\left[(1-R_1)1_{\{\tau>1\}}1_{\{R_1\leqslant 1\}}\right] \\
&= \mathbb{E}\left[(1-R_1)1_{\{\inf_{t\in[0,1]} R_t>\eta\}}1_{\{R_1\leqslant 1\}}\right] \\
&= \mathbb{E}\left[(1-R_1)1_{\{\inf_{t\in[0,1]} R_t>\eta\}}1_{\{\eta<R_1\leqslant 1\}}\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[(1-R_1)\,1_{\{\eta<R_1\leqslant 1\}}\,1_{\{\inf_{t\in[0,1]} R_t>\eta\}}|R_1\right]\right] \\
&= \int_\eta^1 (1-R_1)P(\inf_{t\in[0,1)} R_t > \eta|R_1)\tilde{f}_{R_1} dR_1,
\end{aligned}$$

where $\tilde{f}_{R_1}$ denotes the pdf of the lognormal distribution of R_1. The main step we have used in this calculation is the law of iterated expectations. For each given funding ratio of the insured DB pension plan at year one R_1, the conditional probability of the pension plan not to be terminated before year one $P(\inf_{t\in[0,1)} R_t > \eta|R_1)$ decreases as η increases, hence the integrand $(1-R_1)P(\inf_{t\in[0,1)} R_t > \eta|R_1)$ is nonnegative. Accordingly, $\frac{\partial S(\eta)}{\partial \eta} < 0$.

Next, as the termination ratio approaches $\min\{R_0, 1\}$, the limit of $S(\eta)$ is 0 and as the termination ratio approaches 0 we have

$$\lim_{\eta\to 0} S(\eta) = \Phi\left(-\frac{1}{\sigma_R}\ln(R_0) - B\right) - R_0\exp\{\mu_R\}\Phi\left(-\frac{1}{\sigma_R}\ln(R_0) - B - \sigma_R\right).$$

If $q \geqslant \Phi\left(-\frac{1}{\sigma_R}\ln(R_0) - B\right) - R_0\exp\{\mu_R\}\Phi\left(-\frac{1}{\sigma_R}\ln(R_0) - B - \sigma_R\right)$, the ESC is trivially satisfied for any termination ratio $\eta \in (0, \min\{R_0, 1\})$. If $0 < q < \Phi\left(-\frac{1}{\sigma_R}\ln(R_0) - B\right) - R_0\exp\{\mu_R\}\Phi\left(-\frac{1}{\sigma_R}\ln(R_0) - B - \sigma_R\right)$, one can show analogously to the previous proof that there exists a unique lower bound termination ratio $\underline{\eta_\epsilon} \in (0, \min\{R_0, 1\})$ such that $S(\underline{\eta_\epsilon}) = q$ and admissible termination ratios, which satisfy the ESC, form the set $[\underline{\eta_\epsilon}, \min\{R_0, 1\})$.

Chapter 5

Portability, Salary and Asset Price Risk: A Continuous-Time Expected Utility Comparison of DB and DC Pension Plans[1]

5.1 Introduction

In the last two chapters we have analysed the insurance provided by the PBGC to defined benefit plan sponsors and we have studied how a pension guarantee fund can optimally intervene and terminate underfunded DB pension plans. In chapter 3 we have taken the benefits of the DB policyholder as given and in chapter 4 we have abstracted from the benefits and solely considered the funding ratio of the DB plan. In the present chapter we model the underlying DB plan stochastically and compare it to its main counterpart, the Defined Contribution (DC) plan, from the employee's perspective. Unlike chapter 2 and chapter 3 we do not study the pricing of the insurance contracts, but we work in an expected utility framework as in the last chapter and focus on the expected utility the employee can achieve at his retirement date in either of the two pension contracts.

Defined Benefit (DB) and Defined Contribution (DC) plans are two important types of private retirement plans in developed countries. In a DB plan, the employee's pension benefit is determined by a formula which takes years of service for the employer and wages or salary into account. In a DC plan, sponsoring companies (and often also their employees) pay a promised contribution to an external pension fund, which invests the contributions in financial assets. The pension payment is then simply determined as the market value of the backing assets. The DB plan was the dominant form of plan, but in the last decade the number of DC plans has a steady upward-moving trend. For instance, according to the US Flow of Funds Accounts, the division between assets held in private DB plans and private

[1]This chapter is based on Uzelac and Chen (2014)

DC plans was 60% versus 40% in 1987, and in 2007 this division was reversed. In the UK, Government Actuary's Department observed that final salary DB plans constituted 92% of all pension funds in 1979 and this number was reduced to 41% in 2005.

There are quite some significant tradeoffs between DB and DC plans, particularly when it comes to what risks the employees bear. Bodie et. al. (1985) and Blake (2000) provide a very extensive review on the tradeoffs. In this place, we want to emphasize three tradeoffs with respect to investment, portability and salary risk. In a DB plan, the sponsoring company is responsible for providing promised (future) pension benefits to the employees. In other words, the sponsoring company decides about investment policies in a pension fund and consequently also bears the entire investment risks in a DB plan. In a DC plan, the company does not ensure a promised pension payment to the employees. The employees bear the entire investment risks. From the employees perspective, salary risk is present in both the DB plan and the DC plan. In the former the employee bears the salary risk because the defined benefits are usually directly linked to his salary, while in the latter the amount of contributions the employee can make mainly depends on the development of his salary. Portability risk is the risk, not to have the ability to transfer years of credited service or accumulated benefits from one employer to another. It is widely accepted knowledge that portability risk plays a minor role in DC plans, while it is considered as the driving source of risk in single employer DB pension plans. Since the pension payment of a DC plan mainly depends on the value of the backing assets, a DC plan can be easily ported between job switchings. On the contrary, DB plan holders lose mostly part of their benefits after changing jobs since most DB plans lack portability provisions.

In a DC plan the asset price risk is the most important risk factor since the accumulated pension benefit of the employee is the market value of the contributions made while working and the investment returns earned on the plan balances. Blake, Cairns and Wood (2001) measure risk in the DC plan by computing VaR estimates during the accumulation phase. They find that the asset allocation strategy mainly drives the asset price risk since the VaR estimates are considerably more sensitive to the asset allocation strategy than to the choice of the asset return model.

Portability risk is considered the main risk factor in DB plans, especially because of the huge and increasing workforce mobility and the fact that only few single employer DB programs contain portability provisions[2]. Hall (1982) reports that workers in the US hold 10 or 11 jobs during their working lives. Blake (2000) mentions that fewer than 5% of workers remain with the same employer and that the average worker in the UK changes jobs about six times in a life time. Schrager (2009) further points out that job turnover has increased substantially in the 1990's compared to earlier decades. Blake and Orszag (1997) provide

[2]According to the Bureau of Labor Statistics Employee Benefits Survey in 1991 only 13% of full time workers were covered by portability provisions (see Foster (1994) for the different categories of portability provisions).

a detailed analysis of portability risk. In particular, they quantify different types of portability losses like the cash equivalent loss and the backloading loss for different deterministic wage paths and numbers of job moves in the UK. Specifically they report that even a low number of job moves can cause huge portability losses, for instance someone changing jobs once in a mid-career can lose up to 16% of the full service pension. A typical UK worker moving six times in a career could end up with a pension of only 70 − 75% of a pension of a worker with the same salary experience who remains in the same job for his whole career.

The main contribution of this chapter to the literature about the comparison of DB and DC pension plans is that we provide a formal model for comparing the two major types of private retirement plans by explicitly taking account of their most important risk factors in the presence of stochastic wages, job moving and asset prices. We compare the DB with the DC plan in an expected utility-based framework. Three frequently used utility functions in the pension insurance literature are considered: power utility, mean-shortfall and mean-downside deviation. The latter two utility functions penalize realizations of the terminal pension payment below a threshold – demonstrating the loss-aversion property. Under mean-shortfall, the penalty has a linear form. Mean-downside deviation punishes the loss more severely and the loss takes a quadratic form. Our main means of comparison is to compute the critical job switching intensity (from the DB plan) such that the beneficiary is indifferent between the DB and the DC plan.

Our methodology to compare the two pension retirement plans is similar to that of Siegmann (2011) who computes minimum funding ratios for the DB plan for the above mentioned utility functions. In particular, we also make the pension outcomes comparable by matching contributions in the two retirement plans. The main difference is that the latter focuses on the time diversification effects in a DB plan and models a static pension fund, while we model the DB plan from the perspective of a representative employee and focus on the portability risk.

We confirm some results in the existing literature (e.g. Coco and Lopes (2011), Samwick and Skinner (2004), Poterba, Rauh, Venti and Wise (2007) and Siegmann (2011)). First, a rise in the salary growth rate increases the attractiveness of the DB plan, while a higher salary volatility decreases its attractiveness. This reveals that the salary risk is more pronounced in the (final) salary DB plan. Second, the DB plan is preferred by an older beneficiary. It is mainly due to the fact that the overall portability loss becomes less severe due to the shorter contract duration. Third, adjusting the contribution of the beneficiary to a higher level makes the DC plan more attractive. Fourth, equity holding in a DC plan plays a substantial role in the relative attractiveness of the retirement plans, but there does not exist a clear dominating strategy for all the preferences.

Moreover, our model shows that portability losses substantially decrease the attractiveness of DB plans. In addition, by comparing the plans across utility functions we find that a mean-downside deviation beneficiary prefers the DB plan in most cases relatively more than the mean-shortfall beneficiary. Our model further yields one striking result which is inconsistent with the existing literature: the attractiveness of the DB plan can decrease in the level of risk aversion and the DC plan can become most attractive for the most risk-averse power-beneficiary. The rationale behind this most striking result is twofold. On the one hand, portability risk is modeled as a jump risk which generates much disutility for very risk-averse beneficiaries. On the other hand, the DC plan can offer better diversification because it is not purely driven by the income risk (asset risk plays a decisive role too).

The remainder of the present chapter is organized as follows. Section 5.2 models the pension payment in a DB and a DC pension plan. Additionally, we show how the contributions from these two plans can be matched. Section 5.3 determines analytically the expected utility of the beneficiary in a DB plan (for DC plan we rely on a simulation technique). Three utility functions are addressed: power utility, mean-shortfall and mean-downside deviation utility. In the subsequent section 5.4, the DB and DC plans are compared by mainly determining the indifference job switching intensity. Section 5.5 concludes the chapter and section 5.6 provides a detailed calculation for the propositions in the main text.

5.2 Model Setup

As in the last two chapters we consider a representative employee. The representative employee decides at $t = 0$ which pension plan he enters and he earns a pension benefit in T years from now. For simplicity we assume that the employee keeps this retirement plan until the retirement date.[3] For the DC pension plan we assume that the pension benefits are paid out as a lump sum, while the DB plan pays pension benefits as a life annuity, which is also usually the case in practice. Moreover, we abstract from mortality risk during the accumulation phase, inflation risk and sponsor bankruptcy.

The employee receives a salary which in our model is a continuous stochastic process $(S_t)_{t\geq 0}$. Furthermore, the employee is allowed to change jobs during his career. To simplify the model setup, we assume that the employee changes a job only for exogenous reasons. More precisely, we consider job changes due to personal reasons and exclude unemployment and any kind of endogenous or strategic job moves. In other words, we assume that whenever the employee changes a job, he is capable of finding a comparable job and his salary is not affected by the job move. This rationale justifies the continuous salary process assumption in the presence of job moving. The number of job moves is modeled as an (in)homogenous Poisson process $N(t)_{t\geq 0}$ with intensity $\lambda_t \geq 0$.[4] The expected or average number of job moves between $[0, t]$ is given by $\int_0^t \lambda_u \, du$. The salary process is assumed to follow a diffusion process with a possibly time-varying drift coefficient, which allows to better capture some

[3]Our framework excludes the case in which the employee can switch between the DB and DC plan.

