Essays on intertemporal consumption and portfolio choice

van Bilsen, Servaas

Publication date:
2015

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

Link to publication in Tilburg University Research Portal

Citation for published version (APA):
van Bilsen, S. (2015). *Essays on intertemporal consumption and portfolio choice*. CentER, Center for Economic Research.

Essays on

Intertemporal Consumption and Portfolio Choice

Servaas van Bilsen

Essays on
Intertemporal Consumption and Portfolio Choice

Proefschrift

ter verkrijging van de graad van doctor aan Tilburg University op gezag van de rector magnificus, prof. dr. E.H.L. Aarts, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de Aula van de Universiteit op woensdag 4 november 2015 om 16.15 uur door

Servaas van Bilsen

geboren op 1 augustus 1987 te Rotterdam.

Promotiecommissie:

Promotores:	prof. dr. Roger Laeven
	prof. dr. Theo Nijman
	prof. dr. Lans Bovenberg
Overige Leden:	prof. dr. Olivia Mitchell
	prof. dr. Gur Huberman
	prof. dr. Roel Beetsma
	prof. dr. Frank de Jong
	prof. dr. Hans Schumacher

Acknowledgements

This dissertation is the result of four years of hard work. During this period, I have been supported and encouraged by a number of extraordinary people. Without their support and encouragement, this dissertation would not have been possible. I would like to express my sincere gratitude to all of them, especially the ones mentioned below.

First and foremost, I would like to thank my PhD supervisors Roger Laeven, Theo Nijman and Lans Bovenberg for their dedication and support throughout my PhD journey. I wrote my Research MSc Thesis under the supervision of Roger. After completing my Research Master's degree, Roger offered me a position as a PhD candidate. I owe him many thanks for this. During my PhD period, I have always enjoyed working with Roger. I would like to thank him for the time and energy he has devoted to helping me improve my work. Many thanks to Roger also for offering me a tenure track position at the University of Amsterdam. I have known Theo and Lans since my third year of the Bachelor Econometrics & Operations Research. At that time, I was a research assistant at the network for studies on pensions, aging and retirement (Netspar). They have always been an inspiration to me. Many thanks goes to them for their time, energy, ideas, and intellectual spirit.

A special thanks goes to Olivia Mitchell. She was my host when I visited The Wharton School of the University of Pennsylvania. She made my time at Wharton an enjoyable and unforgettable experience. I am grateful for the time she devoted to me and my work. We met almost every morning. I greatly thank her for giving me the opportunity to present at the AEW Seminar. Many thanks to Olivia also for her willingness to serve as a member of the PhD committee. I also feel indebted to the committee members Gur Huberman, Roel Beetsma, Hans Schumacher and Frank de Jong for their time, energy, and helpful input in creating this dissertation.

Finally, I am deeply grateful to my parents – Dor and Brigitte – and my two younger brothers – Pepijn and Camiel – for their unconditional love, moral support and encouragement. Many thanks also to my friends whose friendship has proven invaluable over the years. Especially I would like to thank Emile van Elen and Vincent Schothuis for being my paranymphs. I thank Emile also for proofreading parts of my dissertation.

Servaas van Bilsen

October 1, 2015

Contents

Chapter 1

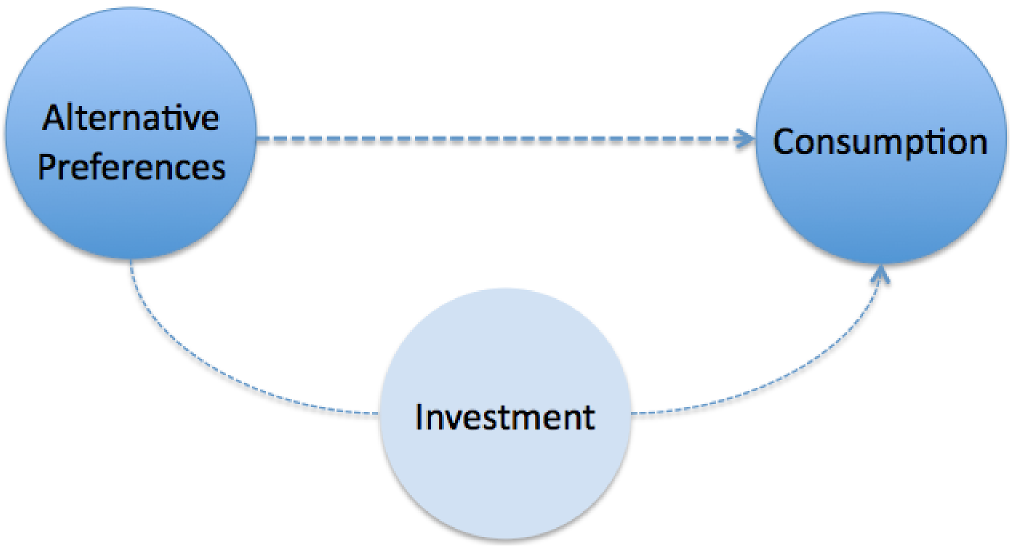
Introduction

This dissertation aims to extend the academic literature on optimal consumption and portfolio choice over the life cycle. In particular, we analyze optimal choice under preference specifications that incorporate loss aversion, internal habit formation and probability weighting. Furthermore, this dissertation formalizes and analyzes a new pension contract, a so-called personal pension plan with risk sharing (PPR), that plays a dominant role in recent policy reform discussions in the Netherlands. This dissertation has implications for a wide variety of real world pension contracts. We analyze (dis)saving and investing in not only the accumulation phase but also the decumulation (payout) phase of defined contribution (DC) pension plans. This is highly relevant as many retirees worry about the lack of guidance and regulation on how to draw-down accumulated wealth in retirement. This dissertation is equally relevant to analyze reform options for defined benefit (DB) pension plans. In many countries, employers are no longer able or willing to absorb the (investment) risks of their pension plans. We analyze pension plans (without external risk sponsors) that aim to retain key attractive features of DB pension plans (such as stable lifelong income streams). Adequate design of consumption and portfolio strategies, which is the central theme of this dissertation, is thus of great importance to many workers and retirees around the world.

Part I

The classical workhorse model for the determination of an agent's optimal consumption and portfolio choice is the Merton model (Merton, 1969). This model advocates to invest a constant fraction of total wealth into risk-bearing assets, and to consume at a constant fraction of wealth. The Merton model implies life cycle investment of financial wealth (net of human capital); see also Bodie, Merton, and Samuelson (1992). These results are driven by strong simplifying assumptions about preferences, labor income, and future investment opportunities. The first part of my dissertation (Chapters 2, 3 and 4) explores novel extensions of the classical Merton model. In particular, we focus on deriving and studying optimal consumption and portfolio choice under alternative preference specifications. To keep the analysis tractable and to isolate the effect of preference specifications, we retain the assumptions of risk-free (tradable) labor income (see, e.g., Cocco, Gomes, and Maenhout, 2005; Benzoni, Collin-Dufresne, and Goldstein, 2007, for extensions), and independent and normally distributed stock returns (see, e.g., Liu, 2007, for extensions). Figure 1.1 illustrates the central idea of the first part of my dissertation: the analysis of optimal consumption and portfolio choice under alternative preferences. By contrast, the second part of my dissertation (Chapters 5, 6 and 7) abstains from explicit preference assumptions, and takes the consumption and portfolio decisions as given.

Figure 1.1.



The first part of my dissertation focuses on deriving and studying optimal consumption and portfolio decisions under alternative preference specifications.

Chapter 2 derives and analyzes the optimal consumption and portfolio choice of a loss averse agent. His reference level, which divides consumption into gains and losses, is endogenously updated over time. Loss aversion and reference dependence constitute two key aspects of prospect theory (PT for short), developed by Tversky and Kahneman (1992). While the PT literature typically considers an exogenous reference level, Chapter

2 assumes that the current reference level is a function of past consumption choices, reflecting internal habit formation.[1] We find that, compared to the Merton model, consumption is shifted from good to bad economic scenarios. As a result, the agent can maintain consumption above the reference level in many economic scenarios; he only consumes below the reference level when the economy is doing really bad. This finding is due to loss aversion, and triggers a demand for "guarantee like" features in pension products. We also find that consumption adjusts gradually (and not directly as in the Merton consumption strategy) to unexpected financial shocks. This finding is due to endogenous updating of the reference level, and justifies a mechanism for smoothing the change in consumption due to financial shocks. The fraction of total wealth invested in risk-bearing assets is low in economic scenarios where consumption is close to the reference level. Indeed, the coefficient of relative risk aversion increases as consumption approaches the reference level. As is well-known, under the Merton model, the individual invests a constant (age independent) fraction of wealth into risk-bearing assets. Chapter 2 shows that the endogenous reference specification triggers life cycle investing, not only in the accumulation phase but also in the decumulation phase (i.e., life cycle investment of not only financial wealth but also total wealth). Intuitively, households with a shorter investment horizon are less flexible in absorbing financial shocks. Hence, older households take less investment risk and thus own smaller investment portfolios. Furthermore, our model is consistent with two stylized facts about consumption data: hump-shaped consumption profiles, and excess smoothness and sensitivity in consumption.

The third key aspect of PT is probability weighting. Chapter 3 analyzes the impact of probability weighting on the agent's optimal consumption and investment decisions. This chapter – which extends Chapter 2 – explores a dynamic consumption and investment choice problem featuring loss aversion, endogenous updating of the reference level, as well as probability weighting. We show that an inverse S-shaped probability weighting function is able to generate an endogenous floor on consumption (i.e., a "hard" guarantee rather than a "soft" guarantee as in Chapter 2).

[1] Our preference model implies that the elasticity of intertemporal substitution and the coefficient of relative risk aversion are not constant (as assumed by the Merton model) but rather depend on age or financial shocks. In particular, in the case of habit formation, the elasticity of intertemporal substitution and the coefficient of relative risk aversion depend on the investment horizon and thus on age. With loss aversion, the coefficient of relative risk aversion depends on financial shocks, giving rise to a dynamic investment policy.

In Chapter 4, we build a consumption and investment choice model that combines the ratio model of (internal) habit formation with stochastic differential utility (i.e., continuous-time limit of recursive utility; see Duffie and Epstein, 1992). These two utility models are particularly popular in the life cycle literature. We obtain closed-form solutions by applying a linearization to the agent's budget constraint. Our results show that the agent gradually adjusts consumption to financial shocks. This justifies a return smoothing mechanism. We are able to fully characterize (in terms of the preference parameters) this return smoothing mechanism. The ratio model of habit formation analyzed in this chapter differs from the additive model of habit formation (analyzed in Chapters 2 and 3), in that relative risk aversion is constant. As a result, the optimal investment strategy is state-independent, and thus easy to implement. This is a clear advantage of the ratio model of habit formation over the additive model of habit formation. While in the Merton model the coefficient of relative risk aversion and the elasticity of intertemporal substitution are intimately related, this is not the case in the model of Chapter 4.

Part II

The pension plans proposed by Bovenberg and Nijman (2015) promise to play a new role in the provision of retirement income in the Netherlands.[2] These pension plans, which are called personal pension plans with risk pooling (PPRs), unbundle the various functions of variable annuities. In particular, a PPR individualizes the (dis)savings and investment functions of variable annuities, and arranges the insurance function (i.e., pooling of idiosyncratic longevity risk) collectively. A PPR defines property rights in terms of a personal investment account, rather than in terms of payouts or annuity units (as in variable annuities). Policyholders can adopt an investment approach or a consumption approach to a PPR. The investment approach takes the investment policy, the assumed rate of return (ARR) and the initial amount of capital as given. The consumption stream is derived endogenously (i.e., volatility of consumption, growth rate

[2] At the time of completion of this dissertation, the Dutch government has proposed new regulation on the decumulation of Dutch DC pension plans based on the PPR concept. Furthermore, the Dutch government has announced to consider PPRs as a very important option for the reform of the current Dutch DB pension plans (e.g., to overcome the conflicts of interest between policyholders and to clarify property rights).

of consumption and initial consumption). By contrast, the consumption approach takes the consumption stream as given, and derives the investment policy, the ARR and the initial amount of capital endogenously.

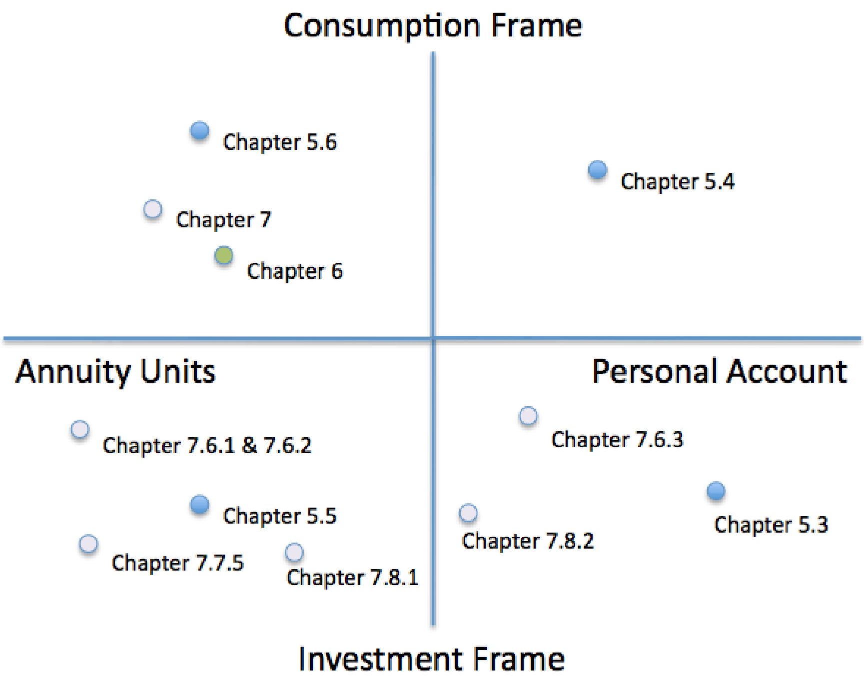
Chapter 5 explores the investment approach and the consumption approach in more detail. This chapter also explores a collective defined contribution (CDC) and a collective defined ambition (CDA) pension system. These collective pension systems define property rights in terms of annuity units, rather than in terms of personal investment accounts. A CDC and a CDA pension system feature one general investment account and thus cannot tailor (dis)saving and investment policies to individual preferences and individual investment beliefs. Furthermore, valuation of annuity units can be difficult as assets are not assigned to individual policyholders. This may result in conflicts of interests between policyholders. An advantage of a collective pension system is that non-traded risks (e.g., systematic longevity risk) can be shared between generations.

The pension plans considered in the second part of my dissertation can be classified along two criteria: the definition of property rights (personal investment account versus annuity units) and the framing of pension plans (investment frame versus consumption frame). Figure 1.2 classifies the various sections of Chapters 5, 6 and 7 along these two criteria. The horizontal axis shows the first criterium, while the vertical axis depicts the second criterium.

The pension plan considered in Chapter 6 adopts a consumption frame and defines property rights in terms of annuity units; see also Figure 1.2. In this chapter, we assume that annuity units respond gradually to financial shocks.Gradual absorption of financial shocks is consistent with internal habit formation (this is formally shown in Chapter 4). This chapter values the annuity units in a market-consistent fashion. In particular, we show that the market-consistent discount rate includes a risk premium that rises with the horizon and that the optimal fraction of accumulated wealth invested in risk-bearing assets decreases as the policyholder ages. Also, we show that gradual absorption of financial shocks leads to predictable changes in annuity units.

Chapter 7 investigates the pricing and risk management of variable annuities. We consider an economy with three risk factors: real interest rate risk, expected inflation risk and stock market risk (Chapters 5 and 6 only consider stock market risk). This chapter shows that the prices of variable real annuities can be less sensitive to the nominal

Figure 1.2.



The figure classifies the sections of Chapters 5, 6 and 7 along two criteria: the definition of property rights (horizontal axis) and the framing of pension plans (vertical axis). For example, the pension plan considered in Chapter 6 adopts the consumption approach and defines property rights in terms of annuity units.

interest rate than the prices of fixed nominal annuities. This finding is of key importance to determine the optimal hedging of interest rate risk, and is driven by three factors. First, the desired growth rate of the annuity payment may change with the interest rate due to intertemporal substitution in consumption. Second, the desired growth rate of the annuity payment increases with the expected inflation rate. Hence, the prices of real annuities depend on the real rather than the nominal interest rate. In an incomplete financial market in which expected inflation risk and real interest rate risk cannot be hedged at the same time, insurers must trade-off hedging expected inflation risk against hedging real interest rate risk. This reduces the nominal interest sensitivity of the annuity factor, especially when fluctuations in the nominal interest rate are driven by changes in the expected inflation rate rather than by changes in the real interest rate. Third, the expected returns on risky securities tend to be less sensitive to the nominal interest rate as compared to the returns on safe securities. Hence, the nominal interest sensitivity of real annuities is relatively low if the policyholder takes speculative stock market risk.

Part I:
Optimal Consumption and Portfolio Choice Using Preference Models

Chapter 2

Consumption and Portfolio Choice under Loss Aversion and Endogenous Updating of the Reference Level[3]

This chapter explicitly derives the optimal dynamic consumption and portfolio choice of a loss averse agent who endogenously updates his reference level. His optimal choice seeks protection against consumption losses due to downside financial shocks. This induces a (soft) guarantee on consumption and is due to loss aversion. Furthermore, his optimal consumption choice gradually adjusts to financial shocks. This resembles the payout streams of financial plans that respond sluggishly, smoothing investment returns to reduce payout volatility, and is due to endogenous updating. The welfare losses associated with various suboptimal consumption and portfolio strategies are also evaluated. They can be substantial.

2.1. Introduction

The pension fund industry has grown dramatically over the past four decades: total U.S. retirement assets rose from 369 billion dollars in 1974 to 23 trillion dollars in 2013 (Investment Company Institute, 2014). During the same period, we have seen in particular a pronounced increase in retirement saving through personal retirement accounts, such as IRAs and DC plans (Poterba, Venti, and Wise, 2009). More specifically, the percentage of total U.S. retirement assets accounted for by IRAs and DC plans grew

[3] This chapter is co-authored with Roger Laeven and Theo Nijman.

from about 18% in 1974 to about 54% in 2013 (Investment Company Institute, 2014). These figures highlight the importance of adequate individual consumption, savings and investment decisions over the life cycle, and of the design of such individual financial plans.

Since the seminal works of Merton (1969, 1971) and Samuelson (1969), a considerable number of authors have studied optimal consumption and portfolio choice over the life cycle in a wide variety of settings. Standard life cycle models assume that preferences are represented by expected utility with constant relative risk aversion (CRRA); see, e.g., Wachter (2002), Cocco et al. (2005), Liu (2007), Gomes, Kotlikoff, and Viceira (2008), to name just a few. With such standard preferences (and without constraints), the optimal log consumption choice is a *linear* function of the log state price density (see, e.g., Karatzas and Shreve, 1998, p. 103). Furthermore, under such standard preferences, financial shocks are *directly* absorbed into the optimal log consumption choice: a CRRA agent chooses to instantaneously adjust consumption to financial shocks.

These predictions of standard life cycle models stand in sharp contrast to *actual* income streams generated by financial and insurance products. Financial fiduciaries have developed a variety of features, options and guarantees so as to make base financial products more attractive for individuals (see, e.g., Van Rooij, Kool, and Prast, 2007; Antolín, Payet, Whitehouse, and Yermo, 2011; Bodie and Taqqu, 2011). These include guaranteed minimum income benefits, guaranteed minimum withdrawal benefits and minimum rate of return guarantees. In addition, many actively traded financial derivative securities have a nonlinear payoff structure, and provide some degree of protection against downside risk. The popularity of these products contradicts the linearity of the standard consumption rule.

