Stakeholders in pension finance

Boon, Ling-Ni

Publication date:
2017

Document Version
Publisher's PDF, also known as Version of record

Link to publication in Tilburg University Research Portal

Citation for published version (APA):
Boon, L-N. (2017). *Stakeholders in pension finance*. CentER, Center for Economic Research.

Ling-Ni Boon

Stakeholders in Pension Finance

Stakeholders in Pension Finance

Proefschrift

ter verkrijging van de graad van doctor aan Tilburg University op gezag van de rector magnificus, prof. dr. E.H.L. Aarts, en Université Paris-Dauphine op gezag van de rector magnificus, prof. dr. I. Huault, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van Tilburg University op woensdag 6 september 2017 om 14.00 uur door

Ling-Ni Boon

geboren op 16 december 1987 te Melaka, Maleisië.

Promotiecommissie:

Promotores:	Prof. dr. M. Brière
	Prof. dr. C. Gresse
	Prof. dr. B.J.M. Werker
Overige leden:	Prof. dr. D. Blake
	Prof. dr. A.M.B. De Waegenaere
	Prof. dr. J. Glachant
	Prof. dr. T.E. Nijman

Samenvatting

Deze dissertatie richt zich op financiële vraagstukken van drie partijen op het gebied van pensioen, namelijk de beleidsmaker, de pensioenuitvoerder (bijvoorbeeld een verzekeraar of een pensioenfonds) en de deelnemer. Vanwege zowel het volatiele karakter van financiële markten als demografische ontwikkelingen is een herevaluatie van de rol van deze drie genoemde partijen noodzakelijk. In het bijzonder is er aandacht voor de vraag wie welke risico's draagt. De uitkomst van dit onderzoek omvat advies voor deelnemers over het omgaan met langlevenrisico, een evaluatie van de aantrekkelijkheid van beleggen in langlevenrisico voor investeerders, inzichten in het opstellen van pensioencontracten door verzekeraars, en voorstellen omtrent regelgeving voor beleidsmakers om een toekomstbestendig pensioen te creëren.

Preface

This dissertation examines three stakeholders in pension finance: the individual, the policymaker, and the pension provider (e.g., an insurer or a pension fund). In a setting beset by unforeseen financial market circumstances and demographic changes that disfavor financial security in retirement, a re-evaluation of these stakeholders' role is necessary. We explore the regulation and design of retirement plans by incorporating features that characterize the future retirement landscape, such as the increasing burden of risk borne by the individual, and the potential involvement of market investors in the provision of retirement contracts. The implications of our findings encompass guidance for individuals in managing longevity risk, evaluation of the appeal of longevity risk exposure to investors, insights on contract design for the pension provider, and proposals to the policymaker on regulatory measures that foster a sustainable retirement environment.

Acknowledgements

I am indebted to my advisors Marie Brière, Carole Gresse, and Bas J.M. Werker.

I am humbled by having David Blake, Anja M.B. De Waegenaere, Jérôme Glachant, and Theo E. Nijman as members of the committee.

Ling-Ni Boon
Tilburg, September 2017

Contents

Introduction

Retirement financing is susceptible to financial market fluctuations and evolves with demographic trends. The anomalous economic situation after the financial crisis of 2008, coupled with increasing life expectancies, caused the cost of retirement financing to soar, prompting the stakeholders to re-assess their roles.

The employer, for instance, revises the occupational retirement plan to reduce pension costs. These changes shift the bulk of financial and biometric risks onto the individual, who seeks guidance to manage the risks. The individual's task is further complicated by the setting where a market for some of the risks, such as longevity, is not well developed. Overseeing the transition, the policymaker is occupied with the regulatory framework re-design in order to better align the stakeholders' incentives. This dissertation examines the role of pension finance stakeholders in a changing retirement landscape. Its findings elucidate the part that each stakeholder can take on to surmount retirement financing challenges.

Chapter 1 concerns the policymaker and the employer. It is based on Boon et al. (2017a). The chapter investigates Defined Benefit (DB) pension plan investments in relation to regulatory requirements. In DB plans, employers provide lifelong benefits to their employees and assume most of the risks. The investments of DB plans are central not only because financial market gains boost retirement income, but also because they are a source of capital for the long-term financing of the economy, such as for infrastructure building.

The enticement of attaining high financial market returns to minimize the costs of benefit provision could induce a DB plan to take excessive investment risks. Hence, regulatory intervention is warranted. Evidence on the implication of regulatory requirements on DB investments to-date focuses only on a handful of regulatory constraints, and is United-States-centric. By taking advantage of the heterogeneity of the DB plans' regulatory framework within the United States (US), with Canada and the Netherlands, we estimate that while regulatory requirements statistically significantly influence DB funds' risk-taking, the economic significance of the constraints varies. Risk-based capital requirements and mark-to-market asset valuation are associated to 7% less risky investments. In contrast, a 1% higher liability discount rate is related to only 0.8% more allocation to risky assets on average. Fund characteristics (e.g., proportion of retirees, value of assets) also possess low economic significance in accounting for the variation of risk-taking among DB plans. These insights are valuable to a policymaker who balances the desire to safeguard individuals' retirement finances by modulating the plans' investment risk-taking, and to enable the plans' involvement in long-term projects that are essential to economic growth.

Although DB plans constitute the majority of occupational retirement account in asset value, Defined Contribution (DC) pension plans are becoming prevalent. In DC plans, the employer's main role is pared down to satisfying the statutory rate of contribution to the employee's retirement account. Investment management and retirement benefit provision are conducted by a service provider such as an asset manager or an insurer. This describes the basics of the 401(k) retirement savings plan in the US and the Personal Pension Schemes (PPS) in the United Kingdom. An implication of the DB to DC transition is that individuals no longer have immunity from longevity risk, which has conventionally been borne by the DB plan sponsor.

Chapter 2 explores longevity risk in a DC setup. It is based on Boon et al. (2017b). Longevity risk is the risk of survival probabilities misestimation. Unlike mortality risk, this systematic risk cannot be diversified away by pooling. We contrast two options for individuals to manage the risk: bear the risk as a collective, or offload it to an insurer by purchasing an annuity.

When offering insurance against a systematic risk, the insurer requires capital reserve to limit solvency risk. The reserve can be composed of equity capital solicited from investors, who would be willing to supply capital if compensated by a longevity risk premium. Alternately, the insurer can accumulate capital by pricing the annuity above its best estimate cost, but potentially causing individuals to shun the annuity for its high price. Thus, the insurer has to ensure that it is able to financially reward its equity holders, and is precluded from charging too high of a loading as individuals would rather form a collective scheme. Existing literature omits the equity holder and assumes that the full capital reserve is constituted from the loading. This assumption is inconsistent with estimates indicating that individuals' willingness to pay to insure against longevity risk is considerably lower than the level of capital reserve necessary to provide the contract. We introduce equity holders to reconcile these estimates.

We find that individuals have a slight preference for the collective setup, even when the annuity is priced at cost. Moreover, an insurer who sells zero-loading annuities is unable to offer adequate longevity risk premium to its equity holders. Therefore, collective schemes and an annuity market for longevity-risk-hedge only are unlikely to co-exist. This conclusion is under the assumption that the insurer does not take financial market risk and has no advantage in dealing with longevity risk, such as access to reinsurance or product synergy (e.g., selling life insurance as it has the opposite exposure to longevity risk). Under heightened longevity risk, individuals would only prefer annuity contracts if the insurer curtails its solvency risk.

Even if with respect to longevity risk, a collective solution is preferred, annuities could still prove valuable if they mitigate financial market risks. In Chapter 3, we visit the problem of optimal consumption and investment when there are interest rate and stock market risks. The chapter is based on Boon and Werker (2017). Under a Gaussian, mean-reverting interest rate process and constant equity risk premium, the solution to this problem is known. We present an equivalent formulation of the problem, and apply it to demonstrate that a variable annuity fails to optimally hedge interest rate risk.

The variable annuity's deficiency in hedging interest rate risk entails economically significant welfare losses that can be overcome by refining the contract definition according to the equivalent formulation. Our revelation guides insurers to design contracts that individuals have a greater preference for, thus are willing to pay a higher price for. Additionally, our finding informs the policymaker who oversees the permissible types of contracts. With the proliferation of DC plans, individuals' choice of investments and insurance products are paramount to their welfare in retirement. Encouraging the provision of contracts that improves individuals' odds of successfully financing their retirement would be among the policymaker's many intentions.

A holistic view is indispensable to address the pension challenge. This dissertation illuminates the role of the policymaker, the insurer and individuals, and casts insights onto the beleaguered retirement financing situation. The analyses draw upon past knowledge, and add current trends and innovations to characterize the prospective retirement environment. It is hoped that the discoveries enhance our understanding of the predicament, and that their policy implications could assist with tackling it.

CHAPTER 1

Regulation and Pension Fund Risk-Taking

1.1 Introduction

In recent decades, the regulation of financial institutions exhibits the growing prevalence of risk-based mechanisms. Banks, insurance companies and pension funds are to comply with directives such as Basel II and III, Solvency II, and IORP II, respectively, all of which contain elements of risk estimation. Fervent discussions over the effect of risk-based requirements on the institutions' investments, investment performance, financial market at large, and the possible diversion of investment capital away from long-term projects surround these regulatory implementations (European Commission, 2013; Gatti, 2014).

Binding regulatory constraints shape the behavior of financial institutions, possibly in an undesirable manner (Koehn and Santomero, 1980). For example, a non-risk-based solvency requirement reduces the regulated institution's optimal investment in risky assets under all financial market situations. In contrast, under a risk-based solvency requirement, such as a value-at-risk (VaR) limit, the investor's optimal strategy is to take more risk when the potential losses are biggest (Basak and Shapiro, 2001). This is at odds with the regulatory objective of protecting stakeholders in adverse financial market situations. However, this outcome materializes when the VaR at only the terminal date is

This chapter is based on Boon et al. (2017a).

concerned. When the setting is modified to exemplify VaR compliance checks at intervals that are shorter than the investment horizon, the investor would optimally insure against all losses instead (Shi and Werker, 2012; Chen et al., 2017a). Hence, a regulated investor's optimal behavior possibly depends not only on the specifics of the requirement, but also on the financial market situation.

When it concerns pension funds, the balance between regulating risk without overly constraining investments is particularly pertinent. A fund offering guaranteed nominal or inflation-linked benefits could in theory hedge their liabilities by investing their wealth entirely in nominal or inflation-linked bonds respectively (de Jong, 2008; Bodie, 1990). The positive correlation between equity returns and wage growth (Sundaresan and Zapatero, 1997; Peskin, 2001; Lucas and Zeldes, 2006) could justify exposure to more financial market risk if expected future accruals take salary increases into account. Due to the hike in the value of liabilities arising from longer life expectancies, in addition to historically low interest rates, the sustainability of pension plans is under threat. An appropriate level of investment risk-taking could curb the rising cost of retirement income provision, especially since financial market return is a major in-flow for pension funds (Aglietta et al., 2012; Bikker et al., 2012; Chambers et al., 2012).

Defined Benefit (DB) pension plans are an instructive field of investigation because in contrast to banking and insurance industries, pension regulation is less harmonized globally. The North American and Dutch pension systems not only have distinct regulations, but also underwent legislative changes in the recent past (e.g., Pension Protection Act 2006 in the US, Financial Assessment Framework 2007 in the Netherlands). By relying on the heterogeneity in pension regulation, we assess the relative extent that regulatory measures and plan characteristics explain funds' risky asset allocation.

Existing analyses of DB pension regulation mainly focus on the effect of the liability discount rate. In the US, public pension funds discount future retirement payments with the expected rate of investment return, whereas private funds apply the corporate bond rate. Pennacchi and Rastad (2011) and Andonov et al. (2017) point out that among US public funds, those selecting higher discount rates tend to also choose riskier portfolios.

A related branch of the literature seeks to identify plan characteristics, such as size, maturity, and inflation indexation of benefits, that influence the plan's asset allocation. For example, larger plans possess more alternative investments (Chemla, 2004; de Dreu and Bikker, 2012; Dyck and Pomorski, 2016), whereas plans with a higher share of retirees tend to have more investment risk (Rauh, 2009; Bikker et al., 2012; Andonov et al., 2017).

We estimate that regulatory requirements statistically significantly affect funds' exposure to risky assets, and is associated to larger variation in risk-taking than fund characteristics. Risk-based capital requirements and mark-to-market (MtM) asset valuation are associated to an average 7% lower risky investments, whereas the average investment risk variation that is related to differences in fund characteristics are under 1% (Table 1.2).

Additionally, we find evidence that risk-based solvency requirements are associated to less risk-taking under all financial market situations, but a non-risk-based one is related to lower risky investments only during the financial crisis. Besides corroborating the theoretical postulation of Shi and Werker (2012); Chen et al. (2017a), our inference is consistent with evidence from the insurance industry that shows that insurers' demand for risky assets asset is inversely related to the stringency of binding capital requirements (Becker and Opp, 2014; Becker and Ivashina, 2015).

The chapter is organized as follows. In Section 1.2, we present the pension regulatory framework in the US, Canada and the Netherlands and develop our hypothesis. Section 1.3 describes our data and outlines the methodology. We discuss the major drivers of pension asset allocation in Section 1.4. Section 1.5 concludes.

1.2 Regulatory Evolution and Hypothesis Development

In this section, we review the evolution of pension regulation in the US, Canada and the Netherlands, and discuss the hypotheses concerning the influence of regulation on DB funds' risky asset allocation.

1.2.1 Evolution of Pension Regulation

We classify the regulatory requirements into three categories: (1) investment restrictions, (2) valuation requirements, and (3) funding requirements. Pension funds are subject to rules that differ not only among countries but also among fund types within a country. US and Canadian funds are either public, corporate or industry, determined by whether the participants work in the public or private sector. Industry funds are a grouping of funds for individuals working in a specific industry. In the Netherlands, the distinction of the type of funds is between corporate and industry only, because Dutch public sector workers participate in industry funds. Appendix 1.A describes the regulation in more detail whereas Appendix 1.B presents a summary table.

Canada is the sole country with quantitative investment restrictions. Limits on investments in foreign assets, real estate and Canadian resource property[1] existed until 2010.

MtM valuation of assets and liabilities is mandatory in the Netherlands over the full sample period, but is only introduced at a later date in Canada (2000 for liabilities, 2011 for assets) and in the US (2006, fair value with smoothing is allowed for assets).

Due to the convergence to International Accounting Standard (IAS) 19, plan sponsors in the Netherlands and Canada have been required since 2005 and 2011, respectively, to recognize unfunded liabilities on their balance sheet. US corporate plans have similar balance sheet recognition requirements that became more stringent in 2006. The Governmental Accounting Standards Board (GASB) guides but does not dictate the accounting of US public plan sponsors. Hence, US public plan sponsors are exempt from any recognition requirement.

Full funding has always been required for Canadian funds. For Dutch funds, a 100% funding requirement was decreed in 1999, but US corporate funds only followed suit in 2006. If funds fail to meet the minimum funding requirement, they are required to restore it within a fixed period, called the recovery period, that varies from three years (the Netherlands) to ten years (Canada). There is a trend of more stringent minimum funding requirement and recovery period over the years. As for US public pension plans, there is no minimum funding requirement, hence also no maximum recovery period.

In addition to the minimum funding requirement, the Netherlands is the only country among the three with risk-based capital requirements. The capital

[1] Canadian resource property refers to the right, license or privilege to explore for, drill for or take petroleum, natural gas or related hydrocarbons in Canada. These quantitative investment restrictions are set by a federal regulation (Schedule III to the Federal Pension Benefits Standards Regulations), and adopted by all provinces except Quebec and New Brunswick.

reserve is calibrated such that the plan is able withstand adverse financial situations with reasonable confidence (i.e., the probability that the value of assets falls below the value of liabilities value within a year is under 2.5%).

1.2.2 Hypotheses Development

1.2.2.1 Investment Restrictions

If the investment limits are binding, they lead to lower allocations to those assets.

1.2.2.2 Valuation Requirements

MtM asset valuation, possibly with smoothing, is the contemporary norm in DB regulation. Despite that, the virtue of MtM valuation remains contentious. Proponents argue that it provides a more accurate assessment of the institution's risk profile. Opponents claim that it inflates the volatility of the value of assets and liabilities, due to short-term market fluctuations which do not reflect the fundamental value of assets (Allen and Carletti, 2008; Plantin et al., 2008). Recent works offer more refined perspectives on the implication of MtM valuation that depends on the informative content of market prices (Plantin and Tirole, 2015; Otto and Volpin, 2017). We posit that MtM asset valuation could deter a fund that desires to maintain a stable funding ratio from investing in risky assets.

A complementary regulatory measure to MtM asset valuation is the requirement for sponsors to recognize unfunded liabilities on their balance sheets. Such a rule may compel funds to invest in less risky assets to minimize the sponsor's balance sheet volatility (Amir et al., 2010). Allowing for the smoothing of market values partially alleviates the concern of artificial volatility.

Besides the asset valuation method, the choice of the liability discount rate is also crucial (Brown and Pennacchi, 2016). Funds that are allowed to apply a rate that depends on the riskiness of their investments may be encouraged to invest more in risky assets in order to justify a higher discount rate that make them appear better funded (Brown and Wilcox, 2009; Pennacchi and Rastad, 2011; Novy-Marx, 2013; Andonov et al., 2017).

1.2.2.3 Funding Requirements

Pension plans are typically required to meet a minimum funding requirement. Minimum funding requirement reduces the regulated institutions' optimal level of investment risk if it is binding. Furthermore, the worse the financial market situation, the more an optimal-behaving investor reduces risk-taking (Basak and Shapiro, 2001). Thus, we postulate an inverse relationship between risky asset exposure and a binding minimum funding requirement.

In addition to a minimum funding requirement, Dutch funds face risk-based capital requirements since 2007. In theory, the implication of risk-based capital requirements on the investor's risky asset demand depends on the horizon on which the VaR limit is defined. If the horizon coincides with the investment horizon, and if financial markets are doing well, the VaR-constrained investor takes lower risk than an unconstrained investor. When market conditions deteriorate, the constrained investor would instead take more risk, contrary to the regulator's objective (Basak and Shapiro, 2001). However, if the VaR is calculated on a shorter horizon than the investment horizon (i.e., multiple , then the regulated investor would moderate risk-taking under all financial market situations, for all except in the final period (Shi and Werker, 2012; Chen et al., 2017a). As Dutch pension funds' risk-based capital requirements are determined on an annual horizon, a period much shorter than the funds' long investment horizon, we hypothesize that risk-based capital requirements yield less risk exposure under all market conditions.

1.3 Data Description and Methodology

1.3.1 Data Description

The dataset contains annual information on fund characteristics and asset allocations of 978 funds from seven developed countries for 1990-2011. It is provided by CEM Benchmarking Inc. Funds are predominantly from the US and Canada (59% and 25% of all funds, respectively), whereas the rest are from Australia, the Netherlands, New Zealand, Sweden, and the United Kingdom. We focus on three countries: the US, Canada and the Netherlands.

We include only funds that have observations for all the characteristics we elicit from the data. Our sample comprises 589 funds (377 US, 174 Canadian, and 38 Dutch funds), amounting to 3,687 observations from 1992-2011. As there is no particular reason to suspect that funds joining and exiting a sample early or late would react in a systematically different way to regulation, an unbalanced panel is not a drawback to the analysis.

Our sample is representative of DB funds in these countries. As a percentage of each country's total DB assets in 2011, the value of the funds' assets under management in the sample is 35% in the US, 32% in Canada, and 30% in the Netherlands.[2] Additionally, the data exhibit no evidence of self-reporting bias because the difference between the performance of plans that skip reporting for one year and those that continue reporting is small and not statistically different from zero (Dyck and Pomorski, 2016).

We consider three risky asset classes: (1) equities, (2) risky fixed income (mortgages and high yield), and (3) alternatives (commodities, natural resources, infrastructure, real estate, private equity, hedge funds, and tactical asset allocation).[3] These assets are considered risky because the return from investing in them is more volatile than the yield of a government bond. Our annual measure of the funds' risk exposure is the percentage value of their reported investment holding in all risky assets, relative to the total reported value of assets under management in the same year. This measure of pension risk extends that of Bodie et al. (1987) and Rauh (2009), who use the percentage of fund wealth allocated to equities only.[4]

Table 1.1 presents the summary statistics by country and fund type in 1996 and 2011. US and Canadian public funds more than doubled in size on average. Maturity, measured by the percentage of retired members, increased on average by 37% across all categories of funds. The percentage of inflation-indexed contracts decreased for all but US public and Canadian corporate funds. In both 1996 and 2011, North American funds adopted liability discount rates that were twice as high on average as those of Dutch funds. There is considerable dispersion of financial returns across countries and types of funds. Dutch funds outperformed all other funds on average in 2011, but in 1996, their Canadian counterparts achieved higher average returns.

Figure 1.5.1 indicates that funds' asset allocation show diverging trends. Whereas US and Canadian public funds, as well as US industry funds increased their overall risky asset allocation (by 14.7%, 9% and 11.7% respectively), Canadian corporate and industry funds reduced their risk exposure (by 3% and 2.1% respectively). US corporate funds also decreased their average allocation to

[2] Calculated on data from 2011 and information from the Towers Watson Global Pension Asset Study 2012.

[3] Real estate includes REITs. Private equity is comprised of venture capital, leveraged buyout, diversified private equity, and other private equity. Tactical asset allocation refers to fully funded long-only segregated asset pool dedicated to tactical asset allocation.

[4] Another measure of pension risk is the "pension beta" introduced by Jin et al. (2006), and applied by An et al. (2013); Mohan and Zhang (2014). Due to the lack of information on fund liabilities, we are unable to implement it.

risky assets by 8.3%, but only half the reduction for Dutch corporate and industry funds, which declined 22.4% and 16.3% on average. There is a general trend for North American funds to increase their holdings of alternative assets and risky fixed income over the sample period, whereas Dutch funds maintain largely similar exposure to those asset classes over time (Table 1.1).

1.3.2 Explanatory Variables Construction

We consider two categories of explanatory factors: regulatory requirements and fund characteristics. A summary table of the explanatory variable construction, and the expected sign of the coefficients based on the hypotheses presented in Section 1.2.2 is in Appendix 1.C.

1.3.2.1 Regulatory Requirements

Quantitative Investment Restriction is the sum of (100% - the maximum asset weight permissible in percentages) over all restricted asset classes. The higher the variable, the more stringent is the investment limit. These limits existed only prior to 2010 in Canada, and only for alternative assets.[5] All US and Dutch funds have thus a Quantitative Investment Restriction value of zero.

We define three regulatory variables related to valuation requirements. MtM Asset Valuation is an index measuring the degree of adherence to the unadulterated market value. The base of the index is ternary, as determined by whether MtM valuation is strictly imposed (1), smoothing is allowed (0.5), or further discretion is permitted (0). As valuation can be carried out for either funding or accounting purposes, and since there is a specified rate for each purpose, we first construct separate indicators for each purpose, and set MtM Asset Valuation as the average of the two.[6] For example, US corporate funds use fair value with smoothing for funding purposes but apply market value with smoothing for accounting purposes. The index for funding is 0 while the index for account-

[5] Before 2010, Canada imposed maximum limits on Canadian resource property (15%) and on the total of real estate and Canadian resource property (25%). We consider the global restriction. As our data do not contain sufficient granularity on the geographical location of all investments, we omit the foreign asset investment limit that existed in Canada until 2010.

[6] We grant equal weights to them because there is no evident justification for either budgeting or accounting purposes to be more influential in a fund's investment decision. Estimates using only the index for budgeting purposes are of no material difference. The index for accounting purposes is highly correlated with the Minimum Funding Requirement (i.e., both requirements are adjusted to be more stringent in the same years). When omitting Minimum Funding Requirement, the estimates using only the index for accounting purposes are also similar.

ing is 0.5. Hence, we assign 0.25 to MtM Asset Valuation for US corporate funds.

Liability Discount Rate is the spread between the liability discount rate disclosed by the funds and the ten-year government bond yield of the funds' host country. This definition accounts for the different interest rate levels in the countries.

Recognition of Unfunded Liabilities is an index measuring the degree of disclosure. The higher its value, the more likely it is that a DB plan has unfunded liabilities. It is 1 if the liabilities to be recognized on the sponsor's balance sheet include expected increase in accrued benefits, 0.5 if only accrued benefits are taken into account, and 0 otherwise.

Next, we define three variables to reflect funding requirements. Minimum Funding Requirement is the minimum ratio (in percentages) of the value of assets over the value of liabilities that a fund has to maintain each year. When a fund fails to meet the Minimum Funding Requirement, it is required to restore its funded status within a fixed number of years. The maximum number of years allowed for the fund to catch up on its funding ratio is defined as the Recovery Period. A longer Recovery Period implies greater leniency to underfunded DB plans. This may give funds an incentive to take more risk. Finally, the presence of Risk-based Capital Requirements is accounted for with an indicator that is equal to one when risk-based capital reserves are required.

1.3.2.2 Fund Characteristics

To control for the influence of fund characteristics on investments, we include the percentage of retired members (Maturity), the percentage of inflation-indexed benefits (Inflation Indexation), the value of assets under management in billions of US dollar (Size), and the funds' investment return in the previous year (Past Investment Return) as explanatory variables.

Funds with fewer retired members, and those that offer more inflation-indexed contracts, have more incentive to take investment risks (Lucas and Zeldes, 2006; Rauh, 2009; Bikker et al., 2012). Larger funds are able to circumvent the high costs of investing in more complex asset classes, such as by direct investment or by negotiation of better terms with an external manager. Hence, they invest in more alternative assets (Dyck and Pomorski, 2016).

Rauh (2009) and Mohan and Zhang (2014) find that US corporate funds with higher past returns tend to invest in more risky assets, a behavior that they surmise to reflect risk management instead of risk-shifting behavior. In contrast, US public funds tend to shift risk by increasing investment risk when prior returns are low (Pennacchi and Rastad, 2011; Mohan and Zhang, 2014). Besides risk management concerns, past investment returns may affect asset allocation via the lag in rebalancing, which could for example be induced by behavioral biases such as investor inertia (Samuelson and Zeckhauser, 1988). For instance, Dutch pension funds rebalance less than 40% of passive changes in portfolio allocation due to asset price movements within a quarter (Bikker et al., 2010) and a year (Bams et al., 2016). Thus, we include the funds' investment return lagged by a year.

1.3.3 Econometric Specifications

1.3.3.1 Base Model

We regress the percentage allocation to risky assets on regulatory variables and fund characteristics as follows:

$$w_{i,t} = \mathbf{x_{i,t}}\beta + \mathbf{z_{i,t}}\gamma + c_i + u_{i,t} \tag{1.3.1}$$

The observed portfolio share in an asset class (or subclass) is $w_{i,t}$. $\mathbf{x_{i,t}}$ is the vector of regulatory requirements. $\mathbf{z_{i,t}}$ is the vector of fund characteristics that control for observable heterogeneity, c_i is the unobserved fund-specific effect and $u_{i,t}$ is the idiosyncratic error. i is the fund index and t is the year index.

Part of the dispersion in risk exposure may be due to regulation and to fund characteristics, but it may also be attributed to unobserved heterogeneities, such as funding status (Rauh, 2009; Addoum et al., 2010; Mohan and Zhang, 2014), attitude toward risk and governance (Phan and Hegde, 2013; Anantharaman and Lee, 2014; Andonov et al., 2016; Bradley et al., 2016), sponsor characteristics (e.g., profitability, Petersen, 1996; credit ratings, Rauh, 2009; or leverage, Cocco and Volpin, 2007), or other institutional factors (e.g., the existence of a pension benefit guaranty fund for corporate funds in the US and in Ontario, Canada). The fund fixed effect, c_i, mitigates the unobservable cross-sectional heterogeneity. Changes in characteristics such as risk aversion, governance, and funding status are typically gradual. As the median number of years that a fund is in the database is only four, a duration that is likely to be too short for substantial change in the unobserved heterogeneities exemplified, the fund fixed effect is time-invariant.

While we address fund effects parametrically, to mitigate time effects, such as financial market shocks, we make inferences on heterogeneity-robust standard errors that are clustered by year (Petersen, 2009). We allow residuals to be correlated across funds in each year to recognize, for instance, that funds do not rebalance fully (Bams et al., 2016). In a year of outstanding equity market performance, funds' equity allocations increase because of higher equity valuations, resulting in positively correlated residuals across funds. Double-clustering by year and fund would be necessary if we suspect that there exists correlation among the errors of different funds and different years that fades with time (Thompson, 2011). Inference with double-clustered standard errors, however, is of no major difference with that using single-clustered standard errors, except for risky fixed income.[7] The statistical significance of the estimated coefficients is weaker (Table 1.9) because the tradeoff underlying single- and double-clustering is lower bias but higher variance of the coefficient estimates (Thompson, 2011).

1.3.3.2 Censored Model

Under the assumption that pension funds do not take leverage,[8] a regulatory change that makes an asset class more attractive can either prompt funds to start investing in it (increase on the extensive margin), or encourage funds already invested in that asset class to increase their allocation (increase on the intensive margin).

Due to the large proportion of funds holding neither of risky fixed income or alternatives (75% and 18% of the observations, respectively), Equation (1.3.1) may understate the effect of regulation because it ignores changes along the extensive margin. We investigate the extensive margin of risky fixed income and alternatives with a one-sided censored regression model. Censored regression is widely adopted in analyses of individual portfolio holdings, for which data censoring at zero is common (Poterba and Samwick, 2003; Rosen and Wu, 2004).

Let $w_{i,\,t}$ be the observed investment in risky fixed income or alternatives. We define $w_{i,\,t} = \max\left\{0,\, w^*_{i,\,t}\right\}$ to reflect the fact that $w_{i,\,t}$ is censored from below. $w^*_{i,\,t}$ is the unobserved dependent variable in the censored regression model (Chapter 22 of Greene, 2003):

[7] For risky fixed income, the coefficients that become statistically insignificant going from single- to double-clustering have economic significance that are under 1%. Moreover, the censored regression specification (Section 1.3.3.2) is our preferred model for risky fixed income. Therefore, single-clustering by year remains our preferred specification overall.

[8] None of the funds included in this study takes a short position in any of the assets, in any year.

$$\begin{aligned} w^*_{i,\,t} &= \mathbf{x_{i,\,t}}\beta + \mathbf{z_{i,\,t}}\gamma + c_i + u_{i,\,t} \\ w_{i,\,t} &= \max\left\{0,\, w^*_{i,\,t}\right\} \end{aligned} \tag{1.3.2}$$

To obtain consistent estimates of the censored regression model with fixed effects, we apply Honoré's (1992) least absolute deviation estimator for one-sided censored variables.[9] We present and discuss the average marginal effects calculated as per Honoré (2008) to facilitate the economic interpretation of the estimates.

1.3.3.3 Interacted Model

The 2008-09 financial crisis is an opportune case study to evaluate the conjecture that a fund's response to a fixed or VaR-based funding requirements depends on financial market conditions. In poor financial market situations, theory suggests that a non-risk-based funding requirement induces funds to further reduce their risky asset allocation, but a VaR-based constraint could result in more risk exposure instead (Basak and Shapiro, 2001). This outcome does not materialize if the VaR constraint is calculated on a shorter horizon than the investment's time period, as is the case with the capital requirements for Dutch funds (Shi and Werker, 2012; Chen et al., 2017a).

We investigate whether empirics substantiate theory by defining Crisis as an indicator for the years 2008-09, and including it as a multiplicative interacted term with Risk-based Capital Requirements and with Minimum Funding Requirement in separate regressions. We focus on the relative difference in risk-taking by a constrained and unconstrained investor over the crisis and non-crisis periods.

1.4 Determinants of Risky Asset Allocation

Table 1.2 reports the estimates of specification (1.3.1) while Table 1.3 presents the censored regression estimates of specification (1.3.2).

[9] Honoré's (1992) program is available at http://www.princeton.edu/~honore/stata/. Last accessed: 2016-08-05.

1.4.1 Regulatory Requirements and Fund Characteristics

While select fund characteristics and regulatory variables have statistically significant influence on the funds' investments, Risk-based Capital Requirements and MtM Asset Valuation stand out in terms of their economic significance. Risk-based Capital Requirements–unique to the Netherlands since 2007–are associated to 7.1% lower allocation to overall risky assets, with equities composing the bulk of the decrease (-6.5%; Table 1.2, columns 1 and 2). Allocation to risky fixed income is affected to a lesser extent (-0.8% on the extensive margin; Table 1.3, column 1). Our findings concerning capital requirements and fund risk-taking are consistent with Becker and Opp (2014), who show that more stringent capital requirements discourage insurance companies from investing in that asset.

MtM Asset Valuation closely follows Risk-based Capital Requirements in the marginal effect on risk-taking. The closer a fund has to adhere to MtM valuation standards, the less it invests on average in risky assets. If a fund were previously subject to more discretion than valuation with smoothing, and is now required to use MtM valuation (i.e., a 0 to 1 change of the variable), it is estimated to reduce its risky asset allocation by about 6.6% on average, with equities once again constituting almost all of that reduction (Table 1.2, columns 1 and 2). Alternatives are not sensitive to the MtM requirement because these assets are subject to considerable discretion in their valuation. For instance, under the accounting standards FAS 157 and IFRS 13, market-quoted prices or, if unobtainable, the "best information available" is permissible.

Perhaps counter-intuitively, the relation between MtM Asset Valuation and risky fixed income is statistically insignificant. When it concerns bonds with the same credit rating as those defining the liability discount rate (e.g., swaps or high quality corporate bonds), MtM Asset Valuation makes those assets more appealing because their values are correlated with the value of liabilities. By investing in more risky fixed income, a fund is able to maintain a more stable funding ratio while the sponsor can reduce balance sheet volatility. Hence, we expect a positive association between MtM Asset Valuation and risky fixed income investments. However, this reasoning disregards the credit risk premium on risky fixed income (Asvanunt and Richardson, 2016), which could weaken the funds' desire for such assets, and be the underlying reason for the statistically insignificant estimate.

Along the extensive margin, we find that a 1% higher Minimum Funding Requirement is associated with a minute decrease in risky fixed income investments (8.5 bp, Table 1.3, column 1). Theory suggests that funding requirements, when binding, reduce the funds' capacity to invest in risky assets (Basak and Shapiro, 2001). The Minimum Funding Requirement, however, may not be binding for the majority of funds over a large part of the time horizon because the average funding status exceed the minimum required level. For example, US private funds were 102% funded on average between 1992-2007 whereas US state and local funds have an average funding ratio of 100% (Board of Governors of the Federal Reserve System, 2016). Similarly, Dutch funds had in 2007 an average funding ratio of 140% (DNB, 2017).

The subprime crisis that began in late 2007 triggered a downward spiral of the average funding status. It fell from 101% to 74% between 2007 and 2008 for US private funds and from 92% to 65% for US state and local plans (Board of Governors of the Federal Reserve System, 2016). The average funding ratio of Dutch pension funds fell from 150% in June 2007 to about 90% by mid-2008 (De Nederlandsche Bank, 2015). The median funding ratio for Canadian funds slid from 100% in September 2007 to 72% in early 2009 (Marketwired News, 2015). Exploiting this funding ratio shock induced by the crisis, we refine our analysis on the funds' response to Minimum Funding Requirement in Section 1.4.2.

Our estimates concerning the Liability Discount Rate are consistent with the hypothesis that funds with a higher discount rate also tend to invest in more risky assets but the marginal effect is small in comparison to other regulatory requirements. A one standard deviation increase in the liability discount rate is associated to only 1.3% more risky asset holding (0.878 $\times$1.46%; Table 1.2, column 1).

The remaining regulatory requirements have small economic significance. Prolonging the Recovery Period by an additional year is associated with 0.17% more investments in overall risky assets, accounted for by higher allocation to equities (0.44%) (Table 1.2, columns 1 and 2), at the expense of a lower allocation to alternatives along both the intensive (i.e., -0.26%; Table 1.2, column 4) and extensive margins (i.e., -0.24%; Table 1.3, column 2). Funds with a longer recovery period may bear more equity risk because they are able to withstand greater volatility in the value of its assets. Quantitative Investment Restrictions has a statistically significant association to the sole restricted asset class–alternatives. A 10% stricter investment limit is correlated to a 0.31-0.36% lower allocation to alternatives along the intensive and extensive margins (Table 1.2, column 4;

Table 1.3, column 2).

We confirm that the fund characteristics not only have a statistically significant relation with the share invested in risky assets, but also coefficient estimates of the same signs as those found by other studies. For example, $1 billion more in asset under the management is associated with 0.2% higher exposure to risky assets, particularly along the extensive margin of alternatives (i.e., 0.16%; Table 1.2, columns 1 and 4). Alternative assets such as private equity are more attractive for larger funds because these plans are able to invest directly and avoid costly intermediation (Dyck and Pomorski, 2016).

A 10% increase in retired members is associated with a less than 0.1% reduction in equity allocation, slightly lower than the 0.4% estimate of Rauh (2009) for US corporate funds. Along the extensive margin, funds tend to invest 0.16% more in risky fixed income and 0.84% more in alternatives when the number of retired members increases by 10% (Table 1.3, columns 1 and 2). A 1% increase in the portion of members' benefit that is inflation-indexed is linked to 0.02% higher allocation to alternatives along the intensive and extensive margins (Table 1.2, column 4; Table 1.3, column 2), and a 0.02% lower allocation to equities (Table 1.2, column 2). This is consistent with the view that certain alternative assets, such as real assets, have better inflation-hedging potential (Fama and Schwert, 1977; Szymanowska et al., 2014) than equities (Boudoukh and Richardson, 1993; Schotman and Schweitzer, 2000; Ang et al., 2012).

Past Investment Return has low statistical and economic significance on the fund risk-taking. A 10% higher investment return in the previous year is synonymous with 0.3% more risk-taking (Table 1.2, column 1). On a sample of US corporate funds, Rauh (2009) finds that funds invest about 2.2% less in safe assets such as government debt, cash and insurance if the previous year's investment return is 10% higher. Along the extensive margin, a 10% higher investment return last year is associated with 0.3 bp less investment in alternatives (Table 1.3, column 2).

The majority of documented unobserved variables are fund-specific with the exception of funds' participation in a pension benefit protection scheme. The Pension Benefit Guarantee Corporation (PBGC) covers US corporate and industry funds while the Pension Benefit Guarantee Fund (PBGF) covers Canadian funds in Ontario. Such insurance schemes create incentives for funds to increase risk-raking (Sharpe, 1976; Sharpe and Treynor, 1977; Nielson and Chan, 2007;

Crossley and Jametti, 2013). Due the anonymity of the funds in the database, which prohibits identification of Canadian funds in Ontario, our best attempt at investigating this is the inclusion of an indicator for US corporate and industry funds in Equation (1.3.1). The estimates in Table 1.8 indicate that controlling for participation in the PBGC is of no material difference to the base specification estimates in Table 1.2.

1.4.2 Financial Crisis and Funding Requirements

To investigate whether the fund behavior under non-risk-based and risk-based funding requirements depends on the financial market condition, we define Crisis as an indicator for the years 2008-09, and include it in Equation (1.3.1) as an interacted term with Risk Based Capital Requirements and Minimum Funding Requirements (Table 1.4). Table 1.5 provides the corresponding results of the censored regressions for risky fixed income and alternatives (Equation (1.3.2)).

When the financial market is not in crisis, funds facing risk-based capital requirements invest on average 6.9% less in risky assets (Table 1.4, column 1); in 2008-09, this relative difference is similar, -7.2%.[10] The lower risk-taking is mostly accounted for by equities.[11] Thus, the funds' response to risk-based solvency requirements is invariant to the financial market situation, a conclusion that is consistent with Shi and Werker (2012).

The coefficient to the interacted term Risk Based Capital Requirements $\times$ Crisis is statistically insignificant (Table 1.4). This could be due to policies at the Dutch funds' disposal to cope with the funding ratio drop during the crisis (Ponds and Van Riel, 2009). For instance, accrued pension rights and the indexation of benefits are typically conditional on fund solvency. These flexibilities, which US and Canadian funds do not possess, may mute the effect of risk-based capital requirements. Additionally, the Dutch pension regulatory authority introduced broad measures to help funds overcome the crisis-induced challenges (Broeders et al., 2016). Hence, the funds are not limited to respond-

[10] $(-6.947 - 0.376 - 0.283) - (-0.376) = -7.2$. Table 1.4, column 1.

[11] Out of crisis, the relative difference in allocation to equities between constrained and unconstrained funds is -6.2%, whereas in crisis, it is $(-6.207 - 2.727 + 0.304) - (-2.727) = -5.9\%$ (Table 1.4, column 2). On the extensive margin, risk-based capital requirements are associated to less risky fixed income holdings, and estimates are similar to those in the base specification. The relative difference in risky fixed income exposure between constrained and unconstrained funds, during crisis and non-crisis periods, is -0.8% and $(-0.819 - 0.074 + 0.020) - (0.020) = -0.9\%$ respectively (Table 1.5, column 10).

ing to the financial crisis via asset allocation.

Tables 1.6 and 1.7 provide the estimation of the specification with Crisis and Minimum Funding Requirement interaction. Over the non-crisis period, the Minimum Funding Requirement has no statistically significant influence on funds' overall risky asset investments; during the financial crisis, a 100% minimum funding is associated with 3.6%[12] lower investment in risky assets relative to an unconstrained fund. Funds on average satisfy their respective Minimum Funding Requirement prior to the crisis but their funding status markedly decline in the years 2007-08. The financial crisis may have led to the Minimum Funding Requirement that is possibly not binding before 2007 to become binding.

The same is observed for risky fixed income and alternatives. Out-of-crisis, the minimum funding requirement has no statistically significant effect on risky fixed income and alternatives; in-crisis, it is associated with a 2.2% and 4.2% lower allocation to risky fixed income and alternatives, respectively, compensated by a 2.9% higher allocation to equities (Table 1.6, columns 3, 4 and 2). Along the extensive margin, the influence of the minimum funding requirement is associated to a larger difference in risky fixed income allocation when in crisis. A fund facing a 100% minimum funding requirement has on average a 9.4% lower investment in risky fixed income on the extensive margin (Table 1.7, column 1). For alternatives, the effect on the extensive margin is lower than that on the intensive margin (-2.6%; Table 1.7, column 2). Therefore, investors optimally take less risk with a binding non-risk-based solvency requirement (Basak and Shapiro, 2001). Under a risk-based solvency requirement that is estimated over a shorter time period than the investment horizon, investors also optimally have a lower risk exposure (Shi and Werker, 2012; Chen et al., 2017a).

1.5 Conclusion

In the effort to revise the pension regulatory framework, the implication of regulatory measures on the plans' investments is central to the discussion among stakeholders. We present a detailed analysis of a wide range of regulatory requirements in influencing the asset allocation of DB pension plans in the United States, Canada and the Netherlands. These nations differ in their regulatory approaches, and undertook pension reforms in different years.

[12] $(-0.018 \times 100 + 0.851 - 0.018 \times 100) - (0.851) = -3.6$. Table 1.6, column 1.

We find that regulation has at least as much, and for select requirements, much higher marginal effect on funds' risk-taking, relative to fund characteristics such as maturity and size. Among the regulatory measures considered, risk-based capital requirements and MtM asset valuation have the highest economic significance. They are both associated with 7% lower risky asset allocation, even after disentangling the coincident effect of the 2008-09 financial crisis on the funds' investment risk-taking.

Furthermore, a fund's response to risk-based capital and minimum funding requirements depends on the financial market situation. Minimum funding requirement is associated with lower risky asset exposure during the financial crisis, when the constraint was likely to have been binding. In contrast, the relative risk exposure of funds constrained and unconstrained by VaR capital requirements remains similar in and out of the financial crisis.

Amid concerns that risk-based regulation could lead to excessive risk-taking during financial market stress, we provide evidence that allays this apprehension. The relative risk-taking of an investor constrained by a risk-based solvency requirement and an unconstrained investor is similar in crisis and in non-crisis periods. Our quantification of the effect of a wide range of regulatory mechanisms on investment risk-taking may inform policymakers who desire to protect individuals' financial security in retirement by limiting risk-taking, while allowing fund sponsors to take an appropriate amount of financial risk to curb the rising cost of pension provision.

Table 1.1: Summary Statistics

This table provides the summary statistics for returns, asset allocation and fund characteristics, by country and fund type. The total number of funds and observations are presented in Panel A. Panels B and C present the following for 1996 and 2011 respectively: mean (and standard deviation in parenthesis) of the size in billions of USD, maturity (i.e., the % of retired members), the % of inflation-indexed contracts, liability discount rate, total annual return, % allocated to overall risky assets and its subcategories (i.e., equities, risky fixed income and alternative assets).