[4]That is, we allow for a possibly time-varying but deterministic intensity.

empirically observed salary patterns, see the numerical analysis section. Accordingly,

$$\mathrm{d}S_t = \mu_S(t) S_t \mathrm{d}t + \sigma_S S_t \mathrm{d}W_t^S, \quad S_0 = s \tag{5.1}$$

where $\mu_S(t) \geq 0$ denotes the deterministic and possibly time-varying drift (trend in the salary), $\sigma_S > 0$ is the constant volatility and W^S is a standard Brownian motion, which is assumed to be independent of $N(t)_{t\geq 0}$ under the real world probability measure $\mathbb{P}$.

Next, as in chapter 3 we assume that there are two assets in our economy, a riskless asset F with price process $(F_t)_{t\geq 0}$ and a risky non-dividend-paying asset A with price process $(A_t)_{t\geq 0}$, i.e

$$\mathrm{d}F_t = rF_t \mathrm{d}t, \quad F_0 = 1 \tag{5.2}$$

$$\mathrm{d}A_t = \mu_A A_t \mathrm{d}t + \sigma_A A_t \mathrm{d}W_t, \quad A_0 = a. \tag{5.3}$$

The risky asset is modeled as a geometric Brownian motion where the standard Brownian motion W is assumed to be possibly correlated with the standard Brownian motion of salary process with the correlation coefficient ρ. Furthermore, it is independent of the number of job moves, hence we have $d[W, W^S]_t = \rho\, dt$ and $d[W, N]_t = 0$.

In the next subsections we model the pension income processes for the DB and DC pension plan.

5.2.1 DB Pension Plan

The main goal of our modeling framework is to incorporate portability risk into a DB pension plan of a representative employee in the presence of stochastic salaries and stochastic job moving. In practice, portability losses can be of two types. The major type is the cash equivalent loss. DB payments usually depend positively on the product of earnings and tenure. Since each of these tends to increase each year, much of the benefits are accrued in the last years prior to retirement. However, if a worker leaves a firm the final pay used to calculate the retirement benefits is the salary when he left the firm. As this salary is usually lower than the salary prior to retirement, a so-called cash equivalent loss occurs. [5] The second type of portability loss is called the backloading loss. This is an additional portability loss a worker switching jobs may suffer because contributions are backloaded in one scheme but not in another, see Blake and Orszag (1997) for a detailed discussion about the two types of portability losses.

The main factors determining the size of a portability loss are the ages at separation and the estimated real growth rate of wages, see Blake and Orszag (1997). These authors further illustrate that the portability losses are a hump shaped (inverse U shaped) function in the age of the beneficiary. That is, portability losses are increasing in the early career,

[5] Accordingly, this shows that the salary risk and the portability risk are interconnected in practice.

reach a maximum in the mid-career, decrease at the end of the career and are 0 at the retirement date. We provide a simple model at hand which takes the ages at separation into account and thus can capture the inverse U shaped structure of portability losses. However, to keep our model simple we do not link the real growth rate of wages to the size of a portability loss, therefore we do not quantify the size of each portability loss and neither do we exactly distinguish between the two types of portability losses. In other words, the simplifying assumption means that we treat the portability and the salary risk separately. Nevertheless, we can capture average portability losses in different stages of a career and we can also take the feature that portability losses increase with an increasing labor mobility into account.

To do so, we introduce the pension adjusted salary process $(\tilde{S}_t)_{t\geq 0}$ and model this as the jump diffusion

$$\mathrm{d}\tilde{S}_t = \mu_S(t)\tilde{S}_t\mathrm{d}t + \sigma_S\tilde{S}_t\mathrm{d}W_t^S + \tilde{S}_{t-}\mathrm{d}Q_t, \quad \tilde{S}_0 = S_0, \tag{5.4}$$

$t-$ denotes the time immediately before a job move and $Q_t = \sum_{i=1}^{N_t} Y_i$ is a compound Poisson process. $Y_i, i = 1, ...N_t$ are i.i.d. random variables, independent of N_t and the Brownian motions W and W^S. The Y_i's are used to model the percentage changes in the pension adjusted salary process when the employee changes his job. Intuitively, the pension adjusted salary is the salary which is eligible for retirement benefits at time t after taking the accumulated portability losses up to time t into account. More specifically, it contains a continuous part given by the first two terms in (5.4), which describe the changes in the pension income due to changes in the salary. The compound Poisson process captures the portability risk, that is the loss in the pension income due to a job change. Accordingly, we formally need to assume $Y_i < 0, i = 1, ...N_t$. In addition, we assume that whenever the employee changes a job, he loses a deterministic percentage $1 - \beta_i$, of his pension income, i.e $Y_i = \beta_i - 1$ with $0 < \beta_i < 1$.

More specifically, to link the percentage loss to the ages at separation, we allow β to be a deterministic function of time.[6] In particular, we will assume that β is a piecewise constant but time-varying function. Formally we define β as

$$\beta = \Big\{\beta_j, \quad t_j \leq t \leq t_{j+1}, j = 1...J\Big\}, \tag{5.5}$$

where J denotes the number of career periods considered. In addition $t_0 = 0$ and $t_J = T$.

Next, the stochastic differential equation (5.4) has the unique solution at time $t = 0$,

[6] In a more realistic setup, as suggested above the Y_i's would be stochastic and also directly depend on the salary process S, particularly the trend of the salary $\mu_S(t)$. We could also include the trend in a deterministic way to the Y_i's, but in order to keep the impact of the model parameters clear we stick to our simple assumption.

see e.g. Shreve (2004),

$$\begin{aligned}\tilde{S}_T =& s\,\exp\left\{\Big(\int_0^T \mu_S(u)\,du - \frac{1}{2}\sigma_S^2\,T\Big) + \sigma_S\,W_T^S\right\}\prod_{i=1}^{N_T}\beta_i \\ =& s\,\exp\left\{\Big(\int_0^T \mu_S(u)\,du - \frac{1}{2}\sigma_S^2 T\Big) + \sigma_S\,W_T^S\right\}\exp\left\{\sum_{j=1}^{J}(N(t_j) - N(t_{j-1}))\ln(\beta_j)\right\},\end{aligned} \tag{5.6}$$

where in the second equation we have used the piecewise constant property of the jump size β.

The DB plan we consider is a final salary DB plan. That is, we assume that the employee receives a continuous annuity $b(T)$ which is the product of a pre-specified replacement rate α, where $0 < \alpha \leq 1$, and the terminal value of the pension adjusted salary, i.e $b(T) = \alpha\,\tilde{S}_T$. The crucial point is that in order to incorporate portability losses the retirement benefit formula is based on the pension adjusted salary process instead of the salary process. To make the DB plan and the DC plan comparable we are first going to convert the life annuity of the DB plan into a lump sum. Formally the lump sum the beneficiary receives, which we denote $B(T)$, can be determined as

$$B(T) = \int_T^\infty b(T)\,e^{-r\,(\tau-T)}\,p_\tau\,d\tau, \tag{5.7}$$

where p_τ is a continuous survival distribution function and τ is the time of death. We assume that the annuity is paid up to maximum age T^1 and also use the simplifying assumption of a constant mortality intensity μ. Then the lump sum can be computed as

$$\begin{aligned}B(T) &= \int_T^{T^1} b(T)\,e^{-r\,(\tau-T)}\,e^{-\mu(\tau-T)}d\tau \\ &= \frac{b(T)}{r+\mu}\left[1 - e^{-(r+\mu)\,(T^1-T)}\right] := b(T)\,a(T),\end{aligned} \tag{5.8}$$

where $a(T)$ can be interpreted as the annuity factor.

5.2.2 DC Pension Plan

Unlike the DB pension plan, portability risk plays a minor role in DC plans as for the latter the value of pension benefits is simply determined as the market value of the backing assets. Therefore, benefits are easily transferable between jobs, see Zhang (2008). Moreover, portability losses are unlikely to occur since DC plans are not backloaded and the contribution rates are not tied to tenure and age of the workers (see Bodie, Marcus and Merton (1986)). More importantly, the main economic argument for including portability risk into a DC plan is that many moving workers may use their lump sum distributions for spending instead of reinvesting them in another retirement account (see e.g. Schultz (1995)). However, this argument has not been confirmed empirically, see Samwick and Skinner (2004). Accordingly, for the DC plan we assume that job moving will not affect the pension income of the representative employee. Instead as emphasized in the introduction, the employee in a DC pension plan bears mainly the asset price risk which in a DB plan is mainly born by the employer.

As we abstract from portability risk for the DC plan, the DC account value can be modeled as a continuous stochastic process $(X_t)_{t\geq 0}$. We model asset price risk as in chapter 3 and assume that the employee's investment follows a rebalancing strategy. More specifically, the employee chooses at $t = 0$ a constant fraction π, $0 \leq \pi \leq 1$ which will be invested in the risky asset A and the remaining fraction $(1 - \pi)$ is invested in the riskless asset F. Then the DC account value is continuously rebalanced by a DC fund manager, that is at any time $0 < t < T$ the amount πX_t is invested in the risky asset and the remaining amount in the riskless asset.

In order to capture the nature of the DC plan, we need to allow for contributions into the employee's account. We assume that the contributions are made by both the employee and the employer, see section 5.4 for more details. These contributions represent cash inflows into the DC account value. More specifically, we model a stylized DC plan where the employee and the employer contribute continuously the amount $c\, S_t\, \mathrm{d}t$, $0 \leq c \leq 1$, to the employee's pension account and these contributions are also invested continuously over time. In other words the employee and the employer contribute in each time period $\mathrm{d}t$ a predetermined constant percentage c of the current employee's salary to his DC account. This implies that the DC account value evolves according to

$$\mathrm{d}X_t = X_t \left[(r + \pi\, \sigma_A\, \theta)\, \mathrm{d}t + \pi\, \sigma_A\, \mathrm{d}W_t\right] + c\, S_t\, \mathrm{d}t, \quad X_0 = c\, S_0, \tag{5.9}$$

where $\theta = \frac{(\mu_A - r)}{\sigma_A}$ denotes the market price of risk.

As the beneficiary in the DC plan receives a lump sum at the retirement date, the pension benefit simply coincides with the terminal value of the DC account X_T.

5.2.3 Matching the Employee's contributions

In order to make the pension outcomes comparable we need to ensure that the employee bears effectively the same costs in the two pension retirement plans.[7] [8] A way to achieve this requirement is to assume that the employee contributes continuously the amount $q\, S_t\, \mathrm{d}t$,

[7]We do not require the employers costs to be necessarily the same in the two retirement plans, since the pension plans are compared from the employees perspective and in practice the costs the employer bears in the two plans also differ.

[8]Of course this is just a theoretical assumption here. In reality employees in DC often bear higher costs since they need to contribute periodically a fixed rate to the DC account, while the employees in a DB plan often bear less costs as most of the contributions in the DB plan are variable deficit contributions and are mainly covered by the employer.

where $q \leq c$ denotes a constant percentage of the salary, in either retirement plan. In other words, we assume that there is a one-to-one contribution match, i.e the employee makes the same contributions in both retirement plans. The condition $q \leq c$ is needed because in our modeling of the DC plan, c is used to denote the entire contribution rate provided by the employer and the employee.