Also, a substantial body of literature (see, e.g., Sundaresan, 1989; Constantinides, 1990) argues that agents become accustomed to a certain level of consumption. This strand of the literature suggests that agents evaluate and adjust consumption relative to a reference (or a habit) level. The empirical literature (see, e.g., Lupton, 2003) provides evidence of habit persistence in consumption, with consumption being smooth relative to wealth. Moreover, financial fiduciaries (such as life insurers and pension funds) increasingly offer plans with payout streams that are not directly but only sluggishly linked to the performance of the underlying investment portfolio.[4] There have been numerous attempts to reconcile theory and practice of life cycle consumption and portfolio choice. However, to the best of our knowledge, the literature has not yet been able to provide a fully satisfactory answer that accommodates these features — nonlinearity of the consumption rule and smoothing of financial shocks — all together.

This chapter explores consumption and portfolio choice under reference-dependent preferences. More specifically, we analyze optimal consumption and portfolio choice under the utility (or value) function of prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992) and adopt an endogenous updating mechanism for the dynamics of the reference level.[5] The consumption and portfolio choice model we consider is able to generate both a nonlinear consumption rule and smoothing of financial shocks in an integrated framework. The optimal choice seeks protection against consumption losses due to financial shocks inducing a ("soft") guarantee on consumption. Furthermore, the optimal consumption choice exhibits sluggish response to financial shocks.

Following prospect theory, we assume that the agent's instantaneous utility function is represented by the two-part power utility function. This utility function incorporates several behavioral properties, such as reference dependence (i.e., the carriers of utility are gains and losses rather than absolute levels of consumption), loss aversion (i.e., losses hurt more than gains satisfy), and diminishing sensitivity (i.e., the impact of a marginal change in consumption decreases as the agent moves further away from the reference level).[6,7] Diminishing sensitivity implies a convex utility function below the reference level.[8] The empirical literature is, however, inconclusive as to whether the utility function is convex in the loss domain; see, e.g., Abdellaoui, Vossmann, and Weber (2005).[9] Therefore, the current chapter considers not only the case of a convex utility

[4]In many European countries, but also in the US and Japan, the importance of participating (or with profits) annuities is growing (see, e.g., Guillén, Jørgensen, and Nielsen, 2006; Maurer, Mitchell, and Rogalla, 2010). A key characteristic of participating annuities is that investment returns are smoothed so as to reduce payout volatility. For example, in the Netherlands, pension funds are allowed to gradually absorb financial shocks into pension entitlements. Also, life insurers use special smoothing techniques in an attempt to stabilize payouts.

[5]We abstract away from probability weighting.

[6]Kőszegi and Rabin (2006, 2007) develop a class of reference-dependent preferences with endogenous updating (and without probability weighting). See Section 2.4 for further details about the connection between the class of Kőszegi and Rabin (2006, 2007) and our model.

[7]According to Wakker (2010, p. 242), "reference dependence, in combination with loss aversion, is one of the most pronounced empirical phenomena in decision under risk and uncertainty."

[8]We note that, in our context, a convex utility function implies risk-seeking behavior.

[9]Etchart-Vincent (2004) investigated the sensitivity of the utility function to the magnitude of the

function in the loss domain, but also the case of a concave utility function in the loss domain.[10]

Our main results can be summarized as follows. First, we demonstrate that the agent optimally chooses to divide the states of the economy into two categories: insured states (i.e., good to intermediate economic scenarios or, equivalently, low to intermediate state prices) and *un*insured states (i.e., bad economic scenarios or high state prices). In insured states, consumption is guaranteed to be larger than the reference level, while in uninsured states, consumption is smaller than the reference level. If consumption is larger (smaller) than the reference level, then the agent experiences a gain (loss). Because of loss aversion, the agent has a strong preference to maintain consumption above the reference level, but when the state of the economy is really bad, the (soft) guarantee on consumption can no longer be maintained. More specifically, the optimal consumption profile (i.e., the optimal consumption choice as a function of the log state price density) displays a 90° rotated S-shaped pattern.[11] We show that when the agent becomes more afraid of incurring losses, the probability of consumption falling below the reference level decreases. At the same time, the agent must give up some upward potential in order to finance this more conservative consumption profile.

Second, under our preference assumptions, the optimal consumption choice gradually adjusts to financial shocks. Kahneman and Tversky (1979) argue that the status quo, an expectation or an aspiration level can serve as a reference level, but do not specify how the reference level is formed and updated over time. Following the internal habit formation literature (see, e.g., Constantinides, 1990), we assume that the reference level depends on the agent's own past consumption choices. More specifically, we assume that the reference level can be decomposed into two components: a stochastic and a deterministic component.[12] The stochastic component is given by an exponentially weighted average

underlying payoffs. She found that a larger proportion of the subjects exhibited concavity when facing large losses than when facing small losses.

[10] The literature also provides some support for the idea that agents exhibit an inverted S-shaped utility function in the loss domain. For example, Laughhunn, Payne, and Crum (1980) found that a large proportion of the subjects (64%) switched from risk-seeking to risk-averse behavior when facing ruinous losses.

[11] The exact behavior of the agent below the reference level depends on the shape of utility function in the loss domain.

[12] The reference level is characterized by three parameters: the initial reference level, an endogeneity parameter (which measures the degree of endogeneity) and a deprecation parameter (which measures the rate at which the agent depreciates the reference level).

of the agent's own past consumption choices. The specification of the reference level is motivated by the idea that agents become accustomed to a certain level of consumption. A main implication of the consumption and portfolio choice model we consider is that after a financial shock, optimal consumption adjustment is sluggish (at least in the short run). That is, a current financial shock has a larger impact on consumption in the distant future than on consumption in the near future. Part of the financial shock will be directly reflected into gains and losses, another part will smoothly enter through the reference level, which is endogenously updated over time.

Third, the optimal portfolio profile displays a U-shaped pattern: the total dollar amount invested in risk-bearing assets will be lower in intermediate economic scenarios than in good or bad economic scenarios. As a by-product of interest in its own right, the agent implements a life cycle investment strategy, even without taking human capital into account.[13] Since the agent has less time to absorb financial shocks as he grows older, the equity risk exposure, on average, decreases over the life cycle.

Finally, to investigate the impact of implementing suboptimal consumption and portfolio strategies on the agent's welfare, we conduct a welfare analysis. We compute the welfare losses (in terms of the relative decline in certainty equivalent consumption) associated with implementing suboptimal consumption and portfolio strategies. Because of the endogeneity of the reference level, this requires a non-standard computation of certainty equivalents. The results indicate that welfare losses can be substantial. Particularly, for our realistic parameter values, we find that the welfare loss associated with implementing the classical Merton strategy (see Merton, 1969) can be as large as 40%. We also compute the welfare losses of suboptimal behavior due to incorrect assumptions on the underlying agent's preference parameters. We find that consumption and portfolio strategies based on incorrectly assuming a constant exogenous reference level (or only a very limited degree of endogeneity), thus implying no (or only very limited) smoothing of financial shocks, substantially reduce welfare.

In order to solve the consumption and portfolio choice model, we first apply the

[13] Under CRRA utility, the agent has a constant equity risk exposure if the investment opportunity set is assumed to be constant. Bodie et al. (1992) give a justification for adopting a life cycle investment strategy based on human capital considerations. If human capital is risk-free, then agents implicitly hold a risk-free asset. To offset this implicit risk-free asset holding, financial wealth should be tilted toward risky assets. As the share of human capital in total wealth decreases from one to zero during the working period, the optimal proportion of financial wealth invested in risk-free assets increases over the life cycle.

solution technique of Schroder and Skiadas (2002). This method enables us to convert the consumption and portfolio choice model with endogenous updating into a *dual* consumption and portfolio choice model without endogenous updating. The dual utility function is time-additive and separable. This fact facilitates the derivation of the optimal consumption and portfolio choice. Next, we solve the dual problem by using convex duality (or martingale) techniques, and by using techniques proposed by Basak and Shapiro (2001) and Berkelaar, Kouwenberg, and Post (2004) in order to deal with pseudo-concavity and non-differentiability aspects of the problem. We adapt the latter techniques to our setting with intertemporal consumption. Upon transforming our solutions under the dual model back into the primal model, we finally arrive at explicit closed-form solutions to our initial problem under consideration.

The remainder of this chapter is structured as follows. Section 2.2 provides a literature review. The economy is described in Section 2.3. The agent's instantaneous utility function is introduced in Section 2.4. Section 2.5 derives the optimal consumption and portfolio choice. The properties of the optimal strategies are explored in Section 2.6. Section 2.7 considers, as a robustness check, the optimal consumption and portfolio choice under a slightly alternative specification of the agent's instantaneous utility function. Finally, Section 2.8 concludes the chapter. The proofs of the theorems and propositions and the details of the certainty equivalent computations are relegated to the Appendix.

2.2. Literature Review

In this chapter, we extend the existing life cycle literature by analyzing an alternative preference specification that embeds two key aspects of prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992) – loss aversion and reference dependence – in a continuous-time framework.[14] To isolate the effect of preferences, we assume risk-free (tradable) labor income (see, e.g., Cocco et al., 2005; Benzoni et al., 2007; Lynch and Tan, 2011, for extensions), and independent and normally distributed stock returns (see, e.g., Liu, 2007; Buraschi, Porchia, and Trojani, 2010, for extensions). In an extension

[14]Prospect theory has been actively studied in the finance literature to explain the equity premium puzzle (Benartzi and Thaler, 1995), the cross section of average returns (Barberis and Huang, 2008), and the deposition effect (Barberis and Xiong, 2009).

of our model, we also explore the implications of probability weighting as a third key aspect of prospect theory (see Chapter 3).

The literature on optimal consumption *and* portfolio choice under prospect theory type preferences is scarce. Berkelaar et al. (2004) examine analytically optimal portfolio choice under the two-part power utility function. Their model differs from ours in at least two main respects. First and foremost, we assume that the agent is concerned not with terminal wealth, but with intertemporal consumption. This allows us to examine how the agent's consumption strategy evolves as time proceeds and risk resolves, which is our prime focus. Second, in this setting with intertemporal consumption, we allow the agent to not just stochastically but also endogenously update his reference (or habit) level of consumption over time. Guasoni, Huberman, and Ren (2014) explore the optimal consumption (or spending) and portfolio choice of a short fall averse agent. This paper considers a multiplicative habit formation model in which, in contrast to the traditional approach of Abel (1990), the habit level (or reference level) equals past peak spending. By contrast, we assume that the agent's preferences are characterized by the two-part power utility function, and that the reference level is equal to a weighted average of past consumption choices. Jin and Zhou (2008) and He and Zhou (2011, 2014) consider optimal portfolio choice under prospect theory. They focus on the impact of probability weighting on optimal portfolio (not consumption) choice, developing an analytic solution method based on a quantile formulation. They do not consider endogenous updating of the reference level. Our model specification has the attractive feature that it allows to analyze both separately and jointly the effects on consumption and portfolio choice of loss aversion and of endogenous updating of the reference level, which are controlled in the model by separate parameters. Furthermore, our model nests traditional models, such as models with internal habit formation, with an exogenous minimum level of consumption, and with CRRA utility, as special (limiting) cases.

Our first finding is that loss aversion, entailing that negative changes in consumption are perceived more severely than equivalent positive changes in consumption, triggers a demand for "guarantee like" features in the consumption profile that we also encounter in many real life financial plans. This finding is consistent with the related strand of the literature on regret aversion, driven by fears of unfavorable outcomes, initiated by Bell (1982) and Loomes and Sugden (1982). For example, Muermann, Mitchell, and

Volkman (2006) show, in a static portfolio choice problem, that regret aversion has a positive impact on the willingness to pay for a rate of return guarantee on the risky asset; see also Merton and Bodie (2005). Different from the above mentioned papers, our paper generates this implication in a dynamic consumption-portfolio choice setting in which guarantees take the form of a stable consumption profile at (typically) or above (in good states of the world) the reference level of consumption, rather than, e.g., a rate of return guarantee. Only in very bad states of the world, consumption falls below the reference level. Traditional life cycle models (see, e.g., Merton, 1969) cannot explain the demand for "guarantee like" features in the consumption profile.

We combine our model of loss aversion with an endogenous reference level that is a arithmetic function of past consumption choices (Constantinides, 1990). However different from traditional habit formation models, we allow consumption to fall below the reference level (see also Detemple and Karatzas, 2003). Under our endogenous updating mechanism of the reference level, consumption responds gradually to financial shocks. Shocks are absorbed in not only the level of consumption but also future growth rates of consumption.

Building on their earlier work, Kőszegi and Rabin (2009) explore a model that embeds loss aversion and reference dependence into a discrete-period model. In their model, the agent receives utility from the difference between current consumption and last period's expectation of current consumption ("contemporaneous gain-loss utility") and from changes in expectations regarding future consumption ("prospective gain-loss utility"). The agent is loss averse in the sense that losses loom larger than same-sized gains. Also, a contemporaneous loss is more painful than a prospective loss. Kőszegi and Rabin (2009) find that the agent has a first-order precautionary savings motive: the agent increases savings to reduce the marginal utility associated with a future loss. Furthermore, the fact that news about future consumption affects current utility less than news about current consumption creates an immediate incentive to overconsume relative to his optimal pre-committed consumption path. The agent of Kőszegi and Rabin (2009) thus behaves inconsistently while our agent is time consistent.

Pagel (2012) shows that the model of Kőszegi and Rabin (2009) can explain a number of stylized facts about consumption. First, she finds that the precautionary saving motive together with the tendency to overconsume can produce a realistic hump-shaped profile of consumption over the lifetime. Second, she finds that the model of Kőszegi and Rabin (2009) generates excess smoothness and sensitivity in consumption (i.e., consumption adjusts gradually to financial shocks). Intuitively, *un*expected losses today are more painful than expected losses tomorrow. Our model is also able to generate a hump-shaped pattern of consumption as a result of two competing effects: the endogeneity of the reference level (which causes a precautionary savings motive) and an uncertain lifetime (which causes a tendency to consume early in life). Excess smoothness and sensitivity in consumption are also present in our model.

2.3. The Economy

We define a continuous-time financial market following Karatzas and Shreve (1998) and Back (2010). Let $T > 0$ be a fixed finite terminal time. The uncertainty in the economy is represented by a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, on which is defined a standard N-dimensional Brownian motion $\{Z_t\}_{t\in[0,T]}$. Let the filtration $\mathbb{F} \equiv \{\mathcal{F}_t\}_{t\in[0,T]}$ be the augmentation under $\mathbb{P}$ of the natural filtration generated by the standard Brownian motion $\{Z_t\}_{t\in[0,T]}$. Throughout, (in)equalities between random variables are meant to hold $\mathbb{P}$-almost surely.

The financial market consists of an instantaneously risk-free asset and N risky stocks, which are traded continuously on the time horizon $[0, T]$. The price of the risk-free asset, B, evolves according to

$$\frac{\mathrm{d}B_t}{B_t} = r_t \,\mathrm{d}t, \qquad B_0 = 1.$$

The scalar-valued risk-free rate process, r, is assumed to be $\mathcal{F}_t$-progressively measurable and uniformly bounded. The N-dimensional vector of risky stock prices, S, satisfies the following stochastic differential equation:

$$\frac{\mathrm{d}S_t}{S_t} = \mu_t \,\mathrm{d}t + \sigma_t \,\mathrm{d}Z_t, \qquad S_0 = \mathbf{1}_N.$$

Here, $\mathbf{1}_N$ denotes an N-dimensional vector of all ones. The N-dimensional mean rate of return process, μ, and the $(N \times N)$-matrix-valued volatility process, σ, are both assumed to be $\mathcal{F}_t$-progressively measurable and uniformly bounded.

We assume that, for some positive ϵ,

$$\vartheta^\top \sigma_t \sigma_t^\top \vartheta \geq \epsilon ||\vartheta||^2, \qquad \text{for all } \vartheta \in \mathbb{R}^N. \tag{2.3.1}$$

Here, $\top$ denotes the transpose sign. The strong non-degeneracy condition (2.3.1) implies that the inverse of σ_t exists and is bounded. The $\mathcal{F}_t$-progressively measurable market price of risk process, λ, solves the following equation:

$$\sigma_t \lambda_t \equiv \mu_t - r_t \mathbf{1}_N.$$

The unique positive-valued state price density process, M, can now be defined as follows:

$$M_t \equiv \exp\left\{-\int_0^t r_s \,\mathrm{d}s - \int_0^t \lambda_s^\top \,\mathrm{d}Z_s - \frac{1}{2}\int_0^t ||\lambda_s||^2 \,\mathrm{d}s\right\}.$$

The economy is populated by a single price-taking agent endowed with initial wealth $W_0 \geq 0$. The agent's objective is to choose an $\mathcal{F}_t$-progressively measurable N-dimensional process π, referred to as the portfolio process and representing the dollar amounts invested in the N risky stocks, and an $\mathcal{F}_t$-progressively measurable process c, referred to as the consumption process, so as to maximize the expectation of lifetime utility.[15] We impose the following integrability conditions, which we assume throughout to be satisfied for any consumption and portfolio process:

$$\int_0^T \pi_t^\top \sigma_t \sigma_t^\top \pi_t \,\mathrm{d}t < \infty, \qquad \int_0^T \left|\pi_t \left(\mu_t - r_t \mathbf{1}_N\right)\right| \mathrm{d}t < \infty, \qquad \mathbb{E}\left[\int_0^T |c_t|^2 \,\mathrm{d}t\right] < \infty.$$

The wealth process, W, associated with a consumption and portfolio strategy (c, π) satisfies the following *dynamic budget constraint*:

$$\mathrm{d}W_t = \left(r_t W_t + \pi_t^\top \sigma_t \lambda_t - c_t\right) \mathrm{d}t + \pi_t^\top \sigma_t \,\mathrm{d}Z_t, \qquad W_0 \geq 0 \text{ given.} \tag{2.3.2}$$

Equation (2.3.2) reveals that the agent's wealth equals initial wealth, plus trading gains, minus cumulative consumption. The total dollar amount invested in the risk-free asset at time $t \in [0, T]$ is given by $W_t - \pi_t^\top \mathbf{1}_N$. We call a consumption and portfolio strategy

[15]The agent's utility function is introduced in the next section.

admissible if the associated wealth process is uniformly bounded from below. Then the *static budget constraint* is also satisfied; see, e.g., Karatzas and Shreve (1998, p. 91-92) for further details.

2.4. The Agent's Utility Function

This section introduces the agent's (instantaneous) utility function $u(c_t;\theta_t)$. Here, θ_t represents the agent's reference level to which consumption is compared. We assume that the agent derives utility from the difference between consumption c_t and the reference level θ_t. Specifically, following the prospect theory literature (see, e.g., Tversky and Kahneman, 1992), we assume that the agent's utility function $u(c_t;\theta_t)$ is represented by the two-part power utility function:

$$u(c_t;\theta_t) = v(c_t - \theta_t) \equiv \begin{cases} -\kappa(\theta_t - c_t)^{\gamma_1}, & \text{if } c_t < \theta_t; \\ (c_t - \theta_t)^{\gamma_2}, & \text{if } c_t \geq \theta_t. \end{cases} \tag{2.4.1}$$

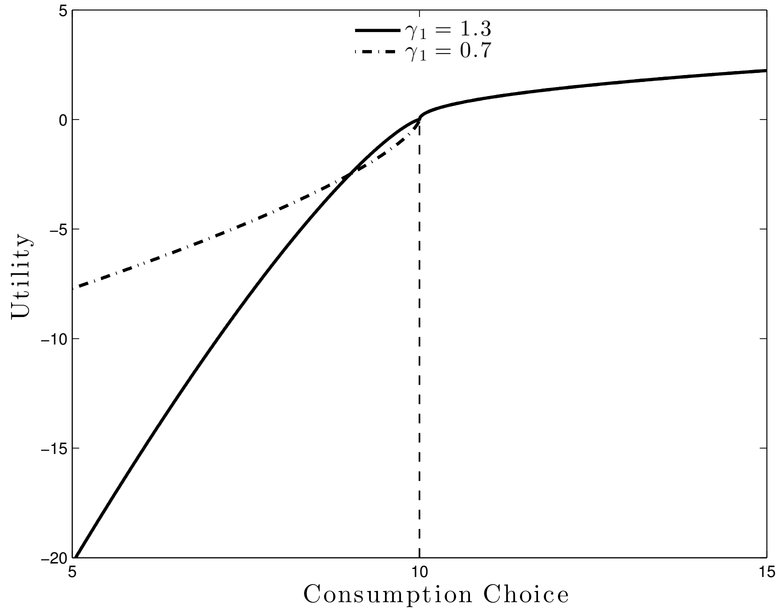
Here, $\gamma_1 > 0$ and $\gamma_2 \in (0,1)$ are curvature parameters, and $\kappa \geq 1$ stands for the loss aversion index. If consumption is larger (smaller) than the reference level, then the agent experiences a gain (loss).