	US			Canada			Netherlands	
	Public	Private		Public	Private			
		Corporate	Industry		Corporate	Industry	Corporate	Industry
Panel A: Total Number of Pension Funds and Observations								
No. of Funds	121	232	24	44	105	25	9	29
No. of Obs	921	1439	127	407	825	188	50	102
Panel B: Summary Statistics in 1996								
No. of Obs	27	62	5	16	39	5	1	2
Size (billions, USD)	15.2 (23.6)	4.5 (8.9)	2.6 (3)	5.5 (11.4)	1.3 (1.5)	2.7 (4.4)	3 (NA)	8 (1.7)
Retired members (%)	30.7 (6.8)	37 (15.8)	32.6 (22.5)	39.5 (25.5)	37.4 (15)	19.4 (8.5)	41.1 (NA)	34.4 (16.2)
Inflation indexation (%)	50.1 (47.9)	8.1 (26.5)	0 (0)	58.8 (43.5)	31 (36.6)	75 (43.3)	0 (NA)	45 (63.6)
Total return (%)	13.2 (2.8)	15 (2.6)	13.4 (1.2)	18.3 (1.6)	18.4 (1.9)	18.7 (1.2)	13.6 (NA)	14.55 (0.2)
Liabilities discount rate (%)	7.7 (0.9)	8.2 (0.7)	7.87 (0.4)	7.7 (0.8)	7.6 (0.5)	7.7 (1.1)	4 (NA)	4 (0)
Asset Allocation (%)								
Risky Assets	60 (15.2)	71.2 (8.7)	63.2 (11.4)	55.6 (10.3)	60.2 (7.3)	66.2 (7.1)	70.8 (NA)	63.4 (5.5)
Equities	53.8 (15)	63.3 (10.6)	60.3 (10.8)	52 (9.6)	56.7 (7.2)	61.9 (8.2)	29.7 (NA)	25.9 (5.4)
Risky Fixed Income	0 (0)	0 (0)	0 (0)	0 (0)	0.1 (0.6)	1 (2.3)	27.5 (NA)	21.5 (3.2)
Alternatives	6.2 (6.1)	7.9 (7.3)	2.9 (1.9)	3.7 (4.9)	3.4 (4.8)	3.2 (3.3)	13.6 (NA)	16.1 (3.3)
Panel C: Summary Statistics in 2011								
No. of Obs	50	102	8	20	32	10	3	19
Size (billions, USD)	31.1 (43.1)	8.2 (12.8)	6.3 (7.4)	14.5 (26.3)	2.9 (3.8)	1.3 (1)	11.7 (5.8)	13.3 (24.2)
Retired members (%)	38.3 (8.3)	57.8 (22.4)	47 (12)	46.3 (21)	54.6 (23.7)	32.8 (13.6)	59.6 (22)	41.9 (20.5)
Inflation idexation (%)	53.4 (48.1)	4.9 (19.8)	25 (46.3)	59 (42.7)	42.2 (45.3)	54 (48.8)	33.3 (57.7)	20 (40)
Total return (%)	1.5 (1.9)	5.6 (4.7)	3.4 (2.8)	2.9 (3.6)	3.8 (3.7)	2.8 (2.6)	6.3 (1)	9.5 (4.6)
Liabilities discount rate (%)	7.3 (1.3)	5 (0.5)	6.8 (1.1)	6.2 (0.5)	5.4 (0.9)	6.1 (0.6)	2.7 (0)	2.8 (0.9)
Asset Allocation (%)								
Risky Assets	74.7 (8.6)	62.9 (13.7)	74.9 (11.5)	64.6 (10.1)	57.2 (11.4)	64.1 (4)	48.4 (14.6)	47.1 (12.8)
Equities	50.5 (11.2)	44.7 (15.9)	48.1 (11.8)	49.9 (9.8)	53.1 (11.2)	53.3 (9)	31.4 (14.2)	28.8 (9.3)
Risky Fixed Income	2.7 (4.1)	1.4 (2.4)	1.9 (2.9)	1.2 (3.2)	0.1 (0.4)	0.6 (0.9)	5.3 (2.5)	3.9 (4.8)
Alternatives	21.5 (12.8)	16.8 (15.8)	24.9 (20.3)	13.5 (12.3)	4 (5.1)	10.2 (7.6)	11.8 (5.9)	14.5 (8.5)

Figure 1.5.1: Time Series of Mean Allocation to Risky Assets By Country and Type of Fund

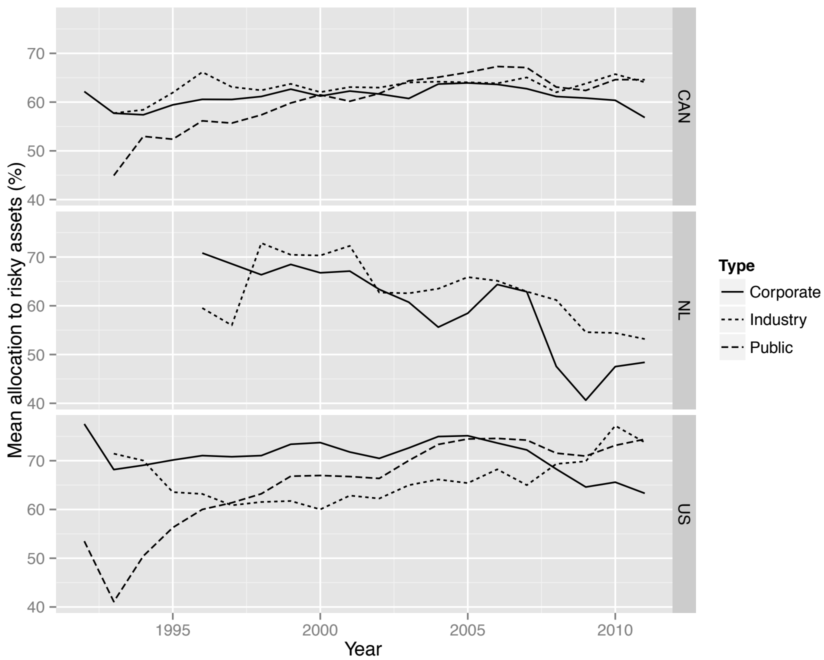


Table 1.2: Pension Funds Risky Asset Allocation and Regulation

This table presents the estimates of the regression of the funds' percentage allocation to risky assets on regulatory requirements and fund characteristics, with fund fixed effects (Equation (1.3.1)). We consider overall risky asset investment (column (1)), and its composing asset classes: equities, risky fixed income (Risky FI) and alternatives (Alt) (columns (2) to (4)). Section 1.3.1 specifies the assets within each asset class. Heteroskecasticity robust standard errors that are clustered by Year are in parentheses.

	Dependent variable:			
	All	Equities	Risky FI	Alt
	(1)	(2)	(3)	(4)
Quantitative Investment Restrictions	-0.023^{**}	0.010	-0.002	-0.031^{***}
	(0.010)	(0.017)	(0.002)	(0.012)
Mark-to-market Asset Valuation	-6.612^{***}	-6.037^{*}	-0.342	-0.233
	(1.891)	(3.224)	(0.422)	(1.780)
Recognition of Unfunded Liabilities	-1.042	-1.921	0.553^{**}	0.325
	(1.197)	(1.936)	(0.256)	(1.917)
Liability Discount Rate	0.878^{***}	0.432	0.110^{*}	0.336
	(0.197)	(0.303)	(0.057)	(0.240)
Minimum Funding Requirement	-0.016	0.008	-0.017	-0.007
	(0.020)	(0.028)	(0.015)	(0.017)
Risk-based Capital Requirements	-7.118^{***}	-6.450^{***}	-0.908	0.240
	(2.139)	(2.200)	(0.584)	(0.655)
Recovery Period	0.169^{***}	0.437^{***}	-0.011	-0.258^{***}
	(0.045)	(0.087)	(0.012)	(0.086)
Maturity	-0.020	-0.096^{**}	0.014^{***}	0.062^{***}
	(0.025)	(0.039)	(0.005)	(0.021)
Inflation Indexation	0.007	-0.018^{**}	0.003	0.022^{***}
	(0.007)	(0.008)	(0.004)	(0.005)
Size	0.205^{***}	0.003	0.038^{***}	0.164^{***}
	(0.023)	(0.042)	(0.005)	(0.029)
Past Investment Return	0.032^{*}	0.065^{*}	-0.007	-0.027
	(0.017)	(0.037)	(0.006)	(0.028)
Observations	3,687	3,687	3,687	3,687
R^2	0.708	0.685	0.679	0.731
Adjusted R^2	0.660	0.634	0.627	0.687
F Statistic (df = 11; 3171)	42.71^{***}	66.27^{***}	24.04^{***}	47.87^{***}

Note: *p<0.1; **p<0.05; ***p<0.01

Table 1.3: Pension Funds Risky Fixed Income and Alternatives Allocation and Regulation, Average Marginal Effect Implied by the Censored Regression

This table presents the average marginal effects implied by the censored regression model (Equation (1.3.2)) that regresses risky fixed income (Risky FI) and alternatives (Alt) on regulatory requirements and fund characteristics. We follow Honoré (2008) by multiplying the coefficient estimates with the proportion of the observations that is not censored. The values displayed here allow us to directly interpret the average economic effect of the censored regression model. For example, if the value of assets under management increases by $1 billion, we estimate that the fund would increase its investment into risky fixed income by on average 0.03% (column (1)).

	Dependent variable:	
	Risky FI	Alt
	(1)	(2)
Quantitative Investment Restrictions	-0.007^{*}	-0.036^{**}
Mark-to-market Asset Valuation	-0.751	-1.114
Recognition of Unfunded Liabilities	0.507^{**}	1.642
Liability Discount Rate	0.168^{***}	0.481^{**}
Minimum Funding Requirement	-0.085^{***}	-0.005
Risk-based Capital Requirements	-0.819^{***}	-1.452
Recovery Period	-0.022^{***}	-0.239^{***}
Maturity	0.016^{**}	0.084^{**}
Inflation Indexation	0.003	$0.021{*}{*}$
Size	0.027^{***}	-0.141^{***}
Past Investment Return	0	0.003^{***}
Observations	3,687	3,687
Proportion Uncensored	0.25	0.82

Note: *p<0.1; **p<0.05; ***p<0.01

Table 1.4: Pension Funds Risky Asset Allocation and Risk-Based Capital Requirements, Interaction with the Crisis Indicator

This table presents the regression estimates of the funds' percentage allocation to risky assets on regulatory requirements and fund characteristics, with fund fixed effects (Equation (1.3.1)), and the interacted term Risk-based Capital Requirements × Crisis. We consider overall risky asset investment (column (1)), and its composing asset classes: equities, risky fixed income (Risky FI) and alternatives (Alt) (columns (2) to (4)). Section 1.3.1 specifies the assets within each asset class. Heteroskecasticity robust standard errors that are clustered by Year are in parentheses.

	Dependent variable:			
	All	Equities	Risky FI	Alt
	(1)	(2)	(3)	(4)
Quantitative Investment Restrictions	-0.024^{**}	0.001	-0.002	-0.023^{**}
	(0.010)	(0.016)	(0.002)	(0.010)
Mark-to-market Asset Valuation	-6.828^{***}	-7.618^{**}	-0.315	1.105
	(1.967)	(3.089)	(0.435)	(1.635)
Recognition of Unfunded Liabilities	-1.033	-1.867	0.553^{**}	0.282
	(1.202)	(1.872)	(0.256)	(1.868)
Liability Discount Rate	0.894^{***}	0.540^{*}	0.108^{*}	0.245
	(0.196)	(0.292)	(0.060)	(0.244)
Minimum Funding Requirement	-0.017	-0.001	-0.017	0.001
	(0.020)	(0.030)	(0.015)	(0.018)
Risk-based Capital Requirements	-6.947^{***}	-6.207^{**}	-0.897	0.157
	(2.587)	(2.698)	(0.596)	(0.774)
Crisis	-0.376	-2.727^{*}	0.047	2.305^{***}
	(0.879)	(1.553)	(0.140)	(0.872)
Recovery Period	0.161^{***}	0.382^{***}	-0.010	-0.211^{**}
	(0.043)	(0.104)	(0.013)	(0.096)
Maturity	-0.021	-0.103^{***}	0.014^{***}	0.068^{***}
	(0.025)	(0.039)	(0.005)	(0.022)
Inflation Indexation	0.008	-0.015^{*}	0.003	0.020^{***}
	(0.007)	(0.008)	(0.003)	(0.005)
Size	0.206^{***}	0.008	0.038^{***}	0.159^{***}
	(0.023)	(0.042)	(0.005)	(0.028)
Past Investment Return	0.027	0.033	-0.006	0.0002
	(0.022)	(0.027)	(0.007)	(0.018)
Risk-based Capital Requirements × Crisis	-0.283	0.304	-0.041	-0.546
	(1.889)	(1.936)	(0.332)	(1.055)
Observations	3,687	3,687	3,687	3,687
R^2	0.708	0.688	0.679	0.734
Adjusted R^2	0.660	0.638	0.627	0.691
F Statistic (df = 11; 3171)	50.87^{***}	285.4^{***}	21.31^{***}	40.82^{***}

Note: *p<0.1; **p<0.05; ***p<0.01

Table 1.5: Pension Funds Risky Fixed Income and Alternatives Allocation and Risk-based Capital Requirements, Average Marginal Effect Implied by the Censored Regression; Interaction with the Crisis Indicator

This table presents the average marginal effects implied by the censored regression model (Equation (1.3.2)) that regresses the percentage allocated to risky fixed income (Risky FI) and alternatives (Alt) on regulatory requirements and fund characteristics, and the interacted term Risk-based Capital Requirements × Crisis. We follow Honoré (2008) by multiplying the coefficient estimates with the proportion of the observations that is not censored. The values displayed here allow us to directly interpret the average economic effect of the censored regression model. For example, if the value of assets under management increases by $1 billion, we estimate that the fund would increase its investment into risky fixed income by on average 0.03% (column (1)).

	Dependent variable:	
	Risky FI	Alt
	(1)	(2)
Quantitative Investment Restrictions	-0.007^{*}	-0.031^{**}
Mark-to-market Asset Valuation	-0.801	0.047
Recognition of Unfunded Liabilities	0.511^{**}	1.529
Liability Discount Rate	0.173^{***}	0.399^{**}
Minimum Funding Requirement	-0.087^{***}	0
Risk-based Capital Requirements	-0.819^{***}	-1.451
Crisis	-0.074	-1.532^{***}
Recovery Period	-0.024^{***}	-0.208^{***}
Maturity	0.016^{**}	0.087^{**}
Inflation Indexation	0.003	$0.019*$
Size	0.026^{***}	-0.128^{***}
Past Investment Return	0	0.002^{***}
Risk-based Capital Requirements × Crisis	0.02	-0.531
Observations	3,687	3,687
Proportion Uncensored	0.25	0.82
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 1.6: Pension Funds Risky Asset Allocation and Minimum Funding Requirement, Interaction with the Crisis Indicator

This table presents the regression estimates of the funds' percentage allocation to risky assets on regulatory requirements and fund characteristics, with fund fixed effects (Equation (1.3.1)), and the interacted term Minimum Funding Requirement × Crisis. We consider overall risky asset investment (column (1)), and its composing asset classes: equities, risky fixed income (Risky FI) and alternatives (Alt) (columns (2) to (4)). Section 1.3.1 specifies the assets within each asset class. Heteroskecasticity robust standard errors that are clustered by Year are in parentheses.

	Dependent variable:			
	All	Equities	Risky FI	Alt
	(1)	(2)	(3)	(4)
Quantitative Investment Restrictions	-0.027^{***}	0.005	-0.003	-0.029^{**}
	(0.010)	(0.017)	(0.002)	(0.011)
Mark-to-market Asset Valuation	-7.223^{***}	-6.981^{**}	-0.423	0.182
	(1.908)	(3.248)	(0.399)	(1.744)
Recognition of Unfunded Liabilities	-0.985	-1.946	0.566^{**}	0.394
	(1.201)	(1.869)	(0.254)	(1.855)
Liability Discount Rate	0.909^{***}	0.516^{*}	0.113^{*}	0.280
	(0.193)	(0.296)	(0.059)	(0.245)
Minimum Funding Requirement	-0.018	-0.0002	-0.017	-0.0001
	(0.021)	(0.029)	(0.015)	(0.018)
Crisis	0.851	-4.725^{***}	0.389^{**}	5.187^{***}
	(0.700)	(1.565)	(0.197)	(1.201)
Risk-based Capital Requirements	-6.887^{***}	-6.369^{***}	-0.865	0.347
	(2.169)	(2.243)	(0.586)	(0.690)
Recovery Period	0.152^{***}	0.397^{***}	-0.012	-0.233^{***}
	(0.046)	(0.102)	(0.012)	(0.090)
Maturity	-0.024	-0.099^{***}	0.014^{**}	0.061^{***}
	(0.025)	(0.038)	(0.005)	(0.020)
Inflation Indexation	0.007	-0.014^{*}	0.002	0.019^{***}
	(0.007)	(0.009)	(0.004)	(0.006)
Size	0.202^{***}	0.014	0.037^{***}	0.151^{***}
	(0.023)	(0.041)	(0.005)	(0.028)
Past Investment Return	0.026	0.034	-0.006	-0.001
	(0.022)	(0.027)	(0.007)	(0.016)
Minimum Funding Requirement × Crisis	-0.018^{**}	0.029^{***}	-0.005^{***}	-0.042^{***}
	(0.009)	(0.004)	(0.002)	(0.007)
Observations	3,687	3,687	3,687	3,687
R^2	0.708	0.688	0.679	0.734
Adjusted R^2	0.660	0.638	0.627	0.691
F Statistic (df = 11; 3171)	50.87^{***}	285.4^{***}	21.31^{***}	40.82^{***}

Note: *p<0.1; **p<0.05; ***p<0.01

Table 1.7: Pension Funds Risky Fixed Income and Alternatives Allocation and Minimum Funding Requirement, Average Marginal Effect Implied by the Censored Regression; Interaction with the Crisis Indicator

This table presents the average marginal effects implied by the censored regression model (Equation (1.3.2)) that regresses the percentage allocated to risky fixed income (Risky FI) and alternatives (Alt) on regulatory requirements and fund characteristics, and the interacted term Minimum Funding Requirement × Crisis. We follow Honoré (2008) by multiplying the coefficient estimates with the proportion of the observations that is not censored. The values displayed here allow us to directly interpret the average economic effect of the censored regression model. For example, if the value of assets under management increases by $1 billion, we estimate that the fund would increase its investment into risky fixed income by on average 0.03% (column (1)).

	Dependent variable:	
	Risky FI	Alt
	(1)	(2)
Quantitative Investment Restrictions	-0.008^{*}	-0.035^{**}
Mark-to-market Asset Valuation	-0.905	-0.738
Recognition of Unfunded Liabilities	0.525^{**}	1.605
Liability Discount Rate	0.178^{***}	0.419^{**}
Minimum Funding Requirement	-0.09^{***}	0
Crisis	-0.169^{*}	3.207^{***}
Risk-based Capital Requirements	-0.768^{***}	-1.376
Recovery Period	-0.028^{***}	-0.224^{***}
Maturity	0.015^{**}	0.082^{**}
Inflation Indexation	0.003	$0.019*$
Size	0.028^{***}	-0.126^{***}
Past Investment Return	0	0.002^{***}
Minimum Funding Requirement × Crisis	-0.004^{***}	-0.026^{***}
Observations	3,687	3,687
Proportion Uncensored	0.25	0.82
Note:	*p<0.1; **p<0.05; ***p<0.01	

Appendix to Chapter 1

1.A Overview of the Pension Regulatory Environment

1.A.1 The United States

In the US, public and corporate pension funds have their own regulatory authority. US private plans are either single employer (corporate funds) or multi-employer (industry funds, also known as Taft-Hartley plans).[13] Corporate funds are subject to strict rules, both for pension plans' budgeting and for sponsors' accounting. Plan budgeting rules impose minimum standards for funding levels, sponsor contributions, recovery periods, and so on. They are set federally under the 1974 Employee Retirement Income Security Act (ERISA). Among ERISA's many amendments, the 2006 Pension Protection Act (PPA) introduced major reforms that came into effect in 2008. PPA required corporate pension plans to target full funding by 2011 (90% before 2008, and a gradual increase from 90% to 100% between 2008 and 2011). Assets should be valued on a market-related basis with at most a two-year average of 90-110% of fair value[14] (compared with the previous five-year average of 90-120%). Liabilities are discounted at corporate bond rates.[15] PPA also required quicker remediation of shortfalls: any funding deficit has to be resolved within a seven-year period (compared with 30 years previously).

The accounting statements of incorporated companies in the US follow the rules set by the Financial Accounting Standards Board (FASB). Over the past decades, the FASB has changed the items that sponsors have to disclose or recognize, and the permissible recognition method. Three standards were in force during our study period, FAS 87 (effective 1986), FAS 132 (2004) and FAS 158

[13]Single-employer plans are retirement plans that are administered by one employer only. Multi-employer plans are collectively bargained plans maintained by labor unions and more than one employer. They are managed by a board of trustees with equal representation of employers and employees. This arrangement is common in industries that are typically unionized, such as construction, entertainment, trucking, and mining.

[14]Fair value requires the assessment of the price that is fair between two specific parties. Market value may meet this criterion, but this is not necessarily the case. In practice, fair value estimations may be based on market prices if they are available and considered reliable. Otherwise, they can be based on an estimate derived from multiple permissble methodologies.

[15]Under PPA, the discount rate for single-employer plans is a two-year average of investment-grade corporate bonds (i.e., AAA, AA and A). The rates are three-tiered (i.e., 5, 5-15, and more than 15 years) to match the duration of plans' liabilities. PPA shortened the averaging period of the discount rate from four to two years.

(2006). Under FAS 87, single-employer fund sponsors have to recognize the cost of providing pensions on their income statement, and to disclose the fair value of pension assets and the present value of pension obligations in the notes to the financial statements. While employers are required to compute their plans' funded status, defined as the fair value of assets less projected benefit obligation (PBO),[16] they do not have to report it on their balance sheet. Only when the accumulated benefit obligation (ABO)[17] exceeds the accrued pension assets must firms recognize the unfunded ABO. FAS 132 adds the requirement to disclose the funds' investment policy, while FAS 158 made it mandatory to always recognize the plans' unfunded liabilities, determined as PBO, on the balance sheet,[18] a stricter requirement than the ABO standard under FAS 87.

US industry funds, in comparison with corporate ones, are subject to more lenient requirements despite being regulated under the same federal act. Historically, industry plans have broad discretion on the valuation assumptions for plan assets and liabilities, as well as on funding methods. The introduction of PPA in 2008 preserved and even expanded these flexibilities. For the purpose of determining annual funding, the only condition on the discount rate is that it has to be actuarially reasonable. Employer and employee contribution rates are decided through a collective bargaining process every three to five years. Due to the lengthy nature of the process, the PPA provides for shortfalls to be amortized over a period of 15 years (previously 30). Multi-employer plans that are under 80% funded have to submit a plan for achieving a one-third improvement in the funded level every ten years. On the accounting side, participating sponsors of industry funds merely have to report the required contributions each year on their financial statements.

In contrast to their private counterparts, US public funds are subject to much laxer rules. Accounting and funding standards were set in 1984 in Governmental Accounting Standards Board (GASB) Statement 25 and in Actuarial Standards of Practice (ASOP) 27. These standards allow funds' assets to be valued on an actuarial basis[19] and their liabilities to be discounted using the expected rate

[16]PBO is the actuarial present value of future pension benefits accrued from past service years. Future events such as compensation increases, turnover and mortality are taken into consideration.

[17]In contrast to PBO, ABO is an estimate of a company's pension liability under the view that the plan is terminated on the date the calculation is performed.

[18]Sponsors of multi-employer plan are required only to report their respective contribution to the plan.

[19]Actuarial valuation recognizes realized and/or unrealized gains and losses in the market value versus book value, typically over a five-year period, rather than immediately.

of return on pension plan assets.

1.A.2 Canada

In Canada, there is much less regulatory distinction between public and corporate pension funds. All registered pension plans are regulated under both federal and provincial pension standards. Tax legislation is set at the federal level whereas minimum standards for funding and other issues are set at the provincial level.[20] Despite having provincial jurisdiction, pension legislation on design, funding, communication and administration is fairly consistent across the country (Pugh, 2006). Regulatory harmonization is among the responsibilities of the Canadian Association of Pension Supervisory Authorities (CAPSA). CAPSA advocates for a funding requirement of 100%, as determined using actuarially acceptable assumptions (e.g., market value of assets, accrued liability discounted using Government of Canada bonds). Until 2010, Canadian funds were prohibited from investing more than 25% of their portfolio in real estate, and 15% in Canadian resource property.

Canadian private pension plans and their sponsors prepare their financial statements under the standards set by the Accounting Standards Board of Canada (AcSB). Between December 1986 and 1999, the effective rules for sponsors were set out in Canadian Institute of Chartered Accountants (CICA) 3460, but many of the key assumptions, such as the liability discount rate, were left to the plan administrator's discretion. Effective in 2000, CICA 3461 revoked some of that discretion. Funds have to discount their liabilities at the AA-rated corporate bond rate. For asset valuation, funds can choose between market and market-related value (the latter allows for valuation smoothing). In January 2006, the AcSB announced its decision to converge to the International Financial Reporting Standards (IFRS). A five-year transition period was planned, with the compliance to International Accounting Standard (IAS) 19 becoming effective in January 2011. IAS 19 requires balance sheet recognition of the present value of estimated total retirement benefits, including future compensation net of the fair value of pension assets, discounted using the interest rate on high quality corporate debt. Plan assets are measured at fair market value with no permissible smoothing. Canadian public pension plans' sponsors followed the same set of CICA accounting standards up to 2012.

[20] An exception is the plans for employees of banks, communications companies etc., who fall under the 1985 Federal Pension Benefits Standards Act.

1.A.3 The Netherlands

The Netherlands makes no regulatory distinction between funds covering public or private sector workers. Prior to 2007, Dutch pension funds were subject to the Pensions and Savings Act (Pensioen- en Spaarwet, PSW), which permits several funding methods, some allowing pension costs to be deferred.[21] Since 1999, PSW requires a 100% funding ratio at all times. Although there is no fixed recovery period for underfunding, failure to meet the funding requirement invites regulatory scrutiny (Pugh, 2006). Liabilities are discounted using a fixed actuarial rate of 4% whereas assets are marked-to-market. Since 2003, the funds' assets have to be valued at at least 105% of the provision for liabilities, of which 5% constitutes a risk reserve. Funds are granted up to eight years to recoup any shortfall on the reserve value. These risk reserves are a precursor to the risk-based capital requirement.

In 2007, a new statute, the Pension Act (Pensioenwet) was introduced. It comprises the Financial Assessment Framework (Financieel Toetsingskader, FTK), which lays down pension funds' financial requirements. The FTK outlines regulations concerning the liability discount rate (i.e., swap rate), and maintains the requirement for MtM asset valuation, along with 100% minimum funding. The FTK also sets capital buffers to ensure, with a 97.5% confidence level, that funds' assets will not be less than the level of liabilities within a year. If funds fail to meet the minimum funding or capital requirements, they are granted a period of three years to meet the minimum solvency requirements and up to 15 years to recoup the buffer requirements. Among the three countries under study, only the Netherlands has risk-based capital requirements that are similar to those that apply in Europe for insurance companies, and are under discussion for European pension plans (EIOPA, 2012).

Companies listed on a market in the European Union (EU) have been required to abide by IAS 19 since 2005. While this standard applies only to listed companies in the EU, the Dutch government approved a bill in 2005 to encourage unlisted companies to adopt it as well. Before IAS 19 was adopted, the Dutch accounting regulation, Raad voor de Jaarverslaggeving RJ 271 (2002 edition) required the previous year's pension contribution premium to be recognized in the income statement as an operating expense and the previous year's premium adjustment paid for salary increments to be shown on the balance

[21] For example, the (65-x) method allowed salary or other pension increases on past service benefits to be funded over the remaining years until retirement age, typically 65. This method allows the deferral of pension costs.

sheet. Because of the stand-alone nature of Dutch occupational pensions,[22] the employer's pension liabilities are not easily determined. This is further complicated by policy mechanisms such as conditional indexing, which is difficult to value. The sponsors of industry funds treat industry plans as DC funds from an accounting perspective, and recognize only the promised contribution due each year on their balance sheet. On the contrary, corporations with their own pension funds have to recognize unfunded pension liabilities on their balance sheets.

[22] Dutch occupational pension funds are independent trusts. Since the governing board has equal representation of employers and unions, the employer does not have exclusive power on decision-making, and is not solely responsible for any underfunding (Bovenberg and Nijman, 2009).

1.B Summary of the Regulatory Environment

Country	U.S.			Canada	Netherlands
Fund Type	**Public**	**Corporate**	**Industry**		
Investment restrictions					
Quantitative Investment Restrictions	None	None	None	Prior to 2005: 30% limit on foreign assets Prior to 2010: 15% limit on Canadian resource property, 25% limit on real estate and Canadian resource property	None
Valuation requirements					
Asset Valuation	GASB: Actuarial valuation allowing five years' smoothing of gains and losses.	*For funding:* Before 2006: ERISA Fair value with smoothing After 2006: PPA (effective in 2009) Fair value. Option to smooth up to 24 months under the PPA. Smoothed value has to be bounded between 90% and 110% of the asset's current market value *For sponsors' accounting:* Since 1986: FAS 87 Market value or market-related value (e.g., 5-year moving average) permitted). In 2006, FAS 157 refined the definition of market value.	Since 1986: ERISA Reasonable actuarial assumptions.	*For funding:* [1] CICA 4600: Fair value of assets *For sponsors' accounting:* Up till 2011: CICA 3460 and 3461 Market value or market-related value (e.g., 5-year moving average permitted) Since 2011: IAS 19 Market value	*For funding:* Before 2007: PSW Market value After 2007: FTK Market value *For sponsors' accounting:* Before 2005: RJ 271 edition 2002 and 2003 2002 ed. did not require the recognition of the value of investment assets. 2003 ed. adopted many of the principles in IAS 19 After 2005: IAS 19 Market value

Country	**U.S.**			**Canada**	**Netherlands**
Fund Type	**Public**	**Corporate**	**Industry**		
Liability Discount Rate	GASB: Expected return of assets	*For funding:* Before 2004: ERISA and subsequent amendments A corridor around the 4-year weighted average[II] of the 30Y T bond. The permissible range above and below the weighted average varied over time 2004-06: PFEA Market rate (corporate bonds), 4-year average Since 2006: PPA Market rate (corporate bonds), with 2-year smoothing allowed *For sponsors' accounting:* FAS 87 Market rate (corporate bonds). 4-year average prior to 2006, 2-year average after	Since 1986: ERISA Discount rate has to be *actuarially reasonable*	*For funding:* Government bond yield (7Y) plus an additional factor (e.g., 0.9%) for the first 10 years, extrapolated after 10 years. Same rule for indexed pension based on Government real yield *For sponsors' accounting:* Before 2000: CICA 3460 Management's "best estimate" of the long-term rate of return on assets After 2000: CICA 3461 Market interest rate at the measurement date on high-quality debt instruments (e.g., AA corporate bonds) with cash flow that matches the timing and amount of the expected benefit payments, or interest rate inherent in the amount at which the accrued benefit obligation could be settled	*For funding:* Before 2007: PSW Fixed actuarial interest rate with a prescribed maximum. If no indexation is provided, then >4% is allowed, otherwise ≤4% Since 2007: FTK Yield curve that is based on the euro swap curve as set by the DNB *For sponsors' accounting:* Since 2005: IAS 19 High quality corporate bond yield only for listed corporate sponsors

Country	U.S.			Canada	Netherlands
Fund Type	**Public**	**Corporate**	**Industry**		
Recognition of Liabilities on the Sponsor's Balance Sheet	Between 1986 and 1994: GASB No. 5 Disclosure but no recognition Since 1994: GASB No. 27 Recognition of Net Pension Obligation, which is the shortfall in the annually required contribution, as a liability	Before 2006: FAS 87 Only unfunded liabilities in excess of ABO are recognized on the balance sheet Since 2006: FAS 158 All over/underfunded liabilities in excess of PBO are recognized on the sponsor's balance sheet	Since 1986: ERISA Participating sponsors merely report contributions on their financial statements but not the plan's long-term financial risks	Up till 2011: CICA 3460 and 3461 Surplus/ insufficiency of funding relative to pension expense recognized Since 2011: IAS 19 The following value is recognized: Present value of ABO less unrecognized past service costs, ± actuarial gains / losses not recognized less fair value of plan assets	Since 2005: IAS 19 The following value is recognized: Present value of ABO less unrecognized past service costs, ± actuarial gains / losses not recognized less fair value of plan assets
Funding requirements					
Minimum Funding Requirements	No min (0%)	Since 1994: Retirement Protection Act Min funding of 90% Since 2006: PPA 100% funding target but phased in over three years beginning 2008, at the rate of 92% (2008), 94% (2009), 96% (2010), 100% after	100%	100%	Before 1999: PSW "65-x" funding standard, 65 is the assumed normal retirement age and "x" is the plan member's current age Since 1999: PSW Assets must cover the present value of the accrued pensions (i.e. 100%) Since 2007: FTK 100%

Country	U.S.			Canada	Netherlands
Fund Type	Public	Corporate	Industry		
Risk-based Capital Requirements	None	None	None	None	Since 2007: FTK Regulatory capital requirement determined by fixed shocks onto the various risks exposure that correspond to 105% at confidence level of 97.5% with a year horizon. For a stylized pension fund with equal investment in equity and bonds, capital requirements amount approximately to a funding ratio of 130%
Recovery Period	None	Before 2006: 30Y Since 2006: PPA 7Y	Before 2006: ERISA No provision. Since 2006: PPA 10 years, 15 years for seriously endangered plans.	Federal plans and provincial plans in Alberta and Ontario have a maximum amortization period of 10 years since 2009, previously 5 years. Other provinces typically set it at 5 years (with a possibility of extension with the consent of plan members)	Between 1999 and 2007: PSW 10-year transition to attain the new minimum funding requirement of 100% Since 2007: FTK 3 years for solvency margin, up to 15 years for buffer depending on continuity analysis

[I] These regulations concern federally regulated plans only. Rates for provincially regulated plans may differ.

[II] Average yield over 48 months with rates for the most recent 12 months having 40% weight, the second most recent 12 months weighs 30%, the third most recent 12 months weighs 20%, and the fourth weighs 10%.

ABO: Accumulated benefit obligation
CICA: Canadian Institute of Chartered Accountants
DNB: De Nederlandsche Bank (Central Bank of the Netherlands)
ERISA: Employee Retirement Income Security Act
FAS: Financial Accounting Standards
FTK: Financieel Toetsingskader (Financial Assessment Framework)
PBO: Projected benefit obligation
PFEA: Pension Funding Equity Act
PPA: Pension Protection Act
PSW: Pensioen- en spaarfondsenwet (Pensions and Savings Fund Act)

1.C Definition of the Explanatory Variables

Variable	Definition	Expected impact	Sign of the expected impact
	Regulatory Factors		
Investment Requirements			
Quantitative Investment Restrictions	Sum of (100%-Investment limit in % for all restricted asset classes).	Quantitative investment restrictions may lead to lower allocation to the restricted asset class if the limits are binding.	+ (if binding)
Valuation Requirements			
Mark-to-market (MtM) Asset Valuation (for funding and accounting purposes)	Average of two categorical variables based on valuation rules for funding and for accounting purposes independently constructed as follows: 1 if market value, 0.5 if smoothing is allowed, 0 in the case of further discretion than smoothing.	MtM valuation may induce more volatility on the funding ratio, and on the sponsor's balance sheet, thus lead to less investment in risky assets.	-
Recognition of Unfunded Liabilities (on the sponsor's balance sheet)	Categorical variable: 1 if unfunded liabilities (as measured by PBO[III] or equivalent) are recognized on the balance sheet, 0.5 if recognition of excess/ deficit relative to liabilities as measured by ABO[IV] or equivalent is necessary, 0 otherwise.	Recognition requirement may lower allocation to risky assets to reduce volatility of the sponsor's balance sheet.	-
Liability Discount Rate	Spread between the discount rate level for funding purposes disclosed by the fund, and the domestic 10-year government bond.	Higher liability discount rate may be positively associated with more investment in risky assets.	+
Funding Requirements			
Minimum Funding Requirement	Regulatory minimum ratio of the value of assets over the value of liabilities.[VI]	Higher funding requirement may decrease risky asset allocation if binding.	- (if binding)
Risk-based Capital Requirements	Dummy variable: 1 if capital requirements based on a quantitative risk-profile assessment are imposed.[VII]	Capital requirements may discourage investment risk-taking.	-

Variable	Definition	Expected impact	Sign of the expected impact
	Regulatory Factors		
Funding Requirements			
Minimum Funding Requirement	Regulatory minimum ratio of the value of assets over the value of liabilities.[VI]	Higher funding requirement may decrease risky asset allocation if binding.	- (if binding)
Risk-based Capital Requirements	Dummy variable: 1 if capital requirements based on a quantitative risk-profile assessment are imposed.[VII]	Capital requirements may discourage investment risk-taking.	-
Recovery Period	Maximum number of years permitted by regulation to restore the Minimum Funding Requirement if a plan's funding status has fallen below the minimum in any year.	Longer recovery periods may encourage funds to take more risk.	+
	Individual Factors		
Maturity	Percentage of retired members.	More mature funds may allocate less to risky assets.	-
Inflation Indexation	Percentage of member's benefits contractually indexed to inflation.	Funds providing more inflation indexation may allocate more to risky assets.	+
Size	Market value of Assets under Management (AUM) in billions of USD.	Funds with a larger AUM may invest in more in alternatives.	+ (for alternatives)
Past Investment Return	Total investment return in the previous year.	Higher past risky assets' returns may lead to a mechanically larger share in the asset allocation.	+

III Projected Benefit Obligation.

IV Accumulated Benefit Obligation.

VI Dutch funds' "65-x" funding requirement is estimated using $\min\{\frac{Maturity}{65}\times 100, 100\}$, with $Maturity$ as the percentage of retired members.

VII As voluntary adoption of the FTK among Dutch funds has been permitted since 2005, we also vary the definition of RBCR to as an indicator variable beginning in 2004, 2005 or 2006, and obtained similar results. Results presented in the tables adopt the official date of FTK implementation, i.e., 2007.

1.D Pension Benefit Guarantee Corporation

Table 1.8: Pension Funds Risky Asset Allocation and Regulation: Indicator for PBGC Participation

This table presents the estimates of the regression of the funds' percentage allocation to risky assets on regulatory requirements and fund characteristics, with fund fixed effects (Equation (1.3.1)), and an indicator for US corporate and industry funds (PBGC). The indicator represents funds that are members of the Pension Benefit Guarantee Corporation. Heteroskecasticity robust standard errors that are clustered by Year are in parentheses.

	Dependent variable:			
	All	Equities	Risky FI	Alt
	(1)	(2)	(3)	(4)
Quantitative Investment Restrictions	−0.024**	0.010	−0.003	−0.031***
	(0.010)	(0.017)	(0.002)	(0.012)
Mark-to-market Asset Valuation	−6.707***	−6.030*	−0.456*	−0.221
	(1.789)	(3.218)	(0.262)	(1.782)
Recognition of Unfunded Liabilities	−1.058	−1.920	0.534**	0.327
	(1.182)	(1.938)	(0.241)	(1.916)
Liability Discount Rate	0.883***	0.431	0.117**	0.335
	(0.197)	(0.302)	(0.056)	(0.240)
Minimum Funding Requirement	−0.136	0.017	−0.161*	0.007
	(0.111)	(0.177)	(0.090)	(0.074)
Risk-based Capital Requirements	−7.322***	−6.435***	−1.151*	0.264
	(2.229)	(2.282)	(0.606)	(0.694)
Recovery Period	0.138***	0.440***	−0.047**	−0.254***
	(0.049)	(0.092)	(0.021)	(0.086)
Maturity	−0.018	−0.097**	0.017***	0.062***
	(0.024)	(0.038)	(0.005)	(0.020)
Inflation Indexation	0.007	−0.018**	0.002	0.022***
	(0.007)	(0.008)	(0.003)	(0.005)
Size	0.206***	0.002	0.040***	0.164***
	(0.023)	(0.041)	(0.005)	(0.029)
Past Investment Return	0.032*	0.065*	−0.006	−0.027
	(0.018)	(0.037)	(0.006)	(0.028)
PBGC	13.262	−0.970	15.803*	−1.570
	(10.611)	(16.908)	(8.674)	(7.187)
Observations	3,687	3,687	3,687	3,687
R^2	0.708	0.685	0.692	0.731
Adjusted R^2	0.661	0.634	0.642	0.687
Residual Std. Error (df = 3170)	6.365	7.372	1.593	5.481

Note: *p<0.1; **p<0.05; ***p<0.01

1.E Double-Clustered Standard Errors

Table 1.9: Pension Funds Risky Asset Allocation and Regulation: Double-clustered Standard Errors

This table presents the estimates of the regression of the funds' percentage allocation to risky assets on regulatory requirements and fund characteristics, with fund fixed effects (Equation (1.3.1)). Heteroskecasticity robust standard errors that are clustered by Year and Fund are in parentheses.

	Dependent variable:			
	All	Equities	Risky FI	Alt
	(1)	(2)	(3)	(4)
Quantitative Investment Restrictions	−0.023	0.010	−0.002	−0.031*
	(0.015)	(0.022)	(0.004)	(0.016)
Mark-to-market Asset Valuation	−6.612***	−6.037*	−0.342	−0.233
	(2.000)	(3.432)	(0.646)	(1.980)
Recognition of Unfunded Liabilities	−1.042	−1.921	0.553	0.325
	(1.513)	(2.217)	(0.338)	(2.071)
Liability Discount Rate	0.878***	0.432	0.110	0.336
	(0.245)	(0.357)	(0.071)	(0.283)
Minimum Funding Requirement	−0.016	0.008	−0.017	−0.007
	(0.023)	(0.033)	(0.016)	(0.019)
Risk-based Capital Requirements	−7.118***	−6.450**	−0.908	0.240
	(2.455)	(3.016)	(0.713)	(1.237)
Recovery Period	0.169***	0.437***	−0.011	−0.258***
	(0.058)	(0.094)	(0.014)	(0.090)
Maturity	−0.020	−0.096*	0.014*	0.062*
	(0.032)	(0.054)	(0.008)	(0.037)
Inflation Indexation	0.007	−0.018	0.003	0.022**
	(0.012)	(0.015)	(0.006)	(0.010)
Size	0.205***	0.003	0.038***	0.164***
	(0.039)	(0.050)	(0.010)	(0.062)
Past Investment Return	0.032*	0.065*	−0.007	−0.027
	(0.017)	(0.037)	(0.006)	(0.027)
Observations	3,687	3,687	3,687	3,687
R^2	0.708	0.685	0.679	0.731
Adjusted R^2	0.660	0.634	0.627	0.687
Residual Std. Error (df = 3171)	6.370	7.371	1.627	5.480

Note: *p<0.1; **p<0.05; ***p<0.01

CHAPTER 2

Longevity Risk: To Bear or to Insure

2.1 Introduction

Longevity risk is a looming threat to pension systems worldwide. In contrast to mortality risk, which is the idiosyncratic risk surrounding an individual's actual date of death given known survival probabilities, longevity risk is the risk of misestimating future survival probabilities.[1] This systematic risk can be distressful for retirement financing because longevity-linked assets are not yet commonplace (Tan et al., 2015).

The global transition of funded pensions from Defined Benefit (DB) to Defined Contribution (DC) plans[2] precipitates the need for sustainable means of managing mortality and longevity risks, which have conventionally been borne by the DB plan sponsor. The essence of a DC setup grants individuals full freedom in managing their retirement capital, which is accumulated at a statutory rate of saving.

This chapter is based on Boon et al. (2017b).

[1]Longevity and mortality risks are also referred to as macro- and micro-longevity risks respectively.

[2]In 1975, close to 70% of all US retirement assets were in DB plans. In 2015, DB assets accounted for only 33% of total retirement assets. Over the same period, assets in DC plans and Individual Retirement Accounts (IRAs) grew from 20% to 59% (Investment Company Institute, 2016). In the UK, 98% of the FTSE 350 companies offer a DC pension plan in 2017 (Towers Watson, 2017).

While the optimal, rational individual response to mortality risk in a frictionless setting is to pool that risk (Yaari, 1965; Davidoff et al., 2005; Reichling and Smetters, 2015), the corresponding response to longevity risk is less evident. Individuals could either bear it under a collective arrangement, or offload it at a cost by purchasing an annuity contract from an equity-backed insurance company. Both options allow individuals to pool mortality risk, but entail different implications with regard to longevity risk. We compare these arrangements to ascertain the option that maximizes individuals' expected utility. We also investigate the viability of the annuity market by evaluating the risk-return tradeoff with respect to longevity risk for the equityholders of the annuity contract provider.

Since the introduction of Group Self-Annuitization (GSA) by Piggott et al. (2005), retirement schemes in which individuals bear systematic risks as a collective, but pool idiosyncratic ones have captured the attention of scholars. The main novelty of our work is to concurrently model individual preferences and the business of an equityholder-backed annuity provider when longevity risk exists. Despite equityholders' critical role in the provision of contracts, comparisons of the GSA and annuity contracts that include longevity risk disregard this aspect (e.g., Denuit et al., 2011; Richter and Weber, 2011; Maurer et al., 2013; Qiao and Sherris, 2013).

In order to credibly offer insurance against a systematic risk, the annuity provider requires reserve capital that is constituted from either equity contribution and/or from contract loading to absorb unexpected shocks.[3] Either of these sources of reserve has a cost. If the annuity provider solicits capital from equityholders, then it would have to compensate equityholders with a longevity risk premium. If the provider charges too high a loading, then individuals would prefer the GSA over the annuity contract (e.g., Hanewald et al., 2013; Boyle et al., 2015).[4] Therefore, the existence of an annuity market hinges on the provider's

[3]It would be equivalent to consider debt issuance to raise capital, and any dividend policy other than a one-off dividend payment to equityholders (i.e., any gains before the end of the investment horizon are re-invested). This is because the Miller-Modigliani propositions on the irrelevance of capital structure (Modigliani and Miller, 1958) and dividend policy (Miller and Modigliani, 1961) on the market value of firms hold in our setup, which excludes taxes, bankruptcy costs, agency costs, and asymmetric information.

[4]While allocating retirement wealth between the annuity contract and the collective scheme is conceptually appealing, for the feasibility of a collective scheme, individuals can select only one option in our setting (e.g., mandatory participation in a collective scheme averts adverse selection, achieves cost reduction, etc., Bovenberg et al., 2007). Weinert and Gründl

ability to set a contract price such that all stakeholders are willing to participate in the market.

Existing estimates on individuals' willingness to pay to insure against longevity risk are low. Individuals are willing to offer a premium of between 0.75% (Weale and van de Ven, 2016) to 1% (Maurer et al., 2013) for an annuity contract that insures them against longevity risk without default risk. In contrast, the capital buffer that the annuity provider would have to possess to restrain its default risk is much larger (e.g., about 18% of the contract's best estimate value to limit the default rate to 1% in Maurer et al., 2013). These estimates suggest that the annuity provider has little capacity to compose its reserve capital only from contract loading, as is commonly assumed (Friedberg and Webb, 2007; Richter and Weber, 2011; Maurer et al., 2013; Boyle et al., 2015). Equity capital is thus necessary. We attempt to reconcile the gap between the maximum loading that individuals are willing to pay, and the minimum capital necessary to provide annuity contracts that individuals are willing to purchase, by introducing equityholders.

While analyses that incorporate both policy and equityholders exist in insurance (e.g., Filipović et al., 2015; Chen and Hieber, 2016), they are unforeseen in the literature on the comparison of the GSA with annuity contracts, which focuses on policyholders only. An exception is Blackburn et al. (2017), who take the equityholders' viewpoint when investigating longevity risk management and the share value of a life annuity provider. Demand for annuities in their model is determined by an exogenous demand function. Instead, we analyze the policy and equityholders concurrently when annuity demand is endogenous.

Consistent with the inchoate market for longevity-hedging instruments, we assume that the annuity provider has no particular advantage in bearing longevity risk.[5] Moreover, the annuity provider is required to maintain the value of its assets above the value of its liabilities–a plausible regulatory requirement for such a for-profit entity. In contrast to the literature on collective schemes, which largely focuses on inter-generational risk-sharing (e.g., issues concerning its fair-

(2017) analyze the optimal share of a default-free nominal annuity and a tontine, a type of collective scheme, whereas Zhang and Li (2017) investigate a contract that is partially-indexed to longevity risk, that similarly explores a risk-sharing spectrum between the contract provider and the individual.

[5]Insurance companies may in practice have a comparative advantage in bearing longevity risk, such as relying on the synergy of product offerings in terms of risk-hedging (Tsai et al., 2010), or the potential of life insurance sales in hedging longevity risk (i.e., natural hedging) (Cox and Lin, 2007; Luciano et al., 2015).

ness and stability with respect to the age groups, see Gollier, 2008; Cui et al., 2011; Beetsma et al., 2012; Chen et al., 2017b, 2016), we focus instead on risk-sharing between individuals and the annuity provider's equityholders within a generation.

We begin by assuming that the annuity provider composes its buffer entirely from equity capital. In return for their capital contribution, equityholders receive the annuity provider's terminal wealth as a lump sum dividend. Due to equity-capital-cushioning, the annuity contract provides retirement benefits that have a lower standard deviation across scenarios. However, as equity capital is finite, there is a positive (albeit small) probability that the annuity provider defaults. We infer the maximum loading that individuals are willing to offer, and equityholders' risk-adjusted investment return.

We find that individuals marginally prefer the collective scheme. The Certainty Equivalent Loading (CEL), i.e., the level of loading on the annuity contract at which individuals would derive the same expected utility under either option, is slightly negative (i.e., -0.35% to -0.052%; Table 2.3). Furthermore, exposure to longevity risk does not enhance the equityholders' risk-return tradeoff if the annuity provider sells zero-loading contracts, because it yields only half of the Sharpe ratio of an identical investment without exposure to longevity risk, as well as a negative Jensen's alpha (Table 2.4). Consequently, the annuity contract would not co-exist with the collective scheme. The implication of our results would be even stronger if there were frictional costs, e.g., financial distress, agency, regulatory capital, and double taxation costs, because equityholders would require a higher financial return from the capital they provide.

To further comprehend the tradeoff that an individual faces when selecting a contract, we carry out sensitivity tests with respect to the individual's characteristics, longevity risk, and the annuity provider's default risk. Our inference is robust to the deferral period (Table 2.7), stock exposure (Table 2.8), and parameter uncertainty surrounding the longevity model's time trend (Table 2.11). Situations characterized by extremities can intensify individual preference for either contract in an intuitive manner. For instance, the annuity contract is attractive to highly risk-averse individuals because its retirement benefits are less volatile (Table 2.6). If the equity capital is halved, the annuity provider's default risk rises markedly, and the annuity contract becomes less desirable to individuals (Table 2.9).