Then we determine the employee's contribution rate q and the total contribution rate c in the DC plan in two steps. In the first step, q is determined in the DB plan by linking the employee's contribution rate to the replacement rate α. This link is important since the terminal payment in the DB plan crucially depends on the replacement rate. We implicitly assume that all the contributions (employee and employer) in the DB plan are incorporated in the replacement rate. This replacement rate is split into a replacement rate α^{ER}, $0 \leq \alpha^{ER} \leq 1$, coming from the employer's contribution and a replacement rate α^{EE}, $0 \leq \alpha^{EE} \leq 1$, coming from the employee's contributions. That is, $\alpha = \alpha^{ER} + \alpha^{EE}(q)$. More specifically, we assume that the employer first sets the replacement rate α^{EE} by fixing values for the total replacement rate α and the replacement rate coming from his contributions α^{ER}. Then he determines the employee's contribution rate q such that on average the accumulated employee contributions $q \int_0^T S_u \,\mathrm{d}u$ coincide with the self-financed pension income if the employee stays with the employer, which is given by $a(T)\,\alpha^{EE}\,S_T$. In particular, we assume that the employer does not take any potential portability losses into account when setting the employee's contribution rate and therefore it is only the employee who bears the entire costs of the portability losses. Formally, we link the employee's contribution rate to his replacement rate by requiring that

$$\mathbb{E}\big[q \int_0^T S_u \,\mathrm{d}u\big] = a(T)\,\alpha^{EE}\,\mathbb{E}\big[S_T\big], \tag{5.10}$$

which can be interpreted as the fair contribution condition in the DB plan.

In the following, we will assume that the salary trend is also a piecewise constant but time-varying function, i.e $\mu_S = \Big\{\mu_{S,j} \quad t_j \leq t \leq t_{j+1}, j = 1...J\Big\}$. Then the right hand side of (5.10) becomes $\alpha^{EE}\,a(T)\,s\exp\{\sum_{j=1}^{J}\mu_{S,j}\,(t_j - t_{j-1})\}$. The left hand side can be computed as

$$\begin{aligned}\mathbb{E}\big[q \int_0^T S_u \,\mathrm{d}u\big] =& q \int_0^T s\, e^{\int_0^u \mu_S(v)\,\mathrm{d}v}\,\mathrm{d}u \\ =& q \sum_{j=1}^{J} \int_{t_{j-1}}^{t_j} s\, e^{\sum_{k=1}^{j-1}\mu_{S,k}\,(t_k - t_{k-1}) + \mu_{S,j}(u - t_{j-1})}\,\mathrm{d}u \\ =& q\,\cdot s \cdot \sum_{j=1}^{J} \frac{1}{\mu_{S,j}}\Big(e^{\sum_{k=1}^{j}\mu_{S,k}\,(t_k - t_{k-1})} - e^{\sum_{k=1}^{j-1}\mu_{S,k}\,(t_k - t_{k-1})}\Big),\end{aligned}$$

where $\sum_{k=1}^{j-1} \equiv 0$ for $j = 1$. In the computation we have mainly used the Fubini theorem to interchange the order of integration and the piecewise constant property of the drift coefficient. Finally we solve the equation above for q to obtain

$$q = \frac{\alpha^{EE}\, a(T)\, \exp\{\sum_{j=1}^{J} \mu_{S,j}\,(t_j - t_{j-1})\}}{\sum_{j=1}^{J} \frac{1}{\mu_{S,j}} \left(e^{\sum_{k=1}^{j} \mu_{S,k}\,(t_k - t_{k-1})} - e^{\sum_{k=1}^{j-1} \mu_{S,k}\,(t_k - t_{k-1})} \right)}. \tag{5.11}$$

Note that in case of a constant salary drift the matched employee contribution simplifies to

$$q = a(T)\, \alpha^{EE}\, \mu_S \left(1 - \exp(-\mu_S\, T)\right)^{-1}. \tag{5.12}$$

The (fair) employee contribution rate q in equation (5.11) and (5.12) mainly depends on the salary drift parameters and the length of the career periods. In particular, one can show the the matched contribution rate increases with $\mu_{S,j}$ and decreases with T. Moreover, note that our assumptions immediately ensure that $q \geq 0$. The condition that $q \leq 1$ requires that the nominator in (5.11) is smaller than the denominator. This is the case for any reasonable choice for the salary drift vector μ_S and contract maturity T.

In the second step, the total contribution rate c in the DC plan is determined by taking the above specified employee's contribution rate q and assuming that the employer simply matches the employee's contribution in the DC plan, i.e $c = \delta\, q$, where $\delta \geq 1$ denotes the matching factor.

5.3 Utility-Based Comparison

5.3.1 Utility Functions and Certainty Equivalents

We consider three frequently used utility functions in the financial and pension insurance literature in our expected utility analysis: power utility, mean-shortfall and mean-downside deviation.

For a payoff x, the power utility is defined as

$$u(x) = \begin{cases} \frac{1}{1-\gamma}\, x^{1-\gamma}, & \gamma \neq 1 \\ \ln x, & \gamma = 1 \end{cases} \tag{5.13}$$

where γ is the constant coefficient of relative risk aversion. The power utility is abundantly used in both theoretical and empirical research because of its nice analytical tractability. More importantly, the use of the power utility is also motivated economically since the long-run behavior of the economy suggests that relative risk aversion cannot depend strongly on wealth, see Campbell and Viceira (2002). The certainty equivalent, that is the guaranteed amount of money that an economic agent would accept instead of the risky asset, i.e $u(CE) = \mathbb{E}[u(x)]$, for the power utility is simply given $CE(x) = (1-\gamma)\left(\mathbb{E}[u(x)]\right)^{\frac{1}{1-\gamma}}$ for

$\gamma \neq 1$ and $CE(x) = \exp\Big\{\mathbb{E}[u(x)]\Big\}$ for $\gamma = 1$.

The second utility function we consider, mean-shortfall is given by

$$u(x) = \begin{cases} x - R, & x \geq R\,, \\ -\eta_1\,(R - x), & x < R, \end{cases} \tag{5.14}$$

where R is the reference value, in our context this is the desired target pension income of the employee. Loss aversion boils down to penalizing realizations of x below R with a penalty parameter η_1. This specification is a linearized version of that originally proposed by Kahneman and Tversky (1979) and also used for instance by Benartzi and Thaler (2005). The Certainty equivalent for the mean-shortfall utility is obtained by solving the equation $\mathbb{E}[u(x)] = u(CE)$ for the gain and the loss side separately. Then one obtains

$$CE(x) = \begin{cases} \mathbb{E}[u(x)] + R, & \mathbb{E}[u(x)] \geq 0\,, \\ R - \Big(- \frac{\mathbb{E}[u(x)]}{\eta_1}\Big), & \mathbb{E}[u(x)] < 0. \end{cases} \tag{5.15}$$

The last type of utility function, mean-downside deviation[9] is comparable to mean-shortfall with the essential difference that one uses a quadratic penalty specification. Large shortfalls below the reference point R are penalized more severely:

$$u(x) = \begin{cases} x - R, & x \geq R \\ -\eta_2\,(R - x)^2, & x < R. \end{cases} \tag{5.16}$$

The mean-downside deviation utility is proposed by Boender (1997) in the pension fund context and has since then been adopted in the ALM practice in the Netherlands (see e.g. Siegmann (2011)). Finally the certainty equivalent of the mean-downside deviation utility is given by

$$CE(x) = \begin{cases} \mathbb{E}[u(x)] + R, & \mathbb{E}[u(x)] \geq 0\,, \\ R - \sqrt{- \frac{\mathbb{E}[u(x)]}{\eta_2}}, & \mathbb{E}[u(x)] < 0. \end{cases} \tag{5.17}$$

5.3.2 Expected Utility Results

As in the previous chapter we compute the expected utilities under the real world measure $\mathbb{P}$ for the defined benefit pension plan. We are not able to compute the expected utilities for the DC fund since the stochastic differential equation (5.9) is a sum of two stochastic processes which does not admit a closed-form solution. Therefore, we solve this with an Euler discretization scheme and compute the corresponding expected utilities with Monte Carlo simulation. For the DB plan we further allow for a piecewise constant and time-varying job switching intensity,i.e $\lambda = \Big\{\lambda_j \quad t_j \leq t \leq t_{j+1}, j = 1 ... J\Big\}$.

[9] In the sequel we will frequently abbreviate the mean-shortfall utility as LA utility and the mean-downside deviation utility as DD utility.

Proposition 5.3.1 (Expected Utilities under the DB plan.). *The expected utility for the power utility function is given by*

$$\mathbb{E}[u(B_T^{DB})] = \frac{1}{1-\gamma}\,(\alpha\, s\, a(T)\,)^{1-\gamma}\,\exp\left\{(1-\gamma)\,(\sum_{j=1}^{J}\mu_{S,j}(t_j-t_{j-1})-\frac{1}{2}\sigma_S^2\,T)+\frac{1}{2}(1-\gamma)^2\,\sigma_S^2\,T\right\} \tag{5.18}$$

$$\cdot\ \exp\left\{\sum_{j=1}^{J}\lambda_j\,(t_j-t_{j-1})\,(e^{(1-\gamma)\,\ln(\beta_j)}-1)\right\}. \tag{5.19}$$

For $\gamma = 1$ (log utility), we obtain

$$\mathbb{E}[u(B_T^{DB})] = \ln(a(T)\,\alpha\, s\,) + \left(\sum_{j=1}^{J}\mu_S(t_j-t_{j-1})-\frac{1}{2}\,\sigma_S^2\,T\right)+\sum_{j=1}^{J}\lambda_j(t_j-t_{j-1})\,\ln(\beta_j)\,. \tag{5.20}$$

For the mean-shortfall we have

$$\begin{aligned}
&\mathbb{E}[u(B_T^{DB})]\\
&=\sum_{k_1=0}^{\infty}\sum_{k_2=k_1}^{\infty}\cdots\sum_{k_J=k_{J-1}}^{\infty}\Bigg[\alpha\, s\, a(T)\,\exp\left\{\sum_{j=1}^{J}\mu_{S,j}(t_j-t_{j-1})\right\}\prod_{j=1}^{J}\beta_j^{(k_j-k_{j-1})}\,\Phi(d_1(k_1,...,k_J))\\
&\quad - R\,\Phi(d_2(k_1,...,k_J)) - \eta_1\,R\,\Phi(-d_2(k_1,...,k_J))\\
&\quad + \eta_1\,\alpha\, s\, a(T)\,\exp\left\{\sum_{j=1}^{J}\mu_{S,j}(t_j-t_{j-1})\right\}\prod_{j=1}^{J}\beta_j^{(k_j-k_{j-1})}\,\Phi(-d_1(k_1,...,k_J))\Bigg]\\
&\quad \times\prod_{j=1}^{J}\frac{\lambda_j\,(t_j-t_{j-1})^{(k_j-k_{j-1})}}{(k_j-k_{j-1})!}\,e^{-\lambda_j\,(t_j-t_{j-1})},
\end{aligned} \tag{5.21}$$

where Φ again denotes the cumulative distribution function of a standard normal distribution and $d_1(k_1,...,k_J)$ and $d_2(k_1,...,k_J)$ are given by

$$\begin{aligned}
d_1(k_1,...,k_J) &= \frac{\ln\frac{\alpha\, s\, a(T)\,\prod_{j=1}^{J}\beta_j^{(k_j-k_{j-1})}}{R}+(\sum_{j=1}^{J}\mu_{S,j}(t_j-t_{j-1})+\frac{1}{2}\sigma_S^2\,T)}{\sigma_S\sqrt{T}},\\
d_2(k_1,...k_J) &= d_1(k_1,...,k_J)-\sigma_S\sqrt{T}.
\end{aligned}$$