Figure 2.1 illustrates the two-part power utility function (2.4.1) for $\gamma_1 = 1.3$ (solid line) and $\gamma_1 = 0.7$ (dash-dotted line). The figure shows that the two-part power utility function exhibits a kink at the reference level. The kink is due to the different treatment of gains and losses. We note that even in the case of $\kappa = 1$, the agent's utility function displays a kink at the reference level whenever $\gamma_1 \neq \gamma_2$.

A simple calculation shows that the two-part power utility function (2.4.1) is convex below the agent's reference level if $\gamma_1 \leq 1$, and concave otherwise. Convexity corresponds to risk-seeking behavior and concavity to risk-averse behavior.[16] Tversky and Kahneman (1992) found that the agent's utility function is convex in the loss domain. Table 2.1 reviews the empirical literature regarding the shape of the utility function for losses. The

[16] This statement is not true if probabilities are distorted (see Chateauneuf and Cohen, 1994). For example, an S-shaped utility function and overweighting of small probabilities can together explain the fourfold pattern of risk attitudes: risk-averse behavior when gains have large probabilities and losses have small probabilities, and risk-seeking behavior when losses have large probabilities and gains have small probabilities.

Figure 2.1.
Illustration of the two-part power utility function



The figure illustrates the two-part power utility function for $\gamma_1 = 1.3$ (solid line) and $\gamma_1 = 0.7$ (dash-dotted line). The agent's reference level is set equal to 10, the loss aversion index κ to 2.5 and γ_2 to 0.5.

table shows that the literature is inconclusive as to whether the utility function is convex below the reference level. Among the mentioned studies, Etchart-Vincent (2004) explored the sensitivity of the agent's utility function to the magnitude of the underlying payoffs. She found that a larger proportion of the subjects exhibited concavity when facing large losses than when facing small losses. Etchart-Vincent (2004) argued that this finding may be due to the size of the losses at stake. Therefore, the current chapter considers not only the case of a convex utility function in the loss domain ($0 < \gamma_1 \leq 1$), but also the case of a concave utility function in the loss domain ($\gamma_1 > 1$).

Motivated by the literature on internal habit formation (see, e.g., Constantinides, 1990; Detemple and Zapatero, 1992; Detemple and Karatzas, 2003), we assume that the agent's reference level evolves according to:

$$\mathrm{d}\theta_t = (\beta c_t - \alpha\theta_t)\,\mathrm{d}t, \qquad \theta_0 \geq 0 \text{ given.}$$

Here, θ_0 denotes the agent's initial reference level, $\alpha \geq 0$ corresponds to the depreciation

Table 2.1.
Classification of the utility function for losses

	Shape of the utility function for losses			
Study	Convex	Concave	Linear	Mixed
Abdellaoui (2000)	42.5	20.0	25.0	12.5
Abdellaoui et al. (2005)	24.4	22.0	22.0	31.7
Abdellaoui, Bleichrodt, and Paraschiv (2007)	68.8	8.3	22.9	-
Booij and van de Kuilen (2009)	47.1	22.5	30.4	-
Etchart-Vincent (2004)*	37.1	25.7	25.7	11.4

*The reported results are for the case of large losses.

The table reviews the empirical literature regarding the shape of the utility function for losses. Numbers are expressed as a percentage of total subjects. All the mentioned studies use the trade-off method (see Wakker and Deneffe, 1996) to elicit the utility functions of the subjects.

(or persistence) parameter, and $\beta \geq 0$ indexes the extent to which the current reference level responds to current consumption. The agent's reference level exhibits a low degree of depreciation (or a high degree of persistence) if α is low. The impact of current consumption on the current reference level increases as β increases. We can explicitly write the agent's reference level as follows:

$$\theta_s = \beta \int_t^s \exp\{-\alpha(s-u)\}\, c_u \, \mathrm{d}u + \exp\{-\alpha(s-t)\}\, \theta_t, \qquad s \geq t \geq 0. \tag{2.4.2}$$

Equation (2.4.2) shows that the reference level can be decomposed into two components: a stochastic and a deterministic component. The parameter β measures the importance of the stochastic component relative to the deterministic component. In what follows, we refer to β as the endogeneity parameter. The stochastic component becomes more important as β increases. The first component on the right-hand side of equation (2.4.2) is an exponentially weighted integral of the agent's *own* past consumption choices (i.e., the reference level is backward-looking). We observe that the current reference level depends more on consumption in the recent past than on consumption in the distant past. The second component on the right-hand side of equation (2.4.2) is independent of past consumption choices and decreases exponentially at a rate of α.

The two-part utility function (see equation (2.4.1)) is a member of the class of reference-dependent preferences introduced by Kőszegi and Rabin (2006, 2007). They assume that the agent's instantaneous utility function can be decomposed into two components. The first component represents classical utility from consumption; that is, utility derived from absolute levels of consumption. The second component captures reference-dependent gain-loss utility; that is, utility derived from the difference between classical consumption utility and the reference level of utility. Specifically, Kőszegi and Rabin (2006, 2007) consider the following agent's utility function:

$$u(c_t;\theta_t) = \eta \cdot m(c_t) + (1-\eta) \cdot w(m(c_t) - m(\theta_t)). \tag{2.4.3}$$

Here, m stands for the classical consumption utility function, w denotes the gain-loss utility function and $\eta \in [0,1]$ is a weight parameter controlling the relative importance of the two components. The two-part utility function (2.4.1) emerges as a special case of (2.4.3) if the gain-loss utility function w is represented by the two-part power utility function (2.4.1), the weight parameter η is set equal to zero and $m(c_t) = c_t$. Section 2.7 considers another special case of (2.4.3), where the weight parameter η is unequal to zero. Kőszegi and Rabin (2006, 2007) do not assume that the agent's reference level is a weighted integral of past consumption choices. Instead, they assume that the agent's reference level represents an expectation. Both Kőszegi and Rabin (2006, 2007) and our model assume that the reference level is chosen endogenously.[17]

The two-part power utility function (2.4.1) displays loss aversion in the sense that the disutility of a loss of one unit is κ times larger than the utility of a gain of one unit.[18] There is, however, no agreed-upon definition of loss aversion in the literature. According to Kahneman and Tversky (1979), loss aversion refers to the fact that losses loom larger than same-sized gains, i.e., $-w(-x) > w(x)$ for all $x > 0$. A loss aversion index can then be defined as the mean or median value of $-w(-x)/w(x)$ over relevant x (see Abdellaoui, Bleichrodt, and L'Haridon, 2008). Köbberling and Wakker (2005) define the loss aversion index as the ratio between the left-hand and right-hand derivative of the gain-loss utility function at the reference level. The loss aversion index κ is equal to the loss aversion index proposed by Köbberling and Wakker (2005) if $\gamma_1 = \gamma_2$.

Finally, we note that the two-part power utility function (2.4.1) with reference level dynamics given by (2.4.2) includes several important special (limiting) cases. The

[17] Yogo (2008) analyzes asset pricing implications of reference-dependent preferences, with an exogenously given reference level.

[18] As pointed out by Wakker (2010, p. 267), the degree of loss aversion depends on the monetary unit.

internal habit formation model studied by Constantinides (1990) arises as a special case if the agent is infinitely loss averse. The assumption of infinite loss aversion implies that consumption is not allowed to fall below the reference level. If the reference level is also assumed to be exogenous, then the two-part power utility function reduces to a utility function with an exogenous minimum consumption level. Such a utility function has been studied by Deelstra, Grasselli, and Koehl (2003). The constant relative risk aversion (CRRA) utility function emerges as a special case if the reference level is equal to zero and consumption is non-negative. The CRRA utility function has been widely explored in the economics literature since at least Merton (1969).

2.5. The Consumption and Portfolio Choice Problem

This section derives the agent's optimal consumption and portfolio choice. Section 2.5.1 formulates the agent's maximization problem. To determine the optimal consumption and portfolio choice, we transform the agent's (primal) maximization problem into a dual problem. The technique that solves this dual problem is outlined in Section 2.5.2. Section 2.5.3 presents the optimal consumption choice and Section 2.5.4 gives the optimal portfolio choice.

2.5.1. The Agent's Maximization Problem

The agent's dynamic consumption and portfolio choice problem of Section 2.3 with the agent's utility function given in Section 2.4 can, by virtue of the martingale approach (Pliska, 1986; Karatzas, Lehoczky, and Shreve, 1987; Cox and Huang, 1989, 1991), be transformed into the following equivalent static variational problem:

$$\begin{aligned} &\underset{c}{\text{maximize}} && \mathbb{E}\left[\int_0^T \exp\{-\delta t\}\, v\left(c_t - \theta_t\right) \mathrm{d}t\right] \\ &\text{subject to} && \mathbb{E}\left[\int_0^T M_t c_t \,\mathrm{d}t\right] \leq W_0, \qquad \mathrm{d}\theta_t = \left(\beta c_t - \alpha\theta_t\right)\mathrm{d}t, \\ &&& c_t \geq \theta_t - L_t \quad \text{for all } t \in [0, T]. \end{aligned} \tag{2.5.1}$$

Here, $\delta \geq 0$ stands for the subjective rate of time preference. We require that consumption is not allowed to fall more than $L_t \geq 0$ below the agent's reference level θ_t.[19] In addition, we assume that L_t only depends on time t (and not on the state of nature $\omega \in \Omega$).[20] If $L_t = \exp\{-\alpha t\}\,\theta_0$, then consumption is guaranteed to be non-negative. We can view $\theta_t - L_t$ as the agent's *minimum consumption level.*

2.5.2. The Dual Technique

To derive the optimal consumption and portfolio choice in our model, we first apply the solution technique proposed by Schroder and Skiadas (2002). These authors show that a generic consumption and portfolio choice model with linear internal habit formation can be mechanically transformed into a *dual* consumption and portfolio choice model without linear internal habit formation.[21] Hereinafter, we refer to the solution technique considered by Schroder and Skiadas (2002) as the dual technique. This section sketches the basic ideas underlying the dual technique. The Appendix provides more details.

The dual consumption and portfolio choice model (see (2.9.1) in the Appendix) is solved in a dual financial market. This dual financial market is characterized by the dual state price density $\widehat{M}_t$, the dual (instantaneously) risk-free rate $\widehat{r}_t$, the dual volatility $\widehat{\sigma}_t$ and the dual market price of risk $\widehat{\lambda}_t$:

$$\widehat{M}_t \equiv M_t\left(1 + \beta A_t\right),$$

$$\widehat{r}_t \equiv \beta + \frac{r_t - \alpha\beta A_t}{1 + \beta A_t},$$

$$\widehat{\sigma}_t \equiv \sigma_t,$$

$$\widehat{\lambda}_t \equiv \lambda_t - \frac{\beta}{1 + \beta A_t}\int_t^T \exp\{-(\alpha - \beta)(s - t)\}\, P_{t,s}\Psi_{t,s}\,\mathrm{d}s.$$

Here, $P_{t,s}$ corresponds to the time t price of a default-free unit discount bond that matures at time $s \geq t \geq 0$, and $\Psi_{t,s}$ stands for the time t volatility of the instantaneous return

[19] In the case of risk-seeking behavior in the loss domain, the agent's maximization problem is ill-posed if consumption is not bounded from below (a maximization problem is called ill-posed if its supremum is infinite).

[20] One could argue that L_t should also depend on the agent's past consumption choices. However, this would complicate the agent's maximization problem considerably. We leave it for future research to explore the impact of an endogenous L_t on the agent's optimal consumption and portfolio choice.

[21] The consumption and portfolio choice model considered in the current chapter features a utility specification that incorporates linear internal habit formation.

on such a bond (both in the primal financial market). We can view $A_t \geq 0$ as the time t price of a bond paying a continuous coupon:

$$A_t \equiv \frac{1}{M_t}\mathbb{E}_t\left[\int_t^T M_s \exp\{-(\alpha-\beta)(s-t)\}\,\mathrm{d}s\right].$$

In case the investment opportunity set is constant, A_t only depends on time t. As a consequence, the optimal portfolio choice can be computed explicitly in this case (see Section 2.5.4).

Dual wealth $\widehat{W}_t$ is subject to the following dynamic equation:

$$\mathrm{d}\widehat{W}_t = \left(\widehat{r}_t\widehat{W}_t + \widehat{\pi}_t^\top\widehat{\sigma}_t\widehat{\lambda}_t - \widehat{c}_t\right)\mathrm{d}t + \widehat{\pi}_t^\top\widehat{\sigma}_t\,\mathrm{d}Z_t, \qquad \widehat{W}_0 = \frac{W_0 - A_0\theta_0}{1+\beta A_0}. \tag{2.5.2}$$

Here, $\widehat{c}_t \equiv c_t - \theta_t$ stands for the agent's *surplus consumption choice* and $\widehat{\pi}_t$ denotes the dual portfolio choice. Dual wealth $\widehat{W}_t$ is equal to the discounted value of future surplus consumption choices. Hence, we can view $\widehat{W}_t$ as wealth needed to finance future gains and losses. In what follows, we refer to $\widehat{W}_t$ as *surplus wealth*.

The condition of consumption being bounded from below in (2.5.1) implies that the agent's initial wealth W_0 must be sufficiently large to ensure the existence of an optimal consumption strategy. Specifically, we require

$$W_0 \geq -\mathbb{E}\left[\int_0^T \frac{\widehat{M}_t}{\widehat{M}_0}L_t\,\mathrm{d}t\right] - \beta A_0\mathbb{E}\left[\int_0^T \frac{\widehat{M}_t}{\widehat{M}_0}L_t\,\mathrm{d}t\right] + A_0\theta_0. \tag{2.5.3}$$

The right-hand side of equation (2.5.3) corresponds to initial wealth that is required to finance the minimum consumption stream $\{\theta_t - L_t\}_{t\in[0,T]}$. We note that W_0 is also required to be non-negative; see equation (2.3.2).

2.5.3. The Optimal Consumption Choice

This section derives the optimal consumption choice. We obtain the optimal consumption choice as follows. First, the agent's maximization problem (2.5.1) is converted into its dual problem (Section 2.5.2). The dual utility function is time-additive and separable. This fact facilitates the derivation of the optimal consumption and portfolio choice. Second, the dual problem is solved using martingale techniques and by adapting to our setting with intertemporal consumption the solution technique as described by Basak and Shapiro (2001) and Berkelaar et al. (2004). The central idea of the latter solution technique is to split the agent's (dual, in our case) problem into two maximization problems: a gain part problem and a loss part problem. The optimal solution to each problem represents a *local* maximum of the dual problem. The *global* maximum of the dual problem is determined by comparing, in a particular way, the two local maxima. Finally, the optimal surplus consumption choice $\widehat{c}_t^*$ is translated back into the agent's optimal consumption choice c_t^*. Theorem 1 below presents the optimal consumption choice c_t^*. We note that the theorem distinguishes between risk-averse and risk-seeking behavior in the loss domain. Indeed, in the case of risk-averse behavior in the loss domain, the utility function is concave below the reference level, whereas in the case of risk-seeking behavior in the loss domain, the utility function is convex in the loss domain.

Theorem 1. *Consider an agent with the two-part power utility function* (2.4.1) *and reference level dynamics* (2.4.2) *who solves the consumption and portfolio choice problem* (2.5.1)*. Let θ^* be the agent's optimal reference level implied by substituting the (past) optimal consumption choice in* (2.4.2) *and let y be the Lagrange multiplier associated with the static budget constraint in* (2.5.1)*. Define*

$$k_t \equiv \frac{y \exp\{\delta t\}}{\gamma_2} \quad \text{and} \quad l_t \equiv \frac{y \exp\{\delta t\}}{\kappa \gamma_1}.$$

Then:

- *If the agent is risk-averse in the loss domain, the optimal consumption choice c_t^* at time $t \in [0, T]$ is given by*

$$c_t^* = \begin{cases} \theta_t^* + \left(k_t \widehat{M}_t\right)^{\frac{1}{\gamma_2 - 1}}, & \text{if } \widehat{M}_t \leq \xi_t; \\ \theta_t^* - \left[\left(l_t \widehat{M}_t\right)^{\frac{1}{\gamma_1 - 1}} \wedge L_t\right], & \text{if } \widehat{M}_t > \xi_t. \end{cases}$$

The threshold ξ_t is determined in such a way that $f(\xi_t) = 0$ where the function f is defined as follows:

$$\begin{aligned} f(x) \equiv{} & \exp\{-\delta t\}(1-\gamma_2)(k_t x)^{\frac{\gamma_2}{\gamma_2 - 1}} + \kappa \exp\{-\delta t\}\left[(l_t x)^{\frac{1}{\gamma_1 - 1}} \wedge L_t\right]^{\gamma_1} \\ & - yx\left[(l_t x)^{\frac{1}{\gamma - 1}} \wedge L_t\right]. \end{aligned} \tag{2.5.4}$$

- *If the agent is risk-seeking in the loss domain, the optimal consumption choice c_t^* at time $t \in [0, T]$ is given by*

$$c_t^* = \begin{cases} \theta_t^* + \left(k_t \widehat{M}_t\right)^{\frac{1}{\gamma_2 - 1}}, & \text{if } \widehat{M}_t \leq \xi_t; \\ \theta_t^* - L_t, & \text{if } \widehat{M}_t > \xi_t. \end{cases}$$

The threshold ξ_t is determined in such a way that $g\left(\xi_t\right) = 0$ where the function g is defined as follows:

$$g(x) \equiv \exp\left\{-\delta t\right\}\left(1 - \gamma_2\right)\left(k_t x\right)^{\frac{\gamma_2}{\gamma_2 - 1}} + \kappa \exp\left\{-\delta t\right\} L_t^{\gamma_1} - y x L_t. \tag{2.5.5}$$

The Lagrange multiplier y is chosen such that the static budget constraint holds with equality.

Theorem 1 demonstrates that the agent optimally chooses to divide the states of the economy into two categories: insured states (good to intermediate economic scenarios or, equivalently, low to intermediate state prices) and uninsured states (bad economic scenarios or high state prices). In insured states, consumption is guaranteed to be larger than the reference level, while in uninsured states, consumption is smaller than the reference level. The optimal consumption choice is, however, never equal to the reference level. Section 2.6 further explores the properties of the optimal consumption choice.

2.5.3.1. Comparative Statics

The threshold ξ_t and the Lagrange multiplier y depend on the preference parameters. Proposition 1 summarizes the impact of an increase in the agent's preference parameters on the threshold ξ_t and the Lagrange multiplier y, ceteris paribus.

Proposition 1. *Consider an agent with the two-part power utility function* (2.4.1) *and reference level dynamics* (2.4.2) *who solves the consumption and portfolio choice problem* (2.5.1). *Then:*

- *All else being equal, if the loss aversion index κ increases, then both the threshold ξ_t and the Lagrange multiplier y increase.*

- *All else being equal, if the agent's initial reference level θ_0 increases, then the threshold ξ_t decreases and the Lagrange multiplier y increases.*

Suppose that initial surplus wealth $\widehat{W}_0$ is non-negative.

- *All else being equal, if the depreciation parameter α increases, then the threshold ξ_t increases and the Lagrange multiplier y decreases.*
- *All else being equal, if the endogeneity parameter β increases, then the threshold ξ_t decreases and the Lagrange multiplier y increases.*

Proposition 1 shows that when the agent becomes more afraid of incurring losses, the probability of consumption falling below the reference level decreases. At the same time, the agent must give up some upward potential to finance the new consumption profile. When the agent's initial reference level increases (or the depreciation parameter α decreases or the endogeneity parameter β increases), more wealth is required to finance future reference levels. As a consequence, the probability of incurring a loss increases.