Greater uncertainty surrounding longevity evolution generates preference for the annuity contract, on the condition that the annuity provider restrains default risk by raising more equity capital. If, for example, the standard deviation of the longevity model's time trend is doubled, risk-averse individuals are willing to pay as much as 3.2% in loading for the annuity contract, but only if the provider has no default risk (Table 2.13). Under an alternative longevity model, which exhibits wider variation of survival probabilities at older ages, risk-averse individuals prefer the collective scheme, but only if the provider's default risk is eliminated too (Table 2.15). Despite any positive loading that individuals offer, none of the cases that we analyze show that the level of loading is sufficient to compensate equityholders (Tables 2.13 and 2.15). This is because in situations of heightened longevity risk, the equityholders' dividend is also more volatile, which compromises the financial performance of longevity risk exposure. Thus, there is no compelling support for annuity contract provision when individuals could form a collective scheme.

We present our model in Section 2.2 and calibrate it in Section 2.3. We first discuss the baseline case results from the individual's perspective (Section 2.4), then from the equityholders' point of view (Section 2.5). Section 2.6 is devoted to sensitivity tests on the individuals' traits, stock exposure, the annuity provider's leverage ratio, as well as the longevity model's attributes. We conclude in Section 2.7.

2.2 Model Presentation

We devise a model to investigate the welfare of individuals under a collective retirement scheme and a market-provided deferred variable annuity contract. The setting comprises a financial market with a constant risk-free rate and stochastic stock index, homogenous individuals with stochastic life expectancies, and two contracts for retirement.[6] We define and discuss these elements in detail in this section.

2.2.1 Financial Market

In a continuous-time financial market, the investor is assumed to be able to invest in a money market account and a risky stock index. The financial market is incomplete due to the lack of longevity-linked securities. We assume that annual returns to the risk-free asset are constant, r. The money market

[6] We abstract from model uncertainty by assuming that the stochastic dynamics underlying the financial assets and life expectancies are known.

account is fully invested in the risk-free asset.

The value of the stock index at time t, which is denoted by S_t, follows the diffusion process, $dS_t = S_t(r + \lambda_S\sigma_S)\,\mathrm{d}t + S_t\sigma_S\,\mathrm{d}Z_{S,t}$. Z_S is a standard Brownian motion with respect to the physical probability measure, σ_S is the instantaneous stock price volatility, and $\lambda_S\sigma_S$ is the constant stock risk premium.

2.2.2 Individuals

At time t_0, individuals who are aged $x = 25$ either form a collective scheme or purchase a deferred annuity contract with a lump sum capital that is normalized to one. Both retirement contracts commence retirement benefit payments at age 66, up to the maximum age of 95, conditional on the individual's survival. Individuals' lifespan is determined by survival probabilities that follow the Lee and Carter (1992) model.

2.2.2.1 Life Expectancy

We assume that individual mortality rates evolve independently from the financial market. Although productive capital falls as the population ages, empirical evidence on the link between demographic structure and asset prices is mixed.[7]

We adopt the Lee and Carter (1992) model, which is widely used (e.g., by the US Census Bureau and the US Social Security Administration) and studied. This is a one-factor statistical model for long-run forecasts of age-specific mortality rates. It relies on time-series methods and is fitted to historical data. The log central death rate for an individual of age x in year t, $\log(m_{x,t})$[8] is assumed to linearly depend on an age-specific constant, and an unobserved period-specific intensity index, k_t:

$$\log(m_{x,t}) = a_x + b_x k_t + \epsilon_{x,t} \tag{2.2.1}$$

$\exp(a_x)$ is the general shape of the mortality schedule across age; b_x is the rate of change of the log central death rates in response to changes in k_t, whereas

[7] Erb et al. (1994); Poterba (2001); Ang and Maddaloni (2003); Visco (2006); Schich (2008b); Arnott and Chaves (2012)

[8] $m_{x,t}$ is the ratio of $D_{x,t}$, the number of deaths of an individual aged x in year t, over $E_{x,t}$, the exposure, defined as the number of aged x individuals who were living in year t.

the error term, $\epsilon_{x,\,t}$, is normally distributed with zero mean and variance σ_x^2.

The Lee and Carter (1992) model is defined for the central death rates, $m_{x,\,t}$, but we apply it to model the annual rate of mortality, $q_{x,\,t}$ by the approximation $q_{x,\,t} \simeq 1 - \exp\left(-m_{x,\,t}\right)$. The probability that someone who is aged x at time t_0 is alive in s-year time, ${}_sp_x$, is then ${}_sp_x = \Pi_{l=0}^{s-1}\left(1 - q_{x+l,\,t+l}\right)$. We denote the conditional probability in year $t \geq t_0$ that an individual of age x at time t will survive for at least s more years as ${}_sp_x^{(t)}$, ${}_sp_x^{(t)} = \Pi_{l=0}^{s-1}\left(1 - q_{x+l,\,t}\right) = \exp\left(\sum_{l=0}^{s-1} -m_{x+l,\,t}\right)$.[9]

While many refinements of Lee and Carter (1992) exist (e.g., the two-factor model of Cairns et al., 2006b, the addition of cohort effects in Renshaw and Haberman, 2006), the model is not only reasonably robust to the historical data used, but also produces plausible forecasts that are similar to those from extensions of the model (Cairns et al., 2011).

2.2.2.2 Welfare

Individuals maximize expected utility in retirement.[10] Benefits from the retirement contracts constitute the individual's only source of income. We consider individuals who exhibit Constant Relative Risk Aversion (CRRA), and evaluate their utility in retirement by Equation (2.2.2).

$$U\left(\Xi\right) = \int_{t_R}^{T} \mathrm{e}^{-\beta(\mathrm{t}-\mathrm{t}_0)} \frac{\Xi_t^{1-\gamma}}{1-\gamma} \, {}_{t-t_0}p_{25} \, \mathrm{d}t \tag{2.2.2}$$

$$\begin{aligned}
{}_{t-t_0}p_{25} &= \text{probability that someone who is 25 years old in year } t_0 \\
& \text{ is alive in year } t \geq t_0 \\
\beta &= \text{subjective discount factor} \\
\gamma &= \text{risk aversion parameter, } \gamma > 1 \\
\Xi_t &= \text{retirement income in year } t \\
t_R &= \text{retirement year, } t_R = t_0 + 66 - 25 \\
T &= \text{year of maximum age, } T = t_0 + 95 - 25
\end{aligned}$$

[9]This is an exponentiated finite sum of log-normal random variables that has no known analytical distribution function. Therefore, we resort to simulation for our analysis. Alternate ways to proceed include quantile estimation of random survival probabilities in Denuit et al., 2011, or the Taylor series approximation by Dowd et al. (2011).

[10]We can ignore bequest motives as both contracts provide income only when the individual is still alive.

2.2.3 Contracts for Retirement

There are two retirement contracts. The first is a collective pension called the Group Self-Annuitization (GSA) scheme. The second is a Deferred Variable Annuity (DVA) contract offered by an annuity provider who is backed by equityholders. We describe both contracts in this section. Appendix 2.A elaborates on the rationale of the definition and provision of the contracts.

The contracts specify the distribution of financial and longevity risks between the stakeholders. As the contracts are intended to underscore longevity risk, both treat stock market risk identically–the risk is fully borne by the individuals. The benefits due, henceforth known as entitlements, are fully indexed to the same underlying financial portfolio called the reference portfolio (e.g., a portfolio that is 20% invested in the stock index, and 80% in the money market account). Thus, if the DVA provider adopts the reference portfolio's investment policy, the provider is hedged against financial market risk.

Longevity risk distribution, however, distinguishes the two contracts. Under the GSA, it is shared equally among individuals. Under the DVA, the risk is borne by equityholders up to a limit implied by their equity contribution, beyond which the DVA provider defaults. Both contracts stipulate to distribute mortality credit according to the survival probabilities, conditional on the date of contract sale. The DVA provider's equityholders bear the risk that the survival probability forecast deviates from the realized values. The provider uses its equity capital to finance underestimation of longevity, and disburses any surplus arising from overestimation of longevity to its equityholders as a dividend.

Due to the non-existence of financial assets that are associated with longevity risk, the risk cannot be hedged by the DVA provider. Additionally, we assume that the number of individuals who either purchase the DVA or participate in a GSA is large enough such that by the Law of Large Numbers, the proportion of surviving individuals within each pool coincides with that implied by the realized survival probabilities, so we can eliminate mortality risk.[11]

2.2.3.1 Deferred Variable Annuity (DVA)

The DVA contract is parametrized by an actuarial construct called the assumed interest rate (AIR), $h = \{h(t)\}_{t=t_0}^{T}$. The AIR is a deterministic rate that determines the cost, A, of a contract sold to an individual who is aged x

[11]The GSA in our setting is a specific case of the GSA in Piggott et al. (2005). We omit mortality risk, unlike Piggott et al. (2005) who consider the pooling of this idiosyncratic risk.

at time t_0 as follows:

$$A(h,\, F,\, t_0,\, x) = (1+F) \int_{t=t_R}^{T} {}_{t-t_0}p_x^{(t_0)} \exp\left(-h(t) \times (t - t_R)\right) \mathrm{d}t \tag{2.2.3}$$

$$\begin{aligned}
{}_{t-t_0}p_x^{(t_0)} &= \text{conditional probability in year } t_0 \\
&\quad\ \text{that an individual of age } x \text{ lives} \\
&\quad\ \text{for at least } t - t_0 \text{ more years} \\
h &= \text{AIR} \\
F &= \text{loading factor} \\
t_R &= \text{retirement year}
\end{aligned}$$

The loading factor, F, is a proportional one-off premium that the DVA provider attaches to a contract. A contract that is priced at its best estimate has a loading factor of zero, $F = 0$.

The DVA contract is indexed to a reference investment portfolio that follows a deterministic investment policy, $\theta \equiv \{\theta_t\}_{t=t_0}^{T}$. θ_t is the fraction of portfolio wealth allocated to the risky stock index at time t, while the remaining $1 - \theta_t$ is invested in the money market account. Let $W_t^{Ref}(\theta)$ be the value of the reference portfolio at time t. The dynamics of the reference portfolio are thus $\mathrm{d}W_t^{Ref} = W_t^{Ref}(r + \theta_t \lambda_S \sigma_S)\, \mathrm{d}t + W_t^{Ref} \theta_t \sigma_S\, \mathrm{d}Z_{S,t}$.

Using an annuitization capital that is normalized to one, the individual purchases $A(h,\, F,\, t_0,\, x)^{-1}$ unit(s) of DVA contract(s), and is entitled to Ξ, for every year t in retirement, $t_R \leq t \leq T$.[12]

$$\begin{aligned}
\Xi(h,\, F,\, t,\, x) &= \frac{\exp\left(-h(t) \times (t - t_R)\right)}{A(h,\, F,\, t_0,\, x)} \frac{W_t^{Ref}(\theta)}{W_{t_0}^{Ref}(\theta)} \qquad (2.2.4) \\
W_t^{Ref}(\theta) &= \text{value of the reference portfolio at time } t
\end{aligned}$$

The AIR influences the expectation and dispersion of the benefit payments over time. For instance, the fund units are front- (back-) loaded (i.e., due in the earlier (later) years of retirement) under a higher (lower) AIR.[13]

[12]The benefits adjust instantaneously with the value of the portfolio to which the contract is indexed. Maurer et al. (2016) make the case for smoothing of the benefits, which is advantageous to both the policyholder and the contract provider.

[13]Let $\tilde{r}$ denote the reference portfolio's expected return, and suppose h is time-invariant. Then an annuity contract with $h = \tilde{r}$ has a constant expected benefit payment path. When

We demonstrate in Appendix 2.A that for any given θ, the AIR that maximizes the individual's expected utility in retirement is Equation (2.2.5), which we refer to as the optimal AIR, h^*. h^* depends on the individual's preference and financial market parameters. It serves as the AIR of both the DVA and GSA.

$$h^*(t, \theta_t) = r + \frac{\beta - r}{\gamma} - \frac{1-\gamma}{\gamma}\theta_t \sigma_S \left(\lambda_S - \frac{\gamma \theta_t \sigma_S}{2}\right) \tag{2.2.5}$$

The DVA provider merely serves as a distribution platform for annuity contracts. It acts in the best interest of its equityholders, who outlive the individuals. The equityholders provide a lump sum capital that is proportional to the value of its estimated liabilities in the year t_0.[14] At every date $t \geq t_0$, the DVA provider's asset value has to be at least equal to the value of its estimated liabilities. In any year $t_0 \leq t \leq T$, if the DVA provider fails to meet the 100% solvency requirement, then the DVA provider defaults. Regulatory oversight is introduced for the DVA provider, because as a for-profit entity, the DVA provider may have an incentive to take excessive risk at the individuals' expense (Filipović et al., 2015). We impose a solvency constraint as it is not only the norm in regulatory regimes for insurers (e.g., Solvency II in the European Union), but is also shown to be effective in mitigating risk-shifting (Filipović et al., 2015).

In every year of retirement, the individual receives a benefit that is equal to the DVA entitlement,

$$\Xi^{DVA}(h^*, F, t, x) = \Xi(h^*, F, t, x) \tag{2.2.6}$$

conditional on the individual's survival and the DVA provider's solvency. $\Xi(.)$ is Equation (2.2.4) while h^* is Equation (2.2.5).

In the event of default, the residual wealth of the DVA provider is distributed among all living individuals, in proportion to the value of their contracts that remains unfulfilled. Equityholders receive none of the residual wealth. We impose a resolution mechanism that obliges individuals to use the provider's liquidated wealth to purchase an equally-weighted portfolio of zero-coupon bonds, of maturities from the year of default if the individual is already retired, or from the year of retirement, until the year of maximum age. Assuming that the bond issuer

$h < \tilde{r}$, then the expected benefit stream is upward sloping, with increasing variance as the individual ages. Conversely, when $h > \tilde{r}$, the expected benefit stream is downward sloping, and the variance is higher during the initial payout phase. Horneff et al. (2010) provide an exposition on retirement benefits under numerous AIRs and reference portfolios.

[14] The estimation of the value of liabilities is explained in Appendix 2.B.

poses no default risk, then the individual has a guaranteed income until death, but receives no mortality credit. If the individual dies before the maximum age, the face value of the bonds that mature subsequently is not bequeathed. This resolution to insolvency is harsh on the individuals because it eliminates the mortality credit, but it reflects the empirical evidence that individuals substantially discount the value of an annuity that poses default risk (Wakker et al., 1997; Zimmer et al., 2009).

2.2.3.2 Group Self-Annuitization (GSA)

Similar to the DVA, the GSA is parameterized by the optimal AIR, h^*, and is indexed to a reference portfolio with the investment policy θ. The aged-x individual receives $A\left(h^*,\, 0,\, t,\, x\right)^{-1}$ contract(s) for every unit of contribution at time t. In any year $t \geq t_R$, the GSA's entitlement depends on the reference portfolio's value at time t, $W_t^{Ref}\left(\theta\right)$.

The description of the GSA thus far is identical to a DVA contract with zero loading, $F = 0$. The GSA's distinctive feature is that the entitlements are adjusted according to its funding status. Let the funding ratio at time t, FR_t, be the ratio of the GSA's value of assets, taking into account the investment return from the preceding year, over the best estimated value of its liabilities.[15] For any year t in retirement, $t_R \leq t \leq T$, the individual is entitled to $\Xi^{GSA}(h^*,\, 0,\, t,\, x)$.

$$\begin{aligned} \Xi^{GSA}(h^*,\, 0,\, t,\, x) &= \Xi(h^*,\, 0,\, t,\, x) \times \frac{FR_t}{1} \qquad (2.2.7) \\ &= \frac{\exp\left(-h^*\left(t,\, \theta_t\right) \times \left(t - t_R\right)\right)}{A\left(h^*,\, 0,\, t_0,\, x\right)} \frac{W_t^{Ref}\left(\theta\right)}{W_{t_0}^{Ref}\left(\theta\right)} FR_t \\ FR_t &= \text{Funding Ratio in year } t \end{aligned}$$

The first two terms of Equation (2.2.7) are identical to the entitlement for a DVA contract with zero loading, Equation (2.2.4). The final term of Equation (2.2.7) represents the adjustment. If FR_t is smaller (larger) than 1, then the GSA entitlement, Ξ^{GSA}, is lower (higher) than the DVA entitlement, Ξ^{DVA}, in year t. Equation (2.2.7) ensures that the GSA is 100% funded in any year.

[15] Estimation of the GSA liabilities is identical to the estimation of liabilities of the DVA provider. See Appendix 2.B for details.

2.3 Model Calibration

We consider three groups of individuals, distinguished by their risk aversion levels, $\gamma = 2$, 5, and 8.[16] Individuals are otherwise homogenous. They have an annual subjective discount factor of 3%,[17] are aged 25 at time $t_0 = 0$, and use a lump sum that is normalized to one, to either purchase DVA, or join the GSA at time t_0. Both contracts stipulate payment of annual retirement benefits from age 66 until age 95, conditional on the individual's survival in any year, according to the contract specification in Section 2.2.3.

The portfolio to which the DVA and GSA are indexed is either fully invested in the money market account ($\theta = 0$), or 20% invested in equities and 80% in the money market account ($\theta = 20\%$). These allocations yield the optimal AIR range of 3-4% (Table 2.1) that is not only observed in the annuity market (Brown et al., 2001), but also typically considered in the related literature (Koijen et al., 2011; Maurer et al., 2013). In Section 2.6.3, we explore alternative investment policies and demonstrate that they uphold the same results as when $\theta = 0$, 20%.

We assume that the DVA provider's equityholders provide a lump sum capital at date t_0 that is 10% of the contract's best estimate price. The level of equity capital contribution is set such that the annuity provider's leverage ratio is 90%. It reflects the average leverage ratio of US life insurers between 1998-2011.[18]

To provide descriptive calculations on individual welfare under the GSA and the DVA, we calibrate the financial market and life expectancy models to US data. These parameters constitute our baseline case.

[16]Using survey responses from the Health and Retirement Study on the US population, Kimball et al. (2008) estimate that the mean risk aversion level among individuals is 8.2, with a standard deviation of 6.8.

[17]While field experiments reveal a wide range of implied subjective discount factor (e.g., see Table 1 in Frederick et al., 2002), we choose a value that is commonly adopted in welfare analysis. For example, in similar analyses on retirement income, Feldstein and Ranguelova (2001) and Hanewald et al. (2013) adopt a subjective discount factor of around 2%.

[18]Leverage Ratio $\equiv 1 -$ Value of Equity/Value of Assets. Based on the A.M. Best data used in Koijen and Yogo (2015), the leverage ratio of US life insurers between 1998 to 2011 is 91.36% on average. Assuming that assets are composed of premium and equity capital only, and normalizing Premium $= 1$, we have Leverage Ratio $= 1 -$ Equity/ $(1 +$ Equity), which we use to solve for Equity when the Leverage Ratio $\approx 90\%$.

2.3.1 Financial Market

We adopt a constant risk-free rate of $r = 3.6\%$. The stock index has an annualized standard deviation of $\sigma_S = 15.8\%$, and an instantaneous Sharpe ratio of $\lambda_S = 0.467$. This implies that the stock risk premium is $\lambda_S \sigma_S = 7.39\%$. These parameters reflect the performance of the market-capitalization-weighted index of US stocks and the yield on the three-month US Treasury bill over the recent past.

2.3.2 Life Expectancy

We estimate the Lee and Carter (1992) model using US female death counts, $D_{x,t}$, and the population's exposure to risk, $E_{x,t}$, from 1980 to 2013 from the Human Mortality Database.[19] The mortality rate for age group x in year t is $D_{x,\,t}/E_{x,\,t}$. By relying on population mortality data, we eschew adverse selection that plagues the annuity market, i.e., the individuals who purchase an annuity typically have a longer average lifespan than the general population (Mitchell and McCarthy, 2002; Finkelstein and Poterba, 2004).

Estimation of the Lee and Carter (1992) model proceeds in three steps. First, k_t is estimated using Singular Value Decomposition. In the second step, a_x and b_x are estimated by Ordinary Least Squares on each age group, x. In the third step, k_t is re-estimated by iterative search to ensure that the predicted number of deaths coincides with the data. For identification of the model, we impose the constraints $\sum_x b_x = 1$ and $\sum_t k_t = 0$.

The estimated model is used for forecasting by assuming an ARIMA(0, 1, 0) time series model for the mortality index k_t.

$$\begin{aligned} k_t &= c + k_{t-1} + \delta_t \\ &\delta \sim \mathcal{N}\left(0, \sigma_\delta^2\right) \end{aligned} \tag{2.3.1}$$

Forecasts of the log of the central death rates for any year t', $t' \geq t$, are given by $\mathbb{E}_t\left[\log\left(m_{x,\,t'}\right)\right] = a_x + b_x \hat{k}_{t'}$, with $\hat{k}_{t'} = (t' - t)\, c + k_t$. The realized log of the mortality rate incorporates the independently and identically normally distributed error terms $\epsilon_x \sim \mathcal{N}\left(0,\, \sigma_x^2\right)$ and $\delta \sim \mathcal{N}\left(0, \sigma_\delta^2\right)$, where $\epsilon_{x,\,t_1}$ and δ_{t_2}

[19] This fitting period is selected using the method of Booth et al. (2002). It involves defining fitting periods starting from the first year of data availability till the last year of data availability, and progressively increasing the starting year. A ratio of the mean deviance of fit of the Lee and Carter (1992) model with the overall linear fit is computed for these fitting periods. The period for which this ratio is substantially smaller than that for periods starting in previous years is chosen as the best fitting period.

are uncorrelated for any t_1, $t_2 \in [t_0, T]$ and x. Therefore, the conditional expected forecast error of $\log(m_{x,t})$ is zero.

We estimate that $\hat{c} = -1.047$, which implies a downward trend for k_t, while the estimate of σ_δ is $\widehat{\sigma_\delta} = 1.744$. In Figure 2.3.1, we present the estimates for a_x, b_x and σ_x. a_x is increasing in age. Estimates for b_x suggest that the change in the sensitivity of age groups to the time trend, k, is not monotone across ages. As for σ_x, it decreases in age non-monotonically until around age 85. With these estimates, 83.8% of the variation in the data is explained.

In Figure 2.3.2, we display a fan plot of the fraction of living individuals by age, between 25 and 95, with the population at age 25 normalized to one. The maximum and minimum realizations have a wide range. At its widest at age 88, the difference is as large as 30%.

Figure 2.3.1: Lee and Carter (1992) Mortality Model Parameter Estimates

The top panel shows the estimates for a_x, the middle panel displays the estimates for b_x, whereas the bottom panel presents the estimates of σ_x, for the Lee and Carter (1992) model as specified by Equation (2.2.1). The calibration sample is the US Female Mortality data from 1980 to 2013, from the Human Mortality Database. The estimate of c is -1.047 and that of σ_δ is 1.744. 83.8% of variation of the sample is explained by these estimates.

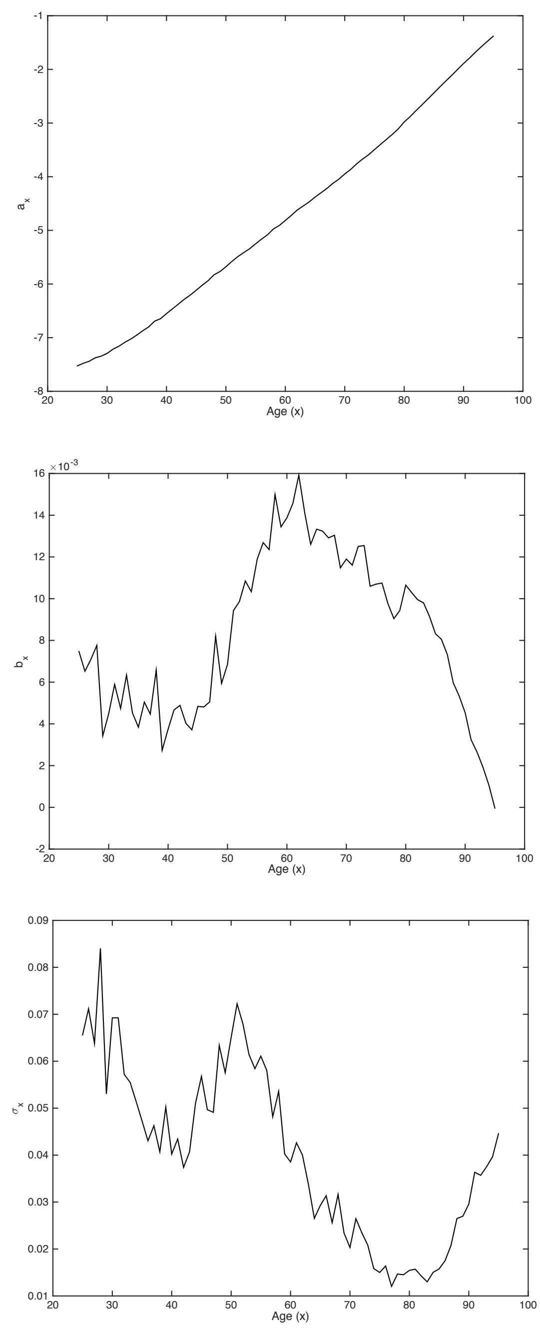
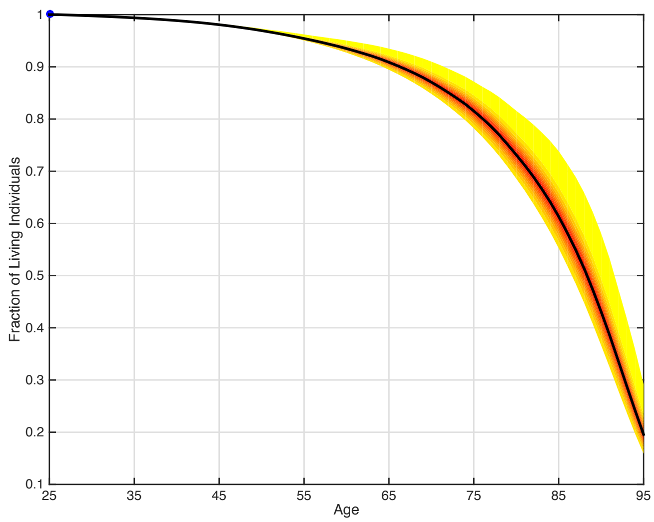


Figure 2.3.2: Lee and Carter (1992)Baseline Case: Fan Plot

This figure presents the fan plot of the simulated fraction of living individuals (i.e., the population of 25-year-olds is normalized to one) over 10,000 replications when longevity is modeled according to Lee and Carter (1992), using estimates in Figure 2.3.1. Darker areas indicate higher probability mass.



2.3.3 Contract Characteristics

In order to develop intuition and grasp the contracts' definition, we discuss the characteristics of the GSA and the DVA under the calibrated parameters. Table 2.1 presents the optimal AIRs as given by Equation (2.2.5), and evaluated at the parameters outlined in Sections 2.3.1 and 2.3.2.

Table 2.1: Baseline Case: Optimal AIR, h^* (%)

This table shows the optimal AIR, Equation (2.2.5), of the DVA and GSA contracts by the individuals' risk aversion parameter, γ. The underlying portfolio to which the contracts are indexed is either 100% invested in the money market account ($\theta = 0$), or 20% in the risky stock index and 80% in the money market account ($\theta = 20\%$).

θ	γ		
(%)	**2**	**5**	**8**
0	3.31	3.50	3.54
20	4.00	4.48	4.48

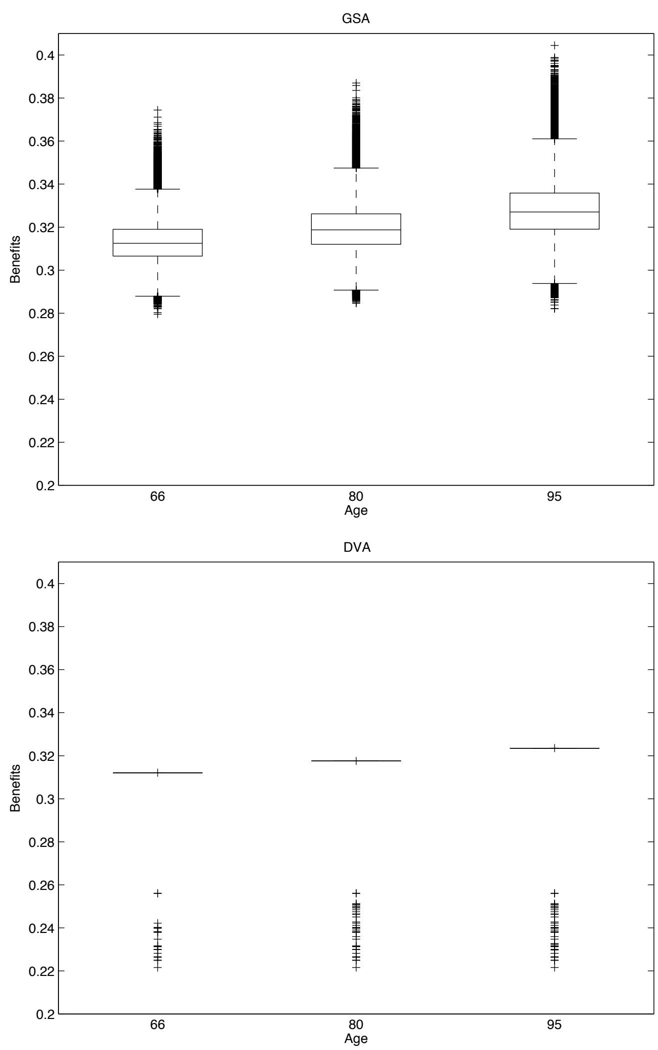
Figure 2.3.3 is a box plot of the benefits that individuals receive under the DVA and the GSA. The median benefits of both contracts grow along the retirement horizon because the optimal AIR is lower than the constant financial market return. For the DVA, the median value is also the maximum, because the surplus from life expectancy misestimates belongs to the equityholders.

The GSA yields more instances of positive than negative adjustments to benefits that are 1.5-time larger than the range between its 75^{th} and 25^{th} percentiles. We infer this from the relative density of "+" symbols above and under the box (Figure 2.3.3, top panel). When the individual attains the maximum age of 95, benefits as large as 25% more than the median could occur. In contrast, in the worse scenario at the same age, the reduction in benefits relative to the median is 12.5% at most. This asymmetric effect on benefits arises from the non-linearity of the Lee and Carter (1992) model. For error terms of the same magnitude (i.e., $\{\epsilon_{x,\,t}\}_{t=t_0}^{T}$ in Equation (2.2.1) and $\{\delta_t\}_{t=t_0}^{T}$ in Equation (2.3.1), for any $x \in \mathbb{Z} \cap [25,\, 95]$), overestimation of the log of the central death rates generates a larger entitlement adjustment than underestimation does. When the DVA provider defaults, the individual is at risk of receiving a much lower benefit. The worst case under the GSA entails up to a 30% lower benefit relative to the median at the maximum age.

The box plots indicate that while both contracts offer comparable benefits at the median, those of the GSA have higher standard deviations across scenarios due to the entitlement adjustments, but upward adjustments are more prevalent than downward ones. The DVA offers less volatile benefits, but is susceptible to severe low benefit outcomes when the provider defaults. These are the main features that the individuals weigh in utility terms.

Figure 2.3.3: Baseline Case: Box Plots of GSA and DVA Benefits

The figure presents the box plot of benefits, for the GSA (top panel), and the DVA (bottom panel), for an individual with a risk aversion level of $\gamma = 5$, at ages 66, 80 and 95. The underlying portfolio is invested in the money market account only. The line in the middle of the box is the median, while the edges of the box represent the 25^{th} and 75^{th} percentiles. The height of the box is the interquartile range, i.e., the interval between the 25^{th} and 75^{th} percentiles. The "+" symbols represent data points that are 1.5 times larger than the interquartile range.



2.4 The Individual's Perspective

We investigate two settings distinguished by the existence of stock market risk. In both, there is longevity risk, but in one instance, there is no investment in the stock market, $\theta = 0$, and so the financial return is constant at r, whereas in the other, $\theta = 20\%$ is invested in the risky stock index while the remaining 80% is allocated to the money market account. All results are based on simulations with 500,000 replications unless specified otherwise. The code that produces all figures and estimates in Sections 2.4 to 2.6 are available from the authors upon request.

2.4.1 Cumulative Default Rate

We measure the DVA provider's default rates with the Cumulative Default Rate, an estimate of the probability that the DVA provider defaults during the individuals' planning horizon.

Let D_t be the indicator function that the DVA provider has defaulted in any year t', $t_0 < t' \leq t \leq T$. For example, if the DVA provider defaults in the year t^*, then $D_t = 1$ for $t \geq t^*$ and $D_t = 0$ for $t < t^*$. Additionally, $D_{t_0} \equiv 0$ because the contracts are sold at their best estimate price, and the equity contribution is non-negative.

The marginal default rate in year t, $d(t)$ is the probability that the annuity provider defaults in year t, conditional on not having defaulted in previous years.

$$\begin{aligned} d(t) &= \text{Marginal Default Rate in year } t \\ &= \frac{\mathbb{E}[D_t]}{1 - \mathbb{E}[D_t]} \end{aligned} \tag{2.4.1}$$

We define the Cumulative Default Rate as

$$\begin{aligned} \text{Cumulative Default Rate} &= 1 - \Pi_{t=t_0}^{T}(1 - d(t)) \qquad (2.4.2) \\ d(t) &= \text{Equation (2.4.1)} \end{aligned}$$

The default rates in the baseline case are at most 0.01% (Table 2.2). As the AIR determines whether the bulk of benefits are due earlier or later in retirement, when combined with the fact that longevity forecast errors are larger at longer horizons, the DVA provider's default rates are inversely related to the AIRs. A higher AIR results in a payment schedule with benefits mostly due earlier in retirement. As such, the longevity estimates are accurate when most of the benefits are paid. Conversely, if the AIR is low, benefit payments are deferred to the end of retirement, when life expectancies are most vulnerable to forecasting errors. Therefore, for a fixed level of equity capital, the DVA provider is less susceptible to defaults when the AIR is higher.[20] For the risk aversion levels $\gamma = 2, 5, 8$, the optimal AIR is increasing in γ (Table 2.1), hence the default rates are decreasing in γ (Table 2.2) for both $\theta = 0, 20\%$. Similarly, the default rates are lower when $\theta = 20\%$ than when $\theta = 0\%$ for all levels of γ because the optimal AIRs are higher under $\theta = 20\%$.

Table 2.2: Baseline Case: Cumulative Default Rates (%)

This table displays the Cumulative Default Rates, Equation (2.4.2), of the DVA provider who sells zero-loading variable annuity contracts with a 40-year deferral period, and has equity capital valued at 10% of the liabilities in the year that the contract was sold. The underlying portfolio to which the DVA and GSA are indexed is either fully invested in the money market account ($\theta = 0$), or 20% in the stock index, and 80% in the money market account ($\theta = 20\%$).

θ	γ		
(%)	**2**	**5**	**8**
0	0.0102	0.0084	0.0082
20	0.0070	0.0038	0.0038

2.4.2 Individual Preference for Contracts

We quantify the individuals' preference for the contracts via the Certainty Equivalent Loading (CEL). This is the level of loading on the DVA (i.e., F in Equation (2.2.3)), that equates an individual's expected utility under the DVA and the GSA. The CEL satisfies Equation (2.4.3). A positive (negative) CEL suggests that the individual prefers the DVA (GSA).

[20]From the regulator's perspective, the notion of an annual probability of default, instead of a cumulative one may be more salient. We explore the "Maximum Annual Conditional Probability of Default", defined as $\max\limits_{\{t = t_0, \ldots, T\}} d(t)$, and find that the maximum annual default rate in the baseline case is 0.0008%. This suggests that the 10% buffer capital is sufficient to restrict default rates of DVA providers who are exposed to only longevity risk to existing regulatory limits (e.g., Solvency II for insurers in Europe).

$$
\begin{aligned}
\mathbb{E}\left[U\left(\frac{\Xi^{DVA}|_{F=0}}{1+CEL}\right)\right] &= \mathbb{E}\left[U\left(\Xi^{GSA}\right)\right] \qquad (2.4.3)\\
\Xi^{DVA}|_{F=0} &= \text{Retirement benefits, Equation (2.2.6),}\\
& \text{of a DVA with zero loading, } F=0\\
\Xi^{GSA} &= \text{Retirement benefits, Equation (2.2.7),}\\
& \text{of a GSA}\\
U(.) &= \text{Utility function, Equation (2.2.2),}
\end{aligned}
$$

Confidence intervals for the CELs are estimated via the Delta Method, for which more details are in Appendix 2.C.

Table 2.3 presents the CEL in the baseline case. The CELs are negative for all risk aversion levels. This implies that individuals prefer the GSA over the DVA, but only marginally. If the DVA contracts were to be sold at a discount of between 0.052% and 0.350%, then individuals would be indifferent between the two contracts. The CEL is increasing in the risk aversion level, γ. This is because more risk-averse individuals have greater preference for the DVA benefits' lower standard deviation across scenarios.

Table 2.3: Baseline Case: Certainty Equivalent Loading (CEL) (%)

This table presents the CEL, Equation (2.4.3), by the risk aversion levels (γ). Individuals aged 25 purchase either the DVA or join the GSA with a lump sum capital normalized to one. The reference portfolio is either fully invested in the money market account ($\theta = 0$), or is $\theta = 20\%$ invested in the stock index and 80% in the money market account. The expected utilities to which the CELs are associated are computed over individuals' retirement between ages 66 and 95. The equityholder's capital is 10% of the present value of liabilities at the date when the contract is sold. The default rates that ensue at this level of equity capitalization are shown in Table 2.2. The 99% confidence intervals estimated by the Delta Method are in parentheses.

θ	γ		
(%)	**2**	**5**	**8**
0	-0.350 [-0.362, -0.339]	-0.200 [-0.211, -0.188]	-0.055 [-0.067, -0.044]
20	-0.349 [-0.361, -0.338]	-0.200 [-0.216, -0.184]	-0.052 [-0.088, -0.016]

2.5 The Equityholders' Perspective

To evaluate the equityholders' risk-return tradeoff on longevity risk exposure, we consider the Sharpe ratio and the Jensen's alpha of providing capital to the annuity provider, against those of investing the same amount of capital in the reference portfolio over the same time period.[21] As in Section 2.4, the annuity provider offers contracts at zero loading.

Equityholders contribute 10% of the DVA provider's best estimate value of liabilities at time t_0, and receive the terminal wealth of the annuity provider, $W_T^{(A)}$, as a dividend. When the value of liabilities is normalized to one, the continuously compounded annualized return of capital provision in excess of the risk-free rate of return is $R^{(A_{exs})} = \log\left(W_T^{(A)}/0.1\right)/(T-t_0) - r$. We evaluate the equityholders' profitability via the Sharpe ratio, $SR = \mathbb{E}\left[R^{(A_{exs})}\right]/\sigma^{(A_{exs})}$, and we compute the Sharpe ratio's confidence intervals in accordance with Mertens (2002).

The Jensen's alpha, α, is given by Equation (2.5.1) (Jensen, 1968).

$$R^{(A_{exs})} = \alpha + \beta R^{(S_{exs})} + u \tag{2.5.1}$$

$R^{(S_{exs})}$ is the annualized return of the stock index in excess of the return on the money market account, and u is the error term. We estimate Equation (2.5.1) by Ordinary Least Squares. α assesses the investment performance of providing capital to the annuity provider, relative to that of the market portfolio, on a risk-adjusted basis. A positive α suggests that longevity risk exposure enhances the equityholders' risk-return tradeoff. When $\theta = 0$, $\beta = 0$ due to the assumption that the mortality evolution is uncorrelated with the financial market dynamics. If in Equation (2.5.1), $R^{(A_{exs})}$ is replaced by the annualized return in excess of the risk-free rate of return for the reference portfolio, then $\alpha = 0$ and $\beta = \theta$. This is because the reference portfolio has identical financial market risk exposure as capital provision, but is not exposed to longevity risk.

[21] The stochastic discount factor, $\{M_t\}_{t=t_0}^{T}$, that follows $\mathrm{d}M_t/M_t = -r\,\mathrm{d}t - \lambda_S\,\mathrm{d}Z_{S,\,t}$, allows us to price any contingent claim exposed to stock market risk only: If X_t is a (random) cash flow generated by a contingent claim at time t, then its price at time t_0 is $\mathbb{E}_{t_0}\left[\int_{t=t_0}^{T}(M_t/M_{t_0})\,X_t\,\mathrm{d}t\right]$. However, when such pricing is carried out for claims due on a long horizon, and the market price of stock risk (i.e., the Sharpe ratio) exceeds its volatility, the price depends on extreme sample paths along which the claim's return explodes (Martin, 2012). As the claims are susceptible to severe underpricing when the Monte Carlo replication sample size is small, we refrain from valuing contingent claims when comparing the equityholders' investment opportunities.

When $\theta = 0$, the annualized excess return of capital provision is between -0.008 and -0.007%, and the standard deviation is 3.9% (Table 2.4, top panel). Relative to the zero excess return from investing in the money market account, equity capital provision is inferior, but the difference is economically insignificant. When $\theta = 20\%$, investing in the DVA provider yields an expected excess return of 1.44% (Table 2.4, bottom panel). This is of no material difference with the expected excess return on the identical financial market portfolio, i.e., $\theta\lambda_S\sigma_S - \theta^2\sigma_S^2/2 = 1.43\%$ when $\theta = 20\%$. However, the standard deviation of excess returns is considerably higher when equityholders are exposed to longevity risk (i.e., $\approx 5\%$, Table 2.4, bottom panel), than when their investment is subject to stock market risk only (i.e., $\theta\sigma_S = 3.17\%$ with $\theta = 20\%$). Consequently, investing in the financial market only is associated with a Sharpe ratio that is around 50% higher than the Sharpe ratio of providing capital to the DVA provider (i.e., 0.29 in Table 2.4, bottom panel, as compared to $\lambda_S - \theta\sigma_S/2 = 0.45$[22] when $\theta = 20\%$). Thus, if equityholders were risk-neutral, then the excess returns imply that they would be indifferent between either investment opportunity. If equityholders were risk averse, then by the Sharpe ratio, investing in longevity risk worsens the equityholders' risk-return tradeoff when the annuity provider sells the contracts at zero loading. The negative Jensen's alpha of -0.0001 corroborates this inference. Any positive loading is infeasible, because it intensifies individuals' preference for the GSA. Therefore, the annuity provider is incapable of adequately compensating its equityholders for exposure to longevity risk.

[22] This is the discrete Sharpe ratio, which is the parameter we estimate using simulation replications, as opposed to the instantaneous Shape ratio, λ_S (Nielsen and Vassalou, 2004).

Table 2.4: Baseline Case: Equityholders' Investment Performance Statistics

This table displays the equityholders' mean annualized return in excess of the risk-free rate of return ($\mathbb{E}\left[R^{(A_{exs})}\right]$, %), standard deviation of annualized excess return ($\sigma^{(A_{exs})}$, %), the Sharpe ratio (SR) and Jensen's alpha ($\mathbb{E}[\alpha]$, %), Equation (2.5.1), of capital provision to the DVA provider. The underlying portfolio is either invested in the money market account only ($\theta = 0$, top panel), or is 20% invested in the risky stock index, and 80% invested in the money market account ($\theta = 20\%$, bottom panel). The 99% confidence intervals are in parentheses.

$\theta = 0$

Statistic	γ		
	2	**5**	**8**
$\mathbb{E}\left[R^{(A_{exs})}\right]$ (%)	-0.008 [-0.010, -0.006]	-0.007 [-0.009, -0.005]	-0.007 [-0.008, -0.005]
$\sigma^{(A_{exs})}$ (%)	3.96 [3.95, 3.40]	3.91 [3.90, 3.91]	3.89 [3.88, 3.90]
SR	-0.002 [-0.0056, 0.0016]	-0.0017 [-0.0054, 0.0019]	-0.0017 [-0.0053, 0.0020]
$\mathbb{E}[\alpha]$ (%)	-0.0001 [-0.0001, -0.0001]	-0.0001 [-0.0001, -0.0001]	-0.0001 [-0.0001, -0.0001]

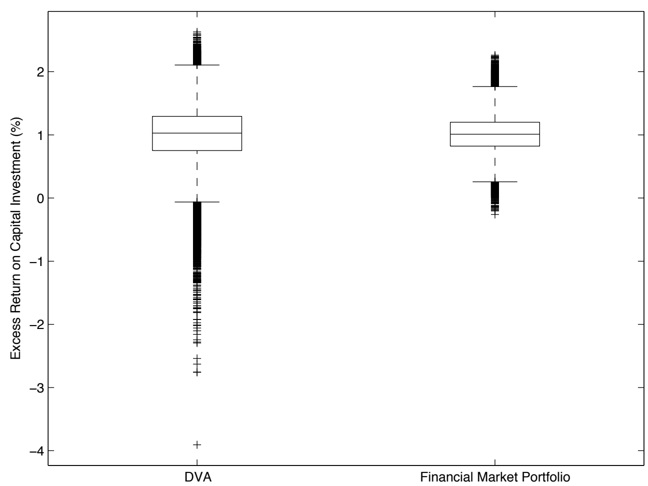
$\theta = 20\%$

Statistic	γ		
	2	**5**	**8**
$\mathbb{E}\left[R^{(A_{exs})}\right]$ (%)	1.44 [1.44, 1.44]	1.44 [1.44, 1.45]	1.44 [1.44, 1.45]
$\sigma^{(A_{exs})}$ (%)	5.04 [5.03, 5.06]	4.95 [4.94, 4.96]	4.95 [4.94, 4.96]
SR	0.29 [0.29, 0.29]	0.29 [0.29, 0.29]	0.29 [0.29, 0.29]
$\mathbb{E}[\alpha]$ (%)	-0.0001 [-0.0001, -0.0001]	-0.0001 [-0.0001, -0.0001]	-0.0001 [-0.0001, -0.0001]

The box plot in Figure 2.5.1 indicates that the medians of the excess returns from either investing in the DVA provider, or in the portfolio having the same investment policy as the DVA contract reference portfolio are comparable. While excess returns on the financial market only are less volatile across scenarios, their maximum is lower than the best excess returns attainable via capital provision. Therefore, longevity risk exposure allows the equityholders to achieve higher excess returns in the best scenario, but entails greater downside risk due to the possible default of the DVA provider.

Figure 2.5.1: Box Plot of Equityholders' Annualized Excess Return (%): $\theta = 20\%$

This figure presents the box plot of the equityholders' annualized return in excess of the risk-free rate (%), from either capital provision to the DVA provider (left), or investing in the reference portfolio (right). The reference portfolio is 20% invested in the risky stock index and 80% in the money market account. The line in the middle of the box is the median, while the edges of the box represent the 25^{th} and 75^{th} percentiles. The height of the box is the interquartile range, i.e., the interval between the 25^{th} and 75^{th} percentiles. The "+" symbols represent data points that are 1.5 times larger than the interquartile range.



2.6 Sensitivity Analysis

In this section, we carry out sensitivity analyses on the individual characteristics, stock exposure, the annuity provider's leverage, and the magnitude of longevity risk. These features influence the annuity provider's default rate and/or the volatility of the GSA benefits across scenarios and they subsequently alter the appeal of the GSA and the DVA to individuals.

2.6.1 Level of Risk Aversion

Individuals' preference for a GSA or a DVA is determined not only by the average level of benefits, but also by the risk on those benefits. We expect more risk-averse individuals to prefer the DVA for the benefits lower volatility, as long as the default rates entailed by their respective optimal AIRs are not too high. We consider highly risk-averse individuals with $\gamma = 10$, 15, and 20.