Finally for the mean-downside deviation utility we obtain

$$
\begin{aligned}
&\mathbb{E}[u(B_T^{DB})] \\
=&\sum_{k_1=0}^{\infty}\sum_{k_2=k_1}^{\infty}\cdots\sum_{k_J=k_{J-1}}^{\infty}\Bigg[\alpha\, s\, a(T)\, e^{\sum_{j=1}^{J}\mu_{S,j}\,(t_j-t_{j-1})}\prod_{j=1}^{J}\beta_j^{(k_j-k_{j-1})}\,\Phi(d_1(k_1,...,k_J)) - R\Phi(d_2(k_1,...,k_J)) \\
&-\eta_2\, R^2\,\Phi(-d_2(k_1,...,k_J)) + 2\eta_2 R\alpha\, s\, a(T)\, e^{\sum_{j=1}^{J}\mu_{S,j}\,(t_j-t_{j-1})}\prod_{j=1}^{J}\beta_j^{(k_j-k_{j-1})}\,\Phi(-d_1(k_1,...,k_J)) \\
&-\eta_2(\alpha\, s\, a(T)\,)^2 e^{(2\sum_{j=1}^{J}\mu_{S,j}\,(t_j-t_{j-1})+\sigma_S^2\,T)}\prod_{j=1}^{J}\beta_j^{2\,(k_j-k_{j-1})}\Phi(-d_3(k_1,...,k_J))\Bigg] \\
&\times\prod_{j=1}^{J}\frac{\lambda_j\,(t_j-t_{j-1})^{(k_j-k_{j-1})}}{(k_j-k_{j-1})!}\,e^{-\lambda_j\,(t_j-t_{j-1})},
\end{aligned}
\tag{5.22}
$$

where

$$
d_3(k_1,...,k_J) = \frac{\ln\frac{\alpha\, s\, a(T)\prod_{j=1}^{J}\beta_j^{(k_j-k_{j-1})}}{R} + (\sum_{j=1}^{J}\mu_{S,j}(t_j-t_{j-1}) + \frac{3}{2}\sigma_S^2\,T)}{\sigma_S\sqrt{T}}.
$$

Proof 5.3.2. *The proof of the proposition is given in Appendix 5.6.*

5.4 Numerical analysis

In our benchmark case we assume that the size of the portability losses β, the salary trend μ_S and the job switching intensity λ are constant. The more realistic case, where these parameters are time-varying is devoted as a sensitivity analysis to the next subsection. We compare the DB and the DC pension plan mainly by computing the indifference job switching intensity with a numerical search algorithm. This is the job switching intensity which makes the employee equally well off in terms of the corresponding expected utility in any of the two pension plans. We denote the indifference job switching intensity as λ^*. The employee receives a higher expected utility from a lower value of λ as the overall portability loss is smaller the less frequently he changes his job. It implies that the DB plan is more attractive than the DC plan for values of $\lambda < \lambda^*$, while the DC plan is favored for values of $\lambda > \lambda^*$. At $\lambda = \lambda^*$, the employee is indifferent between the two plans. Consequently, a higher value of λ^* implies that the DB plan becomes more attractive in more situations. Note that if for a specific parameter combination there does not exist an indifference intensity, it simply means that the DC plan is even preferable to a DB plan without portability losses, which is the so-called cash balance pension plan.[10] For our numerical analysis, we choose

[10]This is a so-called hybrid pension plan, which has the main features of DB plans but with the main difference that pension benefits are portable.

the following benchmark model parameters:

$$\begin{aligned} \alpha = 0.2,\ \alpha^{ER} = 0.15,\ \mu = 0.0005,\ \delta = \frac{3}{2},\ S_0 = 1000, \\ T = 25, T^1 = 30,\ \mu_S = 0.015,\ \sigma_S = 0.13,\ \beta = 0.95, \\ r = 0.02,\ \mu_A = 0.055,\ \sigma_A = 0.25,\ \rho = 0. \end{aligned}$$

Specifically, we assume that the employee makes a decision to enter one of the two retirement plans at the age of 40 and he retires at 65. The replacement rate coming from the employee's contributions is $\alpha^{EE} = 0.05$. This implies that the employee's contribution rate q is approximately 5.2%. The values μ_S=0.015 and $\sigma_S = 0.13$ are empirically estimated by Topel and Ward (1992). For the matching mechanism of the employee's and the employer's contribution, we assume the standard matching mechanism in practice (see e.g Samwick and Skinner (2004)): the employer contributes 0.5 $ on each dollar contributed by the employee. This implies that in our benchmark case the total contribution rate in the DC plan is $c = 7.8\,\%$ $(= q\,\delta)$. Furthermore, the correlation coefficient ρ between the salary process and the risky financial asset is set to 0, following Davis and Willen (2000) who find a low correlation for shocks in earnings and stock market returns. Most importantly, the value for the portability loss size β is chosen to reflect the empirical estimates of Blake and Orszag (1997). In particular, these authors estimated that a typical UK worker moving six times in a career could end up with a pension of only $70 - 75\%$ of a pension of a worker with the same salary experience who remains in the same job for his whole career. The value of $\beta = 0.95$ implies that a worker with six job changes only obtains $73.5\%(= 0.95^6)$ of the retirement income compared to one without job changes in our model.

We consider the following set of coefficient values for the risk aversion parameters, see also Siegmann (2011) for similar values:

$$\gamma \in \{1, 2, 4\},\ \eta_1 = \eta_2 \in \{2.25, 5\}. \tag{5.23}$$

Furthermore, we assume that the reference point for the mean-shortfall utility and the mean-downside deviation utility is a multiple of the value of the investment in the money market account. In our benchmark case we choose $R = 5\ c\,s\ \exp\{r\,T\}$. We consider the values $\pi = 0.4$, $\pi = 0.57$, $\pi = 0.75$, $\pi = 0.75$ and $\pi = 0.9$ for the fraction invested in risky assets. The fraction $\pi = 0.57$ is our benchmark investment strategy, where the value is estimated from DC pension asset allocation data for the US from Broadbent, Palumbo and Woodman (2006).[11] Finally we use the time discretization of $\mathrm{d}t = \frac{1}{12}$ for the Euler scheme and $n = 100000$ simulation runs to evaluate the expected utilities of the DC pension plan.

Table 5.1 displays values for the indifference job switching intensity λ^* for the three utility functions and their corresponding risk aversion parameters under the four considered investment strategies. An economically intuitive way to interpret this and the tables at

[11]The estimation method is the same as described in chapter 3.

Utility	Risk aversion	$\pi = 0.4$	$\pi = 0.57$	$\pi = 0.75$	$\pi = 0.9$
CRRA	$\gamma = 1$	0.3423	0.3169	0.3058	0.3088
	$\gamma = 2$	0.2540	0.2690	0.3087	0.3673
	$\gamma = 4$	0.0897	0.1724	0.3052	0.4309
LA	$\eta_1 = 2.25$	0.3068	0.2532	0.1929	0.1349
	$\eta_1 = 5$	0.2965	0.2504	0.1890	0.1399
DD	$\eta_2 = 2.25$	0.3526	0.3121	0.2831	0.2958
	$\eta_2 = 5$	0.3061	0.2801	0.2821	0.3442

Table 5.1: Values of λ^* for the benchmark case.

hand is to compute the expected (average) number of job moves under which the DB plan is still preferred by the employee, which is given by $T \times \lambda^*$. For our benchmark investment strategy $\pi = 0.57$ the DB plan is on average preferred in ascending order of the risk aversion parameters up to 7, 6 and 4 job moves for the power utility, 6 job moves for the LA utility and 7 and 6 job moves for the DD utility.

Furthermore, the table shows that the investment strategy has a huge impact on the indifference job switching intensities for all utility functions, where the impact is most pronounced for the most risk averse power beneficiary and the LA beneficiary. More specifically, one observes that there is no clear dominating investment strategy. In our context, the best investment strategy (among the four values of π) is the one with the lowest indifference job switching intensity, which implies that the DC plan will be most frequently preferred. The most risky strategy is best for the LA utility maximizer independent of his loss aversion. Intuitively, the LA utility maximizer would choose the most risky strategy because gains in the pension income through gains in the financial portfolio receive the highest weight for this utility function. On the other hand, potential losses are not severely penalized, therefore the LA utility maximizer tolerates the high financial risk. The investment strategy $\pi = 0.75$ is best for a less loss averse DD utility maximizer and the least risk averse power beneficiary. The benchmark investment strategy is preferred by the more loss averse power beneficiary. More risk averse power beneficiaries ($\gamma = 2$ and $\gamma = 4$) find the most conservative investment strategy best.

Next we fix the best investment strategies above and make a comparison across the different utility functions. We can state the following two interesting points. First, the DC plan is most attractive for the most risk averse power beneficiary since this beneficiary would prefer the DC plan after 3 job moves on average. Second, comparing the two utility functions with the loss aversion property we see that the DB plan is considerably more attractive for the DD beneficiary. Compared to the best strategy of a LA beneficiary, he would on average need to have 4 more job moves to prefer the DC pension plan. It is important to note that the latter point holds for any investment strategy. Intuitively this is because the DD beneficiary is more loss averse than the LA beneficiary, therefore he prefers to take less financial risk than his LA counterpart.

Most interestingly, table 5.1 reveals that in our model the DC plan can become significantly more attractive with increasing risk aversion. It is particularly the case for the more conservative investment strategies $\pi = 0.4$ and $\pi = 0.57$ and most pronounced for the power beneficiary, which also accounts for the result above that the DC plan is most attractive for the most risk averse power beneficiary. This result is to some extent inconsistent with the known result in the literature that DB plans relative attractiveness increases with increasing risk aversion of the beneficiaries, see e.g Siegmann (2011). Our result can be explained by two effects. First, and probably most importantly, portability risk is modeled as a jump risk, which represents a substantial source of risk for risk averse pension beneficiaries. Second, the DC plan offers a better diversification than the DB plan because the benefits here do not only depend on the evolution of the salary process. This effect is the more pronounced the lower the investment in the risky asset is and also the lower the correlation between the risky asset and the salary process is. For a very high equity holding like $\pi = 0.9$, however, the volatility of the financial portfolio, which is given by $\pi\,\sigma_A$, becomes fairly high such that the financial risk dominates the portability and the salary risk in the DB plan. Accordingly, for a very risky investment strategy the DB becomes significantly more attractive with increasing risk aversion.

5.4.1 Sensitivity Analysis

In this subsection we investigate how far the more realistic model setup with a time-varying portability loss size, salary trend and job switching intensity affects the above stated results. As suggested in our model setup we assume the corresponding parameters to be piecewise constant. Specifically we consider 3 time periods, where $t = [0 \quad 10 \quad 20 \quad 25]$. The first 10 years are referred to as the early career, the next 10 years as the mid-career and the last 5 years as the end of the career. In each case we let one parameter be time-varying while keeping the other parameters constant.

In table 5.2 we compute the indifference job switching intensity for the more realistic case of hump-shaped portability losses, i.e U-shaped portability loss size β, see Blake and Orszag, Chapter 4 (1997). That is, portability losses are increasing up to the end of the mid-career and reach a maximum there, reflected by the high portability loss size ($\beta_2 = 0.9$), while they are very small at the end of the career, accordingly $\beta_3 = 0.99$. The portability loss size at the early career is assumed to be the same as in our benchmark case, i.e $\beta_1 = 0.95$. We observe that our main results stated above, about the impact of the investment strategy, the effect of the risk aversion and also that DD beneficiaries relatively prefer more DB plans than LA beneficiaries, remain unchanged. More specifically, we observe that the more pronounced portability losses in the mid-career imply that the employee would prefer the DC plan after 1 or 2 less job moves on average for any utility function. This indicates that portability losses play a considerable role in the relative attractiveness of the

DB plan.