2.5.4. The Optimal Portfolio Choice

To derive the optimal portfolio choice, we first need to derive the agent's optimal wealth W_t^*. As pointed out in the Appendix (see Proposition 4), the agent's optimal wealth W_t^* can be decomposed as follows:

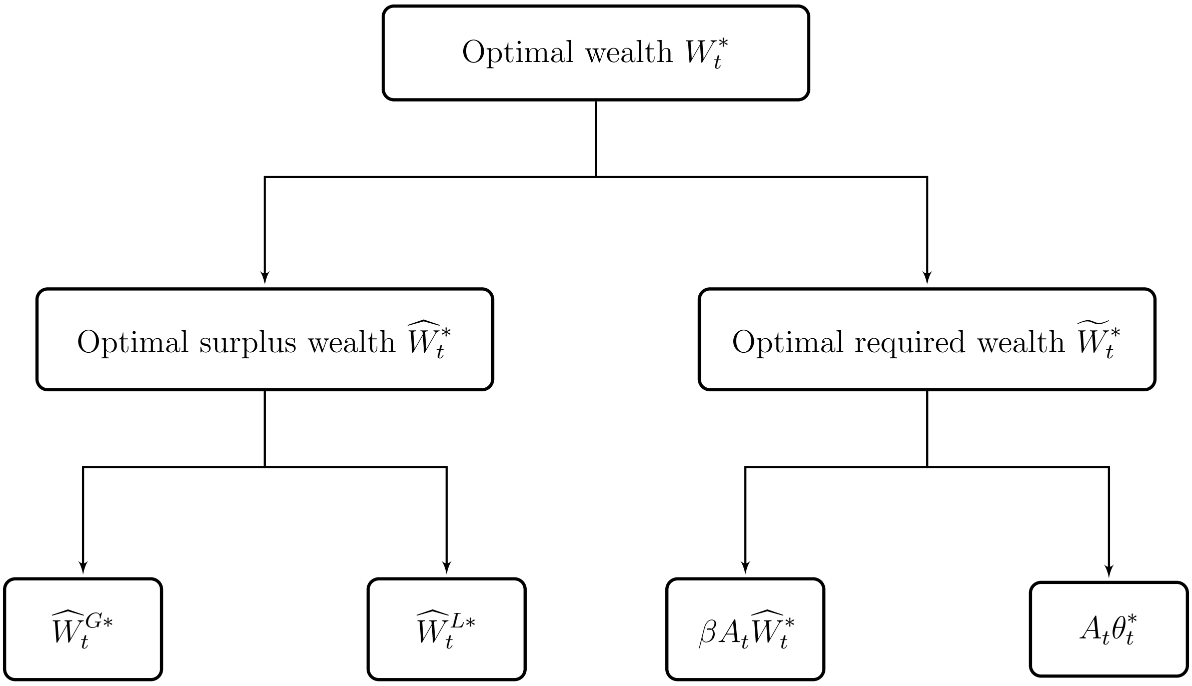
$$W_t^* = \widehat{W}_t^* + \widetilde{W}_t^*. \tag{2.5.6}$$

Here, $\widehat{W}_t^*$ denotes optimal surplus wealth, and $\widetilde{W}_t^*$ stands for wealth required to finance future optimal reference levels. We refer to $\widetilde{W}_t^*$ as optimal *required wealth*. Optimal surplus wealth $\widehat{W}_t^*$ and optimal required wealth $\widetilde{W}_t^*$ can be further decomposed as follows:

$$\widehat{W}_t^* = \widehat{W}_t^{G*} + \widehat{W}_t^{L*} \quad \text{and} \quad \widetilde{W}_t^* = \beta A_t \widehat{W}_t^* + A_t \theta_t^*. \tag{2.5.7}$$

Here, $\widehat{W}_t^{G*}$ denotes wealth required to finance future optimal gains, $\widehat{W}_t^{L*}$ corresponds to wealth required to finance future optimal losses, $\beta A_t \widehat{W}_t^*$ stands for wealth required to finance the stochastic part of future optimal reference levels, and $A_t \theta_t^*$ represents wealth required to finance the deterministic part of future optimal reference levels. Figure 2.2 illustrates the decomposition of the agent's optimal wealth W_t^*.

Figure 2.2.
Decomposition of the agent's optimal wealth W_t^*



The figure illustrates the decomposition of the agent's optimal wealth W_t^*.

Proposition 2 below presents $\widehat{W}_t^{G*}$ and $\widehat{W}_t^{L*}$ for the case of a constant investment opportunity set (i.e., $r_t = r$, $\sigma_t = \sigma$ and $\lambda_t = \lambda$). The general expressions for $\widehat{W}_t^{G*}$ and $\widehat{W}_t^{L*}$ are given in the Appendix.

Proposition 2. *Consider an agent with the two-part power utility function* (2.4.1) *and reference level dynamics* (2.4.2) *who solves the consumption and portfolio choice problem* (2.5.1) *assuming a constant investment opportunity set. Let $\mathcal{N}$ denote the cumulative distribution function of a standard normal random variable. Define Γ_u, Π_u, $d_1(x)$, $d_2(x)$ and $d_3(x)$ as follows:*

$$\Gamma_u = \frac{\delta - \gamma_2 \widehat{r}_u}{1-\gamma_2} - \frac{1}{2}\frac{\gamma_2}{(1-\gamma_2)^2}||\lambda||^2, \qquad \Pi_u = \frac{\delta - \gamma_1 \widehat{r}_u}{1-\gamma_1} - \frac{1}{2}\frac{\gamma_1}{(1-\gamma_1)^2}||\lambda||^2,$$

$$d_1(x) = \frac{1}{||\lambda||\sqrt{s-t}}\left[\log(x) - \log\left(\widehat{M}_t\right) + \int_t^s \widehat{r}_u \,\mathrm{d}u - \frac{1}{2}||\lambda||^2(s-t)\right],$$

$$d_2(x) = d_1(x) + \frac{||\lambda||}{1-\gamma_2}\sqrt{s-t}, \qquad d_3(x) = d_1(x) + \frac{||\lambda||}{1-\gamma_1}\sqrt{s-t}.$$

Then:

- *If the agent is risk-averse in the loss domain, we find*

$$\widehat{W}_t^{G*} = \left(k_t \widehat{M}_t\right)^{\frac{1}{\gamma_2-1}} \int_t^T \exp\left\{-\int_t^s \Gamma_u \,\mathrm{d}u\right\} \mathcal{N}\left[d_2\left(\xi_s\right)\right] \mathrm{d}s,$$

$$\begin{aligned}\widehat{W}_t^{L*} = {} & \left(l_t \widehat{M}_t\right)^{\frac{1}{\gamma_1-1}} \int_t^T \exp\left\{-\int_t^s \Pi_u \,\mathrm{d}u\right\} \left(\mathcal{N}\left[d_3\left(\zeta_s \vee \xi_s\right)\right] - \mathcal{N}\left[d_3\left(\xi_s\right)\right]\right) \mathrm{d}s \\ & - \int_t^T \exp\left\{-\int_t^s \widehat{r}_u \,\mathrm{d}u\right\} L_s \mathcal{N}\left[-d_1\left(\zeta_s \vee \xi_s\right)\right] \mathrm{d}s.\end{aligned}$$

Here, $\zeta_s \equiv \exp\{\delta s\} \gamma_1 \kappa L_s^{\gamma_1-1} y^{-1}$. The threshold ξ_s is determined in such a way that $f\left(\xi_s\right) = 0$ where the function f is given by equation (2.5.4).

- *If the agent is risk-seeking in the loss domain, we find*

$$\widehat{W}_t^{G*} = \left(l_t \widehat{M}_t\right)^{\frac{1}{\gamma_2-1}} \int_t^T \exp\left\{-\int_t^s \Gamma_u \,\mathrm{d}u\right\} \mathcal{N}\left[d_2\left(\xi_s\right)\right] \mathrm{d}s,$$

$$\widehat{W}_t^{L*} = \int_t^T \exp\left\{-\int_t^s \widehat{r}_u \,\mathrm{d}u\right\} L_s \mathcal{N}\left[-d_1\left(\xi_s\right)\right] \mathrm{d}s.$$

The threshold ξ_s is determined in such a way that $g\left(\xi_s\right) = 0$ where the function g is given by equation (2.5.5).

When the dual state price density tends to zero (so that the probability of the dual state price density $\widehat{M}_s$ being smaller than the threshold ξ_s approaches one), optimal surplus wealth $\widehat{W}_t^*$ converges to the optimal wealth of an agent with CRRA utility. Hence, in good economic scenarios, the agent behaves like a CRRA agent.

The optimal dual portfolio choice can be constructed using hedging arguments. We explicitly determine the optimal dual portfolio choice for the case of a constant investment opportunity set. To this end, it is convenient to express $\widehat{W}_t^*$ as a function of time t and the dual state price density $\widehat{M}_t$; that is, $\widehat{W}_t^* = h\left(t, \widehat{M}_t\right)$ for some (regular) function h. Straightforward application of Itô's Lemma to the function h yields

$$\mathrm{d}\widehat{W}_t^* = \left[\frac{\partial h}{\partial t} - \frac{\partial h}{\partial \widehat{M}_t}\widehat{M}_t \widehat{r}_t + \frac{1}{2}\frac{\partial^2 h}{\partial \widehat{M}_t^2}\widehat{M}_t^2 ||\lambda||^2\right] \mathrm{d}t - \frac{\partial h}{\partial \widehat{M}_t}\widehat{M}_t \lambda^\top \mathrm{d}Z_t. \tag{2.5.8}$$

Comparing the diffusion part of the dynamic budget constraint (2.5.2) with the diffusion part of equation (2.5.8) yields the dual optimal portfolio choice:

$$\widehat{\pi}_t^* = -\frac{\partial h}{\partial \widehat{M}_t} \widehat{M}_t \widehat{\lambda}^\top \widehat{\sigma}^{-1}. \tag{2.5.9}$$

The agent's optimal (primal) portfolio choice follows from Schroder and Skiadas (2002):

$$\pi_t^* = \widehat{\pi}_t^* + \beta A_t \widehat{\pi}_t^*. \tag{2.5.10}$$

The optimal dual portfolio choice $\widehat{\pi}_t^*$ can be further decomposed as follows:

$$\widehat{\pi}_t^* = \widehat{\pi}_t^{G*} + \widehat{\pi}_t^{L*}.$$

Here, $\widehat{\pi}_t^{G*}$ denotes the optimal dual portfolio choice that finances gains, and $\widehat{\pi}_t^{L*}$ is the optimal dual portfolio choice that finances losses. Theorem 2 below presents $\widehat{\pi}_t^{G*}$ and $\widehat{\pi}_t^{L*}$ for the case of a constant investment opportunity set. This theorem follows from application of equation (2.5.9). The optimal primal portfolio choice then follows from equation (2.5.10).

Theorem 2. *Consider an agent with the two-part power utility function (2.4.1) and reference level dynamics (2.4.2) who solves the consumption and portfolio choice problem (2.5.1) assuming a constant investment opportunity set. Let ϕ denote the standard normal probability density function. Then:*

- *If the agent is risk-averse in the loss domain, we find*

$$\begin{aligned}\widehat{\pi}_t^{G*} = \widehat{\lambda}^\top \widehat{\sigma}^{-1} \Bigg[&\frac{1}{1-\gamma_2} \widehat{W}_t^{G*} \\ &+ \left(k_t \widehat{M}_t\right)^{\frac{1}{\gamma_2 - 1}} \int_t^T \exp\left\{ - \int_t^s \Gamma_u \, \mathrm{d}u \right\} \frac{\phi\left[d_2\left(\xi_s\right)\right]}{||\lambda|| \sqrt{s-t}} \, \mathrm{d}s \Bigg],\end{aligned}$$

$$\widehat{\pi}_t^{L*} = \widehat{\lambda}^\top \widehat{\sigma}^{-1} \Bigg[\frac{1}{\gamma_1 - 1} \left(\widehat{W}_t^{L*} + \int_t^T \exp\left\{ - \int_t^s \widehat{r}_u \, \mathrm{d}u \right\} L_s \mathcal{N} \left[-d_1 \left(\zeta_s \vee \xi_s \right) \right] \mathrm{d}s \right)$$
$$+ \left(l_t \widehat{M}_t \right)^{\frac{1}{\gamma_1 - 1}} \int_t^T \exp\left\{ - \int_t^s \Pi_u \, \mathrm{d}u \right\} \frac{\phi \left[d_2 \left(\zeta_s \vee \xi_s \right) \right] - \phi \left[d_2 \left(\xi_s \right) \right]}{||\lambda|| \sqrt{s-t}} \mathrm{d}s$$
$$+ \int_t^T \exp\left\{ - \int_t^s \widehat{r}_u \, \mathrm{d}u \right\} L_s \frac{\phi \left[-d_1 \left(\zeta_s \vee \xi_s \right) \right]}{||\lambda|| \sqrt{s-t}} \mathrm{d}s \Bigg].$$

- *If the agent is risk-seeking in the loss domain, we find*

$$\widehat{\pi}_t^{G*} = \widehat{\lambda}^\top \widehat{\sigma}^{-1} \Bigg[\frac{1}{1 - \gamma_2} \widehat{W}_t^{G*}$$
$$+ \left(k_t \widehat{M}_t \right)^{\frac{1}{\gamma_2 - 1}} \int_t^T \exp\left\{ - \int_t^s \Gamma_u \, \mathrm{d}u \right\} \frac{\phi \left[d_2 \left(\xi_s \right) \right]}{||\lambda|| \sqrt{s-t}} \mathrm{d}s \Bigg],$$

$$\widehat{\pi}_t^{L*} = \widehat{\lambda}^\top \widehat{\sigma}^{-1} \int_t^T \exp\left\{ - \int_t^s \widehat{r}_u \, \mathrm{d}u \right\} L_s \frac{\phi \left[-d_1 \left(\xi_s \right) \right]}{||\lambda|| \sqrt{s-t}} \mathrm{d}s.$$

Theorem 2 reveals that in good economic scenarios, the optimal dual portfolio strategy $\widehat{\pi}_t^*$ can be approximated by $\widehat{\lambda}^\top \widehat{\sigma}^{-1} / \left(1 - \gamma_2\right) \widehat{W}_t^*$. In these economic scenarios, the agent behaves like a CRRA agent and invests a *constant* proportion of surplus wealth in risk-bearing assets.

2.6. Analysis of the Solution

With the analytical solutions and comparative statics to the general consumption and portfolio choice problem provided in Section 2.5 (and the Appendix), we proceed in this section to their numerical analysis. Section 2.6.1 introduces the underlying assumptions and discusses the key parameter values used in the numerical analysis. Section 2.6.2 illustrates the agent's optimal consumption and portfolio choice. Finally, Section 2.6.3 conducts a welfare analysis.

2.6.1. Assumptions and Key Parameter Values

We allow the agent to invest his wealth in a risk-free asset and a single risky stock. The investment opportunity set is assumed to be constant (i.e., $r_t = r$, $\sigma_t = \sigma$ and $\lambda_t = \lambda$). The equity premium $\sigma\lambda = \mu - r$ is set at 4%. The risk-free rate r is set at 1%, and the volatility of innovations to the risky stock price σ is set at 20%. These estimates coincide with the estimates reported by Gomes et al. (2008).

The terminal time T equals 20 years. We view T as the total number of years of retirement. Initial wealth W_0 can be viewed as total pension wealth at the age of retirement.[22] For the ease of illustration, we assume that the agent retires at 65.

The loss aversion index κ is set equal to 2.5. The estimates of the median loss aversion index reported in the literature vary from 1 to 5 (see, e.g., Abdellaoui et al., 2008). The degree of loss aversion largely differs among individuals, and typically depends on the model. In the welfare analysis, we consider, among other things, the impact of a change in the loss aversion index κ on the agent's welfare. Finally, the subjective rate of time preference δ is set equal to 1%.

2.6.2. The Optimal Consumption and Portfolio Choice

2.6.2.1. Loss Aversion Only

This section illustrates the optimal consumption and portfolio choice of a loss averse agent without endogenous updating of the agent's reference level (i.e., the endogeneity parameter β is set equal to zero). In addition, we assume that the agent's reference level is constant (i.e., the depreciation parameter α is also set equal to zero). Inspired by Barberis, Huang, and Santos (2001), the agent's constant reference level $\theta_t = \theta$ is assumed to be equal to the level of consumption that would be obtained if the agent's initial wealth W_0 was kept in the money market account for the entire retirement phase.[23] The assumption here is that the agent is likely to be disappointed if consumption is less than the payment he would receive from a fixed annuity. The agent's constant reference level θ solves the following equation:

$$W_0 = \theta \int_0^T \exp\{-rt\}\,\mathrm{d}t \equiv \theta A_0. \tag{2.6.1}$$

Simple algebra yields $\theta = 5.5\% \cdot W_0$; that is, the annuity factor $A_0 \equiv \int_0^T \exp\{-rt\}\,\mathrm{d}t$ is equal to $1/5.5\% \approx 18 < T = 20$. Equation (2.6.1) implies that initial surplus wealth

[22] In the analysis, W_0 equals 500 ($\times$1,000 dollars) units, and we report our results *relative* to W_0.

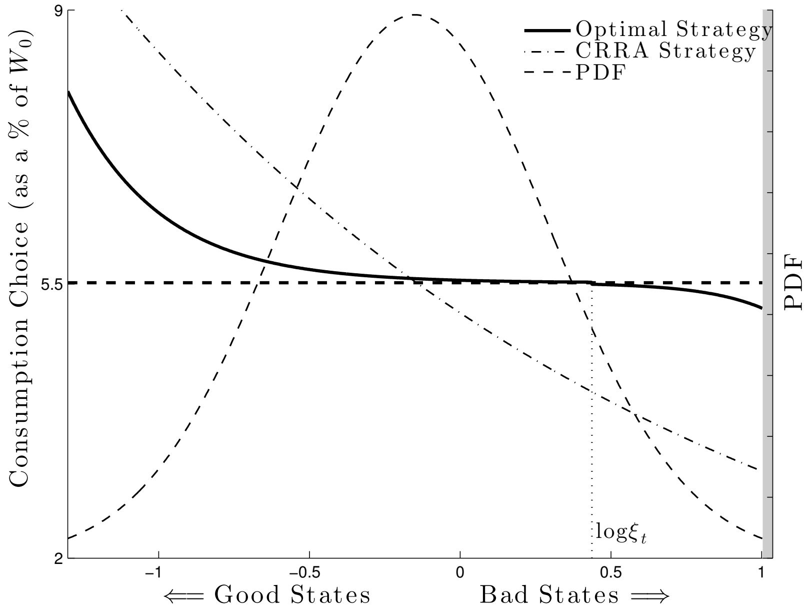
[23] Barberis et al. (2001) argue that the risk-free interest rate serves as a natural benchmark for evaluating gains and losses. In our context, this assumption implies that the agent is likely to be disappointed if consumption is less than the payment he would receive from a fixed annuity.

$\widehat{W}_0 \equiv \widehat{W}_0^G - \widehat{W}_0^L$ is equal to zero. This assertion follows from equations (2.5.6) and (2.5.7) with $\alpha = \beta = 0\%$. Put differently, initial wealth required to finance future gains $\widehat{W}_0^G$ is equal to initial wealth required to finance future losses $\widehat{W}_0^L$. We note that $\widehat{W}_0^G$ and $\widehat{W}_0^L$ are not equal to zero unless the agent is infinitely loss averse.