The optimal AIRs (Table 2.5) for highly risk-averse individuals and the annuity provider's default rates (Table 2.6, top panel) are comparable to those in the baseline case. Yet, in contrast to that case, the CELs are positive (Table 2.6, middle panel). This suggests that individuals who are highly risk-averse prefer the DVA over the GSA, and are willing to pay a one-time loading of between 0.003% and 0.62% for the DVA. Despite that, when the annuity provider charges a loading equal to the CEL, equityholders attain Sharpe ratios that remain inferior to the 0.45 ratio of investing in the reference portfolio, and non-negative Jensen's alphas that are economically insignificant (Table 2.6, bottom panel).

Table 2.5: Highly Risk-Averse Individuals: Optimal AIR, h^* (%)

This table shows the optimal AIR, Equation (2.2.5), of the DVA and GSA for individuals with risk aversion levels of $\gamma = 10$, 15, and 20. All other parameters are identical to those in the baseline case.

θ	γ		
(%)	**10**	**15**	**20**
0	3.56	3.58	3.59
20	4.44	4.25	4.04

Table 2.6: Highly Risk-Averse Individuals: Cumulative Default Rates (%), Certainty Equivalent Loading (CEL) (%) and Investment Performance Statistics

This top panel displays the Cumulative Default Rates, Equation (2.4.2), of the annuity provider. The middle panel shows the CEL, Equation (2.4.3), for individuals with risk aversion levels of $\gamma = 10$, 15, and 20. The bottom panel shows the Sharpe ratio (SR) and Jensen's alpha, Equation (2.5.1), when the loading is set at the CEL estimates in the middle panel. All other parameters are identical to those in the baseline case. The 99% confidence intervals are in parentheses.

Cumulative Default Rates (%)

θ	γ		
(%)	**10**	**15**	**20**
0	0.0106	0.0104	0.0104
20	0.0056	0.0064	0.0086

Certainty Equivalent Loading, CEL (%)

θ	γ		
(%)	**10**	**15**	**20**
0	0.037 [0.025, 0.049]	0.250 [0.233, 0.268]	0.410 [0.356, 0.458]
20	0.003 [-0.062, 0.069]	0.340 [0.095, 0.577]	0.620 [0.087, 1.145]

Sharpe Ratio and Jensen's alpha: Loading = CEL

θ	**Statistic**	γ		
(%)		**10**	**15**	**20**
0	SR	-0.0002 [-0.0039, 0.0034]	0.0083 [0.0046, 0.0119]	0.0146 [0.0110, 0.0183]
	$\mathbb{E}[\alpha]$ (%)	0 [-0.0000, -0.0000]	0.0003 [0.0003, 0.0003]	0.0005 [0.0005, 0.0005]
20	SR	0.292 [0.2920, 0.2920]	0.3062 [0.3062, 0.3062]	0.3171 [0.3171, 0.3171]
	$\mathbb{E}[\alpha]$ (%)	0 [0.0000, 0.0000]	0.0005 [0.0005, 0.0005]	0.0009 [0.0009, 0.0009]

2.6.2 Length of the Deferral Period

As the accuracy of longevity forecast depends on its horizon, the preference for either contract may be sensitive to the age when the individual annuitizes. In the baseline case, individuals are aged 25 when purchasing a DVA contract or participating in the GSA. As retirement benefit payments commence at age 66, the deferral period is 40 years. Here, we shorten the deferral period by considering the situations where individuals decide between the DVA and the GSA at ages 45 and 65 instead (i.e., deferral periods of 20 years, and a year respectively).

When the deferral period is shorter, survival probability forecasts are more accurate. Thus, we expect smaller differences in the average level and standard deviation of benefits between contracts. However, this does not necessarily imply that the CEL estimates would be closer to zero, because the time-preference discounting, as governed by the subjective discount factor, β in Equation (2.2.2), plays a larger role when retirement is imminent. Therefore, while the difference between the benefits would be smaller, the effect in terms of utility would be greater.

Table 2.7 reveals that for individuals with risk-aversion levels of $\gamma = 5, 8$, the effect due to shorter time-discounting dominates the more accurate survival probability forecast; the CEL estimates are negative and more economically significant than those in the baseline case (Table 2.3). Thus, despite the smaller threat that longevity risk poses due to more accurate survival probability forecast, the imminence of retirement induces greater preference for the GSA.

The least risk-averse individual, $\gamma = 2$, also has a stronger preference for the GSA than in the baseline case when the deferral period is 20 years, but this observation reverses when the deferral period is only one year. Thus, apart from when the individual is less risk-averse and purchases an immediate annuity, the baseline results hold.

Table 2.7: Deferral Period: Certainty Equivalent Loading (CEL) (%)

This top panel displays the CEL, Equation (2.4.3), for individuals aged 45 at annuitization, whereas the bottom panel corresponds to the CELs for individuals aged 65 at that time. All other parameters are identical to those in the baseline case. The 99% confidence intervals estimated by the Delta Method are in parentheses.

20-year Deferral

θ	γ		
(%)	**2**	**5**	**8**
0	-0.380 [-0.386, -0.367]	-0.260 [-0.271, -0.252]	-0.150 [-0.161, -0.142]
20	-0.370 [-0.391, -0.350]	-0.270 [-0.293, -0.244]	-0.180 [-0.219, -0.140]

One-year Deferral

θ	γ		
(%)	**2**	**5**	**8**
0	-0.270 [-0.274, -0.266]	-0.230 [-0.234, -0.226]	-0.190 [-0.198, -0.190]
20	-0.260 [-0.262, -0.254]	-0.220 [-0.222, -0.213]	-0.190 [-0.192, -0.182]

2.6.3 Stock Market Risk Exposure

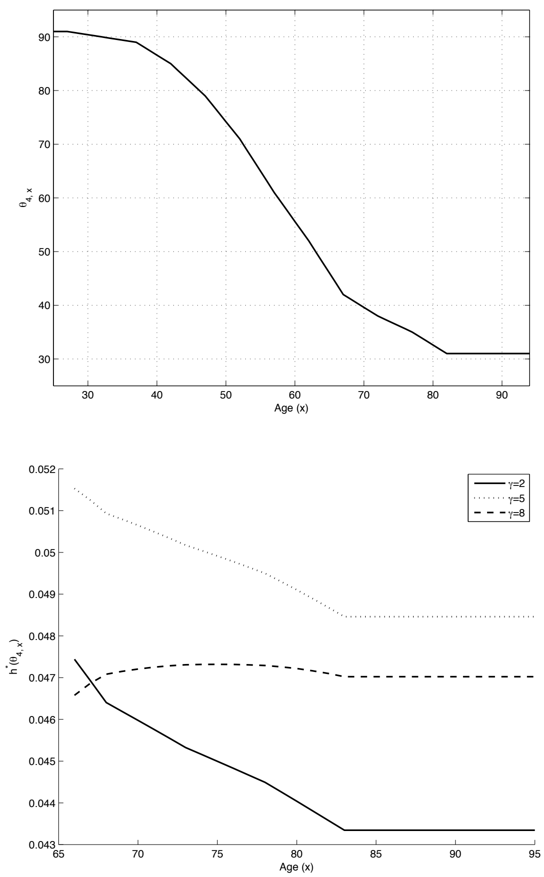
Both the GSA and the DVA contracts offer the AIR that maximizes individuals' expected utility given a fixed proportion invested in the stock index, θ, which we set to either 0% or 20%. We demonstrate that the implication of the baseline case, where individuals marginally prefer the GSA, holds as long as the allocation to the stock index corresponds to an optimal AIR with similar default rates as those in the baseline case.

We consider four alternative exposures to the stock index. The first three are constant allocations over the planning horizon: $\theta_1 = 40\%$, $\theta_2 = 60\%$, $\theta_3 = \frac{\lambda_S}{\gamma\sigma_S}$. θ_3 corresponds to the individual's optimal exposure to stocks (Appendix 2.A). For the least risk-averse individual ($\gamma = 2$), θ_3 is 147.2%. The moderately risk-averse individual ($\gamma = 5$) optimally invests 58.9% in the stock index whereas the most risk-averse individual ($\gamma = 8$) optimally invests 36.8% in stocks.

The fourth exposure that we consider, $\theta_4 = \{\theta_{4,\,x}\}_{x=25}^{95}$, is an age-dependent allocation that begins with around 90% allocation to stocks at age 25, and gradually diminishes to a minimum of about 30% post-retirement, until the maximum age (Figure 2.6.1, top panel). This glidepath allocation is based on the 2014 Target-Date Fund industry average (Yang et al., 2016). A decreasing exposure to stocks as the individual grows older is consistent with popular financial advice (Viceira, 2001). In theory, when the investment opportunity set is constant, horizon-dependent investment strategies are optimal in situations where, for instance, the individual receives labor income (Viceira, 2001; Cocco et al., 2005), or where the individual's risk aversion parameter is time dependent (Steffensen, 2011). For all θs, the optimal AIR is set according to Equation (2.2.5), and summarized in Table 2.8. For the age-dependent θ_4, the optimal AIR also varies over the individual's life-span (Figure 2.6.1, bottom panel).

Figure 2.6.1: Glidepath Allocation to the Stock Index, θ_4 and the Optimal AIR (%)

The top panel shows the age-dependent allocation to stocks, defined on the industry average of Target-Date Funds in the US in 2014 (Yang et al., 2016). The bottom panel displays the corresponding age-dependent optimal AIR, Equation (2.2.5), when the allocation to the stock index is θ_4.



For all θ_i, $i = 1, 2, 3, 4$, the optimal AIRs that correspond to these levels of stock investments (Table 2.8) are higher than those in the baseline case (Table 2.1), where $\theta = 20\%$. Due to the inverse relationship between the default rate and the AIR, the default rates are smaller than those in the baseline case. The positive CEL estimates in Table 2.8 imply that when investing according to θ_i, $i = 1, 2, 3, 4$, individuals prefer the GSA to a similar extent as in the baseline case.

Table 2.8: Exposure to Stock Market Risk: Certainty Equivalent Loading (CEL) (%)

The table displays the the optimal AIR, $h^*(\theta_i)$, Equation (2.2.5), the Cumulative Default Rates, Equation (2.4.2), and the CEL, Equation (2.4.3), when the underlying portfolio is θ_i invested in the stock index and $100-\theta_i$ invested in the money market account, for $i = 1, 2, 3, 4$. Estimates are calculated on 5 million replications. All other parameters are identical to those in the baseline case. The 99% confidence intervals estimated by the Delta Method are in parentheses.

$\theta_1 = 40\%$

Statistics	γ		
(%)	**2**	**5**	**8**
$h^*(\theta_1)$	4.59	5.06	4.72
Default Rates	0.0034	0.0020	0.0030
CEL	-0.340 [-0.344, -0.336]	-0.211 [-0.227, -0.195]	-0.080 [-0.218, 0.058]

$\theta_2 = 60\%$

Statistics	γ		
(%)	**2**	**5**	**8**
$h^*(\theta_2)$	5.08	5.24	4.26
Default Rates	0.0025	0.0021	0.0039
CEL	-0.344 [-0.348, -0.340]	-0.189 [-0.222, -0.155]	0.104 [-0.032, 0.240]

$\theta_3 = \frac{\lambda_S}{\gamma \sigma_S}$, optimal exposure

Statistics	γ		
(%)	**2**	**5**	**8**
θ_3	147.2	58.9	36.8
$h^*(\theta_3)$	0.0060	0.0052	0.0047
Default Rates	0.0003	0.0019	0.0026
CEL	-0.332 [-0.342, -0.321]	-0.186 [-0.235, -0.137]	-0.031 [-0.106, 0.043]

θ_4 = glidepath

Statistics	γ		
(%)	**2**	**5**	**8**
θ_4	Figure 2.6.1, top panel		
$h^*(\theta_4)$	Figure 2.6.1, bottom panel		
Default Rates	0.0052	0.0020	0.0022
CEL	-0.341 [-0.345, -0.336]	-0.307 [-0.390, -0.224]	-0.563 [-0.773, -0.352]

2.6.4 Level of Equity Capital

Individuals' preference for the DVA depends on the provider's default rates which are determined by the level of the provider's capital buffer. We investigate the implication that the DVA provider's default rates has on individual preference by increasing the baseline case's leverage ratio by one standard deviation, comparable to halving the baseline case equity capital to 5%.[23]

When the equity capital is 5% of the value of liabilities instead of 10%, the default rates rise from 0.004-0.01% to 4.96-6.78% (Tables 2.2 and 2.9). The CELs that accompany these high rates are economically significantly negative. For example, when $\theta = 20\%$, the most risk averse individual ($\gamma = 8$) is essentially indifferent between the DVA and GSA in the baseline case, but now values the DVA much less, CEL = -24% (Table 2.9).

While the CELs in the baseline case are similar for $\theta = 0$ and $\theta = 20\%$, they are noticeably more negative for $\theta = 20\%$ when the annuity provider has higher leverage. The amplified preference for the GSA in the presence of stock market risk is due to the resolution when a default occurs. When the annuity provider defaults, individuals recover the provider's residual wealth to purchase an equally weighted portfolio of bonds that mature in every remaining year of retirement until maximum age. This implies that individuals forgo mortality credit. In the case when $\theta = 20\%$, individuals additionally relinquish all potential reward from investing in the stock market. Relative to having the underlying portfolio fully invested in the money market account, when only mortality credit is lost, the consequence of default is more severe when the underlying portfolio is invested in the stock market. Therefore, in Table 2.9, the CELs when $\theta = 20\%$ are considerably more negative than those when $\theta = 0$.

[23] The standard deviation of US life insurers' leverage ratio between 1998-2011 is 3.7% whereas the average is around 90% (A.M. Best data from Koijen and Yogo, 2015). Using the definition of Leverage Ratio $\equiv 1 -$ Value of Equity/Value of Assets, the assumption that Value of Assets = Premium Collected + Value of Equity, and that the Premium Collected is normalized to 1, a 93.7% leverage ratio is equivalent to an equity capital of around 5%.

Table 2.9: Lower Annuity Provider Buffer Capital: Default Rates (%) and Certainty Equivalent Loading (CEL) (%)

This top panel displays the Cumulative Default Rates, Equation (2.4.2), of the annuity provider, whereas the bottom panel shows the CEL, Equation (2.4.3), when the level of equity capital is 5% of the present value of liabilities at the date of contract sale. All other parameters are identical to those in the baseline case. The 99% confidence intervals for the CEL are in parentheses.

Cumulative Default Rates (%)

θ	γ		
(%)	**2**	**5**	**8**
0	6.7826	6.4874	6.4092
20	5.6558	4.9730	4.9634

Certainty Equivalent Loading, CEL (%)

θ	γ		
(%)	**2**	**5**	**8**
0	-3.4 [-3.5, -3.4]	-5.5 [-5.6, -5.5]	-9.4 [-9.5, -9.3]
20	-5.6 [-5.7, -5.5]	-12.9 [-13.2, -12.7]	-24.0 [-24.3, -23.6]

When the annuity provider is more leveraged, the increased occurrence of defaults adversely affects the equityholders' excess return and its standard deviation, and makes exposure to longevity risk even less attractive for both $\theta = 0$ and $\theta = 20\%$. When $\theta = 0$, excess returns on the equityholders' investment is negative (Table 2.10, top panel). When the underlying portfolio is $\theta = 20\%$ invested in the stock market, higher leverage yields an excess return of 1.3%, lower than the 1.44% excess return of the baseline case (Tables 2.10 and 2.4, bottom panels). Due to the higher frequency of defaults, the standard deviation of excess return is around twice that of the baseline case (7.6% vs. 3.9% for $\theta = 0$; 9% vs. 5% for $\theta = 20\%$; Tables 2.4 and 2.10). As a result, the Sharpe ratio is halved whereas the Jensen's alphas are more negative than those in the baseline case.

Table 2.10: Lower Annuity Provider Buffer Capital: Equityholders' Investment Performance Statistics

This table displays the equityholders' mean annualized return in excess of the risk-free rate of return ($\mathbb{E}\left[R^{(A_{exs})}\right]$, %), standard deviation of annualized excess return ($\sigma^{(A_{exs})}$, %), the Sharpe ratio (SR) and Jensen's alpha ($\mathbb{E}[\alpha]$, %), Equation (2.5.1), of capital provision to the DVA provider, when the level of equity capital is 5% of the present value of liabilities at the date of contract sale. The underlying portfolio is either invested in the money market account only ($\theta = 0$, top panel), or is 20% invested in the risky stock index, and 80% in the money market account ($\theta = 20\%$, bottom panel). The 99% confidence intervals are in parentheses.

$\theta = 0$

Statistic	γ		
	2	**5**	**8**
$R^{(A_{exs})}$ (%)	-0.085 [-0.089, -0.082]	-0.085 [-0.088, -0.082]	-0.085 [-0.088, -0.082]
$\sigma^{(A_{exs})}$ (%)	7.59 [7.57, 7.61]	7.56 [7.54, 7.58]	7.55 [7.53, 7.57]
SR	-0.011 [-0.015, -0.008]	-0.011 [-0.015, -0.008]	-0.011 [-0.0145, -0.008]
$\mathbb{E}[\alpha]$ (%)	-0.0009 [-0.0009, -0.0009]	-0.0009 [-0.0009, -0.0009]	-0.0009 [-0.0009, -0.0009]

$\theta = 20\%$

Statistic	γ		
	2	**5**	**8**
$R^{(A_{exs})}$ (%)	1.28 [1.28, 1.28]	1.29 [1.29, 1.30]	1.29 [1.29, 1.30]
$\sigma^{(A_{exs})}$ (%)	9.45 [9.42, 9.47]	9.22 [9.20, 9.25]	9.22 [9.19, 9.24]
SR	0.136 [0.132, 0.139]	0.1401 [0.137, 0.143]	0.140 [0.137, 0.143]
$\mathbb{E}[\alpha]$ (%)	-0.0009 [-0.0009, -0.0009]	-0.0009 [-0.0009, -0.0009]	-0.0009 [-0.0009, -0.0009]

2.6.5 Longevity Risk

We investigate the effect of longevity risk on individual preference between the GSA and DVA via three other scenarios. We introduce parameter uncertainty surrounding the drift of the longevity time trend, increase the standard deviation of the longevity time trend error terms, and adopt the Cairns et al. (2006b) longevity model in replacement of the Lee and Carter (1992) model.

2.6.5.1 Drift Parameter Uncertainty

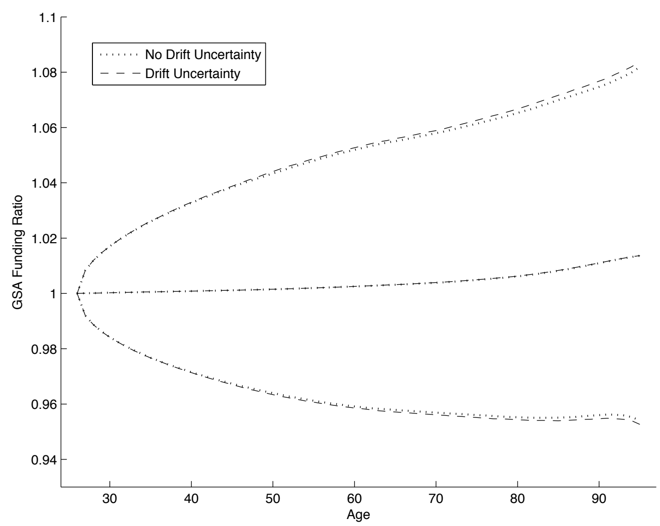
One way to depict the challenge of accurately estimating future survival probabilities is to introduce parameter uncertainty on the estimate of the drift term, c in Equation (2.3.1).

The maximum likelihood estimate for the drift term of the longevity model is normally distributed, $\hat{c} \sim \mathcal{N}\left(c,\, \sigma_c^2\right)$. Based on the sample used for the model calibration, we obtain $\hat{c} = -1.0689$ and $\widehat{\sigma_c}$=0.0521. Without parameter uncertainty, the best m-year-ahead forecast at time t is $\widehat{k_{t+m}} = m\hat{c} + k_t$. To incorporate parameter uncertainty, we draw c_l from the distribution $\mathcal{N}\left(\hat{c},\, \widehat{\sigma_c^2}\right)$ for the l^{th} simulation replication. The time trend governing longevity is thus $k_{t+m,\,l} = mc_l + k_{t,\,l} + \sum_{i=1}^{m} \epsilon_{\delta,\,l}$, $\epsilon_{\delta,\,l} \sim \mathcal{N}\left(0,\, \widehat{\sigma_\delta^2}\right)$, while the best m-year-ahead forecast relies on $\hat{c}$ as c_l is unobserved, i.e., $\widehat{k_{t+m,\,l}} = m\hat{c} + k_{t,\,l}$.

In Figure 2.6.2, we plot the mean, 5^{th} and 95^{th} percentiles of the GSA funding ratio prior to entitlement adjustments, with and without parameter uncertainty (Figure 2.6.2). The GSA funding ratio reflects the entitlement adjustments. For instance, if the funding ratio is 1.02, then the GSA offers a benefit that is 2% higher than the entitlement in that year. Figure 2.6.2 suggests that parameter uncertainty has a faint effect on the benefits. The average entitlement adjustments are comparable to when the drift term is known with certainty. The only noticeable difference is that with parameter uncertainty, the 5^{th} and 95^{th} percentiles are slightly farther apart in the final years of retirement.

Figure 2.6.2: Drift Parameter Uncertainty: GSA Funding Ratio

This figure presents the mean, 5^{th} and 95^{th} percentiles of the funding ratio of a GSA prior to entitlement adjustments, for when longevity is modeled according to Lee and Carter (1992). When there is no parameter uncertainty surrounding the drift term of the longevity time trend, $c = \hat{c}$. When there is parameter uncertainty, $c \sim \mathcal{N}\left(\hat{c}, \widehat{\sigma_c^2}\right)$. The GSA is composed of individuals with a risk-aversion level of $\gamma = 5$ and the underlying portfolio is fully invested in the money market account. All other parameters are identical to those in the baseline case.



When there is uncertainty around the drift parameter, the DVA is disadvantaged by a higher default probability. However, the GSA's appeal also diminishes as entitlement adjustments have a wider variation, especially toward the end of retirement (Figure 2.6.2). Neither of these drawbacks is sufficiently decisive to sway individual preferences. Therefore, the CELs deviate only marginally from those in the baseline case (Table 2.11).

Table 2.11: Drift Parameter Uncertainty: Cumulative Default Rates (%) and Certainty Equivalent Loading (CEL) (%)

The top panel presents the Cumulative Default Rates, Equation (2.4.2), of the annuity provider when there is parameter uncertainty surrounding the drift term of the longevity model's time trend. The bottom panel displays the CEL, Equation (2.4.3). All other parameters are identical to those in the baseline case. The 99% confidence intervals estimated by the Delta Method are in parentheses.

Cumulative Default Rates (%)

θ	γ		
(%)	**2**	**5**	**8**
0	0.0174	0.0144	0.0140
20	0.0066	0.0030	0.0030

Certainty Equivalent Loading, CEL (%)

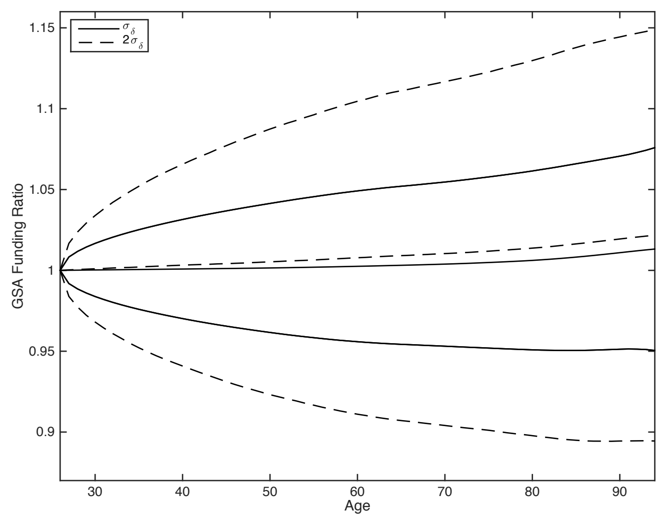
θ	γ		
(%)	**2**	**5**	**8**
0	-0.351 [-0.362, -0.339]	-0.193 [-0.205, -0.182]	-0.045 [-0.056, -0.033]
20	-0.345 [-0.356, -0.333]	-0.194 [-0.210, -0.178]	-0.046 [-0.089, -0.004]

2.6.5.2 Standard Deviation of the Time Trend Errors

We consider another prospect of longevity evolution with the longevity time trend errors that has twice the standard deviation estimated from historical mortality data, i.e., σ_δ of Equation (2.3.1) is replaced by $2\sigma_\delta = 3.488$. At a higher time trend standard deviation, the survival probabilities not only become more variable, but their conditional probabilities also decline (Denuit, 2009). Higher variability of survival outcomes is unfavorable both to the GSA participants, who bear larger variations in benefits, and to the DVA contractholders, due to the greater probability of default. On average, the entitlement adjustments under increased longevity risk are positive and higher than those in the baseline case, but there is also a wider variation in entitlement adjustments that rises in age (Figure 2.6.3).

Figure 2.6.3: Doubled Longevity Time Trend Error Standard Deviation: GSA Funding Ratio

This figure presents the mean, 5^{th} and 95^{th} percentiles of the funding ratio of a GSA prior to entitlement adjustments, for when longevity is modeled according to Lee and Carter (1992). Parameters for the longevity model are either those in Figure 2.3.1 (σ_δ), or with the standard errors of the time trend error terms doubled ($2\sigma_\delta$). Individuals have a risk-aversion level of $\gamma = 5$ and the underlying portfolio is fully invested in the money market account. All other parameters are identical to those in the baseline case.



The default rates are considerably higher when the longevity time trend standard deviation is doubled (i.e., 3.39-5.17%, Table 2.12). Consequently, individuals prefer the GSA to a large degree (Table 2.13, top panel).

Table 2.12: Doubled Longevity Time Trend Error Standard Deviation: Cumulative Default Rates (%)

This table displays the Cumulative Default Rates, Equation (2.4.2), of the annuity provider when the standard deviation to the longevity model time trend error term is doubled. All other parameters are identical to those in the baseline case.

θ	γ		
(%)	**2**	**5**	**8**
0	5.1650	4.8704	4.8008
20	4.0596	3.4066	3.3928

Two factors govern individual preference. The first is the effect on the level and standard deviation of benefits; the second is the annuity provider's higher default risk due to less accurate longevity forecasts. To separately identify the two effects, we eliminate default risk by assuming a sufficiently high level of equity capital. In the absence of default, the least risk-averse individual (i.e., $\gamma = 2$) marginally prefers the GSA, to a similar extent as she did in the baseline case (Table 2.13, middle panel). Thus, the least risk-averse individual's preference is invariant to the size of the standard deviation of the error terms, as long as the provider's default risk is unaffected. As for the more risk-averse individuals (i.e., $\gamma = 5$, 8), they prefer the DVA with no default risk and are willing to offer between 0.2% and 3.2% in loading for it. Thus, a higher standard deviation to the longevity time trend errors transpires to more volatile GSA benefit payment, making the GSA less appealing to individuals overall. However, individuals who are at least moderately risk-averse would prefer the DVA only if the annuity provider's default risk were eliminated.

Despite the seemingly high loading that the annuity provider could charge on a DVA contract with no default risk, the loading is insufficient to yield equityholders a Sharpe ratio superior to the 0.45 ratio of investment without longevity risk exposure (Table 2.13, bottom panel). The Jensen's alpha is positive but economically insignificant. Therefore, longevity risk exposure does not improve the equityholders' risk-return tradeoff.

Table 2.13: Doubled Longevity Time Trend Error Standard Deviation: Certainty Equivalent Loading (CEL) (%) and Investment Performance Statistics

The top panel presents the CEL, Equation (2.4.3), when the standard deviation of the longevity time trend error terms is doubled, and the equity capital is 10% of the value of liabilities on the contract's date of sale. The middle panel displays the CEL in the same setting as in the top panel, but with the equity capital raised sufficiently to eliminate default risk. The bottom panel shows the Sharpe ratio (SR) and Jensen's alpha (α), Equation (2.5.1), when the loading is set at the CEL estimates in the middle panel. All other parameters are identical to those in the baseline case. The 99% confidence intervals are in parentheses.

CEL (%): With Default Risk

θ (%)	γ		
	2	**5**	**8**
0	-2.5 [-2.5, -2.5]	-3.2 [-3.2, -3.1]	-5.0 [-5.1, -4.9]
20	-3.9 [-3.9, -3.8]	-7.7 [-7.8, -7.5]	-15.9 [-16.4, -15.5]

CEL (%): No Default Risk

θ (%)	γ		
	2	**5**	**8**
0	-0.4 [-0.4, -0.4]	0.2 [0.1, 0.2]	0.7 [0.7, 0.7]
20	-0.3 [-0.4, -0.3]	3.2 [2.1, 4.2]	3.2 [3.1, 3.4]

Sharpe Ratio and Jensen's Alpha: No Default Risk, Loading = CEL

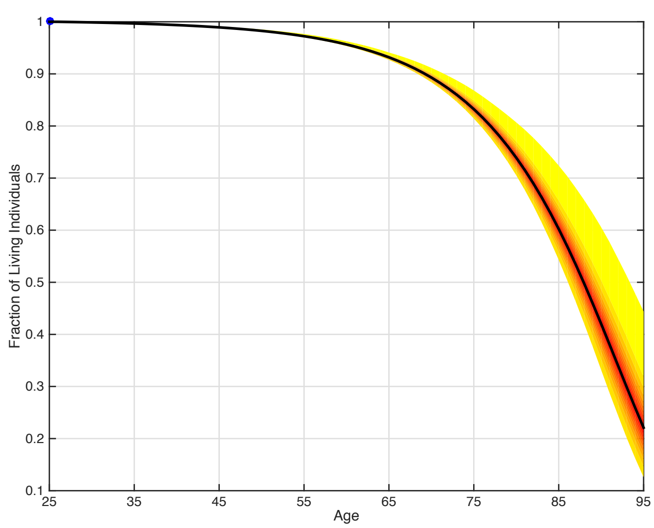
θ (%)	**Statistic**	γ		
		2	**5**	**8**
0	SR	0.0046 [0.0009, 0.0082]	0.0164 [0.0127, 0.0200]	0.0263 [0.0226, 0.0299]
	$\mathbb{E}[\alpha]$ (%)	0 [0.0000, 0.0000]	0.0001 [0.0001, 0.0001]	0.0002 [0.0002, 0.0002]
20	SR	0.4243 [0.4243, 0.4243]	0.4397 [0.4397, 0.4397]	0.4397 [0.4397, 0.4397]
	$\mathbb{E}[\alpha]$ (%)	0.0001 [0.0001, 0.0001]	0.0005 [0.0005, 0.0005]	0.0005 [0.0005, 0.0005]

2.6.5.3 Alternate Longevity Model

We next explore the choice of the longevity model by replacing the Lee and Carter (1992) model with the Cairns et al. (2006b) model, which produces a wider range of survival probabilities at old age. We calibrate the Cairns et al. (2006b) model over the same sample of mortality data as that in Section 2.3.2. Figure 2.6.4 presents the fan plot of the simulated fraction of living individuals under the Cairns et al. (2006b) model. The maximum range of the fraction of 25-year-olds still alive at older ages is 45% (i.e., at age 91), 50% more than the maximum range under the Lee and Carter (1992) model (i.e., 30% interval at age 88; Figure 2.3.2). This wider range translates into greater variability in benefits for the GSA, and higher default rates for the DVA provider.

Figure 2.6.4: Cairns et al. (2006b) Mortality Model: Fan Plot

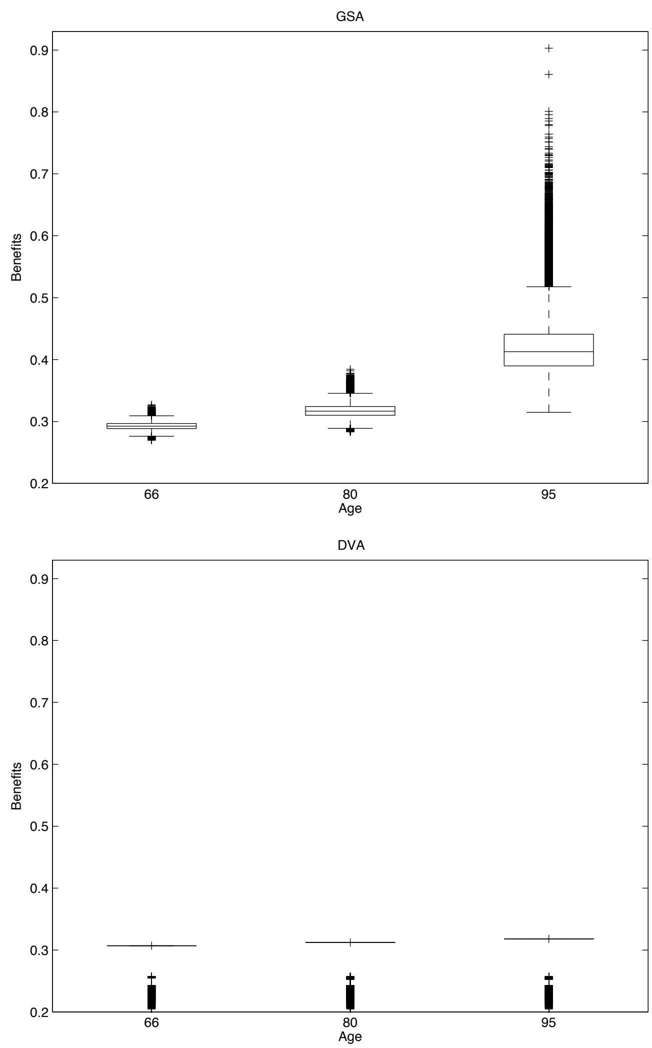
This figure presents the fan plot of the simulated fraction of living individuals (i.e., the population of 25-year-olds is normalized to one) over 10,000 replications when longevity is modeled according to the Cairns et al. (2006b) model. The model is calibrated on US female death counts from 1980 to 2013 taken from the Human Mortality Database. Darker areas indicate higher probability mass.



With either the Lee and Carter (1992) or the Cairns et al. (2006b) model, the rise in GSA benefits with age is accompanied by more uncertainty surrounding the benefits. However, the Cairns et al. (2006b) model produces greater uncertainty as the individual ages, as seen by comparing the top panels in Figures 2.3.3 and 2.6.5. This generates greater individual preference for the DVA under the Cairns et al. (2006b) model.

Figure 2.6.5: Cairns et al. (2006b) Mortality Model: Box Plots of GSA and DVA Benefits

The figure presents the box plots of benefits for the GSA (top panel) and the DVA (bottom panel), for an individual with a risk aversion level of $\gamma = 5$, at ages 66, 80 and 95. The underlying portfolio is invested in the money market account only. The line in the middle of the box is the median, while the edges of the box represent the 25^{th} and 75^{th} percentiles. The height of the box is the interquartile range, i.e., the interval between the 25^{th} and 75^{th} percentiles. The "+" symbols represent data points 1.5 times larger than the interquartile range.



For a fixed level of equity capital, the Cairns et al. (2006b) model yields higher default rates because of the heightened uncertainty surrounding old age survival. If we maintain the baseline case's 90% leverage ratio, the default rates under the Cairns et al. (2006b) model are between 0.48% to 2.21% (Table 2.14), substantially higher than the at-most 0.01% default rates when the Lee and Carter (1992) model is adopted (Table 2.2). Consequent to more defaults, individuals have a lower preference for the DVA (Table 2.14, bottom panel), as the CEL estimates are more negative than those in the baseline case (Table 2.3). Therefore, individuals prefer the DVA contract under the Cairns et al. (2006b) model only if the associated default risk is curtailed. Regardless of whether equityholders provide enough capital to eliminate default risk, the Sharpe ratio of equity provision is lower than the ratio of abstaining from longevity risk exposure. The Jensen's alpha of equity provision is positive but economically insignificant.

Table 2.14: Cairns et al. (2006b) Mortality Model with Default: Cumulative Default Rates (%) and CEL (%)

The top panel presents the Cumulative Default Rates, Equation (2.4.2), whereas the bottom panel displays the CEL, Equation (2.4.3), when life expectancy follows the Cairns et al. (2006b) model, calibrated to the same sample as the Lee and Carter (1992) model. All other parameters are identical to those in the baseline case. The 99% confidence intervals are in parentheses.

Cumulative Default Rates (%)

θ	γ		
(%)	**2**	**5**	**8**
0	2.2120	1.8082	1.7120
20	0.9676	0.4808	0.4756

CEL (%)

θ	γ		
(%)	**2**	**5**	**8**
0	-0.950 [-0.970, -0.930]	-0.660 [-0.690, -0.630]	-0.975 [-1.025, -0.924]
20	-0.877 [-0.906, -0.847]	-0.503 [-0.571, -0.436]	-1.515 [-1.763, -1.268]

Additionally, the choice of the longevity model underlies the inference of Maurer et al. (2013). While we find that individuals marginally prefer the GSA, Maurer et al. (2013) observe the opposite (positive CEL for the contract indexed to longevity; Table 7 of Maurer et al., 2013). When we assume that no default occurs, as do Maurer et al. (2013), we are able to reconcile our results to theirs. For instance, individuals who are moderately risk-averse to risk-averse, $\gamma = 5$ and 8, prefer the DVA; Table 2.15, top panel. The most risk-averse individual is willing to pay as much as 1% in loading to shed longevity risk. Despite that, when the annuity provider sets the loading to be equal to the CEL, the accompanying Sharpe ratio remains inferior to the Sharpe ratio of investing in only the financial market, i.e., 0.45 when $\theta = 20\%$, whereas the Jensen's alpha is positive but economically insignificant (Table 2.15, bottom panel). Therefore, while individual preference is sensitive to the choice of the longevity model, the extent that individuals are willing to pay to insure against longevity risk is insufficient to entice equityholders to gain longevity risk exposure.

Table 2.15: Cairns et al. (2006b) Mortality Model with No Default: Certainty Equivalent Loading (CEL) (%) and Investment Performance Statistics

The top panel presents the CEL, Equation (2.4.3), when life expectancy follows the Cairns et al. (2006b) model, calibrated to the same sample as the Lee and Carter (1992) model. The bottom panel shows the Sharpe ratio (SR) and Jensen's alpha (α), Equation (2.5.1), when the loading is set at the CEL estimates in the top panel. Equity capital is sufficiently high such that no default occurs. All other parameters are identical to those in the baseline case. The 99% confidence intervals are in parentheses.

CEL (%)

θ (%)	γ **2**	**5**	**8**
0	-0.089 [-0.099, -0.079]	0.528 [0.519, 0.537]	1.019 [1.011, 1.028]
20	-0.092 [-0.101, -0.082]	0.461 [0.448, 0.475]	0.874 [0.835, 0.913]

Sharpe Ratio and Jensen's Alpha: No Default Risk, Loading = CEL

θ (%)	**Statistic**	γ **2**	**5**	**8**
0	SR	0.0206 [0.0170, 0.0242]	0.0481 [0.0444, 0.0517]	0.0701 [0.0665, 0.0738]
	$\mathbb{E}[\alpha]$ (%)	0.0001 [0.0001, 0.0001]	0.0002 [0.0002, 0.0002]	0.0002 [0.0002, 0.0002]
20	SR	0.4337 [0.4337, 0.4337]	0.4362 [0.4362, 0.4362]	0.4379 [0.4379, 0.4379]
	$\mathbb{E}[\alpha]$ (%)	0.0001 [0.0001, 0.0001]	0.0001 [0.0001, 0.0001]	0.0002 [0.0002, 0.0002]

2.7 Conclusion

We investigate longevity risk management in retirement planning in the presence of two alternatives: individuals participate in a collective scheme that adjusts retirement income according to longevity evolution, or purchase a variable annuity contract offered by an equityholder-backed annuity provider. Our model features the perspective of not only the individuals, who evaluate their welfare in retirement, but also of the equityholders, who weigh their risk-return tradeoff from longevity risk exposure.

Due to the entitlement adjustments arising from errors in survival probability forecasts, the collective scheme provides more volatile benefits than those of an annuity contract. However, the collective scheme also offers a slightly higher average level of benefits, because for errors of the same magnitude, over- and under-estimating the log central death rates produce asymmetric effects.

The annuity contract provider relies on limited equity capital to subsume forecasting errors, and so is subject to default risk. Although the annuity contract shields individuals from downward entitlement adjustments up to a limit, it deprives individuals of any upward adjustments, as these gains belong to the equityholders.

We find that individuals marginally prefer the collective scheme over the annuity contract priced at its best estimate. This implies that the annuity provider is unable to charge a positive loading on the contract, subsequently failing to compensate its equityholders who bear longevity risk. Therefore, when individuals have the choice to form a collective scheme, the annuity provider who has no advantage at managing longevity risk, and who has to fully hedge financial market risk would not exist in equilibrium. Our finding is robust to numerous individual characteristics, stock market risk exposure, and heightened uncertainty surrounding life expectancy.

The results advocate for collective mechanisms in pension provision, which exist in a handful of countries (e.g., Collective Defined Contribution in the Netherlands, Target Benefit Plans in Canada). The pressing issue of population aging, and the gradual maturation of the longevity risk derivatives market, is likely to spur reform. For example, the US Chamber of Commerce (2016) recommends new plan designs to enhance the private retirement system. Collective schemes may serve as a benchmark that the annuity contract has to match or surpass with respect to the individuals' expected utility.

A limitation of our work is the exclusion of channels that may reduce the insurer's effective longevity exposure, such as synergies in product offering (e.g., natural hedging of longevity risk via the sale of annuities and life insurance contracts; Wong et al., 2017), access to reinsurance (Baione et al., 2016) and shadow insurance (Koijen and Yogo, 2016b). There are also alternative resolution mechanisms in the case of default, and other factors that may influence annuitization decisions, such as bequest motives, medical expenses, social security, uninsurable income, etc. (Lockwood, 2012; Pashchenko, 2013; Peijnenburg et al., 2017; Ai et al., 2016; Yogo, 2016). Examining these features in future research would enrich our knowledge of retirement planning.

Appendix to Chapter 2

2.A Rationale of the Contract Definition

The DVA and GSA contracts are not only modeled along the variable annuity contracts studied in the literature (Koijen et al., 2011; Maurer et al., 2013), but are also relatable to an individual's optimal consumption and investment.

The problem of optimal consumption and investment is composed of two separate parts: the allocation of initial wealth over each retirement year, and the investment strategy. Aase (2015) shows that for an expected-CRRA-utility-maximizing individual facing mortality and stock market risks, the optimal allocation of initial wealth decays geometrically in the retirement horizon. The AIR in our setting represents precisely this rate of decay.

When individuals are subject to longevity risk, its existence would not change the optimal wealth and asset allocation; what would complicate the solution is the ability to react to longevity evolution (Huang et al., 2012). We, however, assume that the contract's parameters are deterministic (i.e., fixed in the year when it is sold, and the incorporation of new information thereafter is prohibited). Therefore, by an appropriate choice of the AIR, h^*, the contract described by Equations (2.2.3) and (2.2.4) coincides with the optimal decumulation path of the individual.

We next solve the utility maximization problem, (2.A.1), to obtain the optimal AIR and investment strategy for a contract defined by Equations (2.2.3) and (2.2.4).

At time t_0, the individual purchases the maximum number of variable annuity contracts affordable with a lump sum capital normalized to one. The annuity contract commences benefit payment in year t_R, until the year T, conditional on the individual's survival. In the financial market setting as described in Section 2.2.1, with a deterministic fraction of wealth $\theta = \{\theta_t\}_{t=t_0}^{T}$ invested in the risky stock index, and $1-\theta$ invested in the money market account, the value of the reference portfolio evolves according to $dW_t/W_t = (r + \theta_t \lambda \sigma_S)\,\mathrm{d}t + \theta_t \sigma_S\,\mathrm{d}Z_{S,\,t}$.

$$\begin{aligned}
\max_{\{\theta_t,\, h(t,\theta_t)\}_{t=t_R}^{T}} \quad & \mathbb{E}_{t_0}\left[U\left(\Xi\right)\right] && (2.A.1)\\
= \quad & \mathbb{E}_{t_0}\left[\int_{t_R}^{T} \mathrm{e}^{-\beta(t-t_0)}\,\frac{\Xi_t^{1-\gamma}}{1-\gamma}\left(\Pi_{s=t_0\ 1}^{t}p_{x+(s-t_0)}^{(s)}\right)\mathrm{d}t\right]\\
\Xi_t = \quad & \begin{cases} \frac{1}{A(h)}\exp\left(-h\left(t,\,\theta_t\right)\left(t-t_R\right)\right)\frac{W_t}{W_{t_0}} & \text{if alive in year } t\\ 0 & \text{otherwise}\end{cases}\\
A\left(h\right) = \quad & \int_{t_R}^{T}\exp\left(-h\left(t,\,\theta_t\right)\left(t-t_R\right)\right)\times\\
& \mathbb{E}_{t_0}\left[\left(\Pi_{s=t_0\ 1}^{t}p_{x+(s-t_0)}^{(s)}\right)\right]\mathrm{d}t\\
h\left(t,\,\theta_t\right) = \quad & \text{AIR}\\
\beta = \quad & \text{subjective discount factor}\\
\gamma = \quad & \text{risk aversion parameter}\\
W_t = \quad & \text{value of the reference portfolio with}\\
& \text{the investment policy } \theta\\
\mathbb{E}_{t_0}\left[\Pi_{s=t_0\ 1}^{t}p_{x+(s-t_0)}^{(s)}\right] = \quad & {}_{t-t_0}p_x^{(t_0)}
\end{aligned}$$

$A\left(h\right)$ is the cost per unit of a zero-loading contract. It is straightforward to verify that the contract has a present expected value of one for any $h \in \mathbb{R}^{T-t_R}$, and thus satisfies the budget constraint. Given any θ, the first order condition, $\partial\mathbb{E}_{t_0}\left[U\left(\Xi\right)\right]/\partial h = 0$ yields the optimal AIR, Equation (2.A.2).

$$\begin{aligned}
h^*\left(t,\,\theta_t\right) &= r+\frac{\beta-r}{\gamma}-\frac{1-\gamma}{\gamma}\theta_t\sigma_S\left(\lambda_S-\frac{\gamma\theta_t\sigma_S}{2}\right) && (2.A.2)\\
r &= \text{constant short rate}\\
\beta &= \text{subjective discount factor}\\
\gamma &= \text{risk aversion parameter}\\
\theta_t &= \text{fraction of wealth allocated to the stock index}\\
& \quad\ \text{at time } t,\, t_R \le t \le T\\
\sigma_S &= \text{standard deviation governing the stock index's dynamics}\\
\lambda_S &= \text{instantaneous Sharpe ratio of the stock index}
\end{aligned}$$

Equation (2.A.2) is composed of the risk-free rate, the difference between the subjective discount factor and the risk-free rate, adjusted by the risk aversion parameter, and a term concerning the exposure to the stock index, weighted by the risk aversion level.

If the returns on the investment were constant at r (e.g., either $\theta = 0$ or $\sigma_S = 0$), for any given level of risk aversion, γ, the shape of the optimal consumption path depends on the relative magnitude of β and r. An individual who discounts future consumption at a higher rate than the constant interest rate (i.e., $\beta > r$, an impatient individual) prefers a downward sloping consumption path whereas a more patient person (i.e., $\beta < r$) optimally chooses an upward sloping path. When $\theta \neq 0$ and $\sigma_S \neq 0$, then the risk aversion level, the standard deviation and the market price of stocks also have a role in determining the optimal consumption path.

The first-order condition corresponding to the allocation to the stock index, $\partial \mathbb{E}_{t_0} [U(\Xi)] / \partial \theta = 0$, implies the optimal allocation to the risky asset:

$$\theta^* = \frac{\lambda_S}{\gamma \sigma_S} \tag{2.A.3}$$

The optimal allocation to the stock index, θ^*, is independent of time and wealth, and is identical to the optimal investment policy of Merton (1969).