Utility	Risk aversion	$\pi = 0.4$	$\pi = 0.57$	$\pi = 0.75$	$\pi = 0.9$
CRRA	$\gamma = 1$	0.2721	0.2533	0.2413	0.2441
	$\gamma = 2$	0.1989	0.2089	0.2431	0.2857
	$\gamma = 4$	0.0631	0.1268	0.2288	0.3374
LA	$\eta_1 = 2.25$	0.2503	0.2032	0.1531	0.1102
	$\eta_1 = 5$	0.2501	0.2019	0.1528	0.1145
DD	$\eta_2 = 2.25$	0.2742	0.2447	0.2260	0.2321
	$\eta_2 = 5$	0.2360	0.2147	0.2179	0.2583

Table 5.2: Values of λ^* for a piecewise constant and U-shaped portability loss size with $\beta = [0.95 \quad 0.9 \quad 0.99]$ (original $\beta = 0.95$).

In table 5.3 we compute the indifference job switching intensities for a piecewise constant and time-decreasing salary trend. Specifically we assume that the employee's salary growth has the highest trend in the early career $\mu_{S,1} = 2.25\%$ and then this trend decreases gradually in the mid-career ($\mu_{S,2} = 1.75\%$) and the late career ($\mu_{S,3} = 1\%$). With a piecewise constant and time-decreasing salary drift we can capture the often, i.e for many workers, empirically observed concave shape of the salary curve, see Blake and Orszag, Chapter 5 (1997). The main results observed in our benchmark case of a constant salary drift still carry over. In addition we see that the higher salary growth rate in the early and mid-career mainly leads to a slight increase in the relative attractiveness of the DB plan. The corresponding economic effects are discussed in the next subsection in figure 5.1.

Utility	Risk aversion	$\pi = 0.4$	$\pi = 0.57$	$\pi = 0.75$	$\pi = 0.9$
CRRA	$\gamma = 1$	0.3500	0.3257	0.3118	0.3148
	$\gamma = 2$	0.2601	0.2733	0.3164	0.3692
	$\gamma = 4$	0.0974	0.1797	0.3068	0.4370
LA	$\eta_1 = 2.25$	0.3247	0.2695	0.2032	0.1506
	$\eta_1 = 5$	0.3195	0.2687	0.2021	0.1595
DD	$\eta_2 = 2.25$	0.3773	0.3324	0.2972	0.2917
	$\eta_2 = 5$	0.3344	0.3050	0.2895	0.3442

Table 5.3: Values of λ^* for a piecewise constant and decreasing salary trend with $\mu_S = [0.0225 \quad 0.0175 \quad 0.01]$ (original $\mu_S = 0.015$).

As a last robustness check we consider a deterministic and time-decreasing job switching intensity λ. It is empirically confirmed that workers change jobs much more frequently when they are younger than when they are older, see for instance Booth, Francesconi and Garcia-Serrano (1997). Accordingly, we set $\lambda = [0.3 \; 0.2 \; 0.1]$. As we have fixed the job switching intensity, we can now not report the indifference job switching intensity. Therefore we consider here the ratio of the certainty equivalents of the DB plan CE^{DB} and that of the DC plan CE^{DC} as the relevant statistic. It has qualitatively the same meaning as the indifference job switching intensity. That is, a ratio above 1 indicates that the DB plan is relatively preferred to the DC plan and the opposite holds if this ratio is less than 1. Table 5.4 again confirms our benchmark results. Specifically, given the best investment strategy is taken, less risk averse power beneficiaries and the DD beneficiaries relatively prefer the DB plan, while the more risk averse power beneficiary and LA utility maximizers would opt for the DC plan.

Utility	Risk aversion	$\pi = 0.4$	$\pi = 0.57$	$\pi = 0.75$	$\pi = 0.9$
CRRA	$\gamma = 1$	1.1707	1.1339	1.1216	1.1185
	$\gamma = 2$	1.0430	1.0615	1.1254	1.2099
	$\gamma = 4$	0.8257	0.9379	1.1188	1.3342
LA	$\eta_1 = 2.25$	1.1179	1.0388	0.9632	0.9088
	$\eta_1 = 5$	1.1105	1.0256	0.9604	0.9150
DD	$\eta_2 = 2.25$	1.2814	1.1812	1.1223	1.1583
	$\eta_2 = 5$	1.2350	1.1616	1.1494	1.3493

Table 5.4: Values for the certainty equivalent ratio $\frac{CE^{DB}}{CE^{DC}}$ for a piecewise constant and decreasing job switching intensity $\lambda = [0.3 \quad 0.2 \quad 0.1]$.

5.4.2 Comparative Statics

In the following subsection we investigate the impact of the crucial contract parameters more closely. To do so we fix our benchmark investment strategy $\pi = 0.57$, and set the risk aversion parameters $\gamma = 2$ and $\eta_1 = \eta_2 = 2.25$. We investigate the impact of the salary process, the career length and the employee's contributions more closely. To better see the effects of the parameters we consider our benchmark case with a constant salary trend, portability loss size and constant job switching intensity.

Figure 5.1 shows values for the indifference job switching intensities λ^* for different levels of the salary drift μ_S. Note first that the employee's contribution rate q will be adjusted according to equation (5.12) for each value of μ_S. We observe the standard result in the literature that an increase in the salary drift, i.e salary growth rate, makes the DB plan more attractive, see e.g Coco and Lopes (2011). The salary drift has a higher impact on the relative attractiveness of the DB plan for the two utility functions with loss aversion. Intuitively, the impact of the salary drift depends on 3 effects in our model. First, for any utility function a higher salary drift implies that it becomes more likely that the beneficiary will receive a higher final salary, which leads to a higher retirement benefit in the DB plan. Second, a higher salary drift also leads to higher contributions in absolute terms in the DC plan. Third, the comparison matching condition (5.12) implies that the matched contribution increases with an increase in the salary drift. As the first effect slightly (moderately) dominates the second and third effect the DB plan becomes relatively more attractive for any utility function.

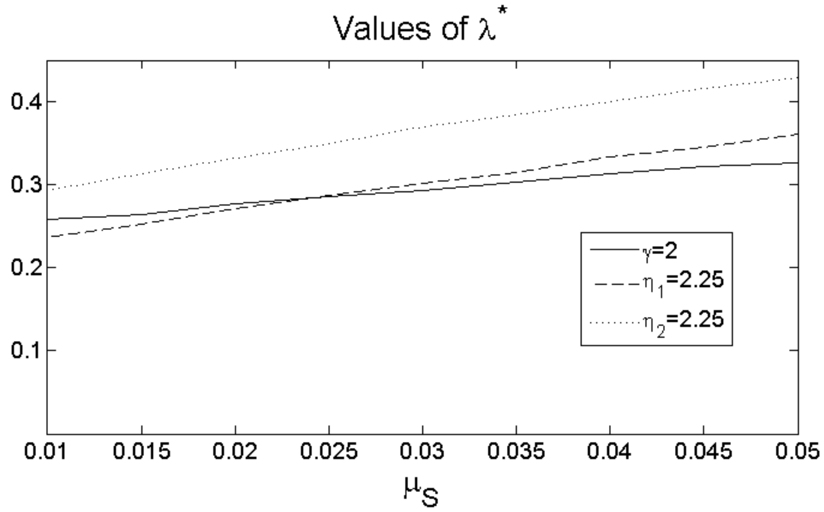

Figure 5.1: Values for the indifference job switching intensity λ^* for different levels of the salary drift μ_S.

Figure 5.2 again shows a standard result in the literature. A higher salary volatility decreases the relative attractiveness of the DB pension plan. This is most pronounced for the power utility function, for the LA utility the salary volatility has a negligible impact while for the DD utility the impact becomes fairly pronounced for higher levels of the salary volatility. Intuitively, in general a higher salary volatility makes the final salary more uncertain, in particular it increases the likelihood of a lower final salary and thus of a lower pension benefit. This effect dominates the effect of more uncertain contributions in the DC plan. More specifically, for the LA utility losses, i.e shortfalls below the target pension income R, are not severely penalized. This implies that the LA beneficiary almost ignores the higher risk of a loss which comes with an increase of the salary volatility, therefore the effect of σ_S is negligible for him. For the DD utility however, losses are penalized more severely, therefore for higher levels of σ_S ($\geq 10\%$) an increase also considerably increases the likelihood that the benefits fall below the target pension income. Accordingly, the relative attractiveness of the DB plan substantially decreases for higher values of the salary volatility. Interestingly, there is a critical volatility level, here $\sigma_S \geq 15\%$, where the DD beneficiary prefers the DC plan more than the LA beneficiary. In other words if the salary risk is very high the DC plan can become more attractive for the more loss averse DD beneficiary.

Comparing figures 5.1 and 5.2 we can state that the evolution of the salary process affects the DB retirements more than the DC pension retirements. This is intuitive as the DB formula is solely based on the final salary whereas the DC pension plan also considerably depends on the investment performance of the financial portfolio.

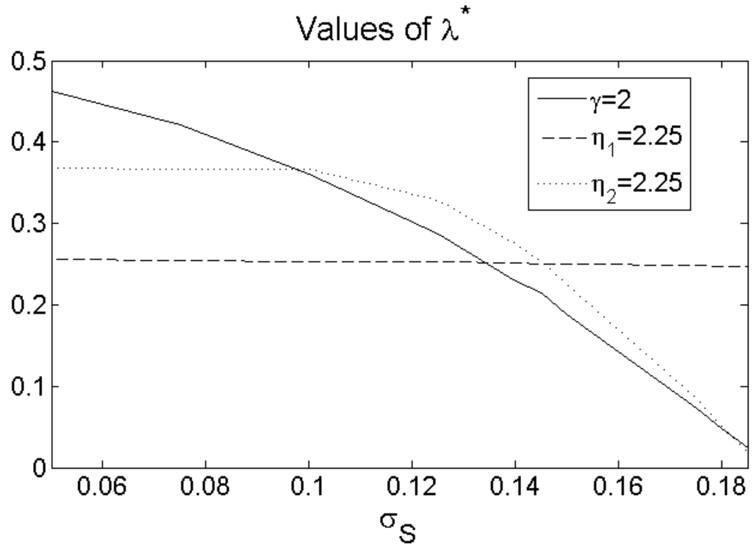

Figure 5.2: Values for the indifference job switching intensity λ^* for different levels of the salary volatility σ_S.

In figure 5.3 we investigate the impact of the career length, or the maturity of the pension retirement contracts, on the relative attractiveness of the two pension retirement plans. A longer (shorter) career length means that the employee enters into the retirement plan when he is younger (older). Again recall that the employee's contribution rate q is computed for each T according to equation (5.12). One clearly sees that the longer the career length of the employee, or the longer the maturity of the pension plan is, the considerably more attractive the DC plan becomes for any utility function. Intuitively, for the older worker the DB plan is considerably more attractive since the overall portability loss is also substantially lower. On the other hand, the DC plan is less attractive because the employee has less time to benefit from the equity premium and to contribute sufficient funds. These two effects substantially dominate the effect that through the contribution matching condition (5.12) the matched contribution rate decreases with the contract maturity. The line of reasoning reverts for the younger employee.

Finally, we investigate how a change in the employee's contribution rate q affects the relative attractiveness of the two pension retirement plans. Therefore, we revert equation (5.12) and compute for each level of q the corresponding employee's replacement rate α^{EE}. α^{EE} increases in q for given μ_S. We observe that the DC plan becomes substantially more attractive with an increasing employee contribution rate q for any utility function as the effect of a higher contribution dominates the effect of a higher replacement rate. This is particularly the case because the employer contribution also increases with the employee contribution in the DC pension plan, which is due to the matching mechanism.