Figure 2.3 illustrates the optimal consumption choice (expressed as a percentage of the agent's initial wealth W_0) at age 70 (i.e., $t = 5$) as a function of the then-current log state price density for the case of risk-averse behavior in the loss domain. Here, consumption is constrained to be non-negative; that is, $L = \theta$. Under the optimal choice, the loss averse agent seeks protection against consumption losses due to financial shocks, thus inducing a (soft) guarantee on consumption. The agent optimally desires to maintain consumption above the reference level, but under really adverse circumstances this (soft) guarantee on consumption cannot be maintained. As a direct consequence, we can divide the states of the economy into two categories: good to intermediate states (i.e., $\log M_t \leq \log \xi_t$) and bad states (i.e., $\log M_t > \log \xi_t$). In good to intermediate states, optimal consumption is guaranteed to be larger than the reference level, while in bad states, optimal consumption is smaller than the reference level. The dotted line shows the probability density function (PDF) of the then-current log state price density conditional upon information available at the age of retirement. The probability of consumption being smaller than the reference level can be controlled by choosing appropriate values for the preference parameters. We observe that the optimal consumption profile displays a 90° rotated S-shaped pattern with a discontinuity at the point $\log M_t = \log \xi_t$. Hence, optimal consumption is never equal to the reference level. The dash-dotted line illustrates the consumption choice of an agent with CRRA utility. The relative risk aversion coefficient γ is set equal to two. The (log) consumption choice of a CRRA agent varies linearly with the (log) state price density. As a consequence, for typical values of the relative risk aversion coefficient γ, a CRRA agent incurs more frequently a loss than a loss averse agent (where we define gains and losses relative to the reference level).

Next, Figure 2.4 displays the optimal consumption choice (expressed as a percentage of the agent's initial wealth W_0) at age 70 as a function of the then-current log state price density for the case of risk-seeking behavior in the loss domain. Here, consumption is allowed to fall 2% point below the (normalized) reference level θ/W_0. We observe again that, because of loss aversion, the agent has a strong preference to maintain consumption

Figure 2.3.
Consumption profile for the case of risk-averse behavior in the loss domain

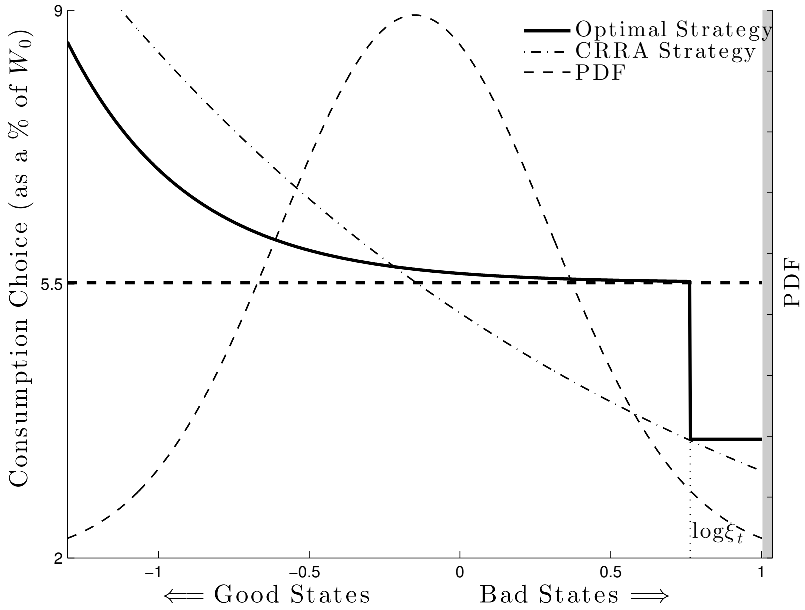


The figure shows the optimal consumption choice (expressed as a percentage of the agent's initial wealth W_0) at age 70 as a function of the then-current log state price density. The curvature parameter γ_1 (γ_2) is set equal to 1.2 (0.7). Consumption is constrained to be non-negative by taking $L = \theta$. The dashed line corresponds to the agent's reference level (expressed as a percentage of W_0). The dotted line shows the probability density function (PDF) of the then-current log state price density conditional upon information available at the age of retirement. The dash-dotted line illustrates the consumption choice (expressed as a percentage of W_0) of an agent with CRRA utility. The relative risk aversion coefficient γ is set equal to two.

above the reference level. As in the case of risk-averse behavior in the loss domain, we can divide the states of the economy into two categories: good to intermediate states (i.e., $\log M_t \leq \log \xi_t$) and bad states (i.e., $\log M_t > \log \xi_t$). In good to intermediate states, optimal consumption is guaranteed to be larger than the reference level, while in bad states, optimal consumption is equal to the minimum consumption level $\theta - L$. We also observe that at the threshold $\log M_t = \log \xi_t$, optimal consumption jumps to the lower bound $\theta - L$. This behavior can be explained by the fact that the agent is risk-seeking in the loss domain.

Figure 2.5 shows the optimal portfolio choice (i.e., the total dollar amount invested in the risky stock) at age 70 as a function of the then-current log state price density for the case of risk-averse behavior in the loss domain. The optimal portfolio choice

Figure 2.4.
Consumption profile for the case of risk-seeking behavior in the loss domain



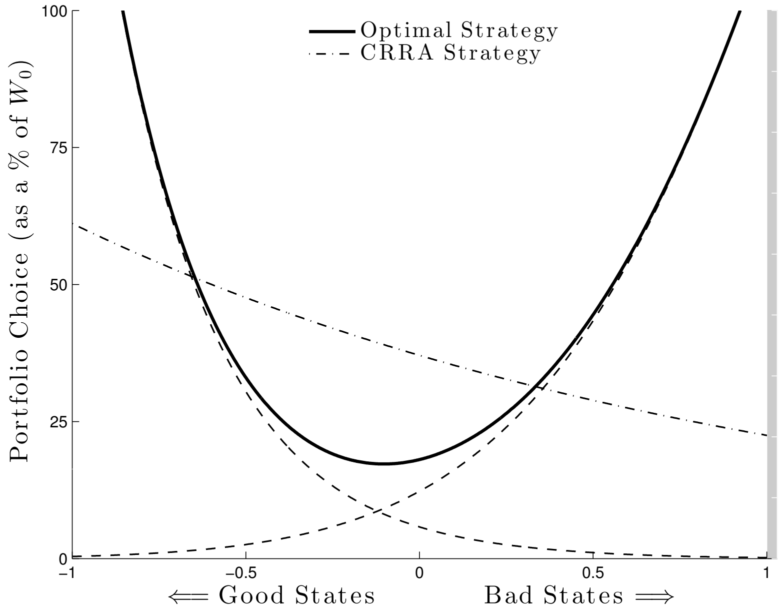
The figure shows the optimal consumption choice (expressed as a percentage of the agent's initial wealth W_0) at age 70 as a function of the then-current log state price density. The curvature parameter γ_1 (γ_2) is set equal to 0.8 (0.6). Consumption is allowed to fall 2% point below the (normalized) reference level θ/W_0. The dashed line corresponds to the agent's reference level (expressed as a percentage of W_0). The dotted line shows the probability density function (PDF) of the then-current log state price density conditional upon information available at the age of retirement. The dash-dotted line illustrates the consumption choice (expressed as a percentage of W_0) of an agent with CRRA utility. The relative risk aversion coefficient γ is set equal to two.

is expressed as a percentage of the agent's initial wealth W_0. We observe that the optimal portfolio profile displays a U-shaped pattern: the total dollar amount invested in the risky stock will be lower in intermediate economic scenarios than in good or bad economic scenarios. When the (non-log) state price density tends to zero, the fraction of surplus wealth $\widehat{W}_t^*$ invested in the risky stock converges to the constant $\lambda/\left[\sigma\left(1-\gamma_2\right)\right]$. Hence, in good economic scenarios, the optimal portfolio choice behaves in a similar fashion as the portfolio choice of a CRRA agent.[24] We note that $W_t^* - \widehat{W}_t^* = A_t\theta$ is fully invested in the money market account. When the state price density is relatively high, the fraction of surplus wealth invested in the risky stock can be approximated by

[24] This is not directly visible in Figure 2.5, where the portfolio choice of the CRRA agent does not match the portfolio choice of the loss averse agent in good (or bad) states, because the relative risk aversion coefficient γ (CRRA agent) differs from its counterpart $1-\gamma_i$ (loss averse agent) specified by the curvature parameters γ_i, $i = 1, 2$, and because total wealth differs from surplus wealth.

the constant $\lambda / [\sigma(1-\gamma_1)] < 0$.[25] Not only in good but also in bad economic scenarios, a loss averse agent behaves like a CRRA agent. In intermediate economic scenarios, the total dollar amount invested in the risky stock is relatively small.

Figure 2.5.
Portfolio profile for the case of risk-averse behavior in the loss domain



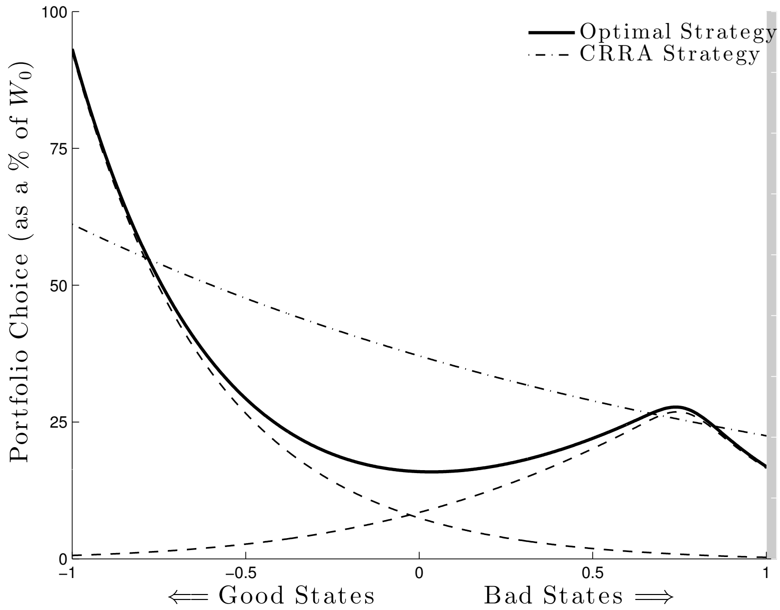
The figure shows the optimal portfolio choice (i.e., the total dollar amount invested in the risky stock) at age 70 as a function of the then-current log state price density. The portfolio choice is expressed as a percentage of the agent's initial wealth W_0. The curvature parameter γ_1 (γ_2) is set equal to 1.2 (0.7). Consumption is constrained to be non-negative; that is, $L = \theta$. The dash-dotted line illustrates the portfolio choice (expressed as a percentage of W_0) of an agent with CRRA utility. The relative risk aversion coefficient γ is set equal to two. The increasing dotted line represents $\widehat{\pi}_t^{L*}/W_0$ while the decreasing dotted line corresponds to $\widehat{\pi}_t^{G*}/W_0$.

Figure 2.6 shows the optimal portfolio choice (i.e., the total dollar amount invested in the risky stock) at age 70 as a function of the then-current log state price density for the case of risk-seeking behavior in the loss domain. The portfolio choice is expressed as a percentage of the agent's initial wealth W_0. As in the case of risk-averse behavior in the loss domain, the optimal portfolio profile displays (primarily) a U-shaped pattern. When the state price density tends to zero, the fraction of surplus wealth invested in the risky stock converges to the constant $\lambda / [\sigma(1-\gamma_2)]$. Hence, in good economic scenarios, the portfolio choice of a loss averse agent behaves in a similar fashion as the portfolio

[25] We note that in bad states (i.e., high state prices), surplus wealth is negative.

choice of a CRRA agent. When the state price density tends to infinity, the fraction of surplus wealth invested in the risky stock ultimately converges to zero. Indeed, in bad economic scenarios, the minimum consumption level $\theta - L$ must be guaranteed.

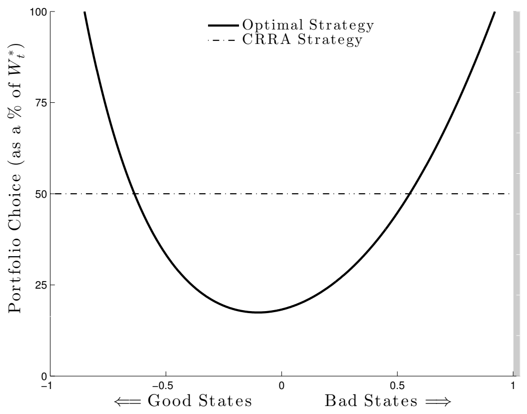
Figure 2.6.
Portfolio profile for the case of risk-seeking behavior in the loss domain



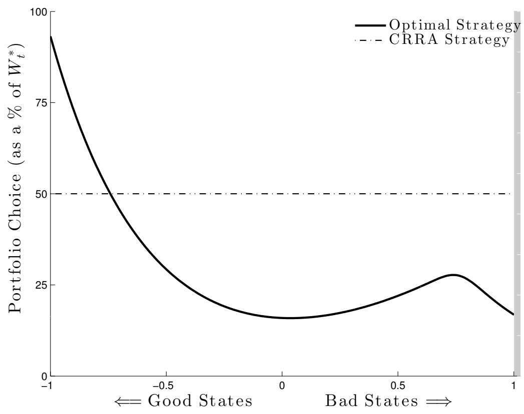
The figure shows the optimal portfolio choice (i.e., the total dollar amount invested in the risky stock) at age 70 as a function of the then-current log state price density. The portfolio choice is expressed as a percentage of the agent's initial wealth W_0. The curvature parameter γ_1 (γ_2) is set equal to 0.8 (0.6). Consumption is allowed to fall 2% point below the (normalized) reference level θ/W_0. The dash-dotted line illustrates the portfolio choice (expressed as a percentage of W_0) of an agent with CRRA utility. The relative risk aversion coefficient γ is set equal to two. The increasing dotted line represents $\widehat{\pi}_t^{L*}/W_0$ while the decreasing dotted line corresponds to $\widehat{\pi}_t^{G*}/W_0$.

Figure 2.7 shows the optimal portfolio choice measured as a fraction of *total wealth* invested in the risky stock at age 70 as a function of the then-current log state price density. We recall that Figures 2.5 and 2.6 show the optimal portfolio choice measured as a fraction of *initial wealth* invested in the risky stock. We observe that the optimal portfolio profile still displays (primarily) a U-shaped pattern. The portfolio choice of a CRRA agent is no longer a decreasing line but a straight line: a CRRA agent always invests a constant fraction $\lambda/(\sigma\gamma)$ of total wealth in the risky stock, irrespective of the state of the economy.

Figure 2.7.
Fraction of total wealth invested in the risky stock



(a) $\gamma_1 = 1.2$ and $\gamma_2 = 0.7$



(b) $\gamma_1 = 0.8$ and $\gamma_2 = 0.6$

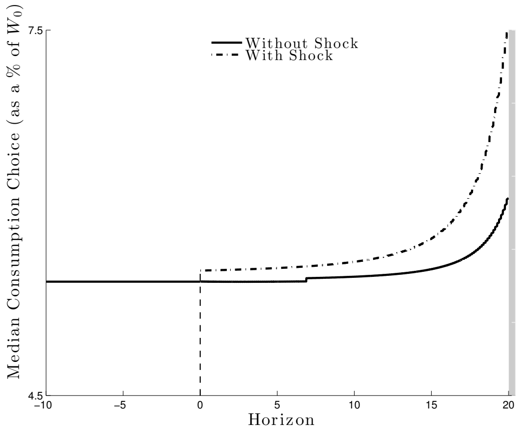
The figure shows the optimal portfolio choice measured as a fraction of total wealth invested in the risky stock at age 70 as a function of the then-current log state price density. Panel (a) displays the case of risk-averse behavior in the loss domain (taking, as before, $\gamma_1 = 1.2$ and $\gamma_2 = 0.7$), while panel (b) displays the case of risk-seeking behavior in the loss domain (taking, as before, $\gamma_1 = 0.8$ and $\gamma_2 = 0.6$). The dash-dotted line illustrates the portfolio choice of a CRRA agent. The relative risk aversion coefficient γ is set equal to two.

2.6.2.2. Loss Aversion and Endogenous Updating

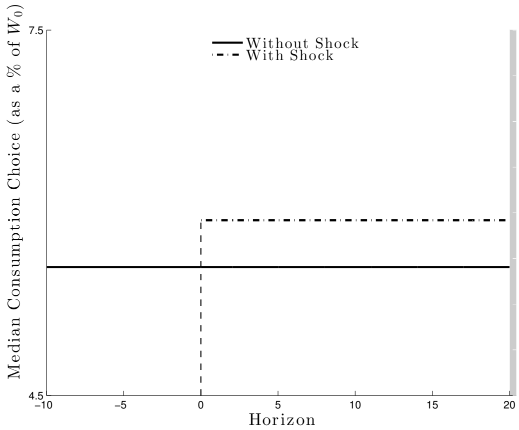
This section considers the case where the loss averse agent endogenously updates his reference level over time. We assume that the endogeneity parameter β as well as the depreciation parameter α are equal to 20%. Also, we assume that the initial reference level θ_0 equals 5.5% of initial wealth W_0, and L_t equals the initial reference level (i.e., $L_t = L = \theta_0$). These parameter values imply that initial surplus wealth $\widehat{W_0}$ equals zero.

Figure 2.8 illustrates the impact of a *positive* shock in initial wealth on median consumption for the case of risk-averse behavior in the loss domain. The left panel applies to the case in which the loss averse agent endogenously updates his reference level over time (i.e., $\alpha = \beta = 20\%$), while the right panel displays the case of no endogenous updating (i.e., $\alpha = \beta = 0\%$). The dash-dotted lines in both panels represent the agent's median consumption choice *with* the shock in initial wealth. We observe that with endogenous updating a financial shock is gradually absorbed into future consumption (i.e., consumption adjusts sluggishly to financial shocks): the impact of a financial shock on consumption is smoothed over time, having a larger impact in the distant future than in the near future. By contrast, in the case of no endogenous updating, a financial shock is directly absorbed into future consumption, leading to an even distribution of the shock's impact on future consumption choice.

Figure 2.8.
Gradual adjustment to financial shocks



(a) With endogenous updating

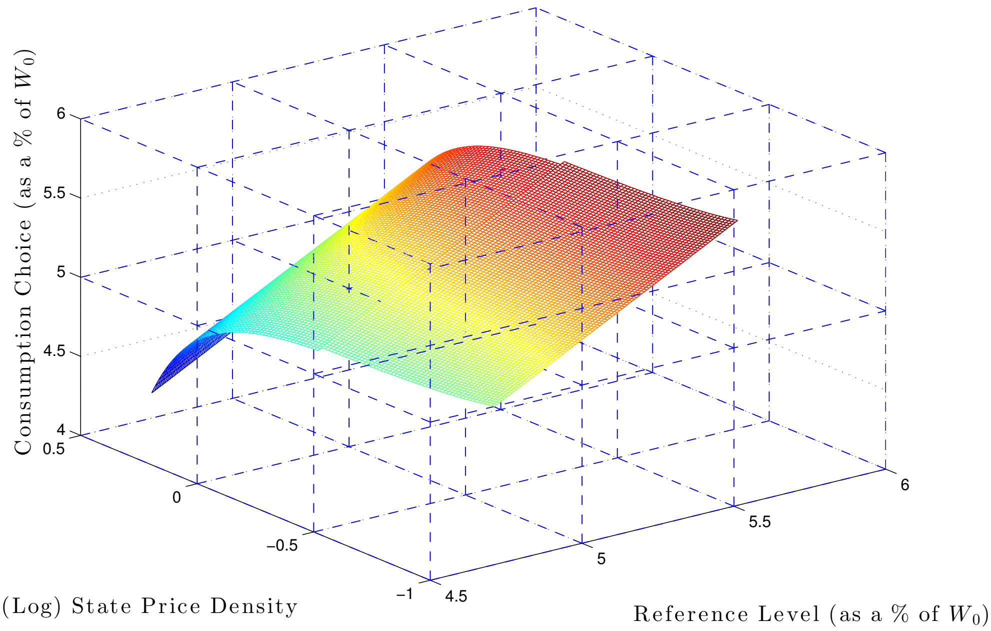


(b) No endogenous updating

The figure illustrates the impact of a positive shock in initial wealth on median consumption (expressed as a percentage of the agent's initial wealth W_0) for the case of risk-averse behavior in the loss domain (i.e., $\gamma_1 = 1.2$). The curvature parameter γ_2 is set equal to 0.7 as before, and L to θ_0. The right panel displays the case of no endogenous updating (i.e., $\alpha = \beta = 0\%$), while the left panel presents the case in which the agent endogenously updates the reference level over time (i.e., $\alpha = \beta = 20\%$). The dash-dotted lines in both panels represent the agent's median consumption choice with a shock in initial wealth from 500 to 750 ($\times$1,000 dollars) units.