The variable annuity contract provides the optimal decumulation path when the AIR is set to $h^*(t, \theta_t^*)$. By prohibiting the incorporation of new information into the contract definition after its date of sale (i.e., enforcing deterministic, but possibly time-varying contract parameters), longevity risk does not influence the optimal AIR and the optimal portfolio choice.

The conception of the GSA as a collective justifies the assumption that it prioritizes individual welfare (i.e., maximizes individuals' expected utility in retirement). Therefore, the GSA offers an AIR that is in the best interest of the individuals, without conflict among its stakeholders. As for the annuity provider, such contracts are also conceivable. For instance, Froot (2007) suggests that insurers should shed all liquid risks for which they have no comparative advantage to outperform (e.g., financial market risk), and devote their entire risk budget to insurance risks (e.g., longevity risk). The selling of variable annuities without any financial guarantee achieves precisely this goal. Besides, Gatzert et al. (2012) demonstrate that if an insurance company sets contract parameters for a participating life insurance contract such that they maximize the contract's value (e.g., expected utility) to the individual, the individual may be willing to pay more for the contract. Therefore, the provision of contracts defined according to Equations (2.2.3) and (2.2.4) under either a cooperative setup or by a for-profit entity is plausible.

2.B Definition of the Book Value of Liabilities

Suppose that the DVA provider or the GSA administrator issues contract(s) to a cohort who is aged x at time t_0, promising entitlements of $\Xi^K(h^*, F, t, x)$, $K \in \{DVA, GSA\}$, in every year t, $t_R \leq t \leq T$, conditional on the individual's survival. The estimate of the entity's book value of liabilities at time t, $t_0 \leq t \leq T$, is:

$$\begin{aligned} L_t \equiv & \Xi^K(h^*, F, t, x) \int_{s=\max\{t_R, t\}}^{T} \exp\left(-h^*(s, \theta)(s-t)\right) \times \\ & {}_{s-t}p_{x+t-t_0}^{(t)} \mathrm{ds} \end{aligned} \tag{2.B.1}$$

$$\begin{aligned} {}_{s-t}p_{x+t-t_0}^{(t)} &= \text{conditional probability in year } t \text{ that a living individual} \\ & \quad \text{of age } x+t \text{ lives for at least } s-t \text{ more years} \\ h^*(t, \theta) &= \text{Optimal } AIR\text{, Equation (2.2.5)} \\ \Xi^K(h^*, F, t, x) &= \text{benefit at time } t \text{ for contract } K \in \{GSA, DVA\} \\ &= \begin{cases} \text{Equation (2.2.6)} & \text{if } K = DVA \\ \text{Equation (2.2.7)} & \text{if } K = GSA \end{cases} \end{aligned}$$

2.B.1 Illustration of the Case with No Risk

To motivate the definition of Equation (2.B.1), let us consider a three-period case ($t = t_0, t_1, t_2$) in the absence of stock market and longevity risks. Assume that the individual buys exactly one unit of the retirement contract at retirement in year t_0, lives with certainty to collect the benefits in year $t_1 = t_0 + 1$, and dies with certainty before the year $t_2 = t_1 + 1$. Suppose that the reference portfolio is fully invested in the money market account, earning an interest rate that is constant at 2%. Furthermore, we adopt a constant AIR, $h = 3\%$, and zero contract loading, $F = 0$. As there is no uncertainty in this example, Equation (2.B.1) should yield precisely the value of liabilities at time t.

By definition of the DVA contract, there are two payments to be made: one in the year t_0 and another in the year t_1. The payment in t_0 is:

$$\begin{aligned}\Xi\left(h,\,0,\,t_0,\,x\right) &= 1\times\frac{W_{t_0}^{Ref}}{W_{t_0}^{Ref}}e^{-h\times(t_0-t_0)}\\ &= 1\end{aligned}$$

The second payment, in present value at time t_1 is:

$$\begin{aligned}\Xi(h,\,0,\,t_1,\,x) &= 1\times\frac{W_{t_1}^{Ref}}{W_{t_0}^{Ref}}e^{-h\times(t_1-t_0)}\\ &=\frac{W_{t_0}^{Ref}e^{0.02}}{W_{t_0}^{Ref}}e^{-h}\\ &=e^{-h+0.02}\\ &=e^{-0.01}\end{aligned}\tag{2.B.2}$$

Discounting Equation (2.B.2) by the constant interest rate, we obtain the present value at time t_0, of the payment due at time t_1:

$$\begin{aligned}PV_{t_0}\left[\Xi(h,\,0,\,t_1,\,x)\right] &= \Xi(h,\,0,\,t_1,\,x)e^{-0.02\times(t_1-t_0)}\\ &= e^{-0.01-0.02}\\ &= e^{-0.03}\end{aligned}$$

The present value of liabilities at time t_0 is

$$\Xi\left(h,\,0,\,t_0,\,x\right)+PV_{t_0}\left[\Xi(h,\,0,\,t_1,\,x)\right] = 1+e^{-0.03}\tag{2.B.3}$$

It remains to show that Equation (2.B.1) yields Equation (2.B.3):

$$\begin{aligned}L_t &= \Xi(h,\,0,\,t_1,\,x)\times\left(e^{-h\times 0}\;{}_0p_t^{(t)}+e^{-h\times 1}\;{}_1p_t^{(t)}\right)\\ &= 1\times\left(1+e^{-h}\right)\\ &= 1+e^{-0.03}\end{aligned}$$

2.B.2 Illustration of the General Case

We price the liabilities of the pension provision entity by constructing a replicating portfolio for its contractual obligation. We demonstrate that the price of the portfolio that replicates all the cash flows of an annuity contract is Equation (2.B.1).

In the setting with longevity but no mortality risk, we consider the liability associated with a contractholder who purchased $1/A$ unit(s) of contracts when aged x in the year $t_0 = 0$, retired in the year $t = t_R$, while being subject to unknown survival probabilities throughout the horizon, until the maximum age in the year $t = T$, when death is certain.

The pension provision entity is contractually obliged to make annual benefit payments from the individual's retirement in the year $t = t_R$ until he or she attains maximum age in the year $t = T$, conditional on her survival. Let W_t^{Ref} be the price at time t of the reference portfolio to which the benefits are indexed, $t \in [t_0, T]$.

Absent longevity risk, by purchasing the sum of all the units of the reference portfolio in Column (2) of Table 2.16 at time t, the annuity provider would be able to fulfill its contractual obligation with certainty. For instance, to meet the payment at time t_R, the annuity provider purchases $1/\left(AW_{t_0}^{Ref}\right) e^{-h\times 0}\ {}_{t_R-t}p_x^{(t_0)}$ units of the reference portfolio at time t_0. When longevity risk is absent, the conditional expectation, made at time t_0, of the individual's survival in year t_R coincides with the realized survival probability, i.e., ${}_{t_R-t}p_x^{(t_0)} = {}_{t_R-t}p_x$. The value of this portfolio will evolve along with the financial market, to be worth exactly $\frac{1}{A}\frac{W_{t_R}^{Ref}}{W_{t_0}^{Ref}} \times \Pi_{l=t_0}^{t_R-1}\ {}_{l-t_0}p_{x+l-t_0}^{(l)}$, the payment due at time t_R. By the same reasoning for the rest of the entries in Column (2), Equation (2.B.4) is thus the total units of the reference portfolio to be held at any time t, such that the pension provision entity fully hedges financial market risk.

$$\int_{s=\max\{t_R,\,t\}}^{T} \frac{1}{A}\frac{1}{W_{t_0}^{Ref}} e^{-h(s-t_R)}\ {}_{s-t}p_{x+t-t_0}^{(t)}\,\mathrm{ds} \tag{2.B.4}$$

Equation (2.B.4) is an estimate of the liabilities at time t, in terms of the *units* of reference portfolio. Each unit is worth W_t^{Ref} at time t. To obtain the *value* of liabilities, we take the portfolio's corresponding value:

$$W_t^{Ref} \quad \times \quad \int_{s=\max\{t_R,\,t\}}^{T} \frac{1}{A}\frac{1}{W_{t_0}^{Ref}} e^{-h(s-t_R)}\ {}_{s-t}p_{x+t-t_0}^{(t)}\,\mathrm{ds} \tag{2.B.5}$$

As $\Xi(h,\, F,\, t,\, x) = W_t^{Ref}/\left(AW_{t_0}^{Ref}\right) e^{-h(t-t_R)}$ by definition, we can substitute it into Equation (2.B.5) to get

Table 2.16: Future Cash Flow and the Best Replicating Portfolio of the Pension Provision Entity

This table shows the value of entitlements due in each year of retirement until maximum age (column (1)), and the corresponding Best Replicating Portfolio in units of the reference portfolio (column (2)). The Best Replicating Portfolio is the conditional expectation of the benefits in future value.

Time	**Benefits in Future Value**	**Best Replicating Portfolio (constructed at time t)**
		Units of the Reference Portfolio to purchase at time t
	(1)	(2)
t_R	$\frac{1}{A}\frac{W^{Ref}_{t_R}}{W^{Ref}_{t_0}}\,\mathrm{e}^{-h(t_R-t_R)}\times\Pi_{l=t_0}^{t_R-1}\ {}_1p_{x+l-t_0}$	$\frac{1}{A}\frac{1}{W^{Ref}_{t_0}}\,\mathrm{e}^{-h(t_R-t_R)}\ {}_{t_R-t}p^{(t)}_{x+t-t_0}$
t_R+1	$\frac{1}{A}\frac{W^{Ref}_{t_R+1}}{W^{Ref}_{t_0}}\,\mathrm{e}^{-h(t_R+1-t_R)}\times\Pi_{l=t_0}^{t_R}\ {}_1p_{x+l-t_0}$	$\frac{1}{A}\frac{1}{W^{Ref}_{t_0}}\,\mathrm{e}^{-h(t_R+1-t_R)}\ {}_{t_R+1-t}p^{(t)}_{x+t-t_0}$
t_R+2	$\frac{1}{A}\frac{W^{Ref}_{t_R+2}}{W^{Ref}_{t_0}}\,\mathrm{e}^{-h(t_R+2-t_R)}\times\Pi_{l=t_0}^{t_R+1}\ {}_1p_{x+l-t_0}$	$\frac{1}{A}\frac{1}{W^{Ref}_{t_0}}\,\mathrm{e}^{-h(t_R+2-t_R)}\ {}_{t_R+2-t}p^{(t)}_{x+t-t_0}$
$\vdots$	$\vdots$	$\vdots$
T	$\frac{1}{A}\frac{W^{Ref}_{T}}{W^{Ref}_{t_0}}\,\mathrm{e}^{-h\times(T-t_R)}\times\Pi_{l=t_0}^{T-1}\ {}_1p_{x+l-t_0}$	$\frac{1}{A}\frac{1}{W^{Ref}_{t_0}}\,\mathrm{e}^{-h(T-t_R)}\ {}_{T-t}p^{(t)}_{x+t-t_0}$

$$\begin{aligned} L_t &:= \Xi(h^*, F, t, x) \int_{s=\max\{t_R, t\}}^{T} \exp\left(-h^*\left(s, \theta\right)\left(s-t\right)\right) \times \\ & \quad {}_{s-t}p_{x+t-t_0}^{(t)} \mathrm{ds} \end{aligned} \tag{2.B.6}$$

Equation (2.B.6) is identical to Equation (2.B.1).

When there is longevity risk, the Best Replicating Portfolio is identical to column (2) of Table 2.16, but this best estimate may not necessarily provide the exact cash flow to meet the annuity provider's contractual obligations because the realized survival probability may deviate from its conditional expectation made at time t, which then triggers the provider's default.

2.C Delta Method

We apply the Delta Method (Theorem 5.5.4 of Casella and Berger, 2002) to estimate the variance of the CELs, which is used to compute their confidence intervals.

Consider the function $g\left(x,\, y\right) = \left(x/y\right)^{1/(\gamma-1)} - 1$. By the definition of Equation (2.4.3), $CEL = g\left(U\left(\Xi^{GSA}\right),\, U\left(\Xi^{DVA}\right)\right)$. We estimate the CEL by plugging the expected utility into $g\left(.\right)$, $g\left(\mathbb{E}_0\left[U\left(\Xi^{GSA}\right)\right],\, \mathbb{E}_0\left[U\left(\Xi^{DVA}\right)\right]\right)$. Theorem 5.5.24 of Casella and Berger (2002) suggests the following estimate for its variance:

$$\begin{aligned} \mathrm{Var}\left\{g\left(\mathbb{E}_0\left[U\left(\Xi^{GSA}\right)\right],\right.\right. &= g_x^2 \mathrm{Var}\left(U\left(\Xi^{GSA}\right)\right) + g_y^2 \mathrm{Var}\left(U\left(\Xi^{DVA}\right)\right) + \\ \left.\left.\mathbb{E}_0\left[U\left(\Xi^{DVA}\right)\right]\right)\right\} & \quad 2 g_x g_y \mathrm{cov}\left(U\left(\Xi^{GSA}\right),\, U\left(\Xi^{DVA}\right)\right) \qquad (2.C.1) \\ g_x &= g_x\left(\mathbb{E}_0\left[U\left(\Xi^{GSA}\right)\right],\, \mathbb{E}_0\left[U\left(\Xi^{DVA}\right)\right]\right) \\ g_y &= g_y\left(\mathbb{E}_0\left[U\left(\Xi^{GSA}\right)\right],\, \mathbb{E}_0\left[U\left(\Xi^{DVA}\right)\right]\right) \end{aligned}$$

g_x and g_y denote the first partial derivative of $g\left(.\right)$ with respect to x and to y respectively. $\mathrm{Var}\left(U\left(\Xi^{K}\right)\right)$ for $K \in \{GSA,\, DVA\}$ and $\mathrm{cov}\left(U\left(\Xi^{GSA}\right),\, U\left(\Xi^{DVA}\right)\right)$ are estimated by the sample variance and sample covariance.

CHAPTER 3

Variable Annuity and Interest Rate Risk

3.1 Introduction

The intertemporal consumption and investment problem in a complete market seeks a consumption stream that maximizes an individual's expected utility, through the support of an investment policy that hedges time-variation of the opportunity set. In a frictionless financial market, an underwriter of insurance contracts that cede financial risks to policyholders (e.g., unit-linked contracts; GDV, 2016) could in principle, embed this knowledge in the contract design at no cost. This is because the underwriter merely executes the investment policy such that the individual receives the optimal income stream with respect to financial risks. This paper shows that an exemplar of such a contract–the variable annuity–precludes the optimal hedge of interest rate risk.

Annuities are contracts that specify periodic payments from the underwriter to the policyholder. A nominal annuity provides a level payment stream determined on the prevailing term structure of interest rate, hence the policyholder is vulnerable to interest rate risk only at the annuitization date. In contrast, a variable annuity (VA) offers payments that are determined on a wider set of financial assets, exposing the policyholder to other financial market risks in addition to interest rate risk, at all payment dates.[1] The objective of maximizing

This chapter is based on Boon and Werker (2017).

[1] We consider a VA without any guarantees on investment returns or death benefits.

the policyholder's expected utility from consuming the VA's payment is analogous to that of the intertemporal consumption and investment problem.

Various scholars point out that the solution to the intertemporal consumption and investment problem is equivalent to the solution from a two-stage problem that first solves for the optimal division of initial capital over the consumption horizon, and then finds the optimal investment policy for each portion of divided capital (Brennan and Xia, 2002; Wachter, 2002; Sharpe, 2007). In the presence of interest rate risk, we prove this equivalent formulation, and apply it to demonstrate that the VA fails to optimally hedge interest rate risk at every consumption date.

With a unique reference portfolio prescribing payments at all consumption dates, the VA confers a poorer tradeoff between the expected level and volatility of consumption relative to the optimal solution, even when the VA is specified to maximize the policyholder's expected utility. Hence, the policyholder experiences a utility loss. This outcome is of no gain to the underwriter, who serves only as an administrator with regard to financial risks.

This chapter overlaps with two strands of work. Technical aspects draw from the literature on optimal investment and consumption, and on intertemporal hedging when investors face a time-varying investment opportunity set. Objective-wise, it pertains to the design of financial and insurance contracts.

Numerous investigations on consumption and investment when there is interest rate risk postulate an individual who maximizes the expected utility derived from wealth at the end of the planning horizon only (Sørensen, 1999; Bajeux-Besnainou et al., 2003). Insights from this case is relevant because the intertemporal consumption problem can be regarded as a series of terminal wealth utility maximization problems. By adopting a Gaussian interest rate term structure model, we follow these works and their extensions.[2]

Merton (1969) and Samuelson (1969) are the pioneers who investigate the setting where individuals derive utility from intertemporal consumption. Under independent and identically distributed returns and constant relative risk

Ledlie et al. (2008) define the types of guarantees, and present an overview of the global variable annuities market.

[2]The extensions include the increment of the number of state variables affecting the investment opportunity set (e.g., Lioui and Poncet, 2001 consider general, multi-dimensional state variables), and the generalization of the interest rate model (e.g., Korn and Kraft, 2002 assume the extended Vasicek model, and allow the market price of interest rate risk to vary).

aversion (CRRA), they show that the optimal investment strategy is constant over time. When the investment opportunity set is time-varying (e.g., interest rate follows the Vasicek (1977) model), Merton (1971, 1973) demonstrate that the optimal dynamic investment portfolio has a time-dependent component to hedge against shifts in the opportunity set.

Contemporary efforts also incorporate mortality risk (Yaari, 1965; Hakansson, 1969; Fischer, 1973; Richard, 1975), that is the risk of an individual's uncertain time of death given known survival probabilities. Mortality risk differs from longevity risk, which is the risk surrounding the misestimation of survival probabilities. Even if the VA poses a drawback, it is hasty to omit it from retirement planning because its major advantage stems from mortality risk pooling.[3] If the mortality credit gain is sufficiently large to offset the loss from imperfect hedging, then the VA maintains its merit in retirement planning. For instance, assuming constant interest rates, Horneff et al. (2010) show that a VA's mortality credit more than compensates for its rigid payout path. We eschew a direct welfare comparison of the gains from mortality credit and the loss from deficient hedging by observation that the VA, being a contract that segregates the treatment of financial and biometric risks, allows us to propose a refinement with regard to its financial risk component only, while preserving its mortality-pooling merit. Therefore, we abstain from mortality and longevity risks, and focus on financial risks only.[4]

As the mitigation of interest rate risk is invaluable to retirement planning, there exists many studies that incorporate interest rate risk in that context, that adopt either the perspective of the individual (Bodie et al., 2004; Horneff et al., 2008; Koijen et al., 2011; Horneff et al., 2015; Chang and Chang, 2017), or that of the pension plan administrator (Boulier et al., 2001; Vigna and Haberman, 2001; Cairns et al., 2006a; Hainaut and Devolder, 2007; Han and Hung, 2012; Guan and Liang, 2014).

[3]Besides that, the VA underwriter's default risk and the incentives provided by the tax system (Brown et al., 1999) also play a role in any comprehensive evaluation of the VA.

[4]Where appropriate, we make remarks on the inclusion of mortality risk when individuals are homogeneous, which is straightforward as long as the risk is uncorrelated with financial market risks, and it does not entail defaults of the VA underwriter, e.g., because the pool of policyholders is large such that their aggregate realized mortality rate coincides with the known mortality rate. The first assumption is tenable as there lacks consensus on the association of demographic structure with asset prices (Erb et al., 1994; Poterba, 2001; Ang and Maddaloni, 2003; Visco, 2006; Schich, 2008b; Arnott and Chaves, 2012). If individuals are heterogeneous, then the incorporation of mortality risk becomes complex, as its distribution among policyholders is unspecified. Heterogeneous policyholders within the pool bear different investment risks, so the mortality credit depends on past financial market returns, and is risky.

Among these strands of works, our analytical objective and methodology resemble Brennan and Xia (2002) and Munk and Sørensen (2004). They analyze the optimal consumption and portfolio allocation over a finite-horizon, when interest rates follow the Vasicek (1977) model, and when individuals exhibit CRRA. The economic implication of our findings, however, relates to unit-linked contract design. This literature predominantly focuses on the VA guarantee, and either takes the underwriter's point of view (Bacinello and Persson, 2002), or assumes a constant interest rate (Boyle and Tian, 2008, 2009; Branger et al., 2010; Bernard et al., 2011; Aase, 2015; Kaltepoth, 2016). In contrast, we examine the policyholder's stance in a setting with interest rate risk, when she is presented with a VA with no guarantee,.

Due to the changing investment opportunities arising from interest rate risk, our article also relates to the literature on intertemporal hedging. The source of time-variation that scholars have considered is not limited to only time-variation of interest rates, but also that of stock and inflation risk premia (Kim and Omberg, 1996; Wachter, 2002; Sangvinatsos and Wachter, 2005; Munk, 2008), and volatility (Brandt, 1999; Chacko and Viceira, 2005; Gomes, 2007). Although the optimal intertemporal hedge demand depends largely on model assumptions,[5] the welfare loss for failing to hedge intertemporally is often large, and increases in the investor's horizon (Sangvinatsos and Wachter, 2005). Therefore, when the individual derives utility from consumption over a long horizon, it is plausible that impaired intertemporal hedging would also generate considerable welfare cost.

Individuals suffer substantial utility loss from deficient intertemporal interest rate risk hedge under the VA. We estimate that they require between 7-19% more initial wealth to purchase a VA that would yield the same expected utility as a contract that provides the optimal outcome with respect to financial risks (Figure 3.5.1). The utility loss is more severe for lengthier consumption horizons, lower subjective discount factor, higher instantaneous standard deviation of the interest rate, lower mean reversion rate and lower level of mean reversion (Figures 3.5.1, 3.6.1 and 3.6.2).

The ascend of Defined Contribution (DC) pension plans as the predomi-

[5]Brandt (1999); Ang and Bekaert (2002) find a small intertemporal hedge demand, whereas Brennan et al. (1997); Barberis (2000); Campbell and Viceira (1999) estimate the demand to be large.

nant type of retirement plan[6] necessitates the provision of financial and insurance contracts to assist individuals with managing financial, mortality and longevity risks. Due to the mounting challenge of hedging liabilities associated to financial guarantees, insurers are more inclined to offer unit-linked contracts such as the VA (Antolín et al., 2011; Goecke, 2013; Koijen and Yogo, 2016a; International Monetary Fund, 2017). Our revelation can improve the design of these contracts.

We present our setup in Section 3.2, then prove the equivalent formulation in Section 3.3. Next, we introduce the VA, and derive its optimal parameters in Section 3.4. In Section 3.5, we evaluate individual welfare under the optimal solution and under the VA with financial market parameters calibrated to US data. We investigate the sensitivity of the welfare losses with respect to the model parameters in Section 3.6. Section 3.7 concludes.

3.2 Setting

3.2.1 Financial Market

We consider a frictionless and complete financial market in continuous time. The planning horizon is of length $T - t_0$, indexed by $t_0 + u$, for $u \in [0, T - t_0]$, whereas the financial market is composed of an instantaneous risk-free asset, a risky stock index, and a constant τ-maturity bond fund, $0 < \tau \le T - t_0$. Expected returns on these securities are determined by the stochastic discount factor of the economy, M_{t_0+u}, that has the dynamics $\mathrm{d}M_{t_0+u} = -r_{t_0+u} M_{t_0+u}\,\mathrm{d}u + \phi_S M_{t_0+u}\,\mathrm{d}Z_{S,\,t_0+u} + \phi_r M_{t_0+u}\,\mathrm{d}Z_{r,\,t_0+u}$. Z_S and Z_r are one-dimensional Brownian motions defined on a complete, filtered probability space, $\left(\Omega,\, \{\mathfrak{F}_u\}_{u=t_0}^{T},\, \mathbb{P}\right)$. The correlation between Z_S and Z_r is ρ_{Sr}. ϕ_S and ϕ_r are the constant loadings on the stochastic innovations in the economy that determine the securities' market price of risk, λ_S and λ_r.[7]

Following Vasicek (1977), we assume that the instantaneous risk-free asset,

[6] In 1975, close to 70% of all US retirement assets were in Defined Benefit (DB) plans. In 2015, DB assets accounted for only 33% of total retirement assets. Over the same period, assets in DC plans and Individual Retirement Accounts (IRAs) grew from 20% to 59% (Investment Company Institute, 2016). In the UK, 98% of the FTSE 350 companies offer a DC pension plan in 2017 (Towers Watson, 2017).

[7] In particular, $\rho = \begin{bmatrix} 1 & \rho_{Sr} \\ \rho_{Sr} & 1 \end{bmatrix}$, $\phi = (\phi_S,\, \phi_r)'$, and $\lambda = (\lambda_S,\, \lambda_r)' = -\rho\phi$. The standard Brownian motion with respect to the risk-neutral probability measure, $\mathbb{Q}$, is $Z^{\mathbb{Q}} = (Z_S^{\mathbb{Q}},\, Z_r^{\mathbb{Q}})$, with $Z_{S,\,t}^{\mathbb{Q}} = Z_{S,\,t} + \lambda_S t$ and $Z_{r,\,t}^{\mathbb{Q}} = Z_{r,\,t} + \lambda_r t$.

whose rate at time $t_0 + u$ is denoted by r_{t_0+u}, follows the Ornstein-Uhlenbeck process, $\mathrm{d}r_{t_0+u} = \kappa(\mu_r - r_{t_0+u})\,\mathrm{d}u + \sigma_r\,\mathrm{d}Z_{r,\,t_0+u}$. κ is the rate of convergence of r_{t_0+u} to the long-term average value of the interest rate, μ_r. σ_r is the instantaneous volatility of the interest rate. The money market account, m_{t_0+u}, accrues at the rate r_{t_0+u}, hence is governed by the dynamics $\mathrm{d}m_{t_0+u} = r_{t_0+u} m_{t_0+u}\,\mathrm{d}u$. The stock index, denoted by S_{t_0+u}, follows the diffusion process $\mathrm{d}S_{t_0+u} = (r_{t_0+u} + \lambda_S\sigma_S)S_{t_0+u}\,\mathrm{d}u + \sigma_S S_{t_0+u}\,\mathrm{d}Z_{S,\,t_0+u}$. σ_S is the instantaneous stock price volatility whereas $\lambda_S\sigma_S$ is the constant equity risk premium.[8]

The stochastic discount factor implies that the price of a zero-coupon bond with t years to maturity at time t_0+u is $P^t_{t_0+u} = \exp\left(A(t) - B(t)r_{t_0+u}\right)$, where

$$A(t) \quad = \quad \mu_r^*[B(t) - t] - \frac{\sigma_r^2}{4\kappa^2}\left[2\left(B(t) - t\right) + \kappa B(t)^2\right]$$

$\mu_r^* = \mu_r - \lambda_r\sigma_r/\kappa$, and $B(t) = \left(1 - \exp\left(-\kappa t\right)\right)/\kappa$.[9] By Itô's Lemma, the dynamics of the zero-coupon bond price is

$$\mathrm{d}P^t_{t_0+u} \quad = \quad [r_{t_0+u} - \lambda_r\sigma_r B(t)]P^t_{t_0+u}\,\mathrm{d}u - \sigma_r B(t) P^t_{t_0+u}\,\mathrm{d}Z_{r,\,t_0+u}$$

Additionally, with the one-factor interest rate model, any fixed-income security, including bonds of any maturity, can be replicated with a dynamic investment strategy involving a single, long-lived, arbitrary bond, and the money market account. In particular, to replicate a bond of maturity t, we invest $\theta_{t_0+l} = -B(t-l)/B(\tau)$ in the constant τ-maturity bond fund, and $1 - \theta_{t_0+l}$ in the money market account at any date $l \in [0,\, t - t_0]$. Therefore, it suffices to include only a constant τ-maturity bond fund among the available bonds.

3.2.2 Individuals

Individuals exhibit CRRA, are endowed with initial wealth W_{t_0}, and derive utility from consumption over a finite horizon. We introduce another time index, $t_0 + h$, $h \in [0,\, T - t_0]$, for the consumption horizon. The reason for the VA's suboptimality hinges on the distinction of the consumption horizon from

[8]Evidence for predictability of the equity risk premium is mixed, e.g., Bossaerts and Hillion (1999); Ang and Bekaert (2007); Welch and Goyal (2008). We concentrate on the setting where changes in the interest rate term structure underlie the time-variation in the investment opportunity set.

[9]Refer to Brennan and Xia (2002) for the derivations. μ_r^* is interpretable as the long-run mean of the interest rate under the risk-neutral measure.

the planning horizon that is indexed by $t_0 + u$, $u \in [0, T - t_0]$.

Let C_{t_0+h} be the individual's consumption at time $t_0 + h$. The individual derives utility from consumption according to Equation (3.2.1), and has no bequest motives.

$$U\left(\{C_{t_0+h}\}_{h=0}^{T-t_0}\right) = \int_0^{T-t_0} \mathrm{e}^{-\beta h}\, u\left(C_{t_0+h}\right)\, \mathrm{d}h \tag{3.2.1}$$

$$\begin{aligned} \beta &= \text{subjective discount factor} \\ \gamma &= \text{risk aversion parameter} \\ C_{t_0+h} &= \text{consumption in year } t_0 + h \end{aligned}$$

where

$$u\left(C_{t_0+h}\right) = \begin{cases} \frac{C_{t_0+h}^{1-\gamma}}{1-\gamma} & \gamma > 0,\, \gamma \neq 1 \\ \log\left(C_{t_0+h}\right) & \gamma = 1 \end{cases}$$

3.3 Optimal Consumption and Investment Problem

3.3.1 General Formulation

The optimal consumption and investment problem is the maximization of the individual's conditional expected utility over the consumption horizon by choosing the consumption and the investment policy, subject to a dynamic budget equation. Merton (1973) solves the problem by dynamic programming. Brennan and Xia (2002) and Munk and Sørensen (2004) apply the martingale method (Karatzas et al., 1987; Cox and Huang, 1989) to formulate it as a static variational problem as given by Equations (3.3.1) and (3.3.2). They first derive the optimal consumption, and then deduce the investment policy which exists due to completeness of the financial market.

$$\sup_{\{C_{t_0+h}\}_{h=0}^{T-t_0}} \quad \mathbb{E}_{t_0}\left[U\left(\{C_{t_0+h}\}_{h=0}^{T-t_0}\right)\right] \tag{3.3.1}$$

$$\text{subject to} \quad \mathbb{E}_{t_0}\left[\int_0^{T-t_0} C_{t_0+h} \frac{M_{t_0+h}}{M_{t_0}}\, \mathrm{d}h\right] \leq W_{t_0} \tag{3.3.2}$$

Theorem 1. *(Brennan and Xia, 2002; Munk and Sørensen, 2004) The optimal consumption at time $t_0 + h$, for any $h \in [0, T - t_0]$, $t_0 \in [0, T]$, that solves (3.3.1) and (3.3.2) is*

$$C^*_{t_0+h} = W_{t_0} X^*_{t_0}(h, T - t_0)\, Y^*_{t_0+h}(h) \tag{3.3.3}$$

where

$$X^*_{t_0}(h, t) \equiv \frac{\exp\left(-\frac{\beta}{\gamma} h\right) \mathbb{E}_{t_0}\left[\left(\frac{M_{t_0+h}}{M_{t_0}}\right)^{1-\frac{1}{\gamma}}\right]}{\int_0^t \exp\left(-\frac{\beta}{\gamma} l\right) \mathbb{E}_{t_0}\left[\left(\frac{M_{t_0+l}}{M_{t_0}}\right)^{1-\frac{1}{\gamma}}\right] \mathrm{d}l} \tag{3.3.4}$$

$$Y^*_{t_0+u}(h) \equiv \frac{\left(\frac{M_{t_0+u}}{M_{t_0}}\right)^{-\frac{1}{\gamma}}}{\mathbb{E}_{t_0}\left[\left(\frac{M_{t_0+h}}{M_{t_0}}\right)^{1-\frac{1}{\gamma}}\right]} \tag{3.3.5}$$

The first term of Equation (3.3.3) is the initial wealth; the second term represents the division of initial wealth over the consumption horizon indexed by h; the third reflects the return of an investment portfolio. This three-term representation alludes to the two-stage formulation in which the individual selects $X^*_{t_0}$ and $Y^*_{t_0}$.

A crucial remark on $C^*_{t_0+h}$ is that although it is indexed only on h, its third component, $Y^*_{t_0+u}(h)$, varies along the planning horizon, u as well. The VA, which we introduce in Section 3.4, despite also having a three-term representation as in Equation (3.3.3), is suboptimal because the investment portfolio depends only on the planning but not the consumption horizon. This constraint induces an incompatibility between the goals of optimally allocating initial wealth to each consumption period, and attaining the desired exposure to financial risks.

The optimal consumption path, $C^*_{t_0+h}$, is feasible in that it can be financed by the dynamic investment policy in Theorem 2.

Theorem 2. *(Theorem 4 of Brennan and Xia, 2002) The optimal investment policy for problem (3.3.1) and (3.3.2) is*

$$\left\{\theta^*_{t_0+u}(h)\right\}_{u=0}^{h} = \left\{\theta^*_{S,\, t_0+u}(h),\, \theta^*_{B,\, t_0+u}(h)\right\}_{u=0}^{h}$$

where

$$\theta^*_{S,\,t_0+u}(h) = -\frac{\phi_S}{\gamma\sigma_S} \tag{3.3.6}$$

$$\theta^*_{B,\,t_0+u}(h) = \frac{\phi_r}{\gamma\sigma_r B(\tau)} + \left(1-\frac{1}{\gamma}\right)\frac{\hat{B}(h-u)}{B(\tau)} \tag{3.3.7}$$

$$\hat{B}(h-u) = \int_0^{h-u} X^*_{t_0}(l,\,h-u)\,B(l)\,\mathrm{d}l \tag{3.3.8}$$

$$X^*_{t_0}(h,\,t) = \textit{Equation } (3.3.4)$$

Equations (3.3.6) and (3.3.7) resemble the optimal investment policy of Merton (1973). The proportion of investment in the risky stock index is constant over the planning and consumption horizons. The investment in the constant maturity bond fund consists of two terms, which Merton (1971) names as the speculative and the hedging term. The constant terms in θ^*_S and θ^*_B are the speculative demand. They depict the relation between the portfolio choice and the expected excess return on the available securities. Both speculative terms are rising in the assets' expected excess return, and are inversely related to the individual's risk aversion level.

The second term of Equation (3.3.7) represents the hedge demand for bonds, through which the individual attempts to hedge against unfavorable and unexpected evolution of interest rates. The hedge demand rises in the individual's risk aversion level. Its presence underlies the dependance of $Y^*_{t_0}$ in Equation (3.3.3) on both the consumption horizon, h, and the planning horizon, u. It implies that the optimal interest rate risk hedge is only attainable when the investment policy varies with both u and h. The relevant hedge portfolio is a weighted-average of bonds with maturities of every consumption date. The set of weights, $X^*_{t_0}$, is identical to the allocation of initial wealth over the consumption horizon.

The log utility individual, who has $\gamma = 1$, has no hedge demand. The definition of the VA's reference portfolio that varies with only u does not constrain her decision. Therefore, she suffers no utility loss if presented with a VA. The infinitely risk averse individual, who has $\gamma \to \infty$, has no speculative demand.

We pursue this line of thought in the next sub-section: The intertemporal consumption and investment problem can be interpreted as a series of terminal wealth problems, one for every consumption date. Furthermore, the optimal investment policy with intertemporal consumption (Theorem 2) is a weighted average of the optimal strategies of the same series of terminal wealth problems (Theorem 4).

3.3.2 Two-Stage (2S) Formulation

We decompose the utility-maximization problem (3.3.1) and (3.3.2) into two stages (2S), as per the "Lockbox Separation" of Sharpe (2007). The individual divides initial wealth into a series of "lockboxes" that is devoted to consumption at a specific date, and that follows a deterministic investment strategy that varies by both the planning and consumption horizons.

On each consumption date, the appropriate box is unlocked and the value of the box, including the financial market gains is consumed. In the absence of interest rate risk, Sharpe (2007) shows the separation of the optimal investment strategy into a series of static strategies that finances consumption at each date. Assuming that interest rate risk exists, we demonstrate that the separation holds with dynamic investment strategies.

At time t_0, the individual determines the fraction of initial wealth, W_{t_0}, to place in each lockbox. The lockboxes are indexed by the consumption horizon, $h \in [0, T - t_0]$. Let $\left\{X_{t_0}^{2S}(h, T - t_0)\right\}_{h=0}^{T-t_0}$ denote the fraction of initial wealth reserved for consumption at date $t_0 + h$, conditional on information at time t_0. It satisfies the constraint $\int_{h=0}^{T-t_0} X_{t_0}^{2S}(h, T - t_0)\, \mathrm{d}h = 1$. Each lockbox is worth $W_{t_0} X_{t_0}^{2S}(h, T - t_0)$ at time t_0.

Lockbox h is indexed to a reference investment portfolio that evolves with the planning horizon.[10] For every consumption date h, there is a dedicated reference portfolio whose value at time $t_0 + u$ is $\tilde{W}_{t_0+u}^{2S}(h)$, $u \in [0, h]$. $\tilde{W}_{t_0+u}^{2S}(h)$ follows an investment policy, $\left\{\theta_{t_0+u}^{2S}(h)\right\}_{u=0}^{h} = \left\{\theta_{S,\, t_0+u}^{2S}(h), \theta_{B,\, t_0+u}^{2S}(h)\right\}_{u=0}^{h}$ that varies with h and u. $\theta_{S,\, t_0+u}^{2S}(h)$ is the proportion of $\tilde{W}_{t_0+u}^{2S}(h)$ invested in the risky stock index, $\theta_{B,\, t_0+u}^{2S}(h)$ is the proportion invested in the constant maturity bond fund, while the residual, $1 - \theta_{S,\, t_0+u}^{2S}(h) - \theta_{B,\, t_0+u}^{2S}(h)$ is invested in the money market account at time $t_0 + u$.

[10] This is a generalization of the term "reference portfolio" in Chapter 2, which follows an investment policy that is dependent only on the planning horizon, u, but not on the consumption horizon, h.

The dynamics of $\tilde{W}^{2S}_{t_0+u}(h)$ with respect to the planning horizon, u, is

$$\begin{aligned}\frac{\mathrm{d}\tilde{W}^{2S}_{t_0+u}(h)}{\tilde{W}^{2S}_{t_0+u}(h)} &= \left(r_{t_0+u} + \theta^{2S}_{S,\,t_0+u}(h)\,\lambda_S\sigma_S - \theta^{2S}_{B,\,t_0+u}(h)\,\lambda_r\sigma_r B(\tau)\right)\mathrm{d}u + \\ &\quad \theta^{2S}_{S,\,t_0+u}(h)\,\sigma_S\,\mathrm{d}Z_{S,\,t_0+u} - \\ &\quad \theta^{2S}_{B,\,t_0+u}(h)\,\sigma_r B(\tau)\,\mathrm{d}Z_{r,\,t_0+u} \qquad (3.3.9)\end{aligned}$$

The 2S formulation comprises three sets of series of the allocation of wealth, $\left\{X^{2S}_{t_0}(h,\, T-t_0)\right\}^{T-t_0}_{h=0}$; the reference portfolio, $\left\{\left\{Y^{2S}_{t_0+u}(h)\right\}^{h}_{u=0}\right\}^{T-t_0}_{h=0}$; and the investment policies, $\left\{\left\{\theta^{2S}_{t_0+u}(h)\right\}^{h}_{u=0}\right\}^{T-t_0}_{h=0}$. At time t_0+h , the individual consumes $C^{2S}_{t_0+h}$. This is the accrued value of the portion of initial wealth that is reserved for consumption at that date–the value of the "lockbox"–when it has been invested according to the investment policy of its reference portfolio since time t_0.

$$C^{2S}_{t_0+h} = W_{t_0} X^{2S}_{t_0}(h,\, T-t_0)\, Y^{2S}_{t_0+h}(h) \qquad (3.3.10)$$

$$Y^{2S}_{t_0+u}(h) \equiv \frac{\tilde{W}^{2S}_{t_0+u}(h)}{\tilde{W}^{2S}_{t_0}(h)} \qquad (3.3.11)$$

The three-term representation of Equation (3.3.10) not only resembles the optimal consumption, Equation (3.3.3), but also has identical interpretation.

For any given series of investment policies for the lockboxes, there is an optimal division of initial wealth. The first stage of 2S is to seek this optimal division of initial wealth. The second stage regards every date t_0+h as a standalone utility maximization problem in which the individual derives utility from consumption only at date t_0+h, by selecting the reference portfolio investment policy. The solution to this terminal wealth problem in our setting is known (Sørensen, 1999).

Under the 2S formulation, we have two static variational problems that we solve by dynamic programming. We first seek the optimal division of initial wealth for any investment policy. After obtaining the optimal division rule to any given investment policy, we find the optimal investment policy. The first stage is formulated as (3.3.12) and (3.3.13), and solved in Theorem 3. The second stage is formulated as (3.3.15) and (3.3.16). Theorem 4 presents its solution.

The first stage of the 2S problem is

$$\sup_{\left\{X_{t_0}^{2S}(h,\, T-t_0)\right\}_{h=0}^{T-t_0}} \quad \mathbb{E}_{t_0}\left[U\left(\left\{C_{t_0+h}^{2S}\right\}_{h=0}^{T-t_0}\right)\right] \tag{3.3.12}$$

$$\text{subject to} \quad \int_{h=0}^{T-t_0} X_{t_0}^{2S}(h,\, T-t_0)\,\mathrm{d}h = 1 \tag{3.3.13}$$

$$\begin{aligned} C_{t_0+h}^{2S} &= \text{Equation (3.3.10)} \\ \frac{\mathrm{d}\tilde{W}_{t_0+u}^{2S}(h)}{\tilde{W}_{t_0+u}^{2S}(h)} &= \text{Equation (3.3.9) with} \\ & \quad \left\{\theta_{t_0+u}^{2S}(h)\right\}_{u=0}^{h} \text{ given.} \end{aligned}$$

Theorem 3. *The optimal division of initial capital,* $\left\{X_{t_0}^{2S}(h,\, T-t_0)\right\}_{h=0}^{T-t_0}$, *that solves (3.3.12) and (3.3.13) is*

$$\frac{\exp\left(-\frac{\beta}{\gamma}h\right)\mathbb{E}_{t_0}\left[\left(Y_{t_0+h}^{2S}(h)\right)^{1-\gamma}\right]^{\frac{1}{\gamma}}}{\int_0^{T-t_0}\exp\left(-\frac{\beta}{\gamma}l\right)\mathbb{E}_{t_0}\left[\left(Y_{t_0+l}^{2S}(l)\right)^{1-\gamma}\right]^{\frac{1}{\gamma}}\,\mathrm{d}l} \tag{3.3.14}$$

where $Y_{t_0+h}^{2S}(h) =$ *Equation* (3.3.11).

Proof. Appendix 3.D. □

Fixing a $h \in [0,\, T-t_0]$, the second stage of the 2S formulation is

$$\sup_{\left\{\theta_{t_0+u}^{2S}(h)\right\}_{u=0}^{h}} \quad \mathbb{E}_{t_0}\left[U\left(\tilde{W}_{t_0+h}^{2S}(h)\right)\right] \tag{3.3.15}$$

$$\frac{\mathrm{d}\tilde{W}_{t_0+u}^{2S}(h)}{\tilde{W}_{t_0+u}^{2S}(h)} = \text{Equation (3.3.9)} \tag{3.3.16}$$

Theorem 4. *(Sørensen, 1999) For any* $h \in [0,\, T-t_0]$, *the optimal investment policy for the portion of capital reserved for consumption at horizon* h, *at time* t_0+u, $u \in [0,\, h]$, *conditional on time* t_0, *that solves (3.3.15) and (3.3.16) is*

$$\left\{\theta_{t_0+u}^{2S}(h)\right\}_{u=0}^{h} = \left\{\theta_{S,\, t_0+u}^{2S}(h),\, \theta_{B,\, t_0+u}^{2S}(h)\right\}_{u=0}^{h}$$

where

$$\theta_{S,\, t_0+u}^{2S}(h) = -\frac{\phi_S}{\gamma\sigma_S} \tag{3.3.17}$$

$$\theta_{B,\, t_0+u}^{2S}(h) = \frac{\phi_r}{\gamma\sigma_r B(\tau)} + \left(1-\frac{1}{\gamma}\right)\frac{B(h-u)}{B(\tau)} \tag{3.3.18}$$

The optimal investment in the stock index, and the speculative demand for the constant maturity bond fund are constant and identical to those in the general formulation (Equations (3.3.6) and (3.3.7)). The optimal investment policies of the general and the 2S formulations differ only in the hedge demand of the constant maturity bond fund. The bond hedge demand in the general formulation, Equation (3.3.7), is a weighted average of that in the 2S formulation, Equation (3.3.18). $\hat{B}\,(h-u)$ in Equation (3.3.8) defines the weights.

3.3.3 Equivalence of the General and 2S Formulations

We demonstrate the equivalence of the consumption paths in the general formulation, $C^{*}_{t_0+h}$ (Section 3.3.1), and in the 2S formulation, $C^{2S}_{t_0+h}$ (Section 3.3.2). They are each composed of three terms

$$\begin{aligned} C^{K}_{t_0+h} &= W_{t_0} X^{K}_{t_0}\,(h,\, T-t_0)\, Y^{K}_{t_0+h}\,(h) \\ K &\in \{*,\, 2S\} \end{aligned} \tag{3.3.19}$$

For $K=*$, $X^{*}_{t_0}\,(h,\, T-t_0) =$ Equation (3.3.4) and $Y^{*}_{t_0+h}\,(h) =$ Equation (3.3.5). For $K=2S$, $X^{2S}_{t_0}\,(h,\, T-t_0) =$ Equation (3.3.14) and $Y^{2S}_{t_0+h}\,(h) =$ Equation (3.3.11). Even if the consumption paths share the mathematical form of Equation (3.3.19), it is not evident that $C^{*}_{t_0+h} = C^{2S}_{t_0+h}$. We prove this by showing that $X^{*}_{t_0}\,(h,\, T-t_0) = X^{2S}_{t_0}\,(h,\, T-t_0)$ and $Y^{*}_{t_0+h}\,(h) = Y^{2S}_{t_0+h}\,(h)$, for all $h \in [0,\, T-t_0]$.

Lemma 5. $X^{*}_{t_0}\,(h,\, T-t_0) = X^{2S}_{t_0}\,(h,\, T-t_0)$.

Proof. Appendix 3.E. The proof involves a straightforward derivation of the conditional expectations in Equations (3.3.4) and (3.3.14). □

As $\left\{Y^{K}_{t_0+u}\,(h)\right\}^{h}_{u\in 0}$ is a random variable, demonstrating that $Y^{*}_{t_0+h}\,(h) = Y^{2S}_{t_0+h}\,(h)$ is more involved. We achieve this in Lemma 6, where we show that for any given $h \in [0,\, T-t_0]$, the terms share the same dynamics over $u \in [0,\, h]$, and verify that they have the same starting values, $Y^{*}_{t_0+0}\,(h) = Y^{2S}_{t_0+0}\,(h)$.

Lemma 6. *For any* $h \in [0,\, T-t_0]$ *and* $u \in [0,\, h]$,

$$\frac{\mathrm{d}Y^{*}_{t_0+u}\,(h)}{Y^{*}_{t_0+u}\,(h)} = \frac{\mathrm{d}Y^{2S}_{t_0+u}\,(h)}{Y^{2S}_{t_0+u}\,(h)} \tag{3.3.20}$$

Additionally, $Y^{*}_{t_0+0}\,(h) = Y^{2S}_{t_0+0}\,(h)$.