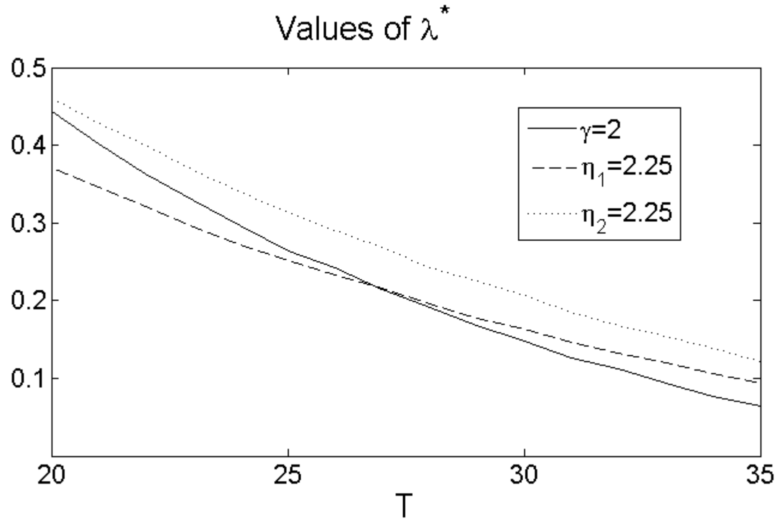

Figure 5.3: Values for the indifference job switching intensity λ^* for different levels of the career length T.

Most interestingly, we see that for contribution rates which converge to 0 the DB pension plan is more attractive for any reasonable number of average job moves, while for higher contributions rates, $q \geq 8.5\%$, the DC plan is always preferred in expected utility terms for any utility function. This result is in line with Samwick and Skinner (2004) who emphasize that primarily inadequate contributions lead to retirement incomes which are on average lower than the DB counterparts, while adequate contribution rates result in a higher median pension income under the DC plan. Samwick and Skinner (1997) even mention that a large number of workers eligible for a DC plan fail to contribute. This figure nicely illustrates that for these workers the DB plan is the better pension plan regardless of their preferences.

5.5 Conclusion

The present chapter models the most important properties from a representative beneficiaries perspective in DB and DC plans: salary risk present in both the DB and DC plan, portability losses in DB plans due to job switchings, and asset price risk born in DC plans. We make comparisons between DB and DC plans by analyzing the expected utility of the pension beneficiary under three preferences: power utility, mean-shortfall and mean-downside deviation preferences. Most of our findings are consistent with the existing literature. Independent of the preferences, the attractiveness of DB plan increases in the salary growth rate and decreases in the salary volatility and the contract maturity. Our model further indicates that portability losses considerably reduce the relative attractiveness of the DB plan. Moreover, we show that for the utility functions with the loss aversion property, a mean-downside deviation beneficiary prefers the DB plan in most cases relatively

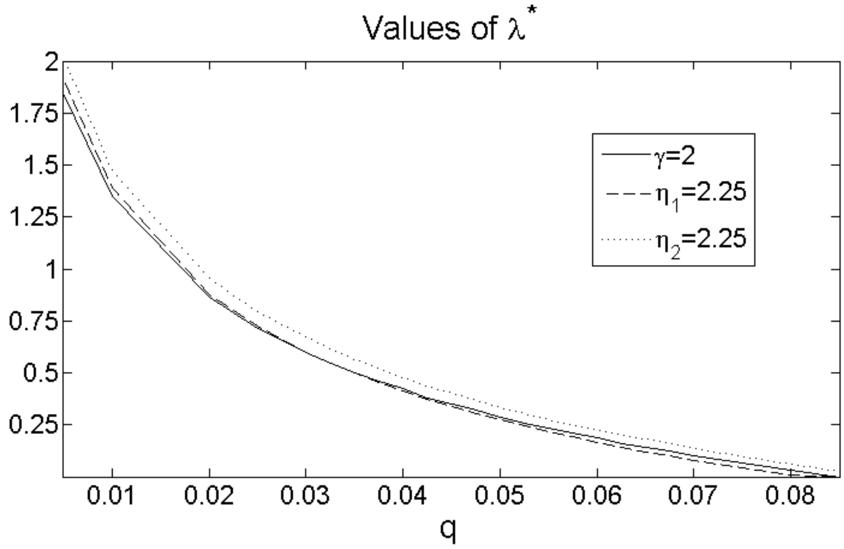

Figure 5.4: Values for the indifference job switching intensity λ^* for different levels of the employee's contribution rates q.

more than the mean-shortfall beneficiary. Finally we have a result, which is inconsistent with existing findings. We find that the attractiveness of the DB plan can decrease in the level of risk aversion. It is justified by the fact that the disutility caused by the portability loss (jump risk) can be particularly severe for very risk averse beneficiaries.

This chapter can be extended by relaxing several assumptions. First, one could also include endogenous or strategic job moves and unemployment in our setup by allowing the salary process to have jumps. Second, the portability risk could be also modeled more realistically by specifying the pension income at retirement as $B(T) = \sum_{i=1}^{N(T)} \tau_i \, S(\tau_i)$, where τ_i denotes the time the employee has worked for employer i. Then the portability losses could be defined as the difference of a pension income without job moves and the pension income with job moves, i.e $T\,S(T) - \sum_{i=1}^{N(T)} \tau_i \, S(\tau_i)$. In this framework we would link the portability risk to the salary risk and thus better capture the major type of portability losses, the cash equivalent losses. Finally, one could allow the beneficiary to have a combination of both, a DC and DB pension plan, or to change the pension plan at some time in his career.

5.6 Appendix: Derivation of Proposition 4.3.1

a) Power Utility: $u(x) = \frac{x^{1-\gamma}}{1-\gamma}$

The expected utility of the power utility is computed by using the independence between W^S and the increments of the Poisson process $N(t_j) - N(t_{j-1})$, $j = 1, ..., J$. [12]. This implies that the expectation in (5.6) factors, thus we can write the expected utility as

$$\mathbb{E}[u(B_T^{DB})] = \frac{1}{1-\gamma}\,(a(T)\,\alpha\, s)^{1-\gamma}\,\mathbb{E}\left[\exp\left\{(1-\gamma)\,(\sum_{j=1}^{J}\mu_{S,j}(t_j - t_{j-1}) - \frac{1}{2}\sigma_S^2\,T) + \sigma_S\,W_T^S\right\}\right]$$
$$\times\ \mathbb{E}\big[\exp\left\{\sum_{j=1}^{J}(N(t_j) - N(t_{j-1}))\,(1-\gamma)\,\ln(\beta_j)\right\}\big]$$

To evaluate the first expectation we just use that the exponent is normally distributed to obtain

$$\mathbb{E}\left[\exp\left\{(1-\gamma)\,(\sum_{j=1}^{J}\mu_{S,j}(t_j - t_{j-1}) - \frac{1}{2}\sigma_S^2\,T) + \sigma_S\,W_T^S\right\}\right]$$
$$= \exp\left\{(1-\gamma)\,(\sum_{j=1}^{J}\mu_{S,j}(t_j - t_{j-1}) - \frac{1}{2}\sigma_S^2\,T) + \frac{1}{2}(1-\gamma)^2\,\sigma_S^2\,T\right\}.$$

Next one can show that

$$\mathbb{E}[e^{c_j\,(N(t_j) - N(t_{j-1}))}] = \exp\left\{\lambda_j\,(t_j - t_{j-1})\,(e^{c_j} - 1)\right\},$$

for any time $t_j \geq 0$ and any piecewise constant c_j, $j = 1, ...J$. Collecting the last two expectations one ends up with (5.18).

b) Mean-Shortfall: $u(x) = x - R$ for $x \geq R$ and $u(x) = -\eta_1(R - x)$ for $x < R$.

To compute the expected utility for the mean shortfall we mainly use the law of iterated expectations, to first condition on the number of job moves $N(t_j)$ in every career period $j = 1, ...J$ and compute the standard Black Scholes expectation. In the second step we derive the joint distribution of all job moves and evaluate the outer expectation. First we have

$$\mathbb{E}[u(B_T^{DB})] = \mathbb{E}\big[\mathbb{E}[1_{\{B_T^{DB} \geq R\}}\,B_T^{DB}|N]\big]$$
$$-\ \mathbb{E}\big[\mathbb{E}[1_{\{B_T^{DB} < R\}}\,\eta_1\,(R - B_T^{DB})|N]\big]$$

Conditioned on the number of job moves being k_j up to career period j, where k_j is an increasing sequence, the two conditional expectations are standard Black Scholes integrals,

[12] Note that the independence of W^S and the Poisson process N immediately implies this independence.

therefore we have

$$\begin{aligned}
&\mathbb{E}[1_{\{B_T^{DB}\geq R\}}\,(B_T^{DB}-R)|N(t_1)=k_1,...N(t_J)=k_J]\\
=&\alpha\, s\, a(T)\, \exp\left\{\sum_{j=1}^{J}\mu_{S,j}(t_j-t_{j-1})\right\}\prod_{j=1}^{J}\beta_j^{(k_j-k_{j-1})}\,\Phi(d_1(k_1,...,k_J))-R\Phi(d_2(k_1,...,k_J));\\
&\mathbb{E}[1_{\{B_T^{DB}<R\}}\eta_1\,(R-B_T^{DB})|N(t_1)=k_1,...,N(t_J)=k_J]\\
=&-\eta_1\, R\,\Phi(-d_2(k_1,...,k_J))+\eta_1\,\alpha\, s\, a(T)\,\exp\left\{\sum_{j=1}^{J}\mu_{S,j}(t_j-t_{j-1})\right\}\\
&\cdot\prod_{j=1}^{J}\beta_j^{(k_j-k_{j-1})}\,\Phi(-d_1(k_1,...,k_J));\\
d_1(k_1,...,k_J)=&\frac{\ln\frac{\alpha s a(T)\prod_{j=1}^{J}\beta_j^{(k_j-k_{j-1})}}{R}+(\sum_{j=1}^{J}\mu_{S,j}(t_j-t_{j-1})+\frac{1}{2}\sigma_S^2\,T)}{\sigma_S\sqrt{T}},\\
d_2(k_1,...k_J)=&d_1(k_1,...,k_J)-\sigma_S\sqrt{T}.
\end{aligned}$$

where Φ denotes the standard normal cdf.

Next we compute the joint distribution of the number of jumps in each career period as

$$\begin{aligned}
&P(N(t_1)=k_1,...,N(t_J)=k_J)\\
=&P(N(t_1)=k_1,N(t_2)-N(t_1)=k_2-k_1,...,N(t_J)-N(t_{J-1})=k_J-k_{J-1})\\
=&\prod_{j=1}^{J}P(N(t_j)-N(t_{j-1})=k_j-k_{j-1})\\
=&\prod_{j=1}^{J}\frac{\lambda_j\,(t_j-t_{j-1})^{(k_j-k_{j-1})}}{(k_j-k_{j-1})!}\,e^{\lambda_j\,(t_j-t_{j-1})}.
\end{aligned}$$

In the second step we have rewritten the number of jumps up to time point j in terms of its increments. In the third equation we have then used the independence of the Poisson process increments. In the last step we have used the law of a (in)homogeneous Poisson process with a piecewise constant intensity.