Figure 2.9 shows the optimal consumption choice (expressed as a percentage of the agent's initial wealth W_0) at age 70 as a function of the then-current log state price density and the then-current reference level. Indeed, we note that the optimal consumption profile depends not only the then-current state price density but also on the then-current reference level (i.e., the optimal consumption profile is path-dependent). The threshold ξ_t is however state independent. The agent is assumed to be risk-averse in the loss domain. The figure shows that the optimal consumption choice increases with the reference level, and decreases with the state price density. Compared to the case of loss aversion only as in the previous subsection, endogeneity of the reference level has the *reinforcing* effect that the agent gives up even more upward potential in then-current consumption to guarantee consumption above the reference level. At the same time, the agent is also willing to accept somewhat larger consumption losses if the state of the economy is really adverse.

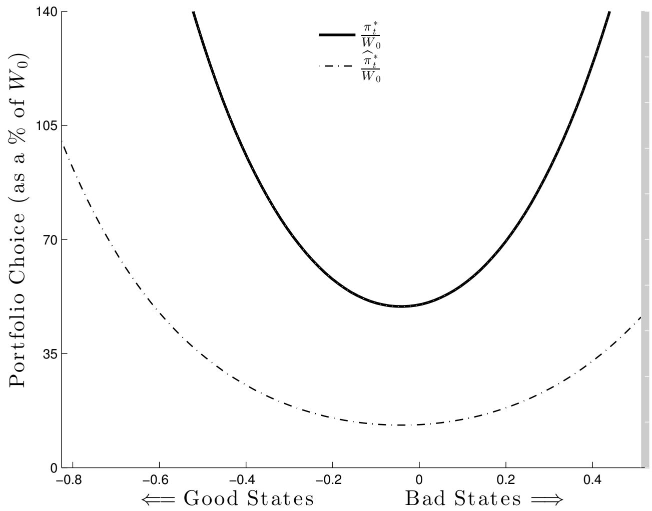
Figure 2.9.
Consumption profile for the case of risk-averse behavior in the loss domain



The figure illustrates the optimal consumption choice (expressed as a percentage of the agent's initial wealth W_0) at age 70 as a function of the then-current log state price density and the then-current reference level. The curvature parameter γ_1 (γ_2) is set equal to 1.2 (0.7).

Figure 2.10 illustrates the optimal portfolio choice (i.e., the total dollar amount invested in the risky stock expressed as a percentage of the agent's initial wealth W_0) at age 70 as a function of the then-current log state price density for the case of risk-averse behavior in the loss domain. As in the case of loss aversion only, the optimal portfolio profile is U-shaped. While the then-current reference level affects the optimal consumption profile (see Figure 2.9), it does not impact the optimal portfolio profile. However, because of endogenous updating, optimal required wealth $\widetilde{W}_t^*$ (i.e., wealth required to finance future optimal reference levels) is partly invested in the risky stock. Put differently, the dual portfolio choice no longer coincides with the agent's optimal (primal) portfolio choice. By contrast, in the case of no endogenous updating as in the previous subsection, optimal required wealth $\widetilde{W}_t^*$ is fully invested in the money market account. Since the reference level depends on the agent's own past consumption choices (i.e., the reference level is path-dependent), the agent typically invests more in the risky stock under endogenous updating.

Figure 2.10.
Portfolio profile for the case of risk-averse behavior in the loss domain



The figure illustrates the optimal portfolio choice (i.e., the total dollar amount invested in the risky stock) at age 70 as a function of the then-current log state price density. The curvature parameter γ_1 (γ_2) is set equal to 1.2 (0.7). The dash-dotted line represents the optimal dual portfolio choice $\widehat{\pi}_t^*/W_0$.

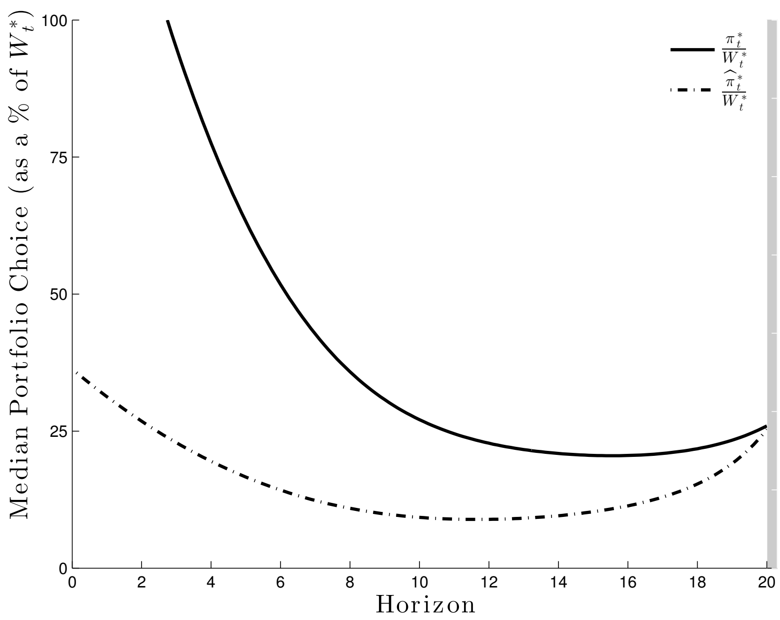
Figure 2.11 illustrates the median optimal portfolio choice measured as a fraction of total wealth invested in the risky stock as a function of the horizon, which represents the number of years spent in retirement. We observe that the agent implements a life cycle investment strategy (i.e., the fraction of wealth invested in the risky stock, on average, decreases as the agent ages). Indeed, since the agent has less time to absorb financial shocks as he grows older, the equity risk exposure, on average, decreases over the life cycle.[26]

2.6.3. Welfare Analysis

This section conducts a welfare analysis. Section 2.6.3.1 reports the welfare losses (in terms of the relative decline in certainty equivalent consumption) associated with

[26] The slight increase in median optimal portfolio choice towards the end of the life span can be explained from the fact that the median optimal dual portfolio choice, which dictates the median optimal (primal) portfolio choice, displays a U-shaped pattern as a function of the horizon. This, in turn, is due to the fact that the absolute difference between optimal median consumption and the reference level as a function of the horizon is U-shaped, being smaller for intermediate horizons than for large and small horizons.

Figure 2.11.
Median portfolio choice



The figure illustrates the median optimal portfolio choice measured as a fraction of total wealth invested in the risky stock as a function of the horizon. The curvature parameter γ_1 (γ_2) is set equal to 1.2 (0.7). The dash-dotted line represents the optimal dual portfolio choice $\widehat{\pi}_t^*/W_0$.

incorrect values of the agent's preference parameters.[27] Precisely, we compute the welfare losses due to implementing suboptimal consumption and portfolio strategies derived by solving the agent's maximization problem on the basis of wrong values of the loss aversion index κ, the depreciation parameter α and the endogeneity parameter β. Section 2.6.3.2 reports the welfare losses associated with implementing alternative (simpler) consumption and portfolio strategies. Throughout the welfare analysis, we assume that the agent's *optimal* consumption and portfolio choice is characterized by the following ("true") values of the preference parameters: $\theta_0 = 5.5\% \cdot W_0$, $\kappa = 2.5$, $\alpha = \beta = 20\%$, $\gamma_1 = 1.2$ and $\gamma_2 = 0.7$. Thus, the agent is risk-averse in the loss domain. The welfare losses are computed *relative* to the agent's optimal consumption and portfolio strategy. The Appendix outlines the numerical procedure employed to compute the welfare losses. This procedure is non-standard due to the endogeneity of the reference level. The numerical procedure is implemented with $\Delta t = 1/8$ and

[27] We define the certainty equivalent of an uncertain consumption strategy to be the constant, certain consumption level that yields indifference to the uncertain consumption strategy.

$\mathcal{S} = 1,000,000$. Here, Δt denotes the time step and $\mathcal{S}$ represents the total number of simulations.

2.6.3.1. Welfare Losses Due to Incorrect Parameter Values

Tables 2.2 and 2.3 report welfare losses due to implementing suboptimal consumption and portfolio strategies derived on the basis of wrong values of the loss aversion index κ, the depreciation parameter α and the endogeneity parameter β. In Table 2.2 we assume that the agent's initial surplus wealth is equal to zero, while in Table 2.3 we assume that the agent has positive initial surplus wealth.[28] Table 2.2 shows that the welfare losses associated with incorrectly assuming a constant reference level (i.e., $\alpha = \beta = 0\%$) are substantial. Specifically, the welfare loss is about 30%. If the agent has positive initial surplus wealth, as in Table 2.3, this welfare loss is even larger. More generally, the tables reveal that consumption and portfolio strategies based on a constant exogenous reference level or on a very limited degree of endogeneity, thus implying no or only very limited smoothing of financial shocks, substantially reduce welfare. At the same time we observe that the impact of a change in the loss aversion index κ is larger when the agent's initial surplus wealth is equal to zero than when the agent's initial surplus wealth is positive. Indeed, κ determines the multiplicity of states in which consumption falls below the reference level. As a consequence, the impact of a change in κ is more pronounced when initial surplus wealth is small.

Table 2.2.
Welfare losses due to incorrect parameter values (zero initial surplus wealth)

Loss aversion index (κ)	Endogeneity parameter (β)			
	0	0.05	0.10	0.20
2.5	31.93	17.50	8.25	0
5	28.58	22.56	18.33	11.85
10	28.00	25.99	24.87	22.40

The table reports the welfare losses (in terms of the relative decline in certainty equivalent consumption) due to implementing suboptimal consumption and portfolio strategies derived on the basis of wrong values of the loss aversion index κ, the depreciation parameter α, and the endogeneity parameter β. The depreciation parameter α always equals the endogeneity parameter β. The agent has zero initial surplus wealth. The numbers represent a percentage.

[28] More specifically, Table 2.2 assumes that the agent's initial wealth W_0 equals 500 (×1,000 dollars) units, while Table 2.3 assumes that W_0 is equal to 750 (×1,000 dollars) units.

Table 2.3.
Welfare losses due to incorrect parameter values (positive initial surplus wealth)

Loss Aversion Index (κ)	Endogeneity Parameter (β)			
	0	0.05	0.10	0.20
2.5	89.33	13.32	0.93	0
5	88.97	13.22	1.14	0.42
10	88.86	13.20	1.17	0.51

The table reports the welfare losses (in terms of the relative decline in certainty equivalent consumption) due to implementing suboptimal consumption and portfolio strategies derived on the basis of wrong values of the loss aversion index κ, the depreciation parameter α, and the endogeneity parameter β. The depreciation parameter α always equals the endogeneity parameter β. The agent has positive initial surplus wealth. The numbers represent a percentage.

2.6.3.2. Welfare Losses Due to Alternative Strategies

Table 2.4 reports the welfare losses, compared to the optimal strategies of a loss averse agent who endogenously updates his reference level, due to implementing the consumption and portfolio strategy of an agent with CRRA utility (i.e., the Merton strategy). The welfare losses are reported for various values of the coefficient of relative risk aversion γ underlying the Merton strategy. The implementation of the Merton strategy, under which log consumption varies linearly with the log state price density and financial shocks are directly absorbed into future consumption, leads to substantial welfare losses of about 40%. The welfare losses are minimal for intermediate values of γ ($\gamma = 5$ in the table). We note that $\gamma = \infty$ corresponds to a risk-free strategy.

Table 2.4.
Welfare losses due to implementing the Merton strategy

Relative Risk Aversion Coefficient (γ)				
1	2	5	10	∞
44.11	37.47	37.39	38.87	40.11

The table reports the welfare losses (in terms of the relative decline in certainty equivalent consumption) due to implementing the consumption and portfolio strategy of an agent with CRRA utility (i.e., the Merton strategy). The table reports the welfare losses for various values of the coefficient of relative risk aversion γ underlying the Merton strategy. The agent has zero initial surplus wealth. The numbers represent a percentage.

Finally, we consider the following practical consumption and portfolio strategy: we assume that the agent consumes a fraction $1/(T - t)$ of wealth W_t. Furthermore, we assume that a constant fraction of wealth is invested in the risky stock (i.e., we assume π_t/W_t to be constant), as under the Merton strategy. Table 2.5 reports the welfare losses for various values of the fraction of wealth invested in the risky stock. We observe that the welfare losses are again substantial, but smaller than when implementing the Merton consumption rule. Indeed, our numerical results reveal that the Merton strategy generates a more volatile consumption profile, with consumption falling below the reference level more often than when implementing the $1/(T-t)$ consumption rule. Thus, from the perspective of a loss averse agent, who strongly prefers to maintain consumption above the reference level, the $1/(T-t)$ consumption rule is less suboptimal than the Merton consumption rule. Furthermore, the welfare losses in Table 2.5 are relatively insensitive to changes in π_t/W_t. The welfare losses are minimal for relatively low fractions of wealth invested in the risky stock ($\pi_t/W_t = 10\%$ in the table). We also computed, under the $1/(T-t)$ consumption rule, the welfare losses associated with implementing various state-independent life cycle investment strategies. We find that the welfare losses do not substantially reduce when implementing a state-independent life cycle investment strategy.

Table 2.5.
Welfare losses due to implementing a practical alternative consumption and portfolio strategy

Fraction of Wealth Invested in the Risky Stock				
0	0.10	0.20	0.30	0.40
30.03	28.71	28.73	29.64	30.90

The table reports the welfare losses (in terms of the relative decline in certainty equivalent consumption) due to implementing a practical alternative consumption and portfolio strategy. The table reports the welfare losses for various values of the fraction of wealth invested in the risky stock (i.e., π_t/W_t). The agent has zero initial surplus wealth. The numbers represent a percentage.

2.7. An Alternative Utility Function

This section explores, as a robustness check, the agent's optimal consumption and portfolio choice under an alternative specification of the agent's instantaneous utility function. More specifically, we assume that the agent's utility function is represented by the kinked HARA utility function. The kinked HARA utility function emerges as a special case of (2.4.3) if (i) classical consumption utility m is represented by the HARA utility function and (ii) the gain-loss utility function w equals the two-part power utility function v with $\gamma_1 = \gamma_2 = 1$. The HARA classical consumption utility function is defined as follows:[29]

$$m(c_t) = \frac{\varphi}{1-\varphi}\left(\frac{\rho}{\varphi}c_t + \psi\right)^{1-\varphi}.$$

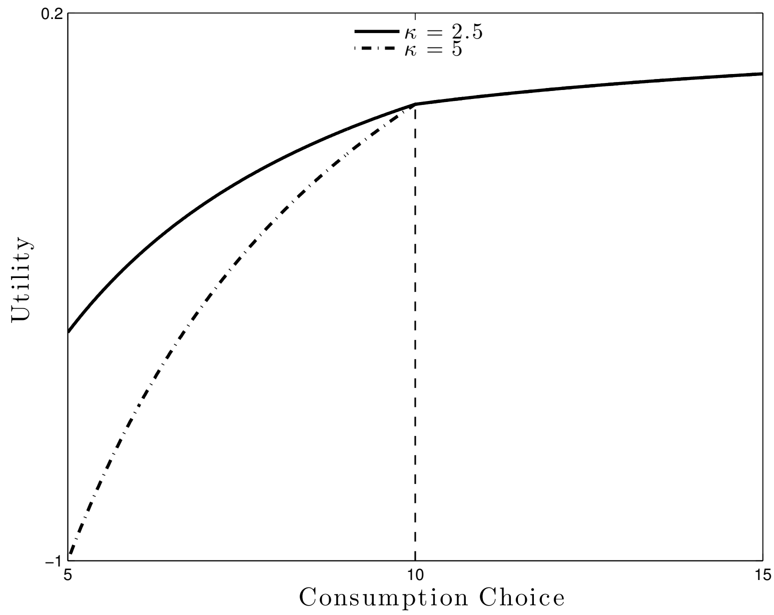
Here, $\varphi \in (0,\infty)\backslash\{1\}$, $\rho > 0$ and $\psi \geq 0$ are preference parameters.

Figure 2.12 illustrates the kinked CRRA utility function, which appears as a special case when $\rho = \varphi$ and $\psi = 0$, for $\kappa = 2.5$ and $\kappa = 5$. The figure shows that the kinked CRRA utility function has a kink at the reference level, with the slope of the utility function over losses being steeper than the slope of the utility function over gains. Furthermore, we observe that the kinked CRRA utility function is concave everywhere. Hence, the agent exhibits risk averse behavior in both the gain and the loss domain. Unfortunately, the kinked HARA utility function cannot be expressed in terms of the agent's surplus consumption choice $\widehat{c}_t \equiv c_t - \theta_t$. As a direct consequence, the solution technique of Schroder and Skiadas (2002) is not applicable here. However, we can still obtain an analytical solution to the optimal consumption and portfolio choice problem if the agent's reference level is exogenously given. The assumption of an exogenous reference level implies that the agent's *own* (past) consumption choices do not affect the reference level. However, factors beyond the control of the agent are allowed to influence the reference level. Hence, the consumption and portfolio choice model considered in this section can be viewed as an external, rather than an internal, habit formation model (see, e.g., Abel, 1990). In what follows, the reader should keep in mind that the reference level is independent of the agent's own (past) consumption choices.

Theorem 3 below presents the optimal consumption choice for an agent with the kinked HARA utility function.

[29] The HARA class of utility functions contains several important special cases. With suitable choice of preference parameters, the HARA utility function can exhibit increasing, decreasing or constant relative risk aversion. Important special cases are the commonly used CRRA ($\rho = \varphi$ and $\psi = 0$), exponential ($\psi = 1$ and $\varphi \to \infty$) and logarithmic ($\rho = 1$ and $\varphi \to 1$) utility functions.

Figure 2.12.
The kinked CRRA utility function



The figure illustrates the kinked CRRA utility function (i.e., $\rho = \varphi$ and $\psi = 0$) for $\kappa = 2.5$ (solid line) and $\kappa = 5$ (dash-dotted line). The reference level θ_t is set equal to 10, the weight parameter η to 0 and the curvature parameter φ to 5.

Theorem 3. *Consider an agent with the kinked HARA utility function and an exogenously given reference level process θ who solves the consumption and portfolio choice problem, with consumption constrained to be non-negative. Then the optimal consumption c_t^* at time $t \in [0, T]$ is given by*

$$c_t^* = \begin{cases} \frac{\varphi}{\rho}\left(\frac{y\exp\{\delta t\}M_t}{\rho}\right)^{-\frac{1}{\varphi}} - \frac{\psi\varphi}{\rho}, & \text{if } M_t < \underline{\xi}_t; \\ \theta_t, & \text{if } \underline{\xi}_t \leq M_t \leq \overline{\xi}_t; \\ \left[\frac{\varphi}{\rho}\left(\frac{y\exp\{\delta t\}M_t}{\rho\bar{\kappa}}\right)^{-\frac{1}{\varphi}} - \frac{\psi\varphi}{\rho}\right] \vee 0, & \text{if } M_t > \overline{\xi}_t. \end{cases}$$

Here, $\bar{\kappa} \equiv \eta + (1-\eta)\kappa$ stands for the adjusted loss aversion index. The thresholds $\underline{\xi}_t$ and $\overline{\xi}_t$ are defined as follows:

$$\underline{\xi}_t = \frac{\rho}{y}\exp\{-\delta t\}\left(\frac{\xi}{\varphi}\theta_t + \psi\right)^{-\varphi}, \qquad \overline{\xi}_t = \frac{\rho\bar{\kappa}}{y}\exp\{-\delta t\}\left(\frac{\xi}{\varphi}\theta_t + \psi\right)^{-\varphi}.$$

The Lagrange multiplier y is chosen such that the static budget constraint holds with equality.