Proof. Appendix 3.F. In the proof of Equation (3.3.20), we derive the expressions of $\mathrm{d}Y^{K}_{t_0+u}\,(h)\,/Y^{K}_{t_0+u}\,(h)$ for $K \in \{*,\, 2S\}$ by applications of Itô's Lemma, deduce the investment policy for the portfolio whose value has the same dynamics as $\mathrm{d}Y^*_{t_0+u}(h)/Y^*_{t_0+u}(h)$, and show that this investment portfolio coincides with that of 2S in Theorem 4. $\square$

Theorem 7. *For all $h \in [0, T - t_0]$, $C^*_{t_0+h} = C^{2S}_{t_0+h}$. The general and two-stage formulations of the optimal consumption and investment problem yield identical consumption paths.*

Proof. This follows from the definition of $C^*_{t_0+h}$ and $C^{2S}_{t_0+h}$ in Equation (3.3.19), Lemmas 5 and 6. $\square$

We focus on the 2S formulation for subsequent analysis.

3.4 Variable Annuity

The VA customizes the expected level and volatility of payments in two parameters: the assumed interest rate (AIR) and the reference portfolio investment policy. This two-pronged approach adheres closely to the intertemporal consumption and investment problem, which we show in Section 3.3.3 to be equivalent to the 2S formulation that first solves for the optimal division of capital over the consumption horizon, and then finds the optimal investment policy for each portion of divided capital.

The AIR is a deterministic rate that dictates the division of initial capital, W_{t_0} over the consumption horizon, $h \in [0, T - t_0]$. Let $a_{t_0} = \{a_{t_0}(h)\}_{h=0}^{T-t_0}$ denote the AIR conditional on information at time t_0. The per unit cost of a VA is[11]

$$A_{t_0}\left(\{a_{t_0}(h)\}_{h=0}^{T-t_0}\right) = \int_0^{T-t_0} \exp\left(-a_{t_0}(h) \times h\right) \mathrm{d}h \qquad (3.4.1)$$

The VA is indexed to a reference investment portfolio that is of value $\tilde{W}^{VA}_{t_0+u}$ at time $t_0 + u$. Together with the AIR, the reference portfolio determines the expectation and dispersion of the VA payments.[12] The VA reference portfolio

[11] As we assumed that individuals have a certain lifetime, the VA has no mortality credit. See Charupat and Milevsky (2002) for the definition of a VA with mortality credit. Blake et al. (2003) define an Equity-Linked-Annuity (ELA) as an annuity contract offering payments indexed to an equity investment portfolio, and an Equity-Linked-Income-Drawdown (ELID) as the ELA's counterpart without mortality credit. The contract we describe here is an ELID.

[12] Let $\tilde{r}$ denote the reference portfolio's expected return, and suppose that the AIR is time-invariant. Then an annuity contract with AIR $= \tilde{r}$ has a constant expected payment path. When AIR $< \tilde{r}$, then the expected payment stream is upward sloping, with increasing riskiness

follows the investment policy that varies with the planing horizon, u, but is independent of the consumption horizon, h, $\left\{\theta^{VA}_{t_0+u}\right\}_{u=0}^{T-t_0} = \left\{\theta^{VA}_{S,\,t_0+u},\, \theta^{VA}_{B,\,t_0+u}\right\}_{u=0}^{T-t_0}$. Unlike the 2S, for which there is one reference portfolio for each consumption date, the VA has a unique reference portfolio for all t_0+h. The independence of the VA's reference portfolio from the consumption horizon, h, is the reason for its suboptimality. The dynamics of $\tilde{W}^{VA}_{t_0+u}$ with respect to u is Equation (3.4.2).

$$\begin{aligned}\frac{\mathrm{d}\tilde{W}^{VA}_{t_0+u}}{\tilde{W}^{VA}_{t_0+u}} &= \left(r_{t_0+u} + \theta^{2S}_{S,\,t_0+u}\lambda_S\sigma_S - \theta^{2S}_{B,\,t_0+u}\lambda_r\sigma_r B(\tau)\right)\,\mathrm{d}u + \\ &\quad \theta^{2S}_{S,\,t_0+u}\sigma_S\,\mathrm{d}Z_{S,\,t_0+u} - \theta^{2S}_{B,\,t_0+u}\sigma_r B(\tau)\,\mathrm{d}Z_{r,\,t_0+u} \qquad (3.4.2)\end{aligned}$$

An individual with an initial capital worth W_{t_0}, who fully annuitizes at time t_0, consumes the VA payment, $C^{VA}_{t_0+h}$ at time t_0+h

$$C^{VA}_{t_0+h} = W_{t_0}\frac{\exp\left(-a_{t_0}(h)\times h\right)}{A_{t_0}\left(\left\{a_{t_0}(h)\right\}_{h=0}^{T-t_0}\right)}Y^{VA}_{t_0+h} \qquad (3.4.3)$$

$$Y^{VA}_{t_0+h} \equiv \frac{\tilde{W}^{VA}_{t_0+h}}{\tilde{W}^{VA}_{t_0}} \qquad (3.4.4)$$

The three-term composition of Equation (3.4.3) is reminiscent of the 2S consumption path, Equation (3.3.10). The second term represents the allocation of initial capital over the consumption horizon, whereas the final term reflects the return on the reference portfolio.

3.4.1 Optimal Variable Annuity

We define the optimal VA contract, VA*, as the VA that maximizes the individual's conditional expected utility, by choice of the AIR and the reference portfolio's investment policy. In particular, VA* is the solution to

$$\sup_{\left\{a_{t_0}(h)\right\}_{h=0}^{T-t_0},\ \left\{\theta^{VA}_{t_0+u}\right\}_{u=0}^{T-t_0}} \mathbb{E}_{t_0}\left[U\left(\left\{C^{VA}_{t_0+h}\right\}_{h=0}^{T-t_0}\right)\right] \qquad (3.4.5)$$

$$\begin{aligned}C^{VA}_{t_0+h} &= \text{Equation (3.4.3)}\\ \frac{\mathrm{d}\tilde{W}^{VA}_{t_0+u}}{\tilde{W}^{VA}_{t_0+u}} &= \text{Equation (3.4.2)}\end{aligned}$$

as the individual ages. Conversely, when AIR $> \tilde{r}$, then the expected payment stream is downward sloping, and the risk is higher during the initial payout phase. Horneff et al. (2010) expound on VA payments under numerous time-invariant AIRs and reference portfolios.

Similar to the utility-maximization problem of the 2S, we have two static variational problems that we solve by dynamic programming. We first find the optimal AIR taking the investment policy, $\left\{\theta_{t_0+u}^{VA}\right\}_{u=0}^{T-t_0}$ as given. Then with the optimal AIR, $a_{t_0}^*(h)$, we solve for the VA* investment policy.

3.4.1.1 Optimal AIR

Derivation of the optimal AIR is similar to solving for the optimal division of initial wealth under 2S, as formulated in (3.3.12) and (3.3.13), because the AIR has the same role. The problem statement is

$$\begin{aligned} \sup_{\{a_{t_0}(h)\}_{h=0}^{T-t_0}} & \quad \mathbb{E}_{t_0}\left[U\left(\left\{C_{t_0+h}^{VA}\right\}_{h=0}^{T-t_0}\right)\right] \\ C_{t_0+h}^{VA} &= \text{Equation (3.4.3)} \\ \frac{\mathrm{d}\tilde{W}_{t_0+u}^{VA}}{\tilde{W}_{t_0+u}^{VA}} &= \text{Equation (3.4.2) with} \\ & \quad \left\{\theta_{t_0+u}^{VA}\right\}_{u=0}^{T-t_0} \text{ is given} \end{aligned} \tag{3.4.6}$$

Theorem 8. *The optimal AIR, $a_{t_0}^*$ that solves (3.4.6) is $a_{t_0}^*(0) = 0$, and for $h \in (0, T - t_0]$,*

$$a_{t_0}^*(h) = \frac{\beta}{\gamma} - \frac{1-\gamma}{\gamma}\frac{g^{VA}(t_0, h)}{h} \tag{3.4.7}$$

where $g^{VA}(t_0, h) =$ Equation (3.C.2).

Proof. Appendix 3.G. □

$a_{t_0}^*(h)$ is a linear combination of the conditional expectation and variance of the log ratios of the reference portfolio value contained in $g^{VA}(t_0, h)$, and the individual's subjective discount factor, weighed by her risk aversion level. The definition of $g^{VA}(t_0, h)$ in Equation (3.C.2) implies

$$\exp\left(-a_{t_0}^*(h)\,h\right) = \exp\left(-\frac{\beta}{\gamma}h\right)\mathbb{E}_{t_0}\left[\left(Y_{t_0+h}^{VA}\right)^{1-\gamma}\right]^{\frac{1}{\gamma}} \tag{3.4.8}$$

Substituting Equation (3.4.8) into Equations (3.4.1) and (3.4.3), we have that a VA with $a_{t_0}^*(h)$ pays at each consumption date,

$$
\begin{aligned}
C^{VA}_{t_0+h} &= W_{t_0} X^{VA}_{t_0}(h,\, T-t_0)\, Y^{VA}_{t_0+h} && (3.4.9)\\
\text{where} & \\
X^{VA}_{t_0}(h,\, T-t_0) &\equiv \frac{\exp\left(-\frac{\beta}{\gamma}h\right)\mathbb{E}_{t_0}\left[\left(Y^{VA}_{t_0+h}\right)^{1-\gamma}\right]^{\frac{1}{\gamma}}}{\int_0^{T-t_0}\exp\left(-\frac{\beta}{\gamma}l\right)\mathbb{E}_{t_0}\left[\left(Y^{VA}_{t_0+l}\right)^{1-\gamma}\right]^{\frac{1}{\gamma}}\,\mathrm{d}l} && (3.4.10)\\
&= \frac{\exp\left(-a^*_{t_0}(h)\times h\right)}{A_{t_0}\left(\left\{a_{t_0}(h)\right\}_{h=0}^{T-t_0}\right)}\\
Y^{VA}_{t_0+h} &= \text{Equation (3.4.4)}
\end{aligned}
$$

3.4.1.2 Optimal Investment Policy

To obtain the optimal VA investment policy, we consider the VA that adopts the optimal AIR and solve the following

$$
\begin{aligned}
\sup_{\left\{\theta^{VA}_{t_0+u}\right\}_{u=0}^{T-t_0}} & \quad \mathbb{E}_{t_0}\left[U\left(\left\{C^{VA}_{t_0+h}\right\}_{h=0}^{T-t_0}\right)\right] && (3.4.11)\\
C^{VA}_{t_0+h} &= \text{Equation (3.4.9)}
\end{aligned}
$$

Theorem 9. *The investment policy for the optimal VA,*

$$\left\{\theta^{VA*}_{t_0+u}\right\}_{u=0}^{T-t_0} = \left\{\theta^{VA*}_{S,\,t_0+u},\, \theta^{VA*}_{B,\,t_0+u}\right\}_{u=0}^{T-t_0}$$

(if it exists) satisfies, for all $u \in [0,\, T-t_0]$, the following extremum condition

$$\int_u^{T-t_0} F\left(l,\, t_0,\, \{y_i(l)\}_{i=1}^{7}\right) G_L\left(u,\, t_0,\, T;\, \theta^{VA*}_{t_0+u}\right)\,\mathrm{d}l = 0 \qquad (3.4.12)$$

for both $L \in \{S,\, B\}$, where

$$
\begin{aligned}
G_S\left(u,\, t_0,\, T;\, \theta^{VA}_{t_0+u}\right) &= \lambda_S\sigma_S - \theta^{VA}_{S,\,t_0+u}\sigma_S^2\gamma + \theta^{VA}_{B,\,t_0+u}\rho_{Sr}\sigma_r\sigma_S B(\tau)\gamma + \\
&\quad (1-\gamma)\rho_{Sr}\sigma_r\sigma_S\int_u^{T-t_0} B(n-u)\,\mathrm{d}n && (3.4.13)\\
G_B\left(u,\, t_0,\, T;\, \theta^{VA}_{t_0+u}\right) &= -\lambda_r\sigma_r B(\tau) - \theta^{VA}_{B,\,t_0+u}\sigma_r^2 B^2(\tau)\gamma + \\
&\quad \theta^{VA}_{S,\,t_0+u}\rho_{Sr}\sigma_r\sigma_S B(\tau)\gamma - \\
&\quad (1-\gamma)\sigma_r^2 B(\tau)\int_u^{T-t_0} B(n-u)\,\mathrm{d}n && (3.4.14)
\end{aligned}
$$

$F\left(l,\, t_0,\, \{y_i(l)\}_{i=1}^{7}\right)$ *is Equation (3.H.2).*

Proof. Appendix 3.H. □

The complication of solving (3.4.11) lies in the constraints being integrals of the state equations, i.e., they are path-dependent. We do not obtain $\left\{\theta_{t_0+u}^{VA*}\right\}_{u=0}^{T-t_0}$ explicitly. The extremum condition (3.4.12) is composed of a system of non-linear equations indexed by $u \in [0, T - t_0]$. We estimate $\left\{\theta_{t_0+u}^{VA*}\right\}_{u=0}^{T-t_0}$ by numerically solving the system of simultaneous equations in Section 3.5.

3.5 Utility Loss of a Variable Annuity

We compare the individual's expected utility when consuming according to 2S and to VA* based on financial market parameters calibrated on US data. The base case adopts the set of financial market parameters in Section 3.5.1, is composed of individuals with a 45-year long consumption horizon (e.g., to exemplify the situation of a 65-year-old in year t_0 who lives with certainty until age 110 in year T), and possess risk aversion levels $\gamma = 2, 3, \ldots, 9, 10$.

3.5.1 Model Calibration

To provide descriptive numerical examples, we calibrate the financial market to monthly yields of US government bonds with maturities of three months, one year, five and ten years from August 1971 to December 2014.[13] As for the stock return, we use the monthly return on the CRSP value-weighted stock return, including dividends, on the NYSE, AMEX, and NASDAQ.[14] We apply the Kalman filtering approach and describe its setup in Appendix 3.I. Table 3.1 presents the estimated parameters and their standard errors obtained by the outer product of gradients.

[13]The choice of maturities is identical to de Jong (2000). He justifies exclusion of bond yields with maturity under 3 months by their exceptionally large one-period change, and those with maturity over 10 years due to the scarcity of bond data, which result in inaccurate interpolation. Yields for bonds of maturities 1, 5, and 10 years are from Gürkaynak et al. (2007), whereas yields for the 3-month Treasury bills are from the Federal Reserve Bank of St. Louis (FRED).

[14]CRSP and FRED data are monthly. To reconcile these with the daily data from Gürkaynak et al. (2007), we take the observation on the last day of the month.

Table 3.1: Estimates of Model Parameters

Maximum likelihood parameter estimates of the interest rate and the stock return, obtained by implementing the Kalman filter on monthly US government bond yields of 3-month, 1, 5, and 10-year maturities, and the return on the CRSP value-weighted stock index, from August 1971 till December 2014. Standard errors are by the outer product of gradients.

Parameter	**Estimate**	**Standard Error**
Stock Return Process: $\mathrm{d}S_{t_0+u} = (r_{t_0+u} + \lambda_S \sigma_S) S_{t_0+u}\,\mathrm{d}u + \sigma_S S_{t_0+u}\,\mathrm{d}Z_{S,\,t_0+u}$		
σ_S	0.158	0.004
λ_S	0.467	0.164
Short Rate Process: $\mathrm{d}r_{t_0+u} = \kappa(\mu_r - r_{t_0+u})\,\mathrm{d}u + \sigma_r\,\mathrm{d}Z_{r,\,t_0+u}$		
μ_r	0.036	0.043
κ	0.067	0.003
σ_r	0.017	0.001
λ_r	-0.350	0.177
Stochastic Discount Factor Process: $\mathrm{d}M_{t_0+u} = -r_{t_0+u} M_{t_0+u}\,\mathrm{d}u + \phi_S M_{t_0+u}\,\mathrm{d}Z_{S,\,t_0+u} + \phi_r M_{t_0+u}\,\mathrm{d}Z_{r,\,t_0+u}$		
ρ_{Sr}	0.120	0.049
ϕ_S	-0.516	
ϕ_r	0.412	

The estimated mean-reversion coefficient of the interest rate, κ, is 0.067. It implies a half-life of $\log(2)/0.067 \approx 10$ years. This slow rate of mean reversion indicates that the fitted term-structure is rather flat. We estimate σ_r to be 1.7%. Our estimates of κ and σ_r are similar to the estimates of 0.06 and of 1.4% respectively by de Jong (2000). The market price of interest rate risk is negative and significant, which means that the bond risk premium is positive, and all bonds yield a Sharpe ratio of $-\lambda_r = 0.35$. The expected excess return on a constant maturity three-month zero-coupon bond fund is 15 bps, and that for a one-, five- and ten-year constant maturity bond fund is 0.5%, 2.5% and 4.3%, respectively. The standard deviation of return to the three-month, one-, five- and ten-year constant maturity bond funds are 0.4%, 1.6%, 7.0% and 12.1% respectively. Additionally, the estimates for the bond yield measurement errors are small (i.e., of magnitudes 10^{-5} to 10^{-6}).

As for the stock index, we find an annualized stock index volatility of close to 16%, as do Brennan and Xia (2002), but our estimate for the stock index Sharpe ratio is higher at 0.47, relative to 0.34 in Brennan and Xia (2002). The estimated stock risk premium is $\lambda_S \sigma_S = 7.39\%$. The stock index is positively correlated with the nominal interest rate, in contrast to the negative correlation estimates of Brennan and Xia (2002); Munk et al. (2004), but aligned with Campbell (1987); Fama and French (1989); Shiller and Beltratti (1992).

3.5.2 Estimation of Utility Loss

3.5.2.1 Certainty Equivalent Wealth Loading (CEWL)

We evaluate individual welfare under 2S and VA* using the Certainty Equivalent Wealth Loading (CEWL). The CEWL is the proportional loading on the initial wealth that the individual requires to purchase a VA* that yields the same expected utility as with the 2S. It satisfies Equation (3.5.1). A positive CEWL suggests that the individual prefers the 2S over the VA*. An individual who is indifferent between the 2S and the VA* has a CEWL of zero.

$$\begin{aligned}
\mathbb{E}_{t_0}\left[U\left((1+CEWL)\,C^{VA*}\right)\right] &= \mathbb{E}_{t_0}\left[U\left(C^{2S}\right)\right] && (3.5.1)\\
C^{VA*} &= \text{Equation (3.4.9),}\\
&\quad \text{where } \left\{\theta^{VA}_{t_0+u}\right\}_{u=0}^{T-t_0} = \left\{\theta^{VA*}_{t_0+u}\right\}_{u=0}^{T-t_0}\\
C^{2S} &= \text{Equation (3.3.10)}
\end{aligned}$$

In particular,

$$CEWL = \left(\frac{\mathbb{E}_{t_0}\left[U\left(C^{2S}\right)\right]}{\mathbb{E}_{t_0}\left[U\left(C^{VA*}\right)\right]}\right)^{\frac{1}{1-\gamma}} - 1 \qquad (3.5.2)$$

To obtain the VA* investment policy for individuals with risk aversion levels $\gamma = 2, \ldots, 10$, we solve Equation (3.4.12) for $L \in \{S, B\}$ numerically with the financial market parameters of Table 3.1. We then compute the CEWL by Equation (3.5.2).

3.5.2.2 Discussion of the CEWL

In the base case, individuals require 7-19% more initial wealth under the VA* than the 2S to attain the same expected utility (Figure 3.5.1, top panel, 45-year horizon).

The utility loss under the VA* is due to the VA's restriction that the reference portfolio investment policy depends only on the planning horizon. By definition of $X_{t_0}^K$, for $K \in \{2S, VA\}$ (Equations (3.3.14) and (3.4.10)), the allocation of initial wealth over the consumption horizon is determined by the investment policy via the conditional expected return on the reference portfolio. The individual with VA* faces the constraint that changing the investment policy for any date $\tilde{u}$ along the planning horizon modifies the allocation of initial wealth to not only the consumption date, h, that coincides with $\tilde{u}$, but also to all other consumption dates. In contrast, the 2S allows the individual to customize the investment policy by the planning and the consumption horizons. This feature is essential to achieve the optimal allocation of initial wealth and optimal investment risk exposure.

The U-shaped curve in Figure 3.5.1 suggests that individuals who are least and most risk-averse have a higher CEWL than an individual who is moderately risk-averse. This non-monotonicity stems from a dichotomy between the objectives of optimally allocating the initial wealth and optimally choosing the level of investment risk under the VA.

To illustrate, consider an individual with high γ. She is risk averse and reluctant to substitute consumption intertemporally. Regardless of the investment opportunities, this individual desires a constant expected consumption growth rate, which is attainable if she consumes the long-run average return of the reference portfolio, and reserve an appropriate portion of wealth that adjusts for risk of each consumption date. Given that the individual is risk averse, this means allocating more wealth for consumption dates farther in the future, when the cumulative financial market performance has higher variability. Another way to mitigate the investment risk associated to distant consumption dates is to decrease investment risk in those periods. Yet, under the base case parameters, decreasing investment risk for the farther future entails a lower VA* allocation of initial wealth to those consumption dates as well.[15] This is incompatible with the individual's coincident desire to reserve more initial wealth for consumption at the end of the horizon. Therefore, she suffers substantial utility loss.

When it concerns a less risk averse individual, an analogous argument applies. This individual is willing to take higher investment risk, and prefers to allocate less wealth to consumption dates farther in the future, because the VA* payment on those dates have more time to accrue gains from the financial market. However, for the parameters that we consider, a higher allocation of initial

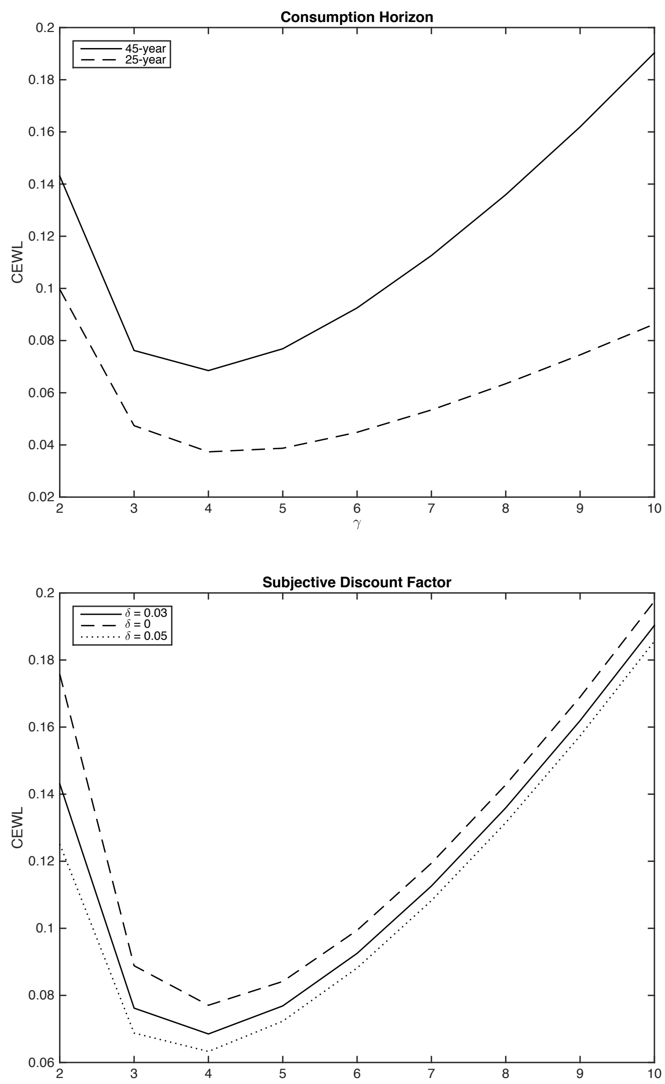
[15] Due to the absence of an analytical solution for the VA* investment policy, we are unable to generalize the relation between the VA* investment policy and the allocation of initial wealth over the horizon. This relation is dependent on the financial market parameters.

wealth to the nearer consumption dates is feasible by decreasing investment risk. The individual is also confronted with the opposing actions necessary to attain the optimal allocation of initial wealth and investment risk exposure.

The moderately risk averse individual has a lower CEWL relative to individuals with more extreme risk aversion levels. She is willing to take some investment risk, and prefers a comparatively uniform allocation of initial wealth over the consumption horizon. Therefore, the VA* yields moderate utility loss.

Figure 3.5.1: CEWL by Risk Aversion Level: Consumption Horizon and Subjective Discount Factor

This figure presents the CEWL by the individual's risk aversion level, γ that ranges from 2 to 10, when the consumption horizon is 25 and 45 years (top panel), and when the individual's subjective discount factor is either $\delta = 0$, 0.03, or 0.05 (bottom panel).



3.6 Robustness

We investigate the sensitivity of the CEWL with respect to parameters that characterize individual preference (i.e., length of the consumption horizon, subjective discount factor), and those governing the term structure of the stochastic interest rates (i.e., volatility, mean reversion level, rate of mean reversion, interest rate level on the date of annuitization relative to the mean reversion level). For the sensitivity with respect to the interest rate model parameters, we compare the CEWLs under the base case parameters in Table 3.1 to the CEWLs in two other situations, where each of the parameter is unilaterally either doubled or halved.

The magnitude of interest rate risk underlies the CEWL's sensitivity with respect to the interest rate model parameters. When interest rate risk is high, the ability to optimally hedge the risk is more valuable. The 2S consequently becomes more desirable. Hence, we first make a few remarks on the influence that select Vasicek (1977) model parameters have on the conditional variance of the interest rate.

The variance for the short rate process following the Vasicek (1977) model at time $t+u$, conditional on time t, $t \in [t_0, T]$, for $u \in [0, T-t_0]$ is

$$\mathrm{Var}_t\left[r_{t+u}\right] = \frac{\sigma_r^2}{2\kappa}\left(1-\exp\left(-2\kappa u\right)\right) \tag{3.6.1}$$

As $\mathrm{Var}_t\left[r_{t+u}\right]$ is increasing in u and σ_r when $\kappa > 0$, we postulate and confirm that the CEWL is higher when the consumption horizon is longer (Figure 3.5.1; top panel), and when the instantaneous volatility is higher (Figure 3.6.1; top panel).

We also observe that a lower subjective discount factor corresponds to a higher CEWL (Figure 3.5.1; bottom panel). This is because the individual with $\delta = 0$ regards consumption in the farther future as equally important as consumption at present. As interest rate risk increases in the horizon, this individual suffers a larger utility loss when unable to optimally hedge it, relative to an individual who prefers to consume more now than later (i.e., $\delta > 0$).

The effect of κ on the interest rates' conditional volatility depends on the horizon, u, and κ. For a horizon and a mean reversion rate that are greater than zero, u, $\kappa > 0$, which is the relevant case to consider in the present retirement context, $\mathrm{Var}_t\left[r_{t+u}\right]$ is decreasing in κ (Proof in Appendix 3.J). When κ is larger, interest rates converge sooner to their mean reversion level. For instance, in the base case, the interest rate's half-life is around 10 years, but at 2κ, the half-life reduces to 30 months only (i.e., $\log(2) / (2 \times 0.067) \approx 2.25$ years). Having any deviation of the interest rate from its mean reversion level be resolved sooner is akin to having lower interest rate risk. Thus, the CEWL is inversely related to κ (Figure 3.6.1; bottom panel).

The top panel of Figure 3.6.2 shows that the CEWL is decreasing in μ_r. μ_r influences the reference portfolio's expected returns but not interest rate risk (Equation (3.6.1)). Hence, an individual consuming C^{2S} has an investment policy that is independent of μ_r (Equations (3.3.17) and (3.3.18)). Suppose μ_r is decreased. The individual with 2S maintains the same investment policy as it remains optimal. To cope with the lower level of expected return, she reserves a larger proportion of initial wealth for consumption at later dates.

In contrast, the individual with a VA* alters the investment policy in response to a decrease in μ_r. A lower μ_r amplifies the extent of the opposing adjustments on the investment policy that is necessary for her to attain the desired allocation of initial wealth over the consumption horizon, and the optimal investment risk exposure. Take the case of a risk averse individual. Her rational response to a lower μ_r is to either allocate more initial wealth to future consumption dates, bear more investment risk, or both. Under the base case parameters, the desired adjustment to initial wealth allocation is possible by making riskier investments. The higher financial market risk exposure necessitated under either response is at odds with the individual's risk aversion. Consequently, she experiences larger utility loss when μ_r is lower.

As with μ_r, r_0 also has an inverse relationship with the CEWL (Figure 3.6.2; bottom panel). The implication of doubling and halving r_0 has a similar underlying explanation as with perturbing μ_r. However, the CEWL difference between the perturbed and the base case is smaller than when μ_r is perturbed by the same magnitude. This is because the effect on r_0 is transient–r_0 eventually converges to μ_r.

Figure 3.6.1: CEWL by Risk Aversion Level: Perturbed σ_r and κ

This figure presents the CEWL by the individual's risk aversion level, γ that ranges from 2 to 10, for the base case and when either σ_r (top panel) or κ (bottom panel) of the Vasicek (1977) short rate process is doubled or halved relative to its value in the base case.

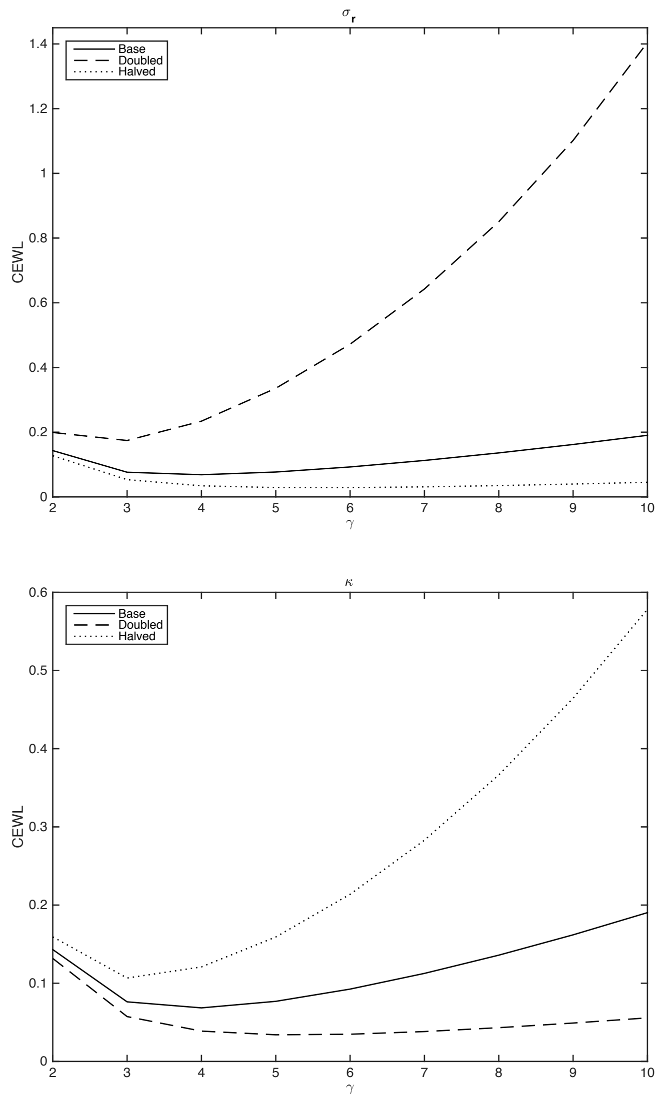
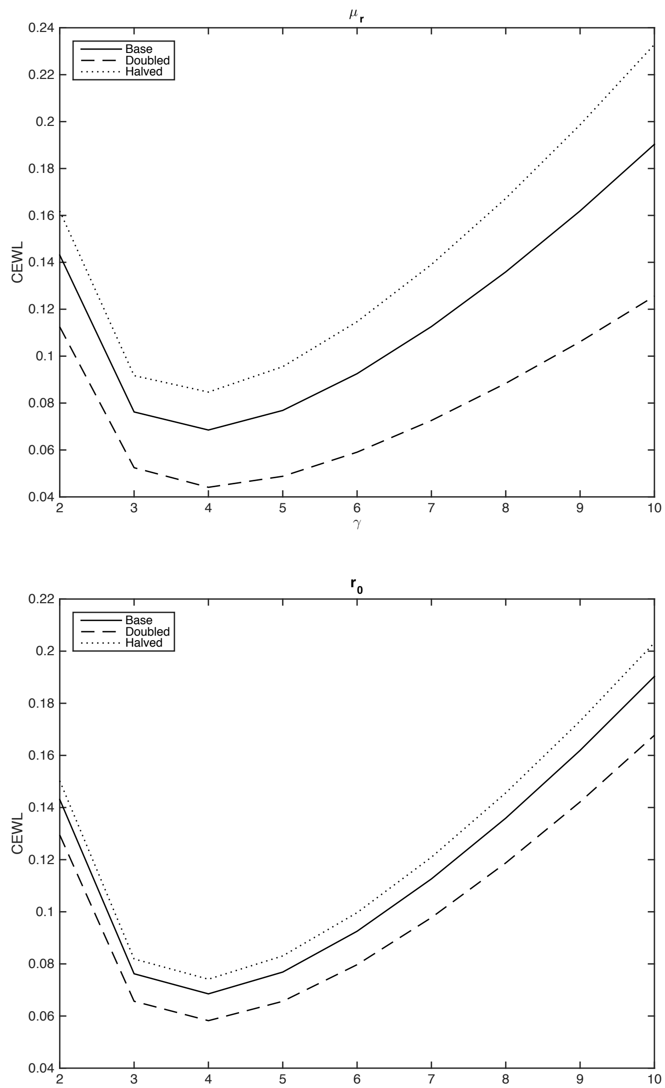


Figure 3.6.2: CEWL by Risk Aversion Level: Perturbed μ_r and r_0

This figure presents the CEWL by the individual's risk aversion level, γ that ranges from 2 to 10, for the base case and when either μ_r (top panel) or r_0 (bottom panel) of the Vasicek (1977) short rate process is doubled or halved relative to its value in the base case.



3.7 Conclusion

We demonstrate the equivalence of the solution to the optimal consumption and investment problem when interest rate risk exists, with the outcome of a two-stage problem that chooses the allocation of initial capital over the consumption horizon, and the investment policy for each portion of the divided capital. We apply the equivalent formulation to demonstrate that the VA entails sizable welfare losses due to the contract's inability to optimally hedge interest rate risk of every consumption date. We illustrate the economic implication of our revelation with a VA, but we emphasize its relevance to unit-linked contracts. Generalizing the contract unit's definition to allow its dependance on both the planning and consumption horizons improves the policyholder's welfare.

Appendix to Chapter 3

3.A $\mathbb{E}_{t_0}\left[\left(M_{t_0+h}/M_{t_0}\right)^{1-\frac{1}{\gamma}}\right]$

Define

$$\Gamma \equiv \phi_r^2 + \phi_S^2 + 2\rho_{Sr}\phi_r\phi_S \tag{3.A.1}$$

For all $h \in [0, T - t_0]$, M_{t_0+h}/M_{t_0} is log-normally distributed, hence

$$\begin{aligned}
\mathbb{E}_{t_0}\left[\left(\frac{M_{t_0+h}}{M_{t_0}}\right)^{1-\frac{1}{\gamma}}\right] &= \exp\left\{\left(1-\frac{1}{\gamma}\right)\mathbb{E}_{t_0}\left[\log\left(\frac{M_{t_0+h}}{M_{t_0}}\right)\right] + \right.\\
&\quad \left.\frac{1}{2}\left(1-\frac{1}{\gamma}\right)^2 \mathrm{Var}_{t_0}\left[\log\left(\frac{M_{t_0+h}}{M_{t_0}}\right)\right]\right\}\\
&= \exp\left\{\left(1-\frac{1}{\gamma}\right)\times\left(\mathbb{E}_{t_0}\left[\log\left(\frac{M_{t_0+h}}{M_{t_0}}\right)\right] + \right.\right.\\
&\quad \left.\left.\frac{1}{2}\left(1-\frac{1}{\gamma}\right)^2 \mathrm{Var}_{t_0}\left[\log\left(\frac{M_{t_0+h}}{M_{t_0}}\right)\right]\right)\right\}\\
&= \exp\left\{\left(1-\frac{1}{\gamma}\right) g^{SDF}(t_0, h)\right\} \qquad (3.A.2)
\end{aligned}$$

$$\begin{aligned}
g^{SDF}(t_0, h) &\equiv \mathbb{E}_{t_0}\left[\log\left(\frac{M_{t_0+h}}{M_{t_0}}\right)\right] + \\
&\quad \frac{1}{2}\left(1-\frac{1}{\gamma}\right)\mathrm{Var}_{t_0}\left[\log\left(\frac{M_{t_0+h}}{M_{t_0}}\right)\right] \qquad (3.A.3)
\end{aligned}$$

$$\mathbb{E}_{t_0}\left[\log\left(\frac{M_{t_0+h}}{M_{t_0}}\right)\right] \equiv (\mu_r - r_{t_0})\,B(h) - h\left(\mu_r + \frac{\Gamma}{2}\right) \tag{3.A.4}$$

$$\begin{aligned}
\mathrm{Var}_{t_0}\left[\log\left(\frac{M_{t_0+h}}{M_{t_0}}\right)\right] &\equiv h\left(\Gamma + \frac{\sigma_r}{\kappa}\left(\frac{\sigma_r}{\kappa} - 2\left(\phi_r + \rho_{Sr}\phi_S\right)\right)\right) + \\
&\quad B(h)\,\frac{2\sigma_r}{\kappa}\left(\phi_r + \rho_{Sr}\phi_S - \frac{\sigma_r}{\kappa}\right) + \\
&\quad \frac{\sigma_r^2}{2\kappa^3}\left(1 - \exp(-2\kappa h)\right) \qquad (3.A.5)
\end{aligned}$$

3.B $\mathbb{E}_{t_0}\left[\left(Y_{t_0+h}^{2S}(h)\right)^{1-\gamma}\right]^{\frac{1}{\gamma}}$

By the log-normality of $Y_{t_0+h}^{2S}(h)$,

$$
\begin{aligned}
\mathbb{E}_{t_0}\left[\left(Y^{2S}_{t_0+h}(h)\right)^{1-\gamma}\right]^{\frac{1}{\gamma}} &= \exp\Bigg\{\frac{1}{\gamma}(1-\gamma)\,\mathbb{E}_{t_0}\left[\log\left(Y^{2S}_{t_0+h}(h)\right)\right]+ \\
&\quad \frac{1}{2\gamma}(1-\gamma)^2\,\mathrm{Var}_{t_0}\left[\log\left(Y^{2S}_{t_0+h}(h)\right)\right]\Bigg\} \\
&= \exp\Bigg\{\left(\frac{1}{\gamma}-1\right)\times\Bigg(\mathbb{E}_{t_0}\left[\log\left(Y^{2S}_{t_0+h}(h)\right)\right]+ \\
&\quad \frac{1}{2}(1-\gamma)\,\mathrm{Var}_{t_0}\left[\log\left(Y^{2S}_{t_0+h}(h)\right)\right]\Bigg)\Bigg\} \\
&= \exp\left\{\left(\frac{1}{\gamma}-1\right)g^{2S}(t_0,\,h)\right\} \qquad (3.B.1)
\end{aligned}
$$

$$
\begin{aligned}
g^{2S}(t_0,\,h) &\equiv \mathbb{E}_{t_0}\left[\log\left(Y^{2S}_{t_0+h}(h)\right)\right]+ \\
&\quad \frac{1}{2}(1-\gamma)\,\mathrm{Var}_{t_0}\left[\log\left(Y^{2S}_{t_0+h}(h)\right)\right] \qquad (3.B.2)
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}_{t_0}\left[\log\left(Y^{2S}_{t_0+h}(h)\right)\right] &\equiv \frac{h\Gamma}{\gamma}\left(1-\frac{1}{2\gamma}\right)+B(h)\,r_{t_0}+(h-B(h))\times \\
&\quad \left[\mu_r+\frac{\sigma_r}{\kappa}\left(1-\frac{1}{\gamma}\right)^2\left(\frac{\sigma_r}{\kappa}-\phi_r-\rho_{Sr}\phi_S\right)\right]- \\
&\quad \frac{\sigma_r^2}{4\kappa^3}\left(1-\frac{1}{\gamma}\right)^2(1-\exp(-2\kappa h)) \qquad (3.B.3) \\
\mathrm{Var}_{t_0}\left[\log\left(Y^{2S}_{t_0+h}(h)\right)\right] &\equiv \frac{h}{\gamma^2}\left(\Gamma+\frac{\sigma_r}{\kappa}\left(\frac{\sigma_r}{\kappa}-2(\phi_r+\rho_{Sr}\phi_S)\right)\right)+ \\
&\quad B(h)\,\frac{2\sigma_r}{\kappa\gamma^2}\left(\phi_r+\rho_{Sr}\phi_S-\frac{\sigma_r}{\kappa}\right)+ \\
&\quad \frac{\sigma_r^2}{2\kappa^3\gamma^2}(1-\exp(-2\kappa h)) \qquad (3.B.4)
\end{aligned}
$$

Γ is defined in Equation (3.A.1).

3.C $\mathbb{E}_{t_0}\left[\left(Y^{VA}_{t_0+h}\right)^{1-\gamma}\right]^{\frac{1}{\gamma}}$

$$\mathbb{E}_{t_0}\left[\left(Y^{VA}_{t_0+h}\right)^{1-\gamma}\right]^{\frac{1}{\gamma}} = \exp\left\{\left(\frac{1}{\gamma}-1\right) g^{VA}\left(t_0,\, h\right)\right\} \tag{3.C.1}$$

$$\begin{aligned} g^{VA}\left(t_0,\, h\right) \equiv\; & \mathbb{E}_{t_0}\left[\log\left(Y^{VA}_{t_0+h}\right)\right] + \\ & \frac{1}{2}\left(1-\gamma\right)\mathrm{Var}_{t_0}\left[\log\left(Y^{VA}_{t_0+h}\right)\right] \end{aligned} \tag{3.C.2}$$

Let $\left\{\theta^{VA}_{t_0+u}\right\}^{T-t_0}_{u=0} = \left\{\theta^{VA}_{S,\,t_0+u},\, \theta^{VA}_{B,\,t_0+u}\right\}^{T-t_0}_{u=0}$ denote the VA's reference portfolio investment policy, conditional on time t_0.

$$\begin{aligned} \mathbb{E}_{t_0}\left[\log\left(Y^{VA}_{t_0+h}\right)\right] \equiv\; & h\mu_r + B\left(h\right)\left(r_{t_0}-\mu_r\right) + \\ & \lambda_S\sigma_S\int_0^h \theta^{VA}_{S,\,t_0+l}\,\mathrm{d}l - \lambda_r\sigma_r B\left(\tau\right)\int_0^h \theta^{VA}_{B,\,t_0+l}\,\mathrm{d}l - \\ & \frac{\sigma_S^2}{2}\int_0^h \theta^{VA^2}_{S,\,t_0+l}\,\mathrm{d}l - \frac{\sigma_r^2 B^2\left(\tau\right)}{2}\int_0^h \theta^{VA^2}_{B,\,t_0+l}\,\mathrm{d}l + \\ & \rho_{Sr}\sigma_r\sigma_S B\left(\tau\right)\int_0^h \theta^{VA}_{S,\,t_0+l}\theta^{VA}_{B,\,t_0+l}\,\mathrm{d}l \end{aligned} \tag{3.C.3}$$

$$\begin{aligned} \mathrm{Var}_{t_0}\left[\log\left(Y^{VA}_{t_0+h}\right)\right] \equiv\; & \frac{\sigma_r^2}{\kappa^2}\left[h - 2B\left(h\right) + \frac{1}{2}B(2h)\right] + \\ & \sigma_S^2\int_0^h \theta^{VA^2}_{S,\,t_0+l}\,\mathrm{d}l + \sigma_r^2 B^2\left(\tau\right)\int_u^h \theta^{VA^2}_{B,\,t_0+l}\,\mathrm{d}l - \\ & 2\rho_{Sr}\sigma_r\sigma_S B\left(\tau\right)\int_0^h \theta^{VA}_{S,\,t_0+l}\theta^{VA}_{B,\,t+l}\,\mathrm{d}l + \\ & 2\rho_{Sr}\sigma_r\sigma_S\int_0^h \theta^{VA}_{S,\,t_0+l}\exp\left(-\kappa\left(h-l\right)\right)\,\mathrm{d}l - \\ & 2\sigma_r^2 B\left(\tau\right)\int_0^h \theta^{VA}_{B,\,t_0+l}\exp\left(-\kappa\left(h-l\right)\right)\,\mathrm{d}l \end{aligned} \tag{3.C.4}$$

Γ is defined in Equation (3.A.1).