Finally we use the distribution of the number of jumps in each career period and evaluate the outer expectation by integrating over the range of possible jumps in each career period and obtain

$$
\begin{aligned}
&\mathbb{E}[u(B_T^{DB})] \\
=&\sum_{k_1=0}^{\infty}\sum_{k_2=k_1}^{\infty}\cdots\sum_{k_J=k_{J-1}}^{\infty}\Bigg[\alpha\, s\, a(T)\, e^{\sum_{j=1}^{J}\mu_{S,j}\,(t_j-t_{j-1})}\prod_{j=1}^{J}\beta_j^{(k_j-k_{j-1})}\,\Phi(d_1(k_1,...,k_J)) - R\Phi(d_2(k_1,...,k_J)) \\
&-\eta_2\, R^2\,\Phi(-d_2(k_1,...,k_J)) + 2\eta_2 R\alpha\, s\, a(T)\, e^{\sum_{j=1}^{J}\mu_{S,j}\,(t_j-t_{j-1})}\prod_{j=1}^{J}\beta_j^{(k_j-k_j-1)}\,\Phi(-d_1(k_1,...,k_J)) \\
&-\eta_2(\alpha\, s\, a(T)\,)^2 e^{(2\sum_{j=1}^{J}\mu_{S,j}\,(t_j-t_{j-1})+\sigma_S^2\,T)}\prod_{j=1}^{J}\beta_j^{2\,(k_j-k_{j-1})}\Phi(-d_3(k_1,...,k_J))\Bigg] \\
&\times\prod_{j=1}^{J}\frac{\lambda_j\,(t_j-t_{j-1})^{(k_j-k_{j-1})}}{(k_j-k_{j-1})!}\,e^{-\lambda_j\,(t_j-t_{j-1})}.
\end{aligned}
$$

c) Mean-Downside Deviation: $u(x)=x$ for $x\geq R$ and $u(x)=-\eta_2(R-x)^2$ for $x<R$

To compute the expected utility for the mean-downside deviation utility we simply need to compute the additional conditional expectation

$$
\begin{aligned}
&\mathbb{E}[1_{\{B_T^{DB}<R\}}\,(B_T^{DB})^2|N(t_1)=k_1,...,N(t_J)=k_J] \\
=&-\eta_2(\alpha\, s\, a(T)\,)^2 e^{(2\sum_{j=1}^{J}\mu_{S,j}\,(t_j-t_{j-1})+\sigma_S^2\,T)}\prod_{j=1}^{J}\beta_j^{2\,(k_j-k_{j-1})}\Phi(-d_3(k_1,...,k_J));
\end{aligned}
$$

$$
d_3(k_1,...,k_J)=\frac{\ln\frac{\alpha\, s\, a(T)\,\prod_{j=1}^{J}\beta_j^{(k_j-k_{j-1})}}{R}+(\sum_{j=1}^{J}\mu_{S,j}(t_j-t_{j-1})+\frac{3}{2}\sigma_S^2\,T)}{\sigma_S\sqrt{T}}.
$$

The other conditional expectations almost carry over from the ones for the mean-shortfall utility, accordingly we eventually have

$$
\begin{aligned}
&\mathbb{E}[u(B_T^{DB})] \\
=&\sum_{k_1=0}^{\infty}\sum_{k_2=k_1}^{\infty}\cdots\sum_{k_J=k_{J-1}}^{\infty}\Bigg[\alpha\, s\, a(T)\, e^{\sum_{j=1}^{J}\mu_{S,j}\,(t_j-t_{j-1})}\prod_{j=1}^{J}\beta_j^{(k_j-k_{j-1})}\,\Phi(d_1(k_1,...,k_J)) - R\Phi(d_2(k_1,...,k_J)) \\
&-\eta_2\, R^2\,\Phi(-d_2(k_1,...,k_J)) + 2\eta_2 R\alpha\, s\, a(T)\, e^{\sum_{j=1}^{J}\mu_{S,j}\,(t_j-t_{j-1})}\prod_{j=1}^{J}\beta_j^{(k_j-k_j-1)}\,\Phi(-d_1(k_1,...,k_J)) \\
&-\eta_2(\alpha\, s\, a(T)\,)^2 e^{(2\sum_{j=1}^{J}\mu_{S,j}\,(t_j-t_{j-1})+\sigma_S^2\,T)}\prod_{j=1}^{J}\beta_j^{2\,(k_j-k_{j-1})}\Phi(-d_3(k_1,...,k_J))\Bigg] \\
&\times\prod_{j=1}^{J}\frac{\lambda_j\,(t_j-t_{j-1})^{(k_j-k_{j-1})}}{(k_j-k_{j-1})!}\,e^{-\lambda_j\,(t_j-t_{j-1})}.
\end{aligned}
$$

Chapter 6

Concluding Remarks

In this dissertation we study three very important types of insurance, equity-linked life, pension insurance and the insurance provided by (pension) insurance guarantee funds.

When dealing with the pricing problem of equity-linked life insurance contracts with surrender guarantees one has to take different sources of risk, as the financial risks like the interest or volatility risk and insurance risks like mortality or longevity risk, into account. In addition one has to model the surrender behavior of the policyholder. Concerning the last point it is important to allow for both exogenous and endogenous surrender. Due to the long maturities of these contracts, it is also important to capture the economic changes which affect both the dynamics of the underlying financial portfolio and the surrender behavior of the policyholder. Our regime-switching rational expectation model in chapter 2, where the regimes represent economic states, captures all these points and hence extends the existing literature on the pricing of surrender options in equity-linked life insurance. The main modeling contribution is that exogenous and endogenous surrender are linked to the economic states. Specifically, we show that the economic state has a significant impact on the contract and particularly surrender option value and that the American style surrender model, which relies on the assumption that policyholders behave monetary optimal, substantially overestimates the surrender option value in any economic state.

Beside the modeling of the financial risk factors more realistically in a specific stochastic model, one could also extend our regime-switching framework, which relies on two simplifying assumptions: firstly, there are no jumps in the financial portfolio value when the regime switches and secondly we can exactly observe the regime-switches. Dealing with the first extension one would have to solve a coupled nonlinear system of partial integro differential equations (PIDE's). For the second extension one would assume that the economic regimes follow a hidden markov chain (HMM) and use filtering techniques to back out the current economic regime.

Insurance guarantee funds, which exist for many insurance types and in many countries, have the shortcoming that their premium charges are not (sufficiently) risk-based. In chapter 3 we consider the largest pension guarantee fund, the US PBGC pension insurance, and derive a risk-based premium in a contingent claim distress termination model. The distress termination is the most common type of termination where the pension fund is terminated by the sponsoring company due to its own financial distress. Apart from providing a formal distress termination model with a closed-form solution for the risk-based premium we also present an empirical analysis. With a dataset of the largest 100 US DB plan sponsors, we illustrate our theoretical pricing formula and show that the premiums paid to the PBGC differ substantially once sponsor and pension fund specific risks are taken into consideration.

As our regime-switching model in chapter 2, our distress termination model could also be extended by modeling further financial risks stochastically. For instance a simple Vasicek model for the short rate would probably still allow for at least semi-closed form solutions. A more interesting theoretical extension would be to incorporate the variable deficit type contributions DB sponsoring companies make into the pension fund and to study their impact on the risk-based premium paid to the pension insurance guarantee fund. It would be also very interesting to extend our empirical part. One could infer further interesting statistics as the shortfall probability of the sponsoring companies and relate this to the premiums they should pay to the PBGC according to our model. In addition, one could also perform a more accurate econometric analysis with the data.

Although a risk-based premium for pension guarantee funds is very appealing from the economic perspective and hence the corresponding authorities are encouraged to introduce such a premium calculation, the law in many countries gives pension guarantee funds only the possibility to intervene by closing underfunded pension funds. Accordingly, optimizing the involuntary termination mechanism is the only means pension guarantee funds have thus far to protect the employees in insured DB pension plans. In chapter 4 we derive the optimal timing of intervention in terms of a critical funding ratio of an insured DB pension plan, that is the funding ratio when the pension guarantee fund prematurely terminates the underfunded pension plan and activates its financial guarantee. To this end, we assume that the pension guarantee fund represents the interests of the policyholders and maximizes their expected utilities subject to two constraints. The pension guarantee fund controls the shortfall probability of insured DB plans and the expected losses of underfunded but not terminated pension plans. By controlling these quantities the pension guarantee fund can better manage its financial guarantee and thus additionally protect the employees. The two main qualitative results of our analysis are: firstly, a premature termination is not beneficial for risk neutral and less risk averse beneficiaries when only the SPC is considered, whereas a premature termination rule is always applicable regardless of the risk aversion of the beneficiaries when the ESC is included. Secondly, additional regulation in terms of adding the ESC leads to disutility for risk neutral and less risk averse beneficiaries, while more risk averse beneficiaries are not harmed by this additional regulation.

To compare (private) DB and DC pension plans one has to model a specific tradeoff since the risks the employee bears differ or are not equally pronounced in the two pension plans. It is also important to model the preferences of the policyholder. In chapter 5 we perform a continuous-time expected utility comparison with different preferences in a model with stochastic wages, stochastic job moving and stochastic asset prices. Our modeling framework takes the driving risk factors into account, that is salary risk in both the DB and the DC plan, portability risk in the DB plan and investment risk in the DC plan. In practice one has observed a rapid demise of DB plans in the last two decades. Our findings that the average portability losses significantly reduce the relative attractiveness of DB plans and that DC pension plans offer better diversification through the participation in the investment portfolio account for this phenomenon. However, we also find that there are some cases where the DB plan is the more appropriate pension plan. This is particularly the case if the employee does not make sufficient contributons into the DC plan or an employee with a not very volatile salary is fairly loss averse.

The last chapter can be extended by following Blake and Orszag (1997) and model portability losses in our framework more realistically taking the two types of portability losses, the cash equivalent losses and the backloading losses, better into account. Moreover one could allow the policyholder to switch between the different pension contracts and for instance compute the utility loss or gain if the employee switches from a DB to a DC pension plan somewhere in the mid-career. Another interesting research question would be to compare the standard DB and DC plan with hybrid pension plans like the cash balance plan. The share of such hybrid contracts has grown in the last years and therefore it would be intersting to illustrate some advantages these plans have over a standard DB plan.

Bibliography

Albizzati, M. and Geman, G.: 1994, Interest rate risk management and valuation of the surrender option in life insurance policies, *The Journal of Risk and Insurance* **61**, 616–637.

Ang, A. and Bekaert, G.: 2002a, Regime switches in interest rates, *Journal of Business and Economic Statistics* **20**, 163–182.

Ang, A. and Bekaert, G.: 2002b, Short rate nonlinearities and regime switches, *Journal of Business and Economic Statistics* **26**, 1243–1274.

Archarya, S. and Dreyfus, J.: 1989, Optimal bank reorganization policies and the pricing of federal deposit insurance, *The Journal of Finance* **44**, 1313–1333.

Bacinello, A.: 2003, Pricing guaranteed life insurance participating policies with annual premiums and surrender option, *North American Actuarial Journal* **37**, 1–17.

Bacinello, A.: 2005, Endogenous model of surrender conditions in equity linked life insurance, *Insurance: Mathematics and Economics* **37**, 270–296.

Benartzi, S. and Thaler, R.: 2005, Myopic loss aversion and the equity premium puzzle, *The Quarterly Journal of Economics* **110**(1), 73–92.

Bernard, C. and Lemieux, C.: 2008, Fast simulation of equity-linked life insurance contracts with a surrender option, *Proceeding of the 2008 Winter Simulation Conference* .

Bielecki, T. and Rutkowski, M.: 2004, *Credit risk: Modelling, Valuation and Hedging*, Springer.

Blake, D.: 2000, Does it matter what type of pension scheme you have?, *The Economic Journal* **110**(2), F46–F81.

Blake, D., Cairns, A. and Wood, K.: 2001, Pensionmetrics: Stochastic pension plan design and value-at-risk during the accumulation phase, *Insurance: Mathematics and Economics* **29**(1), 187–215.