Theorem 3 shows that the state price density can be divided into three regions. In good scenarios (i.e., low state prices), consumption is (strictly) larger than the reference level; in these scenarios, the agent can afford to consume above the reference level. Next, in intermediate economic scenarios (i.e., intermediate state prices), consumption is equal to the reference level. The adjusted loss aversion index $\bar{\kappa}$ determines the multiplicity of states in which consumption is equal to the reference level. Finally, in bad economic scenarios (i.e., high state prices), the agent's wealth is insufficient to finance consumption at the reference level. In the case of two-part power utility (see Section 2.5), similarly, the optimal consumption choice also falls below the reference level in bad states of the world. Figure 2.13 illustrates the optimal consumption profile of an agent with kinked CRRA utility. We observe that, as before, the optimal consumption choice as a function of the log state price density is 90° rotated S-shaped, thus confirming the impact of loss aversion on the optimal consumption profile. We also observe that c_t^* is a continuous function of the state price density. In particular, the optimal consumption profile does not exhibit a jump at the reference level. Indeed, marginal utility at the reference level is finite.

The agent's optimal wealth W_t^* can be decomposed in the same way as in Section 2.5:

$$W_t^* = \widehat{W}_t^* + \widetilde{W}_t^* = \widehat{W}_t^{G*} + \widehat{W}_t^{L*} + \widetilde{W}_t^*$$

Proposition 3 presents $\widehat{W}_t^{G*}$, $\widehat{W}_t^{L*}$ and $\widetilde{W}_t^*$ for the case of a constant investment opportunity set (i.e. $r_t = r$, $\sigma_t = \sigma$ and $\lambda_t = \lambda$).

Proposition 3. *Consider an agent with the kinked HARA utility function and a reference level process θ who solves the consumption and portfolio choice problem, with consumption constrained to be non-negative and assuming a constant investment opportunity set. Let $\mathcal{N}$ be the cumulative distribution function of a standard normal random variable. Define C, $d_1(x)$ and $d_2(x)$ as follows:*

$$C \equiv \frac{\delta + r(\varphi - 1)}{\varphi} + \frac{1}{2}\frac{\varphi - 1}{\varphi^2}||\lambda||^2,$$

$$d_1(x) = \frac{1}{||\lambda||\sqrt{s-t}}\left[\log(x) - \log(M_t) + \left(r - \frac{1}{2}||\lambda||^2\right)(s-t)\right],$$

$$d_2(x) = d_1(x) + \frac{||\lambda||}{\varphi}\sqrt{s-t}.$$

Then:

$$\widehat{W}_t^{G*} = \frac{\varphi}{\rho}\left(\frac{y\exp\{\delta t\}}{\rho}\right)^{-\frac{1}{\varphi}} M_t^{-\frac{1}{\varphi}} \int_t^T \exp\{-C(s-t)\}\mathcal{N}\big[d_2(\underline{\xi}_s)\big]\,\mathrm{d}s - \frac{\psi\varphi}{\rho}\int_t^T \exp\{-r(s-t)\}\mathcal{N}\big[d_1(\underline{\xi}_s)\big]\,\mathrm{d}s,$$

$$\begin{aligned}\widehat{W}_t^{L*} &= \frac{\varphi}{\rho}\left(\frac{y\exp\{\delta t\}}{\rho\bar{\kappa}}\right)^{-\frac{1}{\varphi}} M_t^{-\frac{1}{\varphi}} \\ &\times \int_t^T \exp\{-r(s-t)\}\left(\mathcal{N}\big[d_1(\overline{\xi}_s^*)\big] - \mathcal{N}\left[d_1\left(\overline{\xi}_s\right)\right]\right)\mathrm{d}s \\ &- \frac{\psi\varphi}{\rho}\int_t^T \exp\{-r(s-t)\}\left(\mathcal{N}\big[d_1(\overline{\xi}_s^*)\big] - \mathcal{N}\left[d_1\left(\overline{\xi}_s\right)\right]\right)\mathrm{d}s,\end{aligned}$$

$$\widetilde{W}_t^* = \int_t^T \theta_s \exp\{-r(s-t)\}\left(\mathcal{N}\left[d_1\left(\overline{\xi}_s\right)\right] - \mathcal{N}\big[d_1(\underline{\xi}_s)\big]\right)\mathrm{d}s.$$

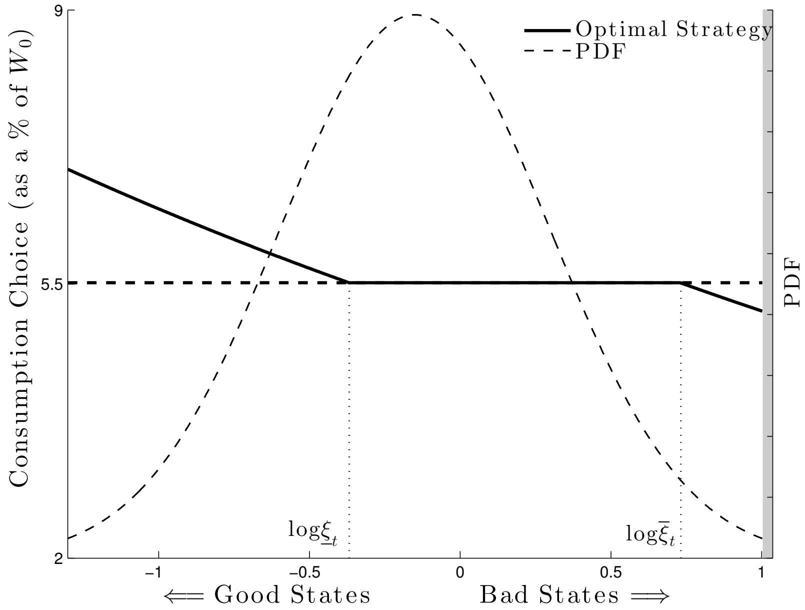
Here, $\overline{\xi}_s^* \equiv \frac{\psi^{-\varphi}\rho\bar{\kappa}}{y\exp\{\delta t\}}$.

The agent's optimal portfolio choice π_t^* can be computed in a similar way as in Section 2.5. Figure 2.14 illustrates the optimal portfolio profile of an agent with kinked CRRA utility. We observe that the optimal portfolio profile displays again a U-shaped pattern. In good as well as in bad states, the agent behaves like a CRRA agent. In particular, in these states, the fraction of wealth invested in the risky stock is equal to the constant $\lambda/(\sigma\varphi)$.

2.8. Conclusion

We have derived the optimal consumption and portfolio choice under the two-part power utility function of Tversky and Kahneman (1992) while allowing the agent to endogenously update his reference level over time. We have shown that loss aversion gives rise to a nonlinear consumption profile, inducing a (soft) guarantee on consumption, and that endogenous updating of the reference level implies smoothing of shocks.

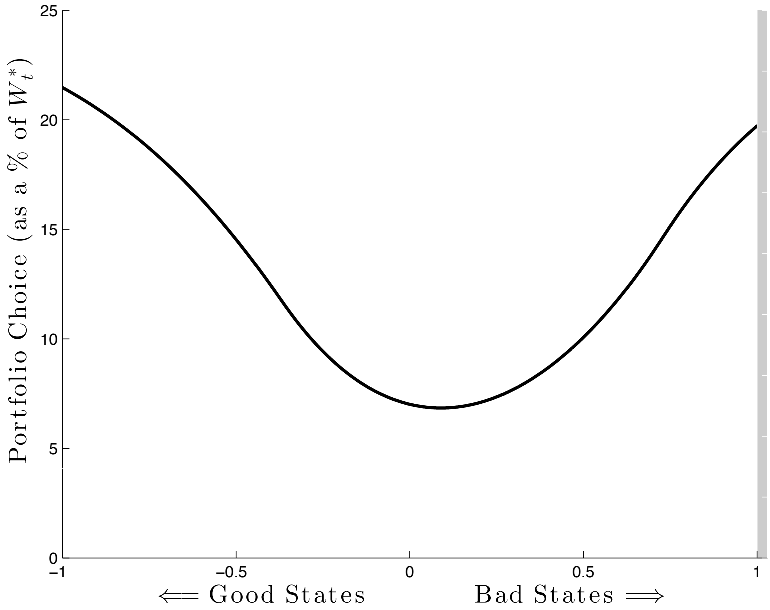
Figure 2.13.
Optimal consumption profile of an agent with kinked CRRA utility



The figure shows the optimal consumption choice (expressed as a percentage of the agent's initial wealth W_0) at age 70 as a function of the then-current log state price density. The curvature parameter φ is set equal to 4. The remaining parameter values are the same as in Section 2.6. The dashed line corresponds to the agent's reference level (expressed as a percentage of W_0). The dotted line shows the probability density function (PDF) of the then-current log state price density conditional upon information available at the age of retirement.

We have assumed that agents can objectively evaluate the probabilities associated with future outcomes. A large body of research suggests that agents subjectively weight probabilities and e.g., have a tendency to overweight unlikely extreme outcomes (see, e.g., Abdellaoui, 2000). Jin and Zhou (2008) and He and Zhou (2011, 2014) consider optimal portfolio choice under subjective probability weighting; see also Laeven and Stadje (2014). However, these authors do not consider intertemporal consumption or endogenous updating of the reference level. In future work we intend to extend our setting with intertemporal consumption and endogenous updating of the reference level to explore the impact of probability weighting on the optimal consumption and portfolio choice. Interestingly, as already shown by He and Zhou (2014), probability weighting may generate an endogenous insurance if small probabilities are sufficiently overweighted.

Figure 2.14.
Optimal portfolio profile of an agent with kinked CRRA utility



The figure shows the optimal portfolio choice measured as a fraction of total wealth invested in the risky stock at age 70 as a function of the then-current log state price density. The curvature parameter φ is set equal to 4. The remaining parameter values are the same as in Section 2.6.

2.9. Appendix

2.9.1. The Dual Technique

Schroder and Skiadas (2002) show that a generic consumption and portfolio choice model with linear internal habit formation can be mechanically transformed into a *dual* consumption and portfolio choice model without linear internal habit formation. The dual technique can be applied to an arbitrary utility function, including the two-part power utility function v (see expression (2.4.1)). To formulate the dual consumption and portfolio choice model, let us define the agent's *surplus consumption choice* $\widehat{c}_t$ as the agent's consumption choice c_t minus the agent's reference level θ_t; that is, $\widehat{c}_t \equiv c_t - \theta_t$. We can view $\widehat{c}$ as a gain process.[30] The agent's maximization problem (2.5.1) is now

[30] We note that a negative gain corresponds to a (positive) loss.

equivalent to the following dual problem:

$$\begin{aligned} &\underset{\widehat{c}}{\text{maximize}} \quad \mathbb{E}\left[\int_0^T \exp\{-\delta t\}\, v(\widehat{c}_t)\,\mathrm{d}t\right] \\ &\text{subject to} \quad \mathbb{E}\left[\int_0^T \widehat{M}_t \widehat{c}_t\,\mathrm{d}t\right] \le \widehat{W}_0 (1+\beta A_0), \qquad \widehat{c}_t \ge -L_t \quad \text{for all } t \in [0,T]. \end{aligned} \tag{2.9.1}$$

Here, $\widehat{M}_t$ and $\widehat{W}_0$ represent the dual counterparts of the state price density M_t and the agent's initial wealth W_0, respectively.

The relationship between the agent's maximization problem (2.5.1) and the dual problem (2.9.1) is characterized in terms of the auxiliary process A:

$$A_t \equiv \frac{1}{M_t}\mathbb{E}_t\left[\int_t^T M_s \exp\{-(\alpha-\beta)(s-t)\}\,\mathrm{d}s\right].$$

We can view A_t as the time t price of a bond paying a continuous coupon. In case the investment opportunity set is constant, A_t only depends on time t. As a direct consequence, the optimal portfolio choice can be computed explicitly in this case. The dual state price density $\widehat{M}_t$ and the dual initial wealth $\widehat{W}_0$ are given by

$$\widehat{M}_t \equiv M_t(1+\beta A_t), \qquad \widehat{W}_0 \equiv \frac{W_0 - A_0\theta_0}{1+\beta A_0}.$$

Furthermore, the dual reference level

$$\widehat{\theta}_s = \beta\int_t^s \exp\{-(\alpha-\beta)(s-u)\}\,\widehat{c}_u\,\mathrm{d}u + \exp\{-(\alpha-\beta)(s-t)\}\,\widehat{\theta}_t, \qquad s \ge t \ge 0,$$

is equal to the agent's reference level θ_s.

Surplus wealth $\widehat{W}_t$ is defined as follows:

$$\widehat{W}_t \equiv \frac{1}{\widehat{M}_t}\mathbb{E}_t\left[\int_t^T \widehat{M}_s \widehat{c}_s\,\mathrm{d}s\right].$$

Surplus wealth $\widehat{W}_t$ is invested in a dual financial market that is characterized by the dual risk-free rate $\widehat{r}_t$, the dual volatility $\widehat{\sigma}_t$ and the dual market price of risk $\widehat{\lambda}_t$:

$$\widehat{r}_t \equiv \beta + \frac{r_t - \alpha\beta A_t}{1+\beta A_t}, \qquad \widehat{\sigma}_t \equiv \sigma_t,$$

$$\widehat{\lambda}_t \equiv \lambda_t - \frac{\beta}{1+\beta A_t}\int_t^T \exp\left\{-(\alpha-\beta)(s-t)\right\} P_{t,s}\Psi_{t,s}\,\mathrm{d}s.$$

Here, $P_{t,s}$ corresponds to the time t price of a default-free unit discount bond that matures at time $s \geq t$ and $\Psi_{t,s}$ stands for the time t volatility of the instantaneous return on such a bond (all in the primal financial market). The optimal dual portfolio choice $\widehat{\pi}_t^*$ is determined such that it finances the optimal surplus consumption choice $\widehat{c}_t^*$.

The next proposition is adapted from Schroder and Skiadas (2002).

Proposition 4. *Suppose that we have solved the dual problem* (2.9.1)*. Let us denote the optimal surplus consumption choice by $\widehat{c}_t^*$, the optimal dual reference level by $\widehat{\theta}_t^*$, the optimal surplus wealth by $\widehat{W}_t^*$ and the optimal dual portfolio choice by $\widehat{\pi}_t^*$. Then:*

- *The optimal consumption for the agent at time $0 \leq t \leq T$ is given by*

$$c_t^* = \widehat{c}_t^* + \widehat{\theta}_t^*.$$

- *The optimal wealth for the agent at time $0 \leq t \leq T$ is given by*

$$W_t^* = \widehat{W}_t^* + \beta A_t \widehat{W}_t^* + A_t \widehat{\theta}_t^*.$$

- *The optimal portfolio choice for the agent at time $0 \leq t \leq T$ is given by*

$$\pi_t^* = \widehat{\pi}_t^* + \beta A_t \widehat{\pi}_t^* + \left(\beta \widehat{W}_t^* + \widehat{\theta}_t^*\right)(\widehat{\sigma}_t)^{-1}\int_t^T \exp\left\{-(\alpha-\beta)(s-t)\right\} P_{t,s}\Psi_{t,s}\,\mathrm{d}s.$$

Proposition 4 shows how to transform the optimal solution to the dual problem (2.9.1) back into the optimal solution to the agent's maximization problem (2.5.1).

2.9.2. Proofs

Proof of Theorem 1

The proof uses some of the techniques developed by Basak and Shapiro (2001) and Berkelaar et al. (2004) to deal with pseudo-concavity and non-differentiability aspects of the problem and adapts these to our setting with intertemporal consumption.

The dual problem, equivalent to the agent's maximization problem (2.5.1), is given by

$$\begin{aligned} &\underset{\widehat{c}}{\text{maximize}} \quad \mathbb{E}\left[\int_0^T \exp\{-\delta t\}\, v\left(\widehat{c}_t\right) \mathrm{d}t\right] \\ &\text{subject to} \quad \mathbb{E}\left[\int_0^T \widehat{M}_t \widehat{c}_t \,\mathrm{d}t\right] \leq \widehat{W}_0\left(1+\beta A_0\right), \qquad \widehat{c}_t \geq -L_t \quad \text{for all } t \in [0,T]. \end{aligned}$$

The corresponding Lagrangian $\mathcal{L}$ is defined as follows:

$$\begin{aligned} \mathcal{L} &= \mathbb{E}\left[\int_0^T \exp\{-\delta t\}\, v\left(\widehat{c}_t\right) \mathrm{d}t\right] + y\left(\widehat{W}_0\left(1+\beta A_0\right) - \mathbb{E}\left[\int_0^T \widehat{M}_t \widehat{c}_t \,\mathrm{d}t\right]\right) \\ &= \int_0^T \mathbb{E}\left[\exp\{-\delta t\}\, v\left(\widehat{c}_t\right) - y\widehat{M}_t\widehat{c}_t\right] \mathrm{d}t + y\widehat{W}_0. \end{aligned}$$

Here, y denotes the Lagrange multiplier associated with the static budget constraint. The agent wishes to maximize $\exp\{-\delta t\}\, v\left(\widehat{c}_t\right) - y\widehat{M_t\widehat{c}_t}$ subject to $\widehat{c}_t \geq -L_t$. Denote the part of the two-part power utility function with domain below zero by v_1, and the part with domain above zero by v_2. Let us denote by c^*_{1t} the agent's optimal surplus consumption choice for utility function v_1, and by c^*_{2t} the agent's optimal surplus consumption choice for utility function v_2.

We first consider the case where the agent is risk-averse in the loss domain. Due to the concavity of v_1 and v_2, the optimal surplus consumption choices c^*_{1t} and c^*_{2t} satisfy the following optimality conditions:[31]

$$\exp\{-\delta t\}\, v'_j\left(c^*_{jt}\right) = y\widehat{M}_t - x_{jt}, \qquad c^*_{jt} \geq -L_t, \qquad \text{for } j = 1,2,$$

$$x_{jt}\left(c^*_{jt} + L_t\right) = 0, \qquad x_{jt} \geq 0, \qquad \text{for } j = 1,2.$$

Here, x_{jt} denotes the Lagrange multiplier associated with the constraint on surplus consumption. After solving the optimality conditions, we obtain the following two local maxima:

$$c^*_{1t} = -\left[\left(l_t\widehat{M}_t\right)^{\frac{1}{\gamma_1-1}} \wedge L_t\right], \qquad c^*_{2t} = \left(k_t\widehat{M}_t\right)^{\frac{1}{\gamma_2-1}}.$$

Here, $l_t \equiv y\exp\{\delta t\} \cdot \frac{1}{\kappa\gamma_1}$ and $k_t \equiv y\exp\{\delta t\} \cdot \frac{1}{\gamma_2}$.

[31] The derivative of a function f at a point a is denoted by $f'(a)$.

To determine the global maximum $\widehat{c}_t^*$, we introduce the following function:

$$\begin{aligned} f\left(\widehat{M}_t\right) &= \exp\{-\delta t\}\, v\left(c_{2t}^*\right) - y\widehat{M}_t c_{2t}^* - \left[\exp\{-\delta t\}\, v\left(c_{1t}^*\right) - y\widehat{M}_t c_{1t}^*\right] \\ &= \exp\{-\delta t\}\left(1-\gamma_2\right)\left(k_t\widehat{M}_t\right)^{\frac{\gamma_2}{\gamma_2-1}} + \kappa\exp\{-\delta t\}\left[\left(l_t\widehat{M}_t\right)^{\frac{1}{\gamma_1-1}} \wedge L_t\right]^{\gamma_1} \\ &\quad - y\widehat{M}_t\left[\left(l_t\widehat{M}_t\right)^{\frac{1}{\gamma_1-1}} \wedge L_t\right]. \end{aligned}$$

The global maximum $\widehat{c}_t^*$ is equal to c_{2t}^* if $f\left(\widehat{M}_t\right) \geq 0$; and equals c_{1t}^* otherwise. It follows that $\lim_{\widehat{M}_t\to\infty} f\left(\widehat{M}_t\right) = -\infty$, $\lim_{\widehat{M}_t\to 0} f\left(\widehat{M}_t\right) = \infty$ and $f'\left(\widehat{M}_t\right) < 0$ for all $\widehat{M}_t$. Hence, $f\left(\widehat{M}_t\right)$ is strictly decreasing. As a direct consequence, $f\left(\widehat{M}_t\right)$ has one zero in the interval $(0,\infty)$. Define ξ_t to be such that $f\left(\xi_t\right) = 0$. The global maximum $\widehat{c}_t^*$ is equal to c_{2t}^* if $\widehat{M}_t \leq \xi_t$; and equals c_{1t}^* otherwise.

We now consider the case where the agent is risk-seeking in the loss domain. Due to the concavity of v_2, the optimal surplus consumption choice c_{2t}^* satisfies the following optimality conditions:

$$\exp\{-\delta t\}\, v_2'\left(c_{jt}^*\right) = y\widehat{M}_t - x_{2t}, \qquad c_{2t}^* \geq -L_t,$$

$$x_{2t}\left(c_{2t}^* + L_t\right) = 0, \qquad x_{2t} \geq 0.$$

After solving the optimality conditions, we obtain the following local maximum:

$$c_{2t}^* = \left(k_t\widehat{M}_t\right)^{\frac{1}{\gamma_2-1}}.$$

Due to the convexity of v_1, the optimal surplus consumption choice c_{1t}^* lies at a corner point of the feasible region. Hence, the only two possible candidates for c_{1t}^* are $-L_t$ and 0.

To determine the global maximum $\widehat{c}_t^*$, we introduce the following function:

$$g\left(\widehat{M}_t\right) = \exp\{-\delta t\}\, v\left(c_{2t}^*\right) - y\widehat{M}_t c_{2t}^* - \left[\exp\{-\delta t\}\, v\left(c_{1t}^*\right) - y\widehat{M}_t c_{1t}^*\right].$$

The global maximum $\widehat{c}_t^*$ is equal to c_{2t}^* if $g\left(\widehat{M}_t\right) \geq 0$; and equals c_{1t}^* otherwise. We distinguish between the following two cases:

- $c_{1t}^* = 0$. Straightforward computations show that $g\left(\widehat{M}_t\right)$ is given by

$$g\left(\widehat{M}_t\right) = \exp\left\{-\delta t\right\}\left(1-\gamma_2\right)\left(k_t \widehat{M}_t\right)^{\frac{\gamma_2}{\gamma_2-1}}.$$

Since $0 < \gamma_2 < 1$ and $y > 0$, it follows that $g\left(\widehat{M}_t\right) > 0$ for all $\widehat{M}_t$. We conclude that $c_{1t}^* = 0$ is never optimal.

- $c_{1t}^* = -L_t$. Straightforward computations show that $g\left(\widehat{M}_t\right)$ is given by

$$g\left(\widehat{M}_t\right) = \exp\left\{-\delta t\right\}\left(1-\gamma_2\right)\left(k_t \widehat{M}_t\right)^{\frac{\gamma_2}{\gamma_2-1}} + \exp\left\{-\delta t\right\}\kappa L_t^{\gamma_1} - y\widehat{M}_t L_t.$$

It follows that $g\left(\widehat{M}_t\right) > 0$ for all $\widehat{M}_t \leq \frac{\kappa}{y}\exp\left\{-\delta t\right\} L_t^{\gamma_1 - 1}$. Also, $\lim_{\widehat{M}_t \to \infty} g\left(\widehat{M}_t\right) = -\infty$ and $g'\left(\widehat{M}_t\right) < 0$ for all $\widehat{M}_t$. Hence, $g\left(\widehat{M}_t\right)$ is strictly decreasing. As a direct consequence, $g\left(\widehat{M}_t\right)$ has one zero in the interval $\left(\frac{\kappa}{y}\exp\left\{-\delta t\right\} L_t^{\gamma_1-1}, \infty\right)$. Define ξ_t to be such that $g\left(\xi_t\right) = 0$. It follows that the global maximum $\widehat{c}_t^*$ is equal to c_{2t}^* if $\widehat{M}_t \leq \xi_t$; and equals c_{1t}^* otherwise.

A standard verification (see, e.g., Karatzas and Shreve, 1998, p. 103) that the optimal solutions obtained from the Lagrangian are the optimal solutions to the dual problem completes the proof. *Q.E.D.*

Proof of Proposition 1

We distinguish between the following two cases:

- Risk-averse behavior in the loss domain. Define the following function:

$$\widetilde{f}(x) \equiv \left(1-\gamma_2\right)\left(\frac{x}{\gamma_2}\right)^{\frac{\gamma_2}{\gamma_2-1}} + \kappa\left[\left(\frac{x}{\gamma_1\kappa}\right)^{\frac{1}{\gamma_1-1}} \wedge L_t\right]^{\gamma_1} - x\left[\left(\frac{x}{\gamma_1\kappa}\right)^{\frac{1}{\gamma_1-1}} \wedge L_t\right].$$

Let $\widetilde{\xi}_t$ be such that $\widetilde{f}\left(\widetilde{\xi}_t\right) = 0$. It follows that $\widetilde{\xi}_t = y\exp\left\{\delta t\right\}\xi_t$. The quantity $\widetilde{\xi}_t$ increases as the loss aversion index κ increases. Furthermore, initial surplus wealth $\widehat{W}_0$ decreases with the initial reference level θ_0, increases with the depreciation parameter α (provided that $\widehat{W}_0$ is non-negative), and decreases with the endogeneity parameter β (provided that $\widehat{W}_0$ is non-negative).

- Risk-seeking behavior in the loss domain. Define the following function:

$$\widetilde{g}\left(x\right) \equiv \left(1-\gamma_2\right)\left(\frac{x}{\gamma_2}\right)^{\frac{\gamma_2}{\gamma_2-1}} + \kappa L_t^{\gamma_1} - xL_t.$$

  Let $\widetilde{\xi}_t$ be such that $\widetilde{g}\left(\widetilde{\xi}_t\right) = 0$. It follows that $\widetilde{\xi}_t = y\exp\left\{\delta t\right\}\xi_t$. The quantity $\widetilde{\xi}_t$ increases as the loss aversion index κ increases. Furthermore, initial surplus wealth $\widehat{W}_0$ decreases with the initial reference level θ_0, increases with the depreciation parameter α (provided that $\widehat{W}_0$ is non-negative), and decreases with the endogeneity parameter β (provided that $\widehat{W}_0$ is non-negative).

The proposition now follows straightforwardly from Berkelaar et al. (2004). *Q.E.D.*

Proof of Proposition 2

Optimal surplus wealth is given by

$$\widehat{W}_t^* = \frac{1}{\widehat{M}_t}\mathbb{E}_t\left[\int_t^T \widehat{M}_s\widehat{c}_s^*\,\mathrm{d}s\right]. \tag{2.9.2}$$

We first consider the case where the agent is risk-averse in the loss domain. Substituting the optimal surplus consumption choice $\widehat{c}_s^*$ into equation (2.9.2) yields

$$\begin{aligned}
\widehat{W}_t^* &= \frac{1}{\widehat{M}_t}\mathbb{E}_t\Bigg[\int_t^T \widehat{M}_s\left(k_s\widehat{M}_s\right)^{\frac{1}{\gamma_2-1}}\mathbb{1}_{\left[\widehat{M}_s\le\xi_s\right]}\,\mathrm{d}s \\
&\quad - \int_t^T \widehat{M}_s\left(l_s\widehat{M}_s\right)^{\frac{1}{\gamma_1-1}}\mathbb{1}_{\left[\xi_s<\widehat{M}_s<\xi_s\vee\zeta_s\right]}\,\mathrm{d}s - \int_t^T \widehat{M}_sL_s\mathbb{1}_{\left[\widehat{M}_s\ge\xi_s\vee\zeta_s\right]}\,\mathrm{d}s\Bigg] \\
&= \left(k_t\widehat{M}_t\right)^{\frac{1}{\gamma_2-1}}\mathbb{E}_t\left[\int_t^T\left(\frac{\widehat{M}_s}{\widehat{M}_t}\right)^{\frac{\gamma_2}{\gamma_2-1}}\exp\left\{\frac{\delta(s-t)}{\gamma_2-1}\right\}\mathbb{1}_{\left[\widehat{M}_s\le\xi_s\right]}\,\mathrm{d}s\right] \\
&\quad - \left(l_t\widehat{M}_t\right)^{\frac{1}{\gamma_1-1}}\mathbb{E}_t\left[\int_t^T\left(\frac{\widehat{M}_s}{\widehat{M}_t}\right)^{\frac{\gamma_1}{\gamma_1-1}}\exp\left\{\frac{\delta(s-t)}{\gamma_1-1}\right\}\mathbb{1}_{\left[\xi_s<\widehat{M}_s<\xi_s\vee\zeta_s\right]}\,\mathrm{d}s\right] \\
&\quad - \mathbb{E}_t\left[\int_t^T\frac{\widehat{M}_s}{\widehat{M}_t}L_s\mathbb{1}_{\left[\widehat{M}_s\ge\xi_s\vee\zeta_s\right]}\,\mathrm{d}s\right].
\end{aligned} \tag{2.9.3}$$

Here, $\zeta_s \equiv \exp\left\{-\delta s\right\}\frac{\gamma_1\kappa}{y}L_s^{\gamma_1-1}$. The closed-form expression for $\widehat{W}_t^*$ can be determined by computing the conditional expectations. In case the investment opportunity set is constant, we find

$$\mathbb{E}_t\left[\frac{\widehat{M}_s}{\widehat{M}_t}L_s\mathbb{1}_{[\widehat{M}_s\geq\xi_s\vee\zeta_s]}\right]=\exp\left\{-\int_t^s\widehat{r}_u\,\mathrm{d}u\right\}L_s\mathcal{N}\left[-d_1\left(\xi_s\vee\zeta_s\right)\right], \tag{2.9.4}$$

$$\mathbb{E}_t\left[\left(\frac{\widehat{M}_s}{\widehat{M}_t}\right)^{\frac{\gamma_2}{\gamma_2-1}}\exp\left\{\frac{\delta(s-t)}{\gamma_2-1}\right\}\mathbb{1}_{[\widehat{M}_s\leq\xi_s]}\right]=\exp\left\{-\int_t^s\Gamma_u\,\mathrm{d}u\right\}\mathcal{N}\left[d_2\left(\xi_s\right)\right], \tag{2.9.5}$$

$$\mathbb{E}_t\left[\left(\frac{\widehat{M}_s}{\widehat{M}_t}\right)^{\frac{\gamma_1}{\gamma_1-1}}\exp\left\{\frac{\delta(s-t)}{\gamma_1-1}\right\}\mathbb{1}_{[\xi_s<\widehat{M}_s<\xi_s\vee\zeta_s]}\right]=\exp\left\{-\int_t^s\Pi_u\,\mathrm{d}u\right\}$$
$$\times\left(\mathcal{N}\left[d_3\left(\xi_s\vee\zeta_s\right)\right]-\mathcal{N}\left[d_3\left(\xi_s\right)\right]\right). \tag{2.9.6}$$

Here, $\mathcal{N}$ is the cumulative distribution function of a standard normal random variable, and Γ_u, Π_u, $d_1(x)$, $d_2(x)$ and $d_3(x)$ are defined as follows:

$$\Gamma_u\equiv\frac{\delta-\gamma_2\widehat{r}_u}{1-\gamma_2}-\frac{1}{2}\frac{\gamma_2}{(1-\gamma_2)^2}||\lambda||^2,\qquad\Pi_u\equiv\frac{\delta-\gamma_1\widehat{r}_u}{1-\gamma_1}-\frac{1}{2}\frac{\gamma_1}{(1-\gamma_1)^2}||\lambda||^2,$$

$$d_1(x)\equiv\frac{1}{||\lambda||\sqrt{s-t}}\cdot\left[\log(x)-\log\left(\widehat{M}_t\right)+\int_t^s\widehat{r}_u\,\mathrm{d}u-\frac{1}{2}||\lambda||^2(s-t)\right],$$

$$d_2(x)\equiv d_1(x)+\frac{||\lambda||}{1-\gamma_2}\sqrt{s-t},\qquad d_3(x)\equiv d_1(x)+\frac{||\lambda||}{1-\gamma_1}\sqrt{s-t}.$$

Substituting the conditional expectations (2.9.4), (2.9.5) and (2.9.6) into equation (2.9.3) yields the optimal surplus wealth.

We now consider the case where the agent is risk-seeking in the domain of losses. Substituting the optimal surplus consumption choice $\widehat{c}_s^*$ into equation (2.9.2) yields

$$\begin{aligned}\widehat{W}_t^* &= \frac{1}{\widehat{M}_t}\mathbb{E}_t\left[\int_t^T \widehat{M}_s\left(k_s\widehat{M}_s\right)^{\frac{1}{\gamma_2-1}}\mathbb{1}_{[\widehat{M}_s\leq\xi_s]}\,\mathrm{d}s - \int_t^T \widehat{M}_s L_s \mathbb{1}_{[\widehat{M}_s>\xi_s]}\,\mathrm{d}s\right]\\ &= \left(k_t\widehat{M}_t\right)^{\frac{1}{\gamma_2-1}}\mathbb{E}_t\left[\int_t^T\left(\frac{\widehat{M}_s}{\widehat{M}_t}\right)^{\frac{\gamma_2}{\gamma_2-1}}\exp\left\{\frac{\delta(s-t)}{\gamma_2-1}\right\}\mathbb{1}_{[\widehat{M}_s\leq\xi_s]}\,\mathrm{d}s\right]\\ &\quad-\mathbb{E}_t\left[\int_t^T\frac{\widehat{M}_s}{\widehat{M}_t}L_s\mathbb{1}_{[\widehat{M}_s>\xi_s]}\,\mathrm{d}s\right].\end{aligned}\tag{2.9.7}$$

The closed-form expression for $\widehat{W}_t^*$ can be determined by computing the conditional expectations. In case the investment opportunity set is constant, we find

$$\mathbb{E}_t\left[\frac{\widehat{M}_s}{\widehat{M}_t}L_s\mathbb{1}_{[\widehat{M}_s>\xi_s]}\right] = \exp\left\{-\int_t^s \widehat{r}_u\,\mathrm{d}u\right\}L_s\mathcal{N}\left[-d_1\left(\xi_s\right)\right],\tag{2.9.8}$$

$$\mathbb{E}_t\left[\left(\frac{\widehat{M}_s}{\widehat{M}_t}\right)^{\frac{\gamma_2}{\gamma_2-1}}\exp\left\{\frac{\delta(s-t)}{\gamma_2-1}\right\}\mathbb{1}_{[\widehat{M}_s\leq\xi_s]}\right] = \exp\left\{-\int_t^s\Gamma_u\,\mathrm{d}u\right\}\mathcal{N}\left[d_2\left(\xi_s\right)\right].\tag{2.9.9}$$

Substituting the conditional expectations (2.9.8) and (2.9.9) into equation (2.9.7) yields the optimal surplus wealth. *Q.E.D.*

Proof of Theorem 3

The proof uses some of the techniques developed by Basak and Shapiro (2001) and Berkelaar et al. (2004) and adapts these to our setting with intertemporal consumption.

The agent's maximization problem is given by

$$\begin{aligned}&\underset{c}{\text{maximize}}\quad \mathbb{E}\left[\int_0^T\exp\{-\delta t\}\,u\left(c_t;\theta_t\right)\mathrm{d}t\right]\\ &\text{subject to}\quad \mathbb{E}\left[\int_0^T M_t c_t\,\mathrm{d}t\right]\leq W_0,\qquad c_t\geq 0\quad\text{for all } t\in[0,T].\end{aligned}$$

The corresponding Lagrangian $\mathcal{L}$ is defined as follows:

$$\begin{aligned}\mathcal{L} &= \mathbb{E}\left[\int_0^T \exp\{-\delta t\}\, u(c_t;\theta_t)\,\mathrm{d}t\right] + y\left(W_0 - \mathbb{E}\left[\int_0^T M_t c_t\,\mathrm{d}t\right]\right)\\ &= \int_0^T \mathbb{E}\left[\exp\{-\delta t\}\, u(c_t;\theta_t) - yM_t c_t\right]\mathrm{d}t + yW_0.\end{aligned}$$

Here, y denotes the Lagrange multiplier associated with the static budget constraint. The agent wishes to maximize $\exp\{-\delta t\}\, u(c_t;\theta_t) - yM_tc_t$ subject to $c_t \geq 0$. Denote the part of the utility function with domain below zero by u_1, and the part with domain above zero by u_2. Let us denote by c_{1t}^* the agent's optimal consumption choice for utility function u_1, and by c_{2t}^* the agent's optimal consumption choice for utility function u_2.

Due to the concavity of u_1 and u_2, the optimal consumption choices c_{1t}^* and c_{2t}^* satisfy the following optimality conditions:

$$\exp\{-\delta t\}\, u_j'\left(c_{jt}^*;\theta_t\right) = yM_t - x_{jt}, \qquad c_{jt}^* \geq 0, \qquad \text{for } j = 1, 2,$$

$$x_{jt}c_{jt}^* = 0, \qquad x_{jt} \geq 0, \qquad \text{for } j = 1, 2.$$

Here, x_{jt} denotes the Lagrange multiplier associated with the non-negativity constraint on consumption. After solving the optimality conditions, we obtain the following two local maxima:

$$c_{1t}^* = \min\left\{\theta_t, \left[\frac{\varphi}{\rho}\left(\frac{y\exp\{\delta t\}\, M_t}{\rho\bar{\kappa}}\right)^{-\frac{1}{\varphi}} - \frac{\psi\varphi}{\rho}\right] \vee 0\right\},$$

$$c_{2t}^* = \max\left\{\theta_t, \frac{\varphi}{\rho}\left(\frac{y\exp\{\delta t\}\, M_t}{\rho}\right)^{-\frac{1}{\varphi}} - \frac{\psi\varphi}{\rho}\right\}.$$

Here, $\bar{\kappa} \equiv \eta + (1-\eta)\cdot\kappa$.

To determine the global maximum c_t^*, we introduce the following function:

$$f(M_t) = \exp\{-\delta t\}\, u(c_{2t}^*;\theta_t) - yM_tc_{2t}^* - \left[\exp\{-\delta t\}\, u(c_{1t}^*;\theta_t) - yM_tc_{1t}^*\right].$$

The global maximum is equal to c_{2t}^* if $f(M_t) \geq 0$; and equals c_{1t}^* otherwise. It follows that $f(M_t)$ changes sign at $\underline{\xi}_t = \frac{\rho}{y}\exp\{-\delta t\}\left(\frac{\rho}{\varphi}\theta_t + \psi\right)^{-\varphi}$ and $\overline{\xi}_t = \frac{\rho\bar{\kappa}}{y}\exp\{-\delta t\}\left(\frac{\rho}{\varphi}\theta_t + \psi\right)^{-\varphi}$.

We consider the following three cases:

- $\underline{\xi}_t \leq M_t \leq \overline{\xi}_t$. It follows that θ_t is the only candidate solution. We conclude that $c_t^* = \theta_t$ is the global maximum.

- $M_t > \overline{\xi}_t$. We compare the candidate solutions $c_{1t}^* = \left[\frac{\varphi}{\rho}\left(\frac{y\exp\{\delta t\}M_t}{\rho\bar{\kappa}}\right)^{-\frac{1}{\varphi}} - \frac{\psi\varphi}{\rho}\right] \vee 0$ and $c_{2t}^* = \theta_t$. Some straightforward computations show that $f\left(\overline{\xi}_t\right) = 0$, $f'\left(\overline{\xi}_t\right) = 0$ and $f''\left(M_t\right) < 0$ for all $M_t > \overline{\xi}_t$. Hence, $f\left(M_t\right) < 0$ for all $M_t > \overline{\xi}_t$. We conclude that $c_t^* = c_{1t}^*$ is the global maximum.

- $M_t < \underline{\xi}_t$. We compare the candidate solutions $c_{1t}^* = \theta_t$ and $c_{2t}^* = \frac{\varphi}{\rho}\left(\frac{y\exp\{\delta t\}M_t}{\rho}\right)^{-\frac{1}{\varphi}} - \frac{\psi\varphi}{\rho}$. Some straightforward computations show that $f\left(\underline{\xi}_t\right) = 0$, $f'\left(\underline{\xi}_t\right) = 0$ and $f''\left(\underline{\xi}_t\right) > 0$ for all $M_t < \underline{\xi}_t$. Hence, $f\left(M_t\right) > 0$ for all $M_t < \underline{\xi}_t$. We conclude that $c_t^* = c_{2t}^*$ is the global maximum.

A standard verification (see, e.