3.D Proof of Theorem 3

Proof. Append the constraint by the Lagrange multiplier λ.

$$\begin{aligned}
\mathfrak{L}\left(\left\{C^{2S}_{t_0+h}\right\}_{h=0}^{T-t_0},\lambda\right) &= \mathbb{E}_{t_0}\left[U\left(\left\{C^{2S}_{t_0+h}\right\}_{h=0}^{T-t_0}\right)\right]+ \\
&\quad \lambda\left(1-\int_0^{T-t_0} X^{2S}_{t_0}(h,T-t_0)\,\mathrm{d}h\right) \\
&= \int_0^{T-t_0} \mathrm{e}^{-\beta \mathrm{h}}\frac{\mathbb{E}_{t_0}\left[C^{2S^{1-\gamma}}_{t_0+h}\right]}{1-\gamma}+\frac{\lambda}{T-t_0}- \\
&\quad \lambda X^{2S}_{t_0}(h,T-t_0)\,\mathrm{d}h \\
&= \int_0^{T-t_0} \frac{\mathrm{e}^{-\beta \mathrm{h}}}{1-\gamma}\times \\
&\quad \mathbb{E}_{t_0}\left[\left(W_{t_0}X^{2S}_{t_0}(h,T-t_0)\,Y^{2S}_{t_0+h}(h)\right)^{1-\gamma}\right]+ \\
&\quad \frac{\lambda}{T-t_0}-\lambda X^{2S}_{t_0}(h,T-t_0)\,\mathrm{d}h \\
&= \int_0^{T-t_0} F\left(t_0,h,X^{2S}_{t_0}(h,T-t_0),\lambda\right)\mathrm{d}h
\end{aligned}$$

where

$$\begin{aligned}
F\left(t_0,h,X^{2S}_{t_0}(h,T-t_0),\lambda\right) &\equiv \frac{\mathrm{e}^{-\beta \mathrm{h}}}{1-\gamma}\left(W_{t_0}X^{2S}_{t_0}(h,T-t_0)\right)^{1-\gamma}\times \\
&\quad \mathbb{E}_{t_0}\left[\left(Y^{2S}_{t_0+h}(h)\right)^{1-\gamma}\right]+ \\
&\quad \frac{\lambda}{T-t_0}-\lambda X^{2S}_{t_0}(h,T-t_0)
\end{aligned}$$

The Lagrangian attains its extremal(s) when the Euler Equation, (3.D.1), holds (Kamien and Schwartz, 2012).

$$\frac{\partial F}{\partial X^{2S}_{t_0}(h,T-t_0)}-\frac{\partial}{\partial h}\frac{\partial F}{\partial X^{2S'}_{t_0}(h)} = 0 \tag{3.D.1}$$

$X^{2S'}_{t_0}(h)$ denotes the derivative of $X^{2S}_{t_0}(h,T-t_0)$ with respect to the consumption horizon h. As $\partial F/\partial X^{2S'}_{t_0}(h)=0$, Equation (3.D.1) is satisfied when

$$\partial F/\partial X^{2S}_{t_0}(h,T-t_0) = 0$$

where

$$\frac{\partial F}{\partial X^{2S}_{t_0}(h,T-t_0)} = \mathrm{e}^{-\beta \mathrm{h}}W^{1-\gamma}_{t_0}\left(X^{2S}_{t_0}(h,T-t_0)\right)^{-\gamma}\mathbb{E}_{t_0}\left[\left(Y^{2S}_{t_0+h}(h)\right)^{1-\gamma}\right]-\lambda$$

$\partial F / \partial X^{2S}_{t_0}(h,\, T - t_0) = 0$ implies that

$$X^{2S}_{t_0}(h,\, T - t_0) \quad = \quad \lambda^{-\frac{1}{\gamma}} \mathrm{e}^{-\frac{\beta}{\gamma}\mathrm{h}} W_{t_0}^{\frac{1-\gamma}{\gamma}} \mathbb{E}_{t_0}\left[\left(Y^{2S}_{t_0+h}(h)\right)^{1-\gamma}\right]^{\frac{1}{\gamma}} \qquad (3.D.2)$$

We recover the Lagrange multiplier by substituting Equation (3.D.2) into the budget constraint, (3.3.13).

$$\lambda^{-\frac{1}{\gamma}} \quad = \quad \frac{1}{\int_0^{T-t_0} \mathrm{e}^{-\frac{\beta}{\gamma}\mathrm{l}} W_{t_0}^{\frac{1-\gamma}{\gamma}} \mathbb{E}_{t_0}\left[\left(Y^{2S}_{t_0+l}(l)\right)^{1-\gamma}\right]^{\frac{1}{\gamma}} \, \mathrm{d}l} \qquad (3.D.3)$$

The proof is complete by substitution of Equation (3.D.3) into Equation (3.D.2). □

3.E Proof of Lemma 5

Proof. It suffices to show that $\forall h \in [0,\, T - t_0]$,

$$\mathbb{E}_{t_0}\left[\left(\frac{M_{t_0+h}}{M_{t_0}}\right)^{1-\frac{1}{\gamma}}\right] \quad = \quad \mathbb{E}_{t_0}\left[\left(Y^{2S}_{t_0+h}(h)\right)^{1-\gamma}\right]^{\frac{1}{\gamma}} \qquad (3.E.1)$$

By comparison of $\mathbb{E}_{t_0}\left[(M_{t_0+h}/M_{t_0})^{1-\frac{1}{\gamma}}\right]$ as defined in Appendix 3.A, and the expression for $\mathbb{E}_{t_0}\left[\left(Y^{2S}_{t_0+h}(h)\right)^{1-\gamma}\right]^{\frac{1}{\gamma}}$ in Appendix 3.B, we verify that Equation (3.E.1) holds. □

3.F Proof of Lemma 6

Proof. $\left\{Y^{K}_{t_0+u}(h)\right\}_{u=0}^{h}$ is the present value at time $t_0 + u$, of the reference portfolio associated to the portion of wealth reserved for consumption at time $t_0 + h$ for $K \in \{*,\, 2S\}$. When $K = 2S$, by definition of $Y^{2S}_{t_0+u}(h)$ in Equation (3.3.11),

$$Y^{2S}_{t_0+u}(h) \quad = \quad \frac{\tilde{W}^{2S}_{t_0+u}(h)}{\tilde{W}^{2S}_{t_0}(h)}$$

This implies $Y^{2S}_{t_0+u}(h)$ has the same dynamics as $\tilde{W}^{2S}_{t_0+u}(h)$ with respect to u. $\tilde{W}^{2S}_{t_0+u}(h)$ is invested according to $\left\{\theta^{2S}_{t_0+u}(h)\right\}_{u=0}^{h} = \left\{\theta^{2S}_{S,\, t_0+u}(h)\,,\, \theta^{2S}_{B,\, t_0+u}(h)\right\}_{u=0}^{h}$, and its dynamics evolve according Equation (3.3.9). Hence,

$$\begin{aligned} \frac{\mathrm{d}Y^{2S}_{t_0+u}(h)}{Y^{2S}_{t_0+u}(h)} \quad &= \quad \frac{\mathrm{d}\tilde{W}^{2S}_{t_0+u}(h)}{\tilde{W}^{2S}_{t_0+u}(h)} \\ &= \quad \text{Equation (3.3.9)} \end{aligned}$$

Next, for $K = *$, in the complete financial market and in the absence of arbitrage, the unique stochastic discount factor in our setting allows us to deduce $Y^*_{t_0+u}(h)$ as follows

$$\begin{aligned}
Y^*_{t_0+u}(h) &= \mathbb{E}_{t_0+u}\left[Y^*_{t_0+h}(h) \times \frac{M_{t_0+h}}{M_{t_0+u}}\right] \\
&= \mathbb{E}_{t_0+u}\left[\left(\frac{M_{t_0+h}}{M_{t_0}}\right)^{-\frac{1}{\gamma}} \mathbb{E}_{t_0}\left[\left(\frac{M_{t_0+h}}{M_{t_0}}\right)^{1-\frac{1}{\gamma}}\right]^{-1} \frac{M_{t_0+h}}{M_{t_0+u}}\right] \\
&= \mathbb{E}_{t_0+u}\left[M^{1-\frac{1}{\gamma}}_{t_0+h} M^{\frac{1}{\gamma}}_{t_0} M^{-1}_{t_0+u}\right] \mathbb{E}_{t_0}\left[\left(\frac{M_{t_0+h}}{M_{t_0}}\right)^{1-\frac{1}{\gamma}}\right]^{-1} \\
&= M^{\frac{1}{\gamma}}_{t_0} \mathbb{E}_{t_0+u}\left[\left(\frac{M_{t_0+h}}{M_{t_0+u}}\right)^{1-\frac{1}{\gamma}} M^{-1+1-\frac{1}{\gamma}}_{t_0+u}\right] \mathbb{E}_{t_0}\left[\left(\frac{M_{t_0+h}}{M_{t_0}}\right)^{1-\frac{1}{\gamma}}\right]^{-1} \\
&= \left(\frac{M_{t_0+u}}{M_{t_0}}\right)^{-\frac{1}{\gamma}} \mathbb{E}_{t_0+u}\left[\left(\frac{M_{t_0+h}}{M_{t_0+u}}\right)^{1-\frac{1}{\gamma}}\right] \times \\
&\quad \mathbb{E}_{t_0}\left[\left(\frac{M_{t_0+h}}{M_{t_0}}\right)^{1-\frac{1}{\gamma}}\right]^{-1} \qquad (3.F.1)
\end{aligned}$$

The final term of Equation (3.F.1) is a constant with respect to u, and serves to normalize the value of $Y^*_{t_0+0}(h)$. The next part of the proof showing that the random variables have the same starting values, $Y^*_{t_0+0}(h) = Y^{2S}_{t_0+0}(h)$, clarifies this point. As this term does not have a role in the dynamics of $Y^*_{t_0+u}(h)$ with respect to u, we focus on the first two terms, which we denote as $\tilde{Y}^*_{t_0+u}(h)$.

$$\begin{aligned}
\tilde{Y}^*_{t_0+u}(h) &\equiv \left(\frac{M_{t_0+u}}{M_{t_0}}\right)^{-\frac{1}{\gamma}} \mathbb{E}_{t_0+u}\left[\left(\frac{M_{t_0+h}}{M_{t_0+u}}\right)^{1-\frac{1}{\gamma}}\right] \qquad (3.F.2) \\
\frac{\mathrm{d}Y^*_{t_0+u}(h)}{Y^*_{t_0+u}(h)} &= \frac{\mathrm{d}\tilde{Y}^*_{t_0+u}(h)}{\tilde{Y}^*_{t_0+u}(h)}
\end{aligned}$$

We define

$$\begin{aligned}
N_u &\equiv \left(\frac{M_{t_0+u}}{M_{t_0}}\right)^{-\frac{1}{\gamma}} \\
V_u(h) &\equiv \mathbb{E}_{t_0+u}\left[\left(\frac{M_{t_0+h}}{M_{t_0+u}}\right)^{1-\frac{1}{\gamma}}\right]
\end{aligned}$$

and write Equation (3.F.2) as $\tilde{Y}^*_{t_0+u}(h) = N_u V_u(h)$. We apply Itô's Lemma to obtain the dynamics of N_u and $V_u\,(h)$ with respect to u.

$$\begin{aligned}\frac{\mathrm{d}N_u}{N_u} &= \frac{1}{\gamma}r_{t_0+u}+\frac{1}{2\gamma}\left(1+\frac{1}{\gamma}\right)\Gamma\,\mathrm{d}u+\\ &\quad -\frac{\phi_S}{\gamma}\,\mathrm{d}Z_{S,\,t_0+u}-\frac{\phi_r}{\gamma}\,\mathrm{d}Z_{r,\,t_0+u} \qquad (3.F.3)\\ \frac{\mathrm{d}V_u\,(h)}{V_u\,(h)} &= \left(1-\frac{1}{\gamma}\right)r_{t_0+u}+\frac{1}{2\gamma}\left(1-\frac{1}{\gamma}\right)\Gamma\,\mathrm{d}u+\\ &\quad \left(1-\frac{1}{\gamma}\right)^2\sigma_r\rho_{Sr}\phi_r\phi_S B\,(h-u)\,\mathrm{d}Z_{r,\,t_0+u} \qquad (3.F.4)\\ \Gamma &= \text{Equation (3.A.1)}\end{aligned}$$

With Equations (3.F.3) and (3.F.4), along with the Vasicek short rate model $\mathrm{d}r_{t_0+u}=\kappa\,(\mu_r-r_{t_0+u})\,\mathrm{d}u+\sigma_r\,\mathrm{d}Z_{r,\,t_0+u}$, we apply Itô's Lemma once again to $f\,(x,\,y)=xy$, with $x=N_u$ and $y=V_u\,(h)$ to obtain

$$\begin{aligned}\frac{\mathrm{d}\tilde{Y}^*_{t_0+u}\,(h)}{\tilde{Y}^*_{t_0+u}\,(h)} &= r_{t_0+u}+\frac{1}{\gamma}\Gamma+\\ &\quad \left(1-\frac{1}{\gamma}\right)\sigma_r B\,(h-u)\,(\phi_r+\rho_{Sr}\phi_S)\,\mathrm{d}u-\\ &\quad \frac{\phi_S}{\gamma}\,\mathrm{d}Z_{S,\,t_0+u}-\\ &\quad \left[\frac{\phi_r}{\gamma}+\left(1-\frac{1}{\gamma}\right)\sigma_r B\,(h-u)\right]\mathrm{d}Z_{r,\,t_0+u} \qquad (3.F.5)\\ \Gamma &= \text{Equation (3.A.1)}\end{aligned}$$

We have thus far derived $\mathrm{d}Y^*_{t_0+u}\,(h)\,/Y^*_{t_0+u}\,(h)$ and $\mathrm{d}Y^{2S}_{t_0+u}\,(h)\,/Y^{2S}_{t_0+u}\,(h)$. To show their equivalence, we deduce the investment policy implied by Equation (3.F.5) by equating the diffusion terms of Equations (3.3.9) and (3.F.5). A necessary and sufficient condition for $Y^*_{t_0+u}$ to have the same dynamics as $Y^{2S}_{t_0+u}$ is that the implied investment policy is identical to Equations (3.3.17) and (3.3.18), and that the investment policy yields identical drift terms for Equations (3.3.9) and (3.F.5).

Equating the diffusion terms of Equations (3.3.9) and (3.F.5),

$$\begin{aligned}\theta^{2S}_{S,\,t_0+u}\,(h)\,\sigma_S &= -\frac{\phi_S}{\gamma}\\ -\theta^{2S}_{B,\,t_0+u}\,(h)\,\sigma_r B\,(\tau) &= -\left[\frac{\phi_r}{\gamma}+\left(1-\frac{1}{\gamma}\right)\sigma_r B\,(h-u)\right]\end{aligned}$$

we obtain the investment policy as given by Equations (3.3.17) and (3.3.18). Next, by substituting Equation (3.3.17) and (3.3.18) into the drift terms of Equation (3.3.9), we get

$$r_{t_0+u} + \frac{1}{\gamma}\Gamma + \ldots$$
$$\left(1 - \frac{1}{\gamma}\right)\sigma_r B\left(h-u\right)\left(\phi_r + \rho_{Sr}\phi_S\right) \tag{3.F.6}$$

Equation (3.F.6) is identical to the drift term of Equation (3.F.5). Therefore, the dynamics of the processes $Y^{*}_{t_0+u}\left(h\right)$ and $Y^{2S}_{t_0+u}\left(h\right)$ are identical.

It remains to show that $Y^{*}_{t_0+0}\left(h\right) = Y^{2S}_{t_0+0}\left(h\right)$. We demonstrate that $\mathbb{E}_{t_0}\left[\left(M_{t_0+h}/M_{t_0}\right)^{1-\frac{1}{\gamma}}\right]^{-1}$ in Equation (3.F.1) serves to normalize the value of $Y^{*}_{t_0+u}\left(h\right)$ to 1 when $u=0$. Consider a $h \geq u$.

$$\begin{aligned}
\tilde{Y}^{*}_{t_0}\left(h\right) &= \mathbb{E}_{t_0}\left[\tilde{Y}^{*}_{t_0+u}\left(h\right)\frac{M_{t_0+u}}{M_{t_0}}\right] \\
&= \mathbb{E}_{t_0}\left[\left(\frac{M_{t_0+u}}{M_{t_0}}\right)^{-\frac{1}{\gamma}}\mathbb{E}_{t_0+u}\left[\left(\frac{M_{t_0+h}}{M_{t_0+u}}\right)^{1-\frac{1}{\gamma}}\right]\frac{M_{t_0+u}}{M_{t_0}}\right] \\
&= \mathbb{E}_{t_0}\left[\mathbb{E}_{t_0+u}\left[\left(\frac{M_{t_0+h}}{M_{t_0+u}}\right)^{1-\frac{1}{\gamma}}\right]\left(\frac{M_{t_0+u}}{M_{t_0}}\right)^{1-\frac{1}{\gamma}}\right] \\
&= \left(\frac{1}{M_{t_0}}\right)^{1-\frac{1}{\gamma}}\mathbb{E}_{t_0}\left[\left(\frac{1}{M_{t_0+u}}\right)^{1-\frac{1}{\gamma}}\mathbb{E}_{t_0+u}\left[\left(M_{t_0+h}\right)^{1-\frac{1}{\gamma}}\right]\left(M_{t_0+u}\right)^{1-\frac{1}{\gamma}}\right] \\
&= \left(\frac{1}{M_{t_0}}\right)^{1-\frac{1}{\gamma}}\mathbb{E}_{t_0}\left[\mathbb{E}_{t_0+u}\left[\left(M_{t_0+h}\right)^{1-\frac{1}{\gamma}}\right]\right] \\
&= \mathbb{E}_{t_0}\left[\left(\frac{M_{t_0+h}}{M_{t_0}}\right)^{1-\frac{1}{\gamma}}\right]
\end{aligned}$$

Hence, when $u=0$,

$$\begin{aligned}
Y^{*}_{t_0}\left(h\right) &= \tilde{Y}^{*}_{t_0}\left(h\right)\mathbb{E}_{t_0}\left[\left(\frac{M_{t_0+h}}{M_{t_0}}\right)^{1-\frac{1}{\gamma}}\right]^{-1} \\
&= 1
\end{aligned}$$

Furthermore, $Y^{2S}_{t_0}\left(h\right) = \tilde{W}^{2S}_{t_0}\left(h\right)/\tilde{W}^{2S}_{t_0}\left(h\right) = 1$. Therefore, the random variables, $\left\{Y^{K}_{t_0+u}\left(h\right)\right\}^{h}_{u\in 0}$, $K \in \{*, 2S\}$, have the same values at time t_0. □

3.G Proof of Theorem 8

Proof. We apply the Calculus of Variations (Gelfand and Fomin, 1963; Kamien and Schwartz, 2012) to solve (3.4.6). We substitute the definition of the utility function, Equation (3.2.1), and of the VA payments, Equation (3.4.3), into the objective function.

$$
\begin{aligned}
\mathbb{E}_{t_0}\left[U\left(\left\{C^{VA}_{t_0+h}\right\}_{h=0}^{T-t_0}\right)\right] &= \mathbb{E}_{t_0}\left[\int\limits_0^{T-t_0} \mathrm{e}^{-\beta \mathrm{h}} \frac{C^{VA}_{t_0+h}{}^{1-\gamma}}{1-\gamma}\,\mathrm{d}h\right] \\
&= \mathbb{E}_{t_0}\left[\int\limits_0^{T-t_0} \frac{\mathrm{e}^{-\beta h}}{1-\gamma} \times \right. \\
&\left.\left(W_{t_0}\frac{\exp\left(-a_{t_0}(h)\times h\right)}{A_{t_0}\left(\left\{a_{t_0}(h)\right\}_{h=0}^{T-t_0}\right)} Y^{VA}_{t_0+h}\right)^{1-\gamma} \mathrm{d}h\right] \\
&= \frac{1}{1-\gamma}\int\limits_0^{T-t_0} \mathrm{e}^{-\beta \mathrm{h}}\left(W_{t_0}\frac{\exp\left(-a_{t_0}(h)\times h\right)}{A_{t_0}\left(\left\{a_{t_0}(h)\right\}_{h=0}^{T-t_0}\right)}\right)^{1-\gamma} \times \\
&\quad \mathbb{E}_{t_0}\left[\left(\mathrm{Y}^{VA}_{t_0+h}\right)^{1-\gamma}\right]\mathrm{d}h \qquad (3.G.1) \\
&= \frac{W_{t_0}^{1-\gamma}}{1-\gamma}\int\limits_0^{T-t_0}\left(\frac{\exp\left(-a_{t_0}(h)\times h\right)}{A_{t_0}\left(\left\{a_{t_0}(h)\right\}_{h=0}^{T-t_0}\right)}\right)^{1-\gamma} \times \\
&\quad \exp\left\{-\beta h + (1-\gamma)\, g^{VA}(t_0,\, h)\right\}\mathrm{d}h \qquad (3.G.2) \\
&= \frac{W_{t_0}^{1-\gamma}}{1-\gamma}\int\limits_0^{T-t_0}\left(\frac{1}{A_{t_0}\left(\left\{a_{t_0}(h)\right\}_{h=0}^{T-t_0}\right)}\right)^{1-\gamma} \times \\
&\quad \exp\left\{-h\left(\beta + (1-\gamma)\, a_{t_0}(h)\right) + \right. \\
&\quad \left.(1-\gamma)\, g^{VA}(t_0,\, h)\right\}\mathrm{d}h \\
&= \frac{W_{t_0}^{1-\gamma}}{1-\gamma} I_1^{\gamma-1} I_2 \qquad (3.G.3)
\end{aligned}
$$

where we obtain Equation (3.G.2) from Equation (3.G.1) by substituting $\mathbb{E}_{t_0}\left[\left(Y^{VA}_{t_0+h}\right)^{1-\gamma}\right]$, which is derived by a similar manner as we obtained Equation (3.C.1). I_1 and

I_2 are two functionals defined as

$$
\begin{aligned}
I_1 &\equiv A_{t_0}\left(\left\{a_{t_0}(h)\right\}_{h=0}^{T-t_0}\right) && (3.G.4)\\
&= \int_0^{T-t_0} F_1(h)\, \mathrm{d}h \\
F_1(h) &\equiv \exp\left(-a_{t_0}(h)\times h\right)
\end{aligned}
$$

$$
\begin{aligned}
I_2 &\equiv \int_0^{T-t_0} F_2(h)\, \mathrm{d}h && (3.G.5)\\
F_2(h) &\equiv \exp\{-h\left(\beta+(1-\gamma)\, a_{t_0}(h)\right)+ \\
&\quad (1-\gamma)\, g^{VA}(t_0,\, h)\} \\
g^{VA}(t_0,\, h) &= \text{Equations (3.C.2)}
\end{aligned}
$$

We re-write Problem (3.4.6) as the maximization of the functional J over the set of functions $a_{t_0}(h)$, for $h \in [0,\, T-t_0]$.

$$
\begin{aligned}
\sup_{\left\{a_{t_0}(h)\right\}_{h=0}^{T-t_0}} \quad & J\left[\left\{a_{t_0}(h)\right\}_{h=0}^{T-t_0}\right] && (3.G.6)\\
J\left[\left\{a_{t_0}(h)\right\}_{h=0}^{T-t_0}\right] &\equiv \frac{W_{t_0}^{1-\gamma}}{1-\gamma} I_1^{\gamma-1} I_2 \\
I_1 &= \text{Equation (3.G.4)} \\
I_2 &= \text{Equation (3.G.5)}
\end{aligned}
$$

Using Lagrange multipliers λ_1 and λ_2, we formulate the constrained Problem (3.G.6) as an unconstrained problem.

$$
\begin{aligned}
\sup_{\left\{a_{t_0}(h)\right\}_{h=0}^{T-t_0}} \quad & J\left[\left\{a_{t_0}(h)\right\}_{h=0}^{T-t_0}\right] + \lambda_1\left(I_1 - \int_0^{T-t_0} F_1(h)\, \mathrm{d}h\right) + \\
& \lambda_2\left(I_2 - \int_0^{T-t_0} F_2(h)\, \mathrm{d}h\right) && (3.G.7)
\end{aligned}
$$

The Euler equation and the first-order condition with respect to the Lagrange multipliers imply the following:

$$
\begin{aligned}
\frac{W_{t_0}^{1-\gamma}}{1-\gamma}\left[(\gamma-1)\, I_1^{\gamma-2} I_2 S(F_1,\, a) + I_1^{\gamma-1} S(F_2,\, a)\right] &= 0 \\
\frac{W_{t_0}^{1-\gamma} I_1^{\gamma-2}}{1-\gamma}\left[(\gamma-1)\, I_2 S(F_1,\, a) + I_1 S(F_2,\, a)\right] &= 0 && (3.G.8)\\
(\gamma-1)\, I_2 S(F_1,\, a) + I_1 S(F_2,\, a) &= 0 && (3.G.9)
\end{aligned}
$$

$a \equiv \{a_{t_0}(h)\}_{h=0}^{T-t_0}$ and the Euler formula, for $k = 1, 2$, is

$$S(F_k, a) \equiv \frac{\partial}{\partial t}\frac{\partial F_k}{\partial a'} - \frac{\partial F_k}{\partial a}$$

a' is the derivative of $a_{t_0}(h)$ with respect to h. To proceed from (3.G.8) to (3.G.9) , we divide both sides of the equality by $W_{t_0}^{1-\gamma} I_1^{\gamma-2} / (1-\gamma) > 0$. As neither F_1 nor F_2 is a function of a', $S(F_1, a) = hF_1$ and $S(F_2, a) = (1-\gamma) hF_2$.

By substitution of $S(F_k, a)$, for $k = 1, 2$, Equation (3.G.9) becomes

$$\begin{aligned} (\gamma-1) I_2 hF_1 + I_1 (1-\gamma) hF_2 &= 0 \\ hI_1F_2 &= hI_2F_1 \end{aligned} \tag{3.G.10}$$

Equation (3.G.10) is trivially satisfied for $h = 0$. For $h \in (0, T-t_0]$, we seek the function $a = a_{t_0}(h)$ which satisfies $I_1F_2 = I_2F_1$. Notice that if $F_1 = F_2$, then $I_1F_2 = I_2F_1$.

$$\begin{aligned} F_1 &= F_2 \\ \exp(-a_{t_0}(h) \times h) &= \exp\left(-h(\beta + (1-\gamma) a_{t_0}(h)) + (1-\gamma) g^{VA}(t_0, h)\right) \\ -a_{t_0}(h) \times h &= -h(\beta + (1-\gamma) a_{t_0}(h)) + (1-\gamma) g^{VA}(t_0, h) \end{aligned}$$

Rearranging, for $h \in (0, T-t_0]$,

$$a_{t_0}^*(h) = \frac{\beta}{\gamma} - \frac{1-\gamma}{\gamma}\frac{g^{VA}(t_0, h)}{h} \tag{3.G.11}$$

The right limit of Equation (3.G.11), $\lim_{h^+ \to 0} a_{t_0}^*(h)$ does not exist. As Equation (3.G.10) is trivially satisfied at $h = 0$, we impose $a_{t_0}^*(h) \equiv 0$. □

3.H Proof of Theorem 9

Proof. A VA contract with the optimal AIR provides payment according to Equation (3.4.9) at time $t_0 + h$. The expected utility from consuming the full amount of benefit at each date is

$$
\begin{aligned}
\mathbb{E}_{t_0}\left[U\left(C^{VA}\right)\right] &= \mathbb{E}_{t_0}\left[\int\limits_0^{T-t_0} \mathrm{e}^{-\beta h}\,\frac{\left(W_{t_0}X_{t_0}^{VA}\left(h,\,T-t_0\right)Y_{t_0+h}^{VA}\right)^{(1-\gamma)}}{1-\gamma}\,\mathrm{d}h\right] \\
&= \frac{W_{t_0}^{1-\gamma}}{1-\gamma}\times \\
&\quad \int\limits_0^{T-t_0} \mathrm{e}^{-\beta h}\left(X_{t_0}^{VA}\left(h,\,T-t_0\right)\right)^{1-\gamma}\mathbb{E}_{t_0}\left[\left(Y_{t_0+h}^{VA}\right)^{1-\gamma}\right]\,\mathrm{d}h \\
&= \frac{W_{t_0}^{1-\gamma}}{1-\gamma}\left(\int\limits_0^{T-t_0} \mathrm{e}^{-\frac{\beta}{\gamma}h}\,\mathbb{E}_{t_0}\left[\left(Y_{t_0+h}^{VA}\right)^{1-\gamma}\right]^{\frac{1}{\gamma}}\,\mathrm{d}h\right)^{\gamma} \qquad (3.H.1) \\
&= \frac{W_{t_0}^{1-\gamma}}{1-\gamma}\times \\
&\quad \left(\int\limits_0^{T-t_0} \exp\left(-\frac{\beta h}{\gamma}+\frac{1-\gamma}{\gamma}g^{VA}\left(t_0,\,h\right)\right)\,\mathrm{d}h\right)^{\gamma}
\end{aligned}
$$

Equation (3.H.1) is obtained by substitution of the definition of $X_{t_0}^{VA}\left(h,\,T-t_0\right)$ as given by Equation (3.4.10).

Define

$$
F\left(u,\,t_0,\,\left\{y_i\left(u\right)\right\}_{i=1}^{7}\right) \equiv \exp\left(-\frac{\beta u}{\gamma}+\frac{1-\gamma}{\gamma}g^{VA}\left(t_0,\,u\right)\right) \qquad (3.H.2)
$$

$g^{VA}\left(t_0,\,u\right)$ as given in Equation (3.C.2), is composed of Equations (3.C.3) and (3.C.4), and can be expressed as

$$
\begin{aligned}
g^{VA}\left(t_0,\,u\right) \equiv\; & u\left(\mu_r+\frac{(1-\gamma)\,\sigma_r^2}{2\kappa^2}\right)+B\left(u\right)\left(r_{t_0}-\mu_r-\frac{(1-\gamma)\,\sigma_r^2}{\kappa^2}\right)+ \\
& \left(1-\exp\left(-2\kappa u\right)\right)\frac{(1-\gamma)\,\sigma_r^2}{4\kappa^3}+\lambda_S\sigma_S y_1\left(u\right)- \\
& \lambda_r\sigma_r B\left(\tau\right)y_2\left(u\right)-\frac{\gamma\sigma_S^2}{2}y_3\left(u\right)-\frac{\gamma\sigma_r^2B^2\left(\tau\right)}{2}y_4\left(u\right)+ \\
& \gamma\rho_{Sr}\sigma_r\sigma_S B\left(\tau\right)y_5\left(u\right)+(1-\gamma)\,\rho_{Sr}\sigma_r\sigma_S y_6\left(u\right)- \\
& (1-\gamma)\,\sigma_r^2 B\left(\tau\right)y_7\left(u\right) \qquad (3.H.3)
\end{aligned}
$$

where we set

$$y_1(u) = \int_0^u \theta^{VA}_{S,\,t_0+l}\,\mathrm{d}l \tag{3.H.4}$$

$$y_2(u) = \int_0^u \theta^{VA}_{B,\,t_0+l}\,\mathrm{d}l \tag{3.H.5}$$

$$y_3(u) = \int_0^u \theta^{VA^2}_{S,\,t_0+l}\,\mathrm{d}l \tag{3.H.6}$$

$$y_4(u) = \int_0^u \theta^{VA^2}_{B,\,t_0+l}\,\mathrm{d}l \tag{3.H.7}$$

$$y_5(u) = \int_0^u \theta^{VA}_{S,\,t_0+l}\theta^{VA}_{B,\,t_0+l}\,\mathrm{d}l \tag{3.H.8}$$

$$y_6(u) = \int_0^u \theta^{VA}_{S,\,t_0+l}B(u-l)\,\mathrm{d}l \tag{3.H.9}$$

$$y_7(u) = \int_0^u \theta^{VA}_{B,\,t_0+l}B(u-l)\,\mathrm{d}l \tag{3.H.10}$$

Hence, solving Problem (3.4.11) is equivalent to finding an extremum to the functional $\int_0^{T-t_0} F\left(u,\, t_0,\, \{y_i(u)\}_{i=1}^7\right)\,\mathrm{d}u$ over the set of functions that depends on the planning horizon, u, $\left\{\theta^{VA}_{t_0+u}\right\}_{u=0}^{T-t_0} = \left\{\theta^{VA}_{S,\,t_0+u},\, \theta^{VA}_{B,\,t_0+u}\right\}_{u=0}^{T-t_0}$, subject to Equations (3.H.4) to (3.H.10).

$$\inf_{\left\{\theta^{VA}_{t_0+u}\right\}_{u=0}^{T-t_0}} \int_0^{T-t_0} F\left(u,\, t_0,\, \{y_i(u)\}_{i=1}^7\right)\,\mathrm{d}u \tag{3.H.11}$$

$$\text{subject to} \quad \text{Equations (3.H.4) to (3.H.10)} \tag{3.H.12}$$

This is an optimal control problem with integral state equations. We proceed by following Kamien and Muller (1976).

We append Equations (3.H.4) to (3.H.10) to the objective function with

Lagrange multipliers $\{\lambda_i(u)\}_{i=1}^{7}$, and obtain the Lagrangian

$$\begin{aligned}
\mathfrak{L}\left(\{y_i(u)\}_{i=1}^{7}, \{\lambda_i(u)\}_{i=1}^{7}\right) &= \int_0^{T-t_0} F\left(u, t_0, \{y_i(u)\}_{i=1}^{7}\right) \mathrm{d}u + \\
&\int_0^{T-t_0} \lambda_1(u) \left(\int_0^u \theta_{S,\, t_0+l}^{VA} \,\mathrm{d}l - y_1(u)\right) \mathrm{d}u + \\
&\int_0^{T-t_0} \lambda_2(u) \left(\int_0^u \theta_{B,\, t_0+l}^{VA} \,\mathrm{d}l - y_2(u)\right) \mathrm{d}u + \\
&+ \ldots + \\
&\int_0^{T-t_0} \lambda_6(u) \left(\int_0^u \theta_{S,\, t_0+l}^{VA} B(u-l) \,\mathrm{d}l - y_6(u)\right) \mathrm{d}u + \\
&\int_0^{T-t_0} \lambda_7(u) \times \\
&\left(\int_0^u \theta_{B,\, t_0+l}^{VA} B(u-l) \,\mathrm{d}l - y_7(u)\right) \mathrm{d}u \qquad (3.H.13)
\end{aligned}$$

By changing the order of integration of every appended constraint, we express Equation (3.H.13) as

$$\begin{aligned}
\mathfrak{L}\left(\{y_i(u)\}_{i=1}^{7}, \{\lambda_i(u)\}_{i=1}^{7}\right) &= \int_0^{T-t_0} H\left(u, \{y_i(u)\}_{i=1}^{7}, \{\lambda_i(u)\}_{i=1}^{7}\right) \mathrm{d}u - \\
&\int_0^{T-t_0} \sum_{i=1}^{7} \lambda_i(u)\, y_i(u) \,\mathrm{d}u \qquad (3.H.14)
\end{aligned}$$

$H\left(u,\,\{y_i\,(u)\}_{i=1}^{7},\,\{\lambda_i\,(u)\}_{i=1}^{7}\right)$ is the Hamiltonian given by

$$\begin{aligned}
H\left(u,\,\{y_i\,(u)\}_{i=1}^{7},\,\{\lambda_i\,(u)\}_{i=1}^{7}\right) \;\equiv\; & F\left(u,\,t_0,\,\{y_i\,(u)\}_{i=1}^{7}\right) + \\
& \theta_{S,\,t_0+u}^{VA}\int\limits_{l=u}^{T-t_0}\lambda_1\,(l)\;\mathrm{d}l + \theta_{B,\,t_0+u}^{VA}\int\limits_{l=u}^{T-t_0}\lambda_2\,(l)\;\mathrm{d}l + \\
& \theta_{S,\,t_0+u}^{VA^2}\int\limits_{l=u}^{T-t_0}\lambda_3\,(l)\;\mathrm{d}l + \theta_{B,\,t_0+u}^{VA^2}\int\limits_{l=u}^{T-t_0}\lambda_4\,(l)\;\mathrm{d}l + \\
& \theta_{S,\,t_0+u}^{VA}\theta_{B,\,t_0+u}^{VA}\int\limits_{l=u}^{T-t_0}\lambda_5\,(l)\;\mathrm{d}l + \\
& \theta_{S,\,t_0+u}^{VA}\int\limits_{l=u}^{T-t_0}\lambda_6\,(l)\,B\,(l-u)\;\mathrm{d}l + \\
& \theta_{B,\,t_0+u}^{VA}\int\limits_{l=u}^{T-t_0}\lambda_7\,(l)\,B\,(l-u)\;\mathrm{d}l \qquad (3.H.15)
\end{aligned}$$

By Theorem 1 of Kamien and Muller (1976), the condition for $\mathfrak{L}\left(\{y_i\,(u)\}_{i=1}^{7},\,\{\lambda_i\,(u)\}_{i=1}^{7}\right)$ to be stationary is

$$\begin{aligned}
\frac{\partial H\left(u,\,\{y_i\,(u)\}_{i=1}^{7},\,\{\lambda_i\,(u)\}_{i=1}^{7}\right)}{\partial\theta_{L,\,t_0+u}^{VA}} &= 0 \text{ for L} \in \{S,\,B\} \qquad (3.H.16) \\
\frac{\partial H\left(u,\,\{y_i\,(u)\}_{i=1}^{7},\,\{\lambda_i\,(u)\}_{i=1}^{7}\right)}{\partial y_i\,(u)} &= \lambda_i\,(u) \\
\text{for} \quad & \text{i=1, 2,. . ., 7} \qquad (3.H.17)
\end{aligned}$$

for all $u \in [0,\,T-t_0]$. We have

$$\begin{aligned}
\frac{\partial H\left(u,\,\{y_i\,(u)\}_{i=1}^{7},\,\{\lambda_i\,(u)\}_{i=1}^{7}\right)}{\partial\theta_{S,\,t_0+u}^{VA}} = & \int\limits_{l=u}^{T-t_0}\lambda_1\,(l)\;\mathrm{d}l + 2\theta_{S,\,t_0+u}^{VA}\int\limits_{l=u}^{T-t_0}\lambda_3\,(l)\;\mathrm{d}l + \\
& \theta_{B,\,t_0+u}^{VA}\int\limits_{l=u}^{T-t_0}\lambda_5\,(l)\;\mathrm{d}l + \\
& \int\limits_{l=u}^{T-t_0}\lambda_6\,(l)\,B\,(l-u)\;\mathrm{d}l \qquad (3.H.18)
\end{aligned}$$

$$\begin{aligned}\frac{\partial H\left(u,\,\{y_i(u)\}_{i=1}^7,\,\{\lambda_i(u)\}_{i=1}^7\right)}{\partial \theta_{B,\,t_0+u}^{VA}} &= \int\limits_{l=u}^{T-t_0} \lambda_2(l)\,\mathrm{d}l + 2\theta_{B,\,t_0+u}^{VA} \int\limits_{l=u}^{T-t_0} \lambda_4(l)\,\mathrm{d}l + \\ &\quad \theta_{S,\,t_0+u}^{VA} \int\limits_{l=u}^{T-t_0} \lambda_5(l)\,\mathrm{d}l + \\ &\quad \int\limits_{l=u}^{T-t_0} \lambda_7(l)\,B(l-u)\,\mathrm{d}l \qquad (3.H.19)\end{aligned}$$

Moreover,

$$\begin{aligned}\frac{\partial H\left(u,\,\{y_i(u)\}_{i=1}^7,\,\{\lambda_i(u)\}_{i=1}^7\right)}{\partial y_i(u)} &= \frac{\partial F\left(u,\,t_0,\,\{y_i(u)\}_{i=1}^7\right)}{\partial y_i(u)} \\ &= \frac{1-\gamma}{\gamma}\frac{\partial g^{VA}(t_0,\,u)}{\partial y_i(u)} \times \\ &\quad F\left(u,\,t_0,\,\{y_i(u)\}_{i=1}^7\right) \qquad (3.H.20)\end{aligned}$$

By substituting Equation (3.H.20) into Equation (3.H.17), we recover the Lagrange multipliers. For $i = 1,\,2,\,\ldots,\,7$,

$$\lambda_i(u) = \frac{1-\gamma}{\gamma}\frac{\partial g^{VA}(t_0,\,u)}{\partial y_i(u)} F\left(u,\,t_0,\,\{y_i(u)\}_{i=1}^7\right) \qquad (3.H.21)$$

We substitute Equation (3.H.21) for $i = 1,\,3,\,5,\,6$ in Equation (3.H.18), whereas in Equation (3.H.19), we substitute Equation (3.H.21) for $i = 2,\,4,\,5,\,7$. By Equation (3.H.16), for all $u \in [0,\,T-t_0]$, and both $L \in \{S,\,B\}$,

$$\int_u^{T-t_0} F\left(l,\,t_0,\,\{y_i(l)\}_{i=1}^7\right) G_L\left(u,\,t_0,\,T;\,\theta_{t_0+u}^{VA}\right)\mathrm{d}l = 0 \qquad (3.H.22)$$

where

$$\begin{aligned}G_S\left(u,\,t_0,\,T;\,\theta_{t_0+u}^{VA}\right) &= \lambda_S\sigma_S - \theta_{S,\,t_0+u}^{VA}\sigma_S^2\gamma + \theta_{B,\,t_0+u}^{VA}\rho_{Sr}\sigma_r\sigma_S B(\tau)\gamma + \\ &\quad (1-\gamma)\,\rho_{Sr}\sigma_r\sigma_S \int_u^{T-t_0} B(n-u)\,\mathrm{d}n \qquad (3.H.23) \\ G_B\left(u,\,t_0,\,T;\,\theta_{t_0+u}^{VA}\right) &= -\lambda_r\sigma_r B(\tau) - \theta_{B,\,t_0+u}^{VA}\sigma_r^2 B^2(\tau)\gamma + \\ &\quad \theta_{S,\,t_0+u}^{VA}\rho_{Sr}\sigma_r\sigma_S B(\tau)\gamma - \\ &\quad (1-\gamma)\,\sigma_r^2 B(\tau) \int_u^{T-t_0} B(n-u)\,\mathrm{d}n \qquad (3.H.24)\end{aligned}$$

□

3.I Setup of the Kalman Filter

To estimate the financial market parameters, we adopt the Kalman filtering approach. This involves expressing the model in state space form, upon which the Kalman filter is applied to obtain the log-likelihood function to be maximized. The Kalman filter is a widely adopted method for the estimation of financial market parameters (e.g., Campbell and Viceira, 2001; Brennan and Xia, 2002; Munk et al., 2004; Koijen et al., 2010).

There are two unobserved state variables, the short rate and the log returns to the stock index. The yields of the bonds of four different maturities (i.e., 3 months, 1, 5 and 10 years), and the log returns to the stock index constitute the measurement equations.

Let $\mathfrak{F}_t$ be an element of the sequence forming a filtration on the states of the world, Ω, in the probability space $(\Omega, \mathfrak{F}, \mathbb{P})$ on which the financial market is defined. The short rate process, r, is discretized as $r_{t+\Delta t} = \mu_r \left(1 - e^{-\kappa \Delta t}\right) + e^{-\kappa \Delta t} r_t + w_{1,\,t}$, $w_{1,\,t} = \int_t^{t+\Delta t} e^{-\kappa(t-l)} \sigma_r \, dZ_{r,l}$, which implies

$$r_{t+\Delta t}|\mathfrak{F}_t \quad \sim \quad \mathcal{N}\left(\mu_r \left(1 - e^{-\kappa \Delta t}\right) + e^{-\kappa \Delta t} r_t,\, \frac{\sigma_r^2}{2\kappa}\left(1 - e^{-2\kappa \Delta t}\right)\right)$$

The stock index diffusion process is discretized as

$$\log\left(\frac{S_{t+\Delta t}}{S_t}\right) = \left(r_t + \lambda_S \sigma_S - \frac{\sigma_S^2}{2}\right)\Delta t + w_{2,\,t}$$

$w_{2,\,t} \sim \mathcal{N}\left(0,\, \sigma_S^2 \Delta t\right)$.[16]

The discretized short rate and log of the stock index diffusion processes define the state equation:

$$\begin{aligned}\begin{bmatrix} r_{t+\Delta t} \\ \log\left(\frac{S_{t+\Delta t}}{S_t}\right) \end{bmatrix} \quad &= \quad \begin{bmatrix} e^{-\kappa \Delta t} & 0 \\ \Delta t & 0 \end{bmatrix} \begin{bmatrix} r_t \\ \log\left(\frac{S_{t+\Delta t}}{S_t}\right) \end{bmatrix} + \\ &\quad \begin{bmatrix} \mu_r \left(1 - e^{-\kappa \Delta t}\right) \\ \Delta t \left(\lambda_S \sigma_S - \frac{\sigma_S^2}{2}\right) \end{bmatrix} + \begin{bmatrix} w_{1,\,t} \\ w_{2,\,t} \end{bmatrix} \end{aligned} \tag{3.I.1}$$

$w_t = (w_{1,\,t}, w_{2,\,t}) \sim N\left(0, Q\right)$.

$$Q \quad = \quad \begin{bmatrix} \frac{\sigma_r^2}{2\kappa}\left(1 - e^{-2\kappa \Delta t}\right) & \bar{\rho} \\ \bar{\rho} & \sigma_S^2 \Delta t \end{bmatrix}$$

[16] Confer Appendix B of Bolder (2001) for details on the derivation.

, $\bar{\rho} = \text{cov}\left(w_{1,\,t},\, w_{2,\,t}\right) = \sigma_r \sigma_S \rho_{Sr}/\kappa\left(1 - e^{\kappa \Delta t}\right)$.[17]

The observation equation is composed of bond yields of maturities 3 months, 1, 5 and 10 years (i.e., $n = 4$ in (3.I.2)) and the return on the stock index.

$$\begin{bmatrix} z(t,\,T_1) \\ z(t,\,T_2) \\ \vdots \\ z(t,\,T_n) \\ log\left(\frac{S_t}{S_{t-\Delta t}}\right) \end{bmatrix} = \begin{bmatrix} \frac{B(t,\,T_1)}{T_1 - t} & 0 \\ \frac{B(t,\,T_2)}{T_2 - t} & 0 \\ \vdots & \vdots \\ \frac{B(t,\,T_n)}{T_n - t} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} r_t \\ \log\left(\frac{S_t}{S_{t-\Delta t}}\right) \end{bmatrix} + \begin{bmatrix} -\frac{A(t,\,T_1)}{T_1 - t} \\ -\frac{A(t,\,T_2)}{T_2 - t} \\ \vdots \\ -\frac{A(t,\,T_n)}{T_n - t} \\ 0 \end{bmatrix} + \begin{bmatrix} v_{1,\,t} \\ v_{2,\,t} \\ \vdots \\ v_{n,\,t} \\ v_{n+1,\,t} \end{bmatrix} \tag{3.I.2}$$

$z(t,T) = -\log\left[P\left(t,T\right)\right]/\left(T-t\right)$ is the yield on a bond of maturity $T-t$; $v = (v_1, v_2, \ldots, v_n, v_{n+1})$, $v \sim N\left(0,\, R\right)$, R is an $(n+1) \times (n+1)$ positive semi-definite matrix with zeros on the last row and last column because we assume that stock return is observed without measurement error.[18]

---

[17] The derivation of $\bar{\rho}$ is provided in Section 3.I.1.

[18] If $v_{n+1,t} \neq 0$ and $w_{2,t} \neq 0$, then they are unidentified; we are unable to determine the measurement error from the innovation of the latent state.

3.I.1 Derivation of $\bar{\rho}$

We have $w_{1,\,t} = \int_t^{t+\Delta t} \mathrm{e}^{-\kappa(t-l)}\,\sigma_r\,\mathrm{d}Z_{r,\,l}$, $w_{2,\,t} = \int_t^{t+\Delta t} \sigma_S\,\mathrm{d}Z_{S,\,l}$, and the correlation between $Z_{r,\,t}$ and $Z_{S,\,t}$ is ρ_{Sr}.

$$\begin{aligned}
\bar{\rho} &= \mathrm{cov}\left(w_{1,\,t},\, w_{2,\,t}\right) \\
&= \mathbb{E}\left[w_{1,\,t} w_{2,\,t}\right] - \mathbb{E}\left[w_{1,\,t}\right]\mathbb{E}\left[w_{2,\,t}\right] \\
&= \mathbb{E}\left[w_{1,\,t} w_{2,\,t}\right] \\
&= \mathbb{E}\left[\int_t^{t+\Delta t} \mathrm{e}^{-\kappa(t-l)}\,\sigma_r\,\mathrm{d}Z_{r,\,l} \times \int_t^{t+\Delta t} \sigma_S\,\mathrm{d}Z_{S,\,l}\right] \\
&= \mathbb{E}\left[\int_t^{t+\Delta t} \mathrm{e}^{-\kappa(t-l)}\,\sigma_r\sigma_S\,\mathrm{d}\left\langle Z_r,\, Z_S\right\rangle_l\right] \\
&= \mathbb{E}\left[\int_t^{t+\Delta t} \mathrm{e}^{-\kappa(t-l)}\,\sigma_r\sigma_S\rho_{Sr}\,\mathrm{d}l\right] \\
&= \sigma_r\sigma_S\rho_{Sr}\mathbb{E}\left[\int_t^{t+\Delta t} \mathrm{e}^{-\kappa(t-l)}\,\mathrm{d}l\right] \\
&= \frac{\sigma_r\sigma_S\rho_{Sr}}{\kappa}\left(1 - \mathrm{e}^{\kappa\Delta t}\right)
\end{aligned}$$

3.J Sign of the derivative of $\mathrm{Var}_t\left[r_{t+u}\right]$ with respect to κ

Lemma 10. *For any $t \in [0, T]$, when $\{r_{t+u}\}_{u=0}^{T-t}$ follows the Vasicek (1977) model and u, $\kappa > 0$, $\partial \mathrm{Var}_t\left[r_{t+u}\right]/\partial\kappa < 0$.*

Proof. By Equation (3.6.1),

$$\frac{\partial \mathrm{Var}_t\left[r_{t+u}\right]}{\partial\kappa} = \frac{\sigma_r^2 u}{\kappa}\exp\left(-2\kappa u\right) - \frac{\sigma_r^2}{2\kappa^2}\left(1 - \exp\left(-2\kappa u\right)\right) \qquad (3.J.1)$$

As $\sigma_r > 0$ and $\kappa > 0$, we divide Equation (3.J.1) by $\sigma_r^2/\kappa > 0$. $\partial \mathrm{Var}_t\left[r_{t+u}\right]/\partial\kappa < 0$ if and only if

$$\begin{aligned}
u\exp\left(-2\kappa u\right) - \frac{1}{2\kappa}\left(1 - \exp\left(-2\kappa u\right)\right) &< 0 \\
\exp\left(-2\kappa u\right)\left(2\kappa u + 1\right) &< 1 \\
\frac{1}{\mathrm{e}^{2\kappa u}} &< \frac{1}{1 + 2\kappa u} \qquad (3.J.2)
\end{aligned}$$

The inequality Equation (3.J.2) is valid if and only if $e^{2\kappa u} > 1 + 2\kappa u$, $\forall u,\ \kappa > 0$. This holds by a known result of the exponential function. Therefore, $\partial \mathrm{Var}_t \left[r_{t+u} \right] / \partial \kappa < 0$, $\forall u,\ \kappa > 0$. □

3.K Mortality Credit

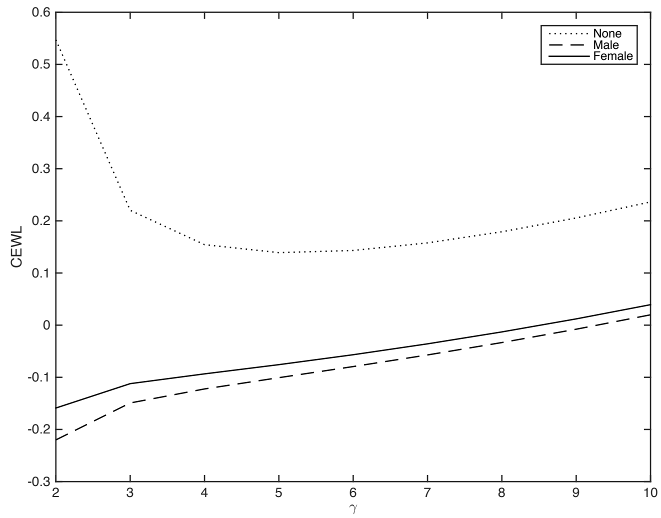
We assess the magnitudes of the utility loss due to a VA*'s impaired interest rate risk hedge and the utility gain from mortality credit, assuming that only the VA* offers mortality credit while the 2S does not.