Blake, D. and Orszag, J.: 1997, Portability and preservation of pension rights in the united kingdom, *Report of the Director General's Inquiry into Pensions* **3**.

Blanchard, O. and Watson, M.: 1986, Are business cycles all alike?, *In Robert Gordon, Ed. The American Business Cycle.,* pp. 123–182.

Blome, S., Fachinger, K., Franzen, D., Scheunstuhl, G. and Yermo, J.: 2007, Pension fund regulation and risk management: Results from an alm optimisation exercise, *OECD Working Papers on Insurance and Private Pensions* **8**.

Bodie, Z.: 2006, On asset and liability matching and federal deposit and pension insurance, *Federal Reserve Bank of St.Louis Review* **88**(4), 323–329.

Bodie, Z., Light, J. O., Morck, R. and Taggart Jr., R.: 1987, Funding and asset allocation in corporate pension plans: An empirical investigation, *Issues in Pension Economics, University of Chicago* **3**.

Bodie, Z., Marcus, A. J. and Merton, R. C.: 1986, Defined benefit versus defined contribution pension plans: What are the real tradeoffs?, *NBER Working paper No. 1719* .

Boender, C.: 1997, A hybrid simulation/optimisation scenario model for asset/liability management, *European Journal of Operational Research* **99**(1), 126–135.

Booth, A., Francesconi, M. and Garcia-Serrano, C.: 1997, Job tenure: Does history matter? Working paper.

Broadbent, J., Palumbo, M. and Woodman, E.: 2006, The shift from defined benefit to defined contribution pension plans - implications for asset allocation and risk management. Working Paper.

Broeders, D. and Chen, A.: 2013, Pension benefit security: A comparison of solvency requirements, a pension guarantee fund and sponsor support, *Journal of Risk and Insurance* **80**(2), 239–272.

Brown, J.: 2008, Guaranteed trouble: The economic effects of the pension benefit guaranty corporation, *Journal of Economic Perspectives* **22**(1), 177–198.

Buffington, J. and Elliot, J.: 2002, American options with regime-switching, *International Journal of Theoretical and Applied Finance* **5**, 497–514.

Campbell, J. and Viceira, L.: 2002, Strategic asset allocation: Portfolio choice for long-term investors, *Oxford University Press, Oxford, UK* .

Chen, A.: 2011, A risk-based model for the valuation of pension insurance, *Insurance: Mathematics and Economics* **49**(3), 401–409.

Chen, A., Ferris, K. and Hsieh, S.: 1994, The valuation of pbgc insurance using an option pricing model, *Journal of Financial and Quantitative Analysis* **29**, 89–99.

Chen, A. and Uzelac, F.: 2014, A risk-based premium: What does it mean for db plan sponsors?, *Insurance: Mathematics and Economics* **54**, 1–11.

Cheng, C. and Uzelac, F.: 2014, Premature termination rule for defined benefit pension guarantee. Working paper.

Coco, J. and Lopes, P.: 2011, Defined benefit or defined contribution? a study of pension choices, *Journal of Risk and Insurance* **78**(4), 931–960.

Cooper, R. and Ross, T.: 2003, Protecting underfunded pensions: The role of guarantee funds, *Journal of Pension Economics and Finance* **2**(3), 247–272.

Cummins, J.: 1988, Risk-based premiums for insurance guarantee funds, *Journal of Finance* **43**(4), 823–839.

Dai, M.and Kwok, Y. and You, H.: 2007, Intensity-based framework and penalty formulation of optimal stopping problems, *Journal of Economic Dynamics & Control* **31**, 3860–3880.

Dar, A. and Dodds, C.: 1989, Interest rates, the emergency fund hypothesis and saving through endowment policies: Some empirical evidence for the u.k, *Journal of Risk and Insurance* **56**, 415–433.

Davis, S. and Willen, P.: 2000, Using financial assets to hedge labor income risks: Estimating the benefits. Working Paper.

De Giovanni, D.: 2010, Lapse rate modeling: A rational expectation approach, *Scandinavian Actuarial Journal* **1**, 56–67.

Doff, R.: 2008, A critical analysis of the solvency ii proposals, *The Geneva Papers* **33**, 193–206.

Duan, J. and Yu, M.: 2005, Fair insurance guarantee premia in the presence of risk-based capital regulations, stochastic interest rate and catastrophe risk, *Journal of Banking&Finance* **29**(10), 2435–2454.

EIOPA: 2011, Eiopa report on the fifth quantitative impact study (qis5) for solvency ii. brussesls.

Elliot, R., Aggoun, L. and Moore, J.: 1994, *Hidden Markov Models- Estimation and Control*, 3 edn, Springer.

Elliot, R., Chan, L. and Siu, T.: 2005, Option pricing and esscher transform under regime switching, *Annals of Finance* **1**, 423–432.

Engle, R., Ghysels, E. and Sohn, B.: 2008, On the economic sources of stock market volatility. Working paper.

Erlwein, C., Mamon, R. and Siu, T.: 2008, The pricing of credit default swaps under a markov-modulated merton's structural model, *North American Actuarial Journal* **12**, 19–46.

Eychenne, K., Martinetti, S. and Roncalli, T.: 2011, Strategic asset allocation and long-term-investment, *Research & Development Lyxor* **6**.

Fama, E. and French, K.: 1988, Dividend yields and expected stock returns, *Journal of Financial Economics* **22**, 3–25.

Forsyth, P. and Vetzal, K.: 2002, Quadratic convergence for valuing american options using a penalty method, *SIAM Journal on Scientific Computing* **23**, 2095–2122.

Foster, A.: 1994, Portability of pension benefits among jobs, *Monthly Labor Review, Bureau of Labor Statistics* .

Gaurav, A. and Kat, H.: 2003, Stocks, bonds and hedge funds: Not a free lunch!, *The Journal of Portfolio Management* **29**(4), 113–120.

Grosen, A. and Jorgensen, P.: 2000, Fair valuation of life insurance liabilities: The impact of interest rate guarantees, surrender options, and bonus policies, *Insurance: Mathematics and Economics* **26**, 37–57.

Guo, X.: 2001, Information and option pricing, *Quantitative Finance* **1**, 38–44.

Hall, E.: 1982, The importance of lifetime jobs in the u.s. economy, *American Economic Review* **9**, 716–724.

Han, L., Lai, G. and Witt, R.: 1997, A financial-economic evaluation of insurance guaranty fund system: An agency cost perspective, *Journal of Banking and Finance* **21**, 1107–1129.

Hardy, M.: 2001, A regime-switching model of long-term stock returns, *North American Actuarial Journal* **5**, 41–53.

Haug, E.: 2007, *The complete Guide to Option Pricing Formulas*, 2 edn, McGraw–Hill.

Jouini, E. and Napp, C.: 2011, Unbiased disagreement in financial markets, waves of pessimism and the risk-return tradeoff, *Review of Finance* **15**, 575–601.

Kahneman, D. and Tversky, A.: 1979, Prospect theory: An analysis of decision under risk, *Econometrica* **47**(2), 263–292.

Kalra, R. and Jain, G.: 1997, A continuous-time model to determine the intervention policy for pbgc, *Journal of Banking and Finance* **40**(1), 1159–1177.

King, R. and Rebelo, S.: 1999, Resuscitating real business cycles, *Handbook of Macroeco nomics* **2**(1), 927–1007.

Kuo, W., Tsai, C. and Chen, W.: 2003, An empirical study on the lapse rate: The cointegration approach, *Journal of Risk and Insurance* **70**, 489–508.

Lewis, C. and Pennacchi, C.: 1994, The value of pension benefit guaranty corporation insurance, *Journal of Money, Credit and Banking* **26**(3), 735–753.

Li, J. and Szimayer, A.: 2014, The effect of policyholder's rationality on unit-linked life insurance contracts with surrender guarantees, *Quantitative Finance* **2**, 327–342.

Love, D., Smith, P. and Wilcox, D.: 2009, Should risky firms offer risk-free db pensions?, *Federal Reserve Board* .

Marcus, A.: 1987, Corporate pension policy and the value of pbgc insurance, *Issues in Pension Economics, University of Chicago Press* .

Martellini, L. and Milhau, V.: 2008, How costly is regulatory short-termism for defined benefit pension funds? Working paper.

Palmer, B.: 2006, Equity-indexed annuities: Fundamental concepts and issues. Working paper.

Poterba, J., Rauh, J., Venti, S. and Wise, D.: 2007, Defined contribution plans, defined benefit plans, and the accumulation of retirement wealth, *Journal of Public Economics* **91**(10), 2062–2086.

Romaniuk, K.: 2011, The optimal investment policy for the pension benefit guarantee corporation. Working paper.

Rymaszewski, P., Schmeisser, H. and Wagner, J.: 2012, Under what conditions is an insurance guaranty fund beneficial for policyholders?, *Journal of Risk and Insurance* **79**(3), 785–815.

Samwick, A. and Skinner, J.: 1997, Abandoning the nest egg? 401(k) plans and inadequate pension saving. Working Paper.

Samwick, A. and Skinner, J.: 2004, How will 401(k) pension plans affect retirement income?, *The American Economic Review* **94**(1), 329–343.

Schmeisser, H. and Wagner, J.: 2013, The impact of introducing insurance guarantee schemes on pricing and capital structure, *Journal of Risk and Insurance* **80**(2), 834–855.

Schrager, A.: 2009, The decline of defined benefit plans and job tenure, *Journal of Pension Economics and Finance* **8**(3), 259–290.

Schultz, E.: 1995, Offered a lump sum, many retirees blow it and risk their future!, *The Wall Street Journal* **A1**.

Schwert, G.: 1989, Why does stock market volatility change over time?, *Journal of Finance.* **44**, 1115–1153.

Sharpe, W.: 1976, Corporate pension funding policy, *Journal of Financial Economics* **3**(3), 183–193.

Shi, Z. and Werker, J.: 2012, Short horizon regulation for long-term investors, *Journal of Banking and Finance* **36**, 3227–3238.

Shreve, S.: 2004, *Stochastic Calculus for Finance II: Continuous-time Models*, 1 edn, Springer.

Siegmann, A.: 2011, Minimum funding ratios for defined-benefit pension funds, *Journal of Pension Economics and Finance* **10**(3), 417–434.

Siu, T.: 2005, Fair valuation of participating policies with surrender options and regime-switching., *Insurance: Mathematics and Economics* **37**, 533–552.

Stanton, R.: 1995, Rational prepayments and the valuation of mortgage backed securities, *Review of Financial Studies* **8**, 677–708.

Stewart, F.: 2007, Benefit security pension fund guarantee schemes, *OECD Working Papers on Insurance and Private Pensions, OECD Publishing* **5**.

Topel, R. and Ward, M.: 1992, Job mobility and the careers of young men, *The Quarterly Journal of Economics* **107**(2), 439–479.

Treynor, J.: 1977, The principles of corporate pension finance, *Journal of Finance* **32**, 627–638.

Uzelac, F. and Chen, A.: 2014, Portability, salary and asset price risk: A continuous-time expected utility comparison of db and dc pension plans. Working Paper.

Uzelac, F. and Szimayer, A.: 2014, Valuation of equity-linked life insurance contracts with surrender guarantees in a regime-switching rational expectation model, *Quantitative Finance* **14**(2), 357–368.

Wilcox, D.: 2006, Reforming the defined-benefit pension system in the united states, *Brookings Papers on Economic Activity* **1**.

Yasui, T.: 2001, Policyholder protection funds: Rationale and structure, *Insurance and Private Pensions Compendium for Emerging Economies, Book 1* .

Zhang, A.: 2008, Optimal investment for a pension fund under inflation risk, *Mathematical Methods of Operations Research* **71**(2), 353–369.