We assume that the VA* underwriter has no default risk and determines the mortality credit based on the Society of Actuaries (SoA) 2014 mortality tables for male and female healthy annuitants. The SoA construct these tables using the actual mortality experience of uninsured private retirement plans. As the SoA mortality rates are defined up till age 120, we extend the horizon of the base case by 10 years as well. We estimate the CEWL (Equation (3.5.2)) for the base case parameters.

Figure 3.K.1 shows the CEWL by the individual's risk aversion level. When there is no mortality credit, all individuals prefer the 2S. Our reasoning on the positive relation between the consumption horizon and the CEL in Section 3.6 applies to the observation that the estimated CEWLs in Figure 3.K.1 are higher than those in the base case. When the VA offers mortality credit, only highly risk-averse individuals prefer the 2S over VA*. Furthermore, due to the higher mortality rates for males relative to those for females, which yields higher mortality credit for males than females, preference for the VA* is greater for males than it is for females. The utility gain from the default-free VA*'s mortality credit compensates for the utility loss from the contract's impaired interest rate risk hedge when individuals are not overly risk-averse.

Figure 3.K.1: CEWL with Mortality Credit

This figure presents the CEWL by the individual's risk aversion level, γ that ranges from 2 to 10, when a default-risk-free VA* provides mortality credit that is determined according to the Society of Actuaries 2014 Mortality Tables for male and female healthy annuitants. All other parameters are identical to those in the base case.



Conclusion

The rising cost of retirement financing that is attributable to lengthier life expectancies and the financial market situation have prompted the stakeholders to re-evaluate their roles. In response, employers transfer the management of financial and biometric risks onto individuals, who are compelled to shoulder greater responsibility for their financial security in retirement. Insurers are coping with the challenge of offering contracts with financial return guarantees by turning their attention to unit-linked contracts, while the policymaker is tasked with revising the regulation that would shape the future retirement environment.

This dissertation explicates the roles of individuals, the policymaker and the insurer in the creation of sustainable retirement solutions. We demonstrate the dominance of regulatory requirements in determining DB pension plans' investment risk-taking in Chapter 1. Our estimate of the extent that rules and plan characteristics account for the variation of investment risk among DB plans can inform the policymaker. In Chapter 2, we contrast two ways for individuals to manage longevity risk, and evaluate the longevity risk premium. We then turn to focus on financial market risk in Chapter 3, in which we illustrate that the variable annuity is unable to optimally hedge interest rate risk. Numerous points for the policymaker's consideration arise from our investigation.

A retirement system is sustainable only when the stakeholders' incentives are compatible. This dissertation attempts to better comprehend of the roles of each stakeholder in the changing landscape. The insights garnered are hoped to contribute toward nurturing a resilient retirement system.

Résumé en Français

Le financement des régimes de retraite est soumis aux fluctuations des marchés financiers et aux évolutions démographiques. La situation économique après la crise financière de 2007-2009, couplée à un allongement de l'espérance de vie, provoque une envolée du coût du financement des régimes de retraite incitant les acteurs à réévaluer leurs rôles.

L'employeur, par exemple, remanie son régime de retraite professionnel pour gérer ses coûts. Ces changements transfèrent à l'individu la majeure partie des risques financiers et biométriques. Chargé de la refonte du système, le législateur doit réorganiser le cadre règlementaire pour une meilleure harmonisation des intérêts de toutes les parties prenantes. La présente thèse examine le rôle des acteurs du financement des régimes de retraite dans un monde en évolution : le législateur, l'individu et l'assureur. Elle clarifie la réponse apportée par chaque partie prenante pour relever les défis du financement des régimes de retraite.

Le Chapitre 1 concerne le législateur et l'employeur. Il étudie les investissements des régimes de retraite à prestations définies (PD) par rapport à la règlementation en matière d'investissement, de valorisation et de comptabilité. Dans le cadre des régimes PD, les employeurs assurent une rente à vie à leurs salariés et assument la plupart des risques. Les investissements des régimes PD font l'objet d'une grande attention, non seulement car les gains financiers stimulent le revenu de retraite, mais aussi car ils constituent une source de capital pour le financement à long terme de l'économie, pour la construction d'infrastructures par exemple.

La compensation du coût des prestations versées par des rendements financiers élevés pourrait, en théorie, induire un comportement plus risqué des promoteurs de régimes. Une intervention règlementaire est donc justifiée. Il convient alors d'étudier l'impact des exigences règlementaires sur les investissements des régimes PD. À ce jour, la littérature traite strictement d'une poignée de contraintes règlementaires et se focalise sur les états-Unis. L'exploitation de l'hétérogénéité du cadre règlementaire relatif aux régimes PD aux états-Unis, au Canada et aux Pays-Bas révèle que les exigences règlementaires influencent de façon statistiquement significative la prise de risque des fonds PD. Cela étant, la portée économique de ces contraintes varie. Les exigences de fonds propres liées aux risques et l'évaluation à la valeur de marché sont associées à une baisse de 7 % des investissements risqués. Par contraste, le taux d'actualisation du passif a une influence moindre ; par exemple, un taux d'actualisation du passif plus élevé de 1 % est associé à une hausse de seulement 0,8 % en moyenne des actifs à risque. Comme pour le taux d'actualisation du passif, les caractéristiques des fonds (par ex., part des retraités, valeur des actifs) ont également un faible impact économique sur la variation de la prise de risque des régimes PD.

Par ailleurs, l'étude révèle que pour les contraintes de solvabilité – l'un des principaux mécanismes employés, c'est-à-dire soit une exigence minimale de financement soit des exigences de fonds propres liées aux risques –, le comportement du fonds dépend de la situation des marchés financiers. Une exigence de solvabilité fondée sur une évaluation du risque est associée à une moindre prise de risque par un fonds règlementé, quelles que soient les conditions du marché. Un niveau de financement minimal défini, en revanche, est associé à des investissements moins risqués pendant la crise financière uniquement. Ces informations sont utiles pour un législateur, qui compense le besoin de protéger les individus par une modulation de la prise de risque des investissements, sans pour autant décourager les investissements de régimes PD dans des projets à long terme essentiels à la croissance économique.

Si les régimes PD représentent, en valeur d'actifs, la majorité des comptes de retraite professionnels, les régimes de retraite à cotisations définies (CD) enregistrent une croissance progressive mais régulière. Dans le cadre des régimes CD, le rôle de l'employeur est réduit au versement du taux de base de cotisations sur le compte de retraite du salarié. La gestion des investissements et le versement des prestations de retraite sont réalisés par un prestataire de services, gestionnaire de fonds ou assureur par exemple. C'est sur ce système que reposent le plan d'épargne-retraite 401(k) aux états-Unis et les régimes de retraite complémentaire individuels (Personal Pension Schemes, PPS) au Royaume-Uni. Conséquence pour les individus du passage d'un régime PD à un régime CD, ils ne sont plus protégés contre le risque de longévité, auparavant classiquement assumé par le promoteur du régime PD. Le risque de longévité est lié à une mauvaise estimation des probabilités de survie. Ce risque systématique ne peut pas être diversifié par une mutualisation. Les individus assument le risque ou le transfèrent, en en supportant le coût, à un assureur.

Le Chapitre 2 étudie le risque de longévité dans le cadre du régime CD. Il présente les différences entre deux options de gestion du risque à la disposition des individus : supporter le risque collectivement ou s'en décharger auprès d'un assureur par l'achat d'une annuité. Le capital de l'assureur peut être composé de fonds propres recueillis auprès d'investisseurs et/ou du produit de la vente de l'annuité à un prix supérieur à son meilleur prix de valorisation. Chaque source de capital est assortie de conséquences spécifiques. Les actionnaires souhaitent être rémunérés par une prime de risque de longévité. À l'opposé, la demande des individus pour ce contrat est déterminée par la comparaison de leur niveau de bien-être avec un régime collectif et avec un contrat d'annuités. L'assureur doit donc garantir sa capacité à rémunérer ses actionnaires, sans pour autant trop augmenter le chargement pour les individus, afin d'éviter qu'ils ne forment un régime collectif. La littérature existante omet l'actionnaire, considérant que la totalité du capital « tampon » est constituée par le chargement payé par les individus. Cette hypothèse est incohérente avec les estimations indiquant que la volonté des individus à payer pour s'assurer contre le risque de longévité est nettement plus faible que le niveau de fonds propres nécessaires pour fournir le contrat (par ex., un manque à gagner de 17 % de la valeur du contrat). Les actionnaires doivent donc être pris en compte pour faire concorder ces estimations.

L'étude révèle que les individus ont une légère préférence pour l'organisation collective. En conséquence, le prestataire de contrat ne peut pas proposer une prime de risque de longévité à ses actionnaires. Dans ce cadre, pour couvrir le risque de longévité, les régimes collectifs et un marché des annuités ont peu de chances de coexister si l'assureur ne tire pas d'avantage à traiter ce risque, au moyen de la réassurance ou de la synergie entre les produits par exemple (par ex., la vente d'assurances vie, qui présentent une exposition inverse au risque de longévité). Les résultats de l'étude prônent la mise en place de régimes collectifs. Même si, vis-à-vis du risque de longévité, une solution collective est privilégiée, les annuités pourraient néanmoins s'avérer utiles si elles permettent d'atténuer les risques des marchés financiers.

Le Chapitre 3 porte sur le problème du niveau optimal de consommation et d'investissement face aux risques liés aux fluctuations des taux d'intérêt et au marché des actions. Une modélisation des taux d'intérêt à l'aide d'un processus gaussien de retour à la moyenne offre la solution à ce problème. Une formulation équivalente de ce problème est présentée et appliquée pour démontrer qu'une annuité variable ne parvient pas à couvrir de façon optimale le risque de taux d'intérêt. Avec un portefeuille indiciel unique, l'annuité variable ne peut pas simultanément couvrir le risque de taux d'intérêt à chaque période de consommation.

La perte de bien-être en résultant est importante sur le plan économique et s'avère d'autant plus grave quand le niveau de retour à la moyenne du taux d'intérêt est bas ou quand les taux d'intérêt sont plus fluctuants. La formulation alternative propose une amélioration éliminant cet inconvénient de l'annuité variable. Cette présentation peut orienter les assureurs dans l'élaboration de contrats plus appréciés des individus, pour lesquels ces derniers seront alors plus enclins à payer un prix plus important. De plus, elle apporte des informations au législateur qui supervise les types de contrats autorisés. Les régimes CD étant de plus en plus courants, le choix des individus en matière d'investissements et de produits d'assurance est d'une importance capitale pour leur bien-être à la retraite. Favoriser l'offre de contrats augmentant les possibilités pour les individus d'améliorer le financement de leur retraite devrait faire partie des nombreuses intentions du législateur.

Une vision holistique est indispensable pour relever le défi des retraites. La présente thèse fait la lumière sur les rôles des législateurs, assureurs et individus, et fournit des informations sur les pressions exercées sur le financement des retraites. S'appuyant sur les connaissances passées, les analyses intègrent les tendances et innovations actuelles pour caractériser l'environnement futur de la retraite. Ces découvertes devraient favoriser la compréhension de cette situation délicate et, grâce à leurs implications, l'élaboration de politiques permettant de l'affronter.

Bibliography

Aase, K. K. (2015). Life insurance and pension contracts I: The time additive life cycle model. *ASTIN Bulletin*, 45(1):1–47.

Addoum, J. M., Van Binsbergen, J. H., and Brandt, M. W. (2010). Asset allocation and managerial assumptions in corporate pension plans. Working paper.

Aglietta, M., Brière, M., Rigot, S., and Signori, O. (2012). Rehabilitating the role of active management for pension funds. *Journal of Banking and Finance*, 36(9):2565–2574.

Ai, J., Brockett, P. L., Golden, L. L., and Zhu, W. (2016). Health state transitions and longevity effects on retirees' optimal annuitization. *Journal of Risk and Insurance*.

Allen, F. and Carletti, E. (2008). Mark-to-market accounting and liquidity pricing. *Journal of Accounting and Economics*, 45(2):358–378.

Amir, E., Guan, Y., and Oswald, D. (2010). The effect of pension accounting on corporate pension asset allocation. *Review of Accounting Studies*, 15(2):345–366.

An, H., Huang, Z., and Zhang, T. (2013). What determines corporate pension fund risk-taking strategy? *Journal of Banking and Finance*, 37(2):597–613.

Anantharaman, D. and Lee, Y. G. (2014). Managerial risk taking incentives and corporate pension policy. *Journal of Financial Economics*, 111(2):328–351.

Andonov, A., Bauer, R. M., and Cremers, M. (2017). Pension fund asset allocation and liability discount rates. *Review of Financial Studies*. forthcoming.

Andonov, A., Hochberg, Y. V., and Rauh, J. D. (2016). Political representation and governance: Evidence from the investment decisions of public pension funds. Working paper.

Ang, A. and Bekaert, G. (2002). International asset allocation with regime shifts. *Review of Financial Studies*, 15(4):1137–1187.

Ang, A. and Bekaert, G. (2007). Stock return predictability: Is it there? *Review of Financial Studies*, 20(3):651–707.

Ang, A., Brière, M., and Signori, O. (2012). Inflation and individual equities. *Financial Analysts Journal*, 68(4):36–55.

Ang, A. and Maddaloni, A. (2003). Do demographic changes affect risk premiums? Evidence from international data. Working paper, National Bureau of Economic Research.

Antolín, P., Schich, S., and Yermo, J. (2011). The economic impact of protracted low interest rates on pension funds and insurance companies. OECD Journal: Financial Market Trends, Organisation for Economic Co-operation and Development.

Arnott, R. D. and Chaves, D. B. (2012). Demographic changes, financial markets, and the economy. *Financial Analysts Journal*, 68(1):23–46.

Asvanunt, A. and Richardson, S. (2016). The credit risk premium. *Journal of Fixed Income*, 26(3):6–24.

Bacinello, A. R. and Persson, S.-A. (2002). Design and pricing of equity-linked life insurance under stochastic interest rates. *Journal of Risk Finance*, 3(2):6–21.

Baione, F., De Angelis, P., Menzietti, M., and Tripodi, A. (2016). A comparison of risk transfer strategies for a portfolio of life annuities based on RORAC. *Journal of Applied Statistics*, pages 1–18.

Bajeux-Besnainou, I., Jordan, J. V., and Portait, R. (2003). Dynamic asset allocation for stocks, bonds, and cash. *Journal of Business*, 76(2):263–287.

Bams, D., Schotman, P. C., and Tyagi, M. (2016). Asset allocation dynamics of pension funds. Netspar discussion paper No. 03/2016-016.

Barberis, N. (2000). Investing for the long run when returns are predictable. *Journal of Finance*, 55(1):225–264.

Basak, S. and Shapiro, A. (2001). Value-at-risk-based risk management: Optimal policies and asset prices. *Review of Financial Studies*, 14(2):371–405.

Becker, B. and Ivashina, V. (2015). Reaching for yield in the bond market. *Journal of Finance*, 70(5):1863–1902.

Becker, B. and Opp, M. M. (2014). Regulatory reform and risk-taking: Replacing ratings. Research paper No. 13-03, Swedish House of Finance.

Beetsma, R. M., Romp, W. E., and Vos, S. J. (2012). Voluntary participation and intergenerational risk sharing in a funded pension system. *European Economic Review*, 56(6):1310–1324.

Bernard, C., Maj, M., and Vanduffel, S. (2011). Improving the design of financial products in a multidimensional Black-Scholes market. *North American Actuarial Journal*, 15(1):77–96.

Bikker, J. A., Broeders, D., and De Dreu, J. (2010). Stock market performance and pension fund investment policy: rebalancing, free float, or market timing. *International Journal of Central Banking*, 6(2):53–79.

Bikker, J. A., Steenbeek, O. W., and Torracchi, F. (2012). The impact of scale, complexity, and service quality on the administrative costs of pension funds: A cross-country comparison. *Journal of Risk and Insurance*, 79(2):477–514.

Blackburn, C., Hanewald, K., Olivieri, A., Sherris, M., et al. (2017). Longevity risk management and shareholder value for a life annuity business. *ASTIN Bulletin*, 47(01):43–77.

Blake, D., Cairns, A. J., and Dowd, K. (2003). Pensionmetrics 2: Stochastic pension plan design during the distribution phase. *Insurance: Mathematics and Economics*, 33(1):29–47.

Board of Governors of the Federal Reserve System Data Download Program (2016). Private defined benefit pension funds: Total funded assets. (Z1/Z1/FL572000075.A) and Private defined benefit pension funds: Pension entitlements (total liabilities) (Z1/Z1/FL574190043.A). Last Accessed: 2016-12-12.

Bodie, Z. (1990). The ABO, the PBO and pension investment policy. *Financial Analysts Journal*, 46(5):27–34.

Bodie, Z., Detemple, J. B., Otruba, S., and Walter, S. (2004). Optimal consumption–portfolio choices and retirement planning. *Journal of Economic Dynamics and Control*, 28(6):1115–1148.

Bodie, Z., Light, J. O., and Morck, R. (1987). Funding and asset allocation in corporate pension plans: An empirical investigation. In Shoven, J., Bodie, Z., and Wise, D. A., editors, *Issues in Pension Economics*, pages 15–48. University of Chicago Press.

Boeri, T., Bovenberg, A. L., Coeure, B., and Roberts, A. (2006). Dealing with the new giants: Rethinking the role of pension funds. *Geneva Reports on the World Economy*, (8).

Bolder, D. J. (2001). Affine term-structure models: Theory and implementation. Working paper 2001-15, Bank of Canada.

Boon, L.-N., Brière, M., and Rigot, S. (2017a). Regulation and pension fund risk-taking. Working paper.

Boon, L.-N., Brière, M., and Werker, B. J. (2017b). Longevity risk: To bear or to insure? Working paper.

Boon, L.-N. and Werker, B. J. (2017). Variable annuity and interest rate risk. Working paper.

Booth, H., Maindonald, J., and Smith, L. (2002). Applying Lee-Carter under conditions of variable mortality decline. *Population Studies*, 56(3):325–336.

Bossaerts, P. and Hillion, P. (1999). Implementing statistical criteria to select return forecasting models: What do we learn? *Review of Financial Studies*, 12(2):405–428.

Boudoukh, J. and Richardson, M. (1993). Stock returns and inflation: A long-horizon perspective. *American Economic Review*, 83(5):1346–1355.

Boulier, J.-F., Huang, S., and Taillard, G. (2001). Optimal management under stochastic interest rates: The case of a protected defined contribution pension fund. *Insurance: Mathematics and Economics*, 28(2):173–189.

Bovenberg, L., Koijen, R. S., Nijman, T., and Teulings, C. (2007). Saving and investing over the life cycle and the role of collective pension funds. *De Economist*, 155(4):347–415.

Bovenberg, L. and Nijman, T. (2009). Developments in pension reform: The case of dutch stand-alone collective pension schemes. *International Tax and Public Finance*, 16(4):443–467.

Boyle, P., Hardy, M., Mackay, A., and Saunders, D. (2015). Variable payout annuities. Working paper, Pension Section Research Committee.

Boyle, P. and Tian, W. (2008). The design of equity-indexed annuities. *Insurance: Mathematics and Economics*, 43(3):303–315.

Boyle, P. and Tian, W. (2009). Optimal design of equity-linked products with a probabilistic constraint. *Scandinavian Actuarial Journal*, 2009(4):253–280.

Bradley, D., Pantzalis, C., and Yuan, X. (2016). The influence of political bias in state pension funds. *Journal of Financial Economics*, 119(1):69–91.

Brandt, M. W. (1999). Estimating portfolio and consumption choice: A conditional euler equations approach. *Journal of Finance*, 54(5):1609–1645.

Branger, N., Mahayni, A., and Schneider, J. C. (2010). On the optimal design of insurance contracts with guarantees. *Insurance: Mathematics and Economics*, 46(3):485–492.

Brennan, M. J., Schwartz, E. S., and Lagnado, R. (1997). Strategic asset allocation. *Journal of Economic Dynamics and Control*, 21(8):1377–1403.

Brennan, M. J. and Xia, Y. (2002). Dynamic asset allocation under inflation. *Journal of Finance*, 57(3):1201–1238.

Broeders, D., Chen, D. H., Minderhoud, P., and Schudel, W. (2016). Pension funds' herding. De Nederlandsche Bank working paper No. 503.

Brown, J. R., Mitchell, O. S., and Poterba, J. M. (2001). The role of real annuities and indexed bonds in an individual accounts retirement program. In *Risk Aspects of Investment-based Social Security Reform*, pages 321–370. University of Chicago Press.

Brown, J. R., Mitchell, O. S., Poterba, J. M., and Warshawsky, M. J. (1999). Taxing retirement income: Nonqualified annuities and distributions from qualified accounts. *National Tax Journal*, 52(3):563–592.

Brown, J. R. and Pennacchi, G. G. (2016). Discounting pension liabilities: Funding versus value. *Journal of Pension Economics and Finance*, 15(3):254–284.

Brown, J. R. and Wilcox, D. W. (2009). Discounting state and local pension liabilities. *American Economic Review*, 99(2):538–542.

Cairns, A. J., Blake, D., and Dowd, K. (2006a). Stochastic lifestyling: Optimal dynamic asset allocation for defined contribution pension plans. *Journal of Economic Dynamics and Control*, 30(5):843–877.

Cairns, A. J., Blake, D., and Dowd, K. (2006b). A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration. *Journal of Risk and Insurance*, 73(4):687–718.

Cairns, A. J., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., and Khalaf-Allah, M. (2011). Mortality density forecasts: An analysis of six stochastic mortality models. *Insurance: Mathematics and Economics*, 48(3):355–367.

Campbell, J. Y. (1987). Stock returns and the term structure. *Journal of Financial Economics*, 18(2):373–399.

Campbell, J. Y. and Viceira, L. M. (1999). Consumption and portfolio decisions when expected returns are time varying. *The Quarterly Journal of Economics*, 114(2):433–495.

Campbell, J. Y. and Viceira, L. M. (2001). Who should buy long-term bonds? *American Economic Review*, pages 99–127.

Carone, G., Eckefeldt, P., Giamboni, L., Laine, V., and Pamies, S. (2016). Pension reforms in the EU since the early 2000's: Achievements and challenges ahead. European Economy Discussion Papers 042, Directorate-General for Economic and Financial Affairs.

Casella, G. and Berger, R. L. (2002). *Statistical Inference*. Duxbury, second edition.

Chacko, G. and Viceira, L. M. (2005). Dynamic consumption and portfolio choice with stochastic volatility in incomplete markets. *Review of Financial Studies*, 18(4):1369–1402.

Chambers, D., Dimson, E., and Ilmanen, A. (2012). The Norway model. *The Journal of Portfolio Management*, 38(2):67–81.

Chang, H. and Chang, K. (2017). Optimal consumption–investment strategy under the vasicek model: HARA utility and legendre transform. *Insurance: Mathematics and Economics*, 72:215–227.

Charupat, N. and Milevsky, M. A. (2002). Optimal asset allocation in life annuities: A note. *Insurance: Mathematics and Economics*, 30(2):199–209.

Chemla, G. (2004). Pension fund investment in private equity and venture capital in the US and canada. *Journal of Private Equity*, 7(2):64–71.

Chen, A. and Hieber, P. (2016). Optimal asset allocation in life insurance: The impact of regulation. *ASTIN Bulletin*, 46(3):605–626.

Chen, A., Nguyen, T., and Stadje, M. (2017a). Risk management with multiple VaR. Working paper.

Chen, D. H., Beetsma, R. M., Broeders, D. W., and Pelsser, A. A. (2017b). Sustainability of participation in collective pension schemes: An option pricing approach. *Insurance: Mathematics and Economics*, 74:182 – 196.

Chen, D. H., Beetsma, R. M., Ponds, E. H., and Romp, W. E. (2016). Intergenerational risk-sharing through funded pensions and public debt. *Journal of Pension Economics and Finance*, 15(2):127–159.

Cocco, J. F., Gomes, F. J., and Maenhout, P. J. (2005). Consumption and portfolio choice over the life cycle. *Review of Financial Studies*, 18(2):491–533.

Cocco, J. F. and Volpin, P. F. (2007). Corporate governance of pension plans: The UK evidence. *Financial Analysts Journal*, 63(1):70–83.

Committee on the Global Financial System (2011). Fixed income strategies of insurance companies and pension funds. CGFS Papers, Bank for International Settlements.

Cox, J. C. and Huang, C.-f. (1989). Optimal consumption and portfolio policies when asset prices follow a diffusion process. *Journal of Economic Theory*, 49(1):33–83.

Cox, S. H. and Lin, Y. (2007). Natural hedging of life and annuity mortality risks. *North American Actuarial Journal*, 11(3):1–15.

Crossley, T. and Jametti, M. (2013). Pension benefit insurance and pension plan portfolio choice. *Review of Economics and Statistics*, 95(1):337–341.

Cui, J., de Jong, F., and Ponds, E. (2011). Intergenerational risk sharing within funded pension schemes. *Journal of Pension Economics and Finance*, 10(1):1–29.

Davidoff, T., Brown, J. R., and Diamond, P. A. (2005). Annuities and individual welfare. *American Economic Review*, 95(5):1573–1590.

de Dreu, J. and Bikker, J. A. (2012). Investor sophistication and risk taking. *Journal of Banking and Finance*, 36(7):2145–2156.

de Jong, F. (2000). Time series and cross-section information in affine term-structure models. *Journal of Business and Economic Statistics*, 18(3):300–314.

de Jong, F. (2008). Pension fund investments and the valuation of liabilities under conditional indexation. *Insurance: Mathematics and Economics*, 42(1):1–13.

De Nederlandsche Bank (2015). Statistieken De Nederlandsche Bank. Last accessed: 2015-04-24.

De Nederlandsche Bank (2017). Table 8.8 Estimated funding ratio of pension funds. Last accessed: 2017-03-16.

Denuit, M. (2009). Life annuities with stochastic survival probabilities: A review. *Methodology and Computing in Applied Probability*, 11(3):463–489.

Denuit, M., Haberman, S., and Renshaw, A. (2011). Longevity-indexed life annuities. *North American Actuarial Journal*, 15(1):97–111.

Donnelly, C. (2017). A discussion of a risk-sharing pension plan. *Risks*, 5(1):12.

Dowd, K., Blake, D., and Cairns, A. J. (2011). A computationally efficient algorithm for estimating the distribution of future annuity values under interest-rate and longevity risks. *North American Actuarial Journal*, 15(2):237–247.

Dyck, A. and Pomorski, L. (2016). Investor scale and performance in private equity investments. *Review of Finance*, 20(3):1081–1106.

Erb, C. B., Harvey, C. R., and Viskanta, T. E. (1994). Forecasting international equity correlations. *Financial Analysts Journal*, pages 32–45.

European Commission (2013). Long-term financing of the European economy. Green Paper 52013DC0150.

European Commission (2015). The 2015 Ageing Report: Economic and budgetary projections for the 28 EU Member States (2013-2060). European Economy series 3, Directorate-General for Economic and Financial Affairs.

Fama, E. F. and French, K. R. (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics*, 25(1):23–49.

Fama, E. F. and Schwert, G. W. (1977). Asset returns and inflation. *Journal of Financial Economics*, 5(2):115–146.

Feldstein, M. and Ranguelova, E. (2001). Individual risk in an investment-based social security system. *American Economic Review*, 91(4):1116–1125.

Filipović, D., Kremslehner, R., and Muermann, A. (2015). Optimal investment and premium policies under risk shifting and solvency regulation. *Journal of Risk and Insurance*, 82(2):261–288.

Finkelstein, A. and Poterba, J. (2004). Adverse selection in insurance markets: Policyholder evidence from the UK annuity market. *Journal of Political Economy*, 112(1):183–208.

Fischer, S. (1973). A life cycle model of life insurance purchases. *International Economic Review*, pages 132–152.

Frederick, S., Loewenstein, G., and O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, pages 351–401.

Friedberg, L. and Webb, A. (2007). Life is cheap: Using mortality bonds to hedge aggregate mortality risk. *The BE Journal of Economic Analysis & Policy*, 7(1).

Froot, K. A. (2007). Risk management, capital budgeting, and capital structure policy for insurers and reinsurers. *Journal of Risk and Insurance*, 74(2):273–299.

Gatti, S. (2014). Private financing and government support to promote long-term investments in infrastructure. Report, Organisation for Economic Co-operation and Development.

Gatzert, N., Holzmüller, I., and Schmeiser, H. (2012). Creating customer value in participating life insurance. *Journal of Risk and Insurance*, 79(3):645–670.

GDV (2016). As to the difference between (unit linked) life insurance and UCITS. Explanatory note, Gesamtverband der Deutschen Versicherungswirtschaft e.V. (GDV), German Insurance Association.

Gelfand, I. and Fomin, S. (1963). Elements of the theory. In Silverman, R. A., editor, *Variational Calculus*, pages 1–31. Prentice-Hall, Inc.

Goecke, O. (2013). Pension saving schemes with return smoothing mechanism. *Insurance: Mathematics and Economics*, 53(3):678–689.

Gollier, C. (2008). Intergenerational risk-sharing and risk-taking of a pension fund. *Journal of Public Economics*, 92(5):1463–1485.

Gomes, F. J. (2007). Exploiting short-run predictability. *Journal of Banking and Finance*, 31(5):1427–1440.

Greene, W. H. (2003). Limited dependent variables and duration model. In *Econometric Analysis*, chapter 22, pages 756–802. Pearson Hall, fifth edition.

Guan, G. and Liang, Z. (2014). Optimal management of dc pension plan in a stochastic interest rate and stochastic volatility framework. *Insurance: Mathematics and Economics*, 57:58–66.

Gürkaynak, R. S., Sack, B., and Wright, J. H. (2007). The us treasury yield curve: 1961 to the present. *Journal of Monetary Economics*, 54(8):2291–2304.

Hainaut, D. and Devolder, P. (2007). Management of a pension fund under mortality and financial risks. *Insurance: Mathematics and Economics*, 41(1):134–155.

Hakansson, N. H. (1969). Optimal investment and consumption strategies under risk, an uncertain lifetime, and insurance. *International Economic Review*, 10(3):443–466.

Han, N.-w. and Hung, M.-w. (2012). Optimal asset allocation for DC pension plans under inflation. *Insurance: Mathematics and Economics*, 51(1):172–181.

Hanewald, K., Piggott, J., and Sherris, M. (2013). Individual post-retirement longevity risk management under systematic mortality risk. *Insurance: Mathematics and Economics*, 52(1):87–97.

Honoré, B. E. (1992). Trimmed lad and least squares estimation of truncated and censored regression models with fixed effects. *Econometrica*, pages 533–565.

Honoré, B. E. (2008). On marginal effects in semiparametric censored regression models. Working paper.

Horneff, V., Maurer, R., Mitchell, O. S., and Rogalla, R. (2015). Optimal life cycle portfolio choice with variable annuities offering liquidity and investment downside protection. *Insurance: Mathematics and Economics*, 63:91–107.

Horneff, W. J., Maurer, R. H., Mitchell, O. S., and Dus, I. (2008). Following the rules: integrating asset allocation and annuitization in retirement portfolios. *Insurance: Mathematics and Economics*, 42(1):396–408.

Horneff, W. J., Maurer, R. H., Mitchell, O. S., and Stamos, M. Z. (2010). Variable payout annuities and dynamic portfolio choice in retirement. *Journal of Pension Economics and Finance*, 9(2):163–183.

Huang, H., Milevsky, M. A., and Salisbury, T. S. (2012). Optimal retirement consumption with a stochastic force of mortality. *Insurance: Mathematics and Economics*, 51(2):282–291.

Human Mortality Database (2015). University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Last Accessed: 2015-03-31.

International Monetary Fund (2017). Low growth, low interest rates, and financial intermediation. In *Global Financial Stability Report*, chapter 2, pages 49–82. International Monetary Fund (IMF).

Investment Company Institute (2016). The US Retirement Market, Second Quarter 2016 (September). Last Accessed: 2016-10-24.

Jensen, M. C. (1968). The performance of mutual funds in the period 1945–1964. *Journal of Finance*, 23(2):389–416.

Jin, L., Merton, R. C., and Bodie, Z. (2006). Do a firm's equity returns reflect the risk of its pension plan? *Journal of Financial Economics*, 81(1):1–26.

Kaltepoth, S. (2016). *How Suboptimal are Retail Products? An Analysis of Different Sources of Suboptimality*. PhD thesis, Duisburg, Essen, Universität Duisburg-Essen.

Kamien, M. I. and Muller, E. (1976). Optimal control with integral state equations. *Review of Economic Studies*, 43(3):469–473.

Kamien, M. I. and Schwartz, N. L. (2012). *Dynamic optimization: the calculus of variations and optimal control in economics and management*. Courier Corporation.

Karatzas, I., Lehoczky, J. P., and Shreve, S. E. (1987). Optimal portfolio and consumption decisions for a small investor on a finite horizon. *SIAM Journal on Control and Optimization*, 25(6):1557–1586.

Kim, T. S. and Omberg, E. (1996). Dynamic nonmyopic portfolio behavior. *Review of Financial Studies*, 9(1):141–161.

Kimball, M. S., Sahm, C. R., and Shapiro, M. D. (2008). Imputing risk tolerance from survey responses. *Journal of the American Statistical Association*, 103(483):1028–1038.

Koehn, M. and Santomero, A. M. (1980). Regulation of bank capital and portfolio risk. *Journal of Finance*, 35(5):1235–1244.

Koijen, R. S., Nijman, T. E., and Werker, B. J. (2010). When can life cycle investors benefit from time-varying bond risk premia? *Review of Financial Studies*, 23(2):741–780.

Koijen, R. S., Nijman, T. E., and Werker, B. J. (2011). Optimal annuity risk management. *Review of Finance*, 15(4):799.

Koijen, R. S. and Yogo, M. (2015). The cost of financial frictions for life insurers. *American Economic Review*, 105(1):445–75.

Koijen, R. S. and Yogo, M. (2016a). Risks of life insurers: Recent trends and transmission mechanisms. In Hufeld, F., Koijen, R. S., and Thimann, C., editors, *The Economics, Regulation, and Systemic Risk of Insurance Markets*, chapter 4, pages 79–99. Oxford University Press, first edition.

Koijen, R. S. and Yogo, M. (2016b). Shadow insurance. *Econometrica*, 84(3):1265–1287.

Korn, R. and Kraft, H. (2002). A stochastic control approach to portfolio problems with stochastic interest rates. *SIAM Journal on Control and Optimization*, 40(4):1250–1269.

Ledlie, M., Corry, D., Finkelstein, G., Ritchie, A., Su, K., and Wilson, D. (2008). Variable annuities. *British Actuarial Journal*, 14(2):327–389.

Lee, R. D. and Carter, L. R. (1992). Modeling and forecasting US mortality. *Journal of the American Statistical Association*, 87(419):659–671.

Lioui, A. and Poncet, P. (2001). On optimal portfolio choice under stochastic interest rates. *Journal of Economic Dynamics and Control*, 25(11):1841–1865.

Lockwood, L. M. (2012). Bequest motives and the annuity puzzle. *Review of Economic Dynamics*, 15(2):226–243.

Lucas, D. and Zeldes, S. P. (2006). Valuing and hedging defined benefit pension obligations: The role of stocks revisited. Working paper, Northwestern University and Columbia University.

Luciano, E., Regis, L., and Vigna, E. (2015). Single-and cross-generation natural hedging of longevity and financial risk. *Journal of Risk and Insurance*.

Marketwired News (2015). Solvency of Canadian pension plans falls to lowest level since 2013, according to Aon Hewitt's latest quarterly survey. Last Accessed: 2016-12-13.

Martin, I. (2012). On the valuation of long-dated assets. *Journal of Political Economy*, 120(2):346–358.

Maurer, R., Mitchell, O. S., Rogalla, R., and Kartashov, V. (2013). Lifecycle portfolio choice with systematic longevity risk and variable investment-linked deferred annuities. *Journal of Risk and Insurance*, 80(3):649–676.

Maurer, R., Mitchell, O. S., Rogalla, R., and Siegelin, I. (2016). Accounting and actuarial smoothing of retirement payouts in participating life annuities. *Insurance: Mathematics and Economics*, 71:268–283.

Mertens, E. (2002). Comments on variance of the IID estimator in Lo (2002). Working paper.

Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *Review of Economics and Statistics*, pages 247–257.

Merton, R. C. (1971). Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory*, 3(4):373–413.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica*, pages 867–887.

Miller, M. H. and Modigliani, F. (1961). Dividend policy, growth, and the valuation of shares. *Journal of Business*, 34(4):411–433.

Mitchell, O. S. and McCarthy, D. (2002). Estimating international adverse selection in annuities. *North American Actuarial Journal*, 6(4):38–54.

Modigliani, F. and Miller, M. H. (1958). The cost of capital, corporation finance and the theory of investment. *American Economic Review*, pages 261–297.

Mohan, N. and Zhang, T. (2014). An analysis of risk-taking behavior for public defined benefit pension plans. *Journal of Banking and Finance*, 40:403–419.

Munk, C. (2008). Portfolio and consumption choice with stochastic investment opportunities and habit formation in preferences. *Journal of Economic Dynamics and Control*, 32(11):3560–3589.

Munk, C. and Sørensen, C. (2004). Optimal consumption and investment strategies with stochastic interest rates. *Journal of Banking and Finance*, 28(8):1987–2013.

Munk, C., Sørensen, C., and Vinther, T. N. (2004). Dynamic asset allocation under mean-reverting returns, stochastic interest rates, and inflation uncertainty: Are popular recommendations consistent with rational behavior? *International Review of Economics and Finance*, 13(2):141–166.

Nielsen, L. T. and Vassalou, M. (2004). Sharpe ratios and alphas in continuous time. *Journal of Financial and Quantitative Analysis*, 39(1):103–114.

Nielson, N. L. and Chan, D. K. W. (2007). Private pensions and government guarantees: clues from canada. *Journal of Pension Economics and Finance*, 6:45–66.

Novy-Marx, R. (2013). Logical implications of the gasb?s methodology for valuing pension liabilities. *Financial Analysts Journal*, 69(1).

Otto, C. A. and Volpin, P. F. (2017). Marking to market and inefficient investment decisions. *Management Science*. Articles in Advance.

Pashchenko, S. (2013). Accounting for non-annuitization. *Journal of Public Economics*, 98:53–67.

Peijnenburg, K., Nijman, T., and Werker, B. J. (2017). Health cost risk: A potential solution to the annuity puzzle. *Economic Journal*.

Pennacchi, G. and Rastad, M. (2011). Portfolio allocation for public pension funds. *Journal of Pension Economics and Finance*, 10(02):221–245.

Peskin, M. (2001). Asset/liability management in the public sector. In Mitchell, O. and Hustead, E., editors, *Pensions in the Public Sector*, chapter 9, pages 195–217. University of Pennsylvania Press Philadelphia.

Petersen, M. A. (1996). Allocating assets and discounting cash flows: pension plan finance. *Pensions, Savings and Capital Markets*.

Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies*, 22(1):435–480.

Phan, H. V. and Hegde, S. P. (2013). Corporate governance and risk taking in pension plans: Evidence from defined benefit asset allocations. *Journal of Financial and Quantitative Analysis*, 48(03):919–946.

Piggott, J., Valdez, E. A., and Detzel, B. (2005). The simple analytics of a pooled annuity fund. *Journal of Risk and Insurance*, 72(3):497–520.

Plantin, G., Sapra, H., and Shin, H. S. (2008). Marking-to-market: Panacea or pandora's box? *Journal of Accounting Research*, 46(2):435–460.

Plantin, G. and Tirole, J. (2015). Marking to market versus taking to market. SRC Discussion Paper No. 51.

Ponds, E. H. and Van Riel, B. (2009). Sharing risk: the netherlands' new approach to pensions. *Journal of Pension Economics and Finance*, 8(1):91–105.

Poterba, J. M. (2001). Demographic structure and asset returns. *Review of Economics and Statistics*, 83(4):565–584.

Poterba, J. M. and Samwick, A. A. (2003). Taxation and household portfolio composition: Us evidence from the 1980s and 1990s. *Journal of Public Economics*, 87(1):5–38.

Qiao, C. and Sherris, M. (2013). Managing systematic mortality risk with group self-pooling and annuitization schemes. *Journal of Risk and Insurance*, 80(4):949–974.

Rauh, J. D. (2009). Risk shifting versus risk management: investment policy in corporate pension plans. *Review of Financial Studies*, 22(7):2687–2733.

Reichling, F. and Smetters, K. (2015). Optimal annuitization with stochastic mortality and correlated medical costs. *American Economic Review*, 105(11):3273–3320.

Renshaw, A. E. and Haberman, S. (2006). A cohort-based extension to the Lee–Carter model for mortality reduction factors. *Insurance: Mathematics and Economics*, 38(3):556–570.

Richard, S. F. (1975). Optimal consumption, portfolio and life insurance rules for an uncertain lived individual in a continuous time model. *Journal of Financial Economics*, 2(2):187–203.

Richter, A. and Weber, F. (2011). Mortality-indexed annuities: Managing longevity risk via product design. *North American Actuarial Journal*, 15(2):212–236.

Rosen, H. S. and Wu, S. (2004). Portfolio choice and health status. *Journal of Financial Economics*, 72(3):457–484.

Samuelson, P. A. (1969). Lifetime portfolio selection by dynamic stochastic programming. *Review of Economics and Statistics*, pages 239–246.

Samuelson, W. and Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty*, 1(1):7–59.

Sangvinatsos, A. and Wachter, J. A. (2005). Does the failure of the expectations hypothesis matter for long-term investors? *Journal of Finance*, 60(1):179–230.

Schich, S. (2008a). Challenges for financial intermediaries offering asset decumulation products. OECD Journal: Financial Market Trends 15, Organisation for Economic Co-operation and Development.

Schich, S. (2008b). Revisiting the asset-meltdown hypothesis. OECD Journal: Financial Market Trends 15, Organisation for Economic Co-operation and Development.

Schotman, P. C. and Schweitzer, M. (2000). Horizon sensitivity of the inflation hedge of stocks. *Journal of Empirical Finance*, 7(3):301–315.

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

Sharpe, W. F. (2007). Lockbox separation. Working paper.

Sharpe, W. F. and Treynor, J. L. (1977). The principles of corporate pension finance. *Journal of Finance*, 32(2):627–638.

Shi, Z. and Werker, B. J. (2012). Short-horizon regulation for long-term investors. *Journal of Banking and Finance*, 36(12):3227–3238.

Shiller, R. J. and Beltratti, A. E. (1992). Stock prices and bond yields: Can their comovements be explained in terms of present value models? *Journal of Monetary Economics*, 30(1):25–46.

Society of Actuaries (2014). RP-2014 Mortality Tables. Last Accessed: 2016-08-24.

Sørensen, C. (1999). Dynamic asset allocation and fixed income management. *Journal of Financial and Quantitative Analysis*, 34(4):513–531.

Steffensen, M. (2011). Optimal consumption and investment under time-varying relative risk aversion. *Journal of Economic Dynamics and Control*, 35(5):659–667.

Sundaresan, S. and Zapatero, F. (1997). Valuation, optimal asset allocation and retirement incentives of pension plans. *Review of Financial Studies*, 10(3):631–660.

Szymanowska, M., Roon, F., Nijman, T., and Goorbergh, R. (2014). An anatomy of commodity futures risk premia. *Journal of Finance*, 69(1):453–482.

Tan, K. S., Blake, D. P., and MacMinn, R. D. (2015). Longevity risk and capital markets: The 2013-14 update. *Insurance: Mathematics and Economics*, 63:1–11.

Thompson, S. B. (2011). Simple formulas for standard errors that cluster by both firm and time. *Journal of Financial Economics*, 99(1):1–10.

Tsai, J. T., Wang, J. L., and Tzeng, L. Y. (2010). On the optimal product mix in life insurance companies using conditional value at risk. *Insurance: Mathematics and Economics*, 46(1):235–241.

U.S. Chamber of Commerce (2016). Private retirement benefits in the 21st century: Achieving retirement security. Report, United States Chamber of Commerce, Washington, DC.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2):177–188.

Viceira, L. M. (2001). Optimal portfolio choice for long-horizon investors with nontradable labor income. *Journal of Finance*, 56(2):433–470.

Vigna, E. and Haberman, S. (2001). Optimal investment strategy for defined contribution pension schemes. *Insurance: Mathematics and Economics*, 28(2):233–262.

Visco, I. (2006). Longevity risk and financial markets. Keynote speech to the 26th SUERF Colloquium, Lisbon, 12-14 October 2006. Last accessed: 2017-04-20.

Wachter, J. A. (2002). Portfolio and consumption decisions under mean-reverting returns: An exact solution for complete markets. *Journal of Financial and Quantitative Analysis*, 37(1):63–91.

Wakker, P., Thaler, R., and Tversky, A. (1997). Probabilistic insurance. *Journal of Risk and Uncertainty*, 15(1):7–28.

Weale, M. and van de Ven, J. (2016). Variable annuities and aggregate mortality risk. *National Institute Economic Review*, 237(1):55–61.

Weinert, J.-H. and Gründl, H. (2017). The modern tontine: An innovative instrument for longevity risk management in an aging society. ICIR working paper series, International Center for Insurance Regulation (ICIR), Goethe University.

Welch, I. and Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, 21(4):1455–1508.

Willis Towers Watson (2017). FTSE 350 defined contribution pension scheme survey 2017. Infographic. Last Accessed: 2017-02-24.

Wong, A., Sherris, M., and Stevens, R. (2017). Natural hedging strategies for life insurers: impact of product design and risk measure. *Journal of Risk and Insurance*, 84:153–175.

World Economic Forum (2017). We'll Live to 100 - How Can We Afford It? White paper.

Yaari, M. E. (1965). Uncertain lifetime, life insurance, and the theory of the consumer. *Review of Economic Studies*, 32(2):137–150.

Yang, J., Acheson, L., Holt, J., Rupp, G., and Spica, K. (2016). 2015 Target-Date Fund Landscape. Report, Vanguard Group, Inc. Last Accessed: 2017-03-06.

Yogo, M. (2016). Portfolio choice in retirement: Health risk and the demand for annuities, housing, and risky assets. *Journal of Monetary Economics*, 80:17–34.

Zhang, S. and Li, J. S.-H. (2017). Longevity risk-sharing annuities: Partial indexation in mortality experience. *Asia-Pacific Journal of Risk and Insurance*, 11(1).

Zimmer, A., Schade, C., and Gründl, H. (2009). Is default risk acceptable when purchasing insurance? Experimental evidence for different probability representations, reasons for default, and framings. *Journal of Economic Psychology*, 30(1):11–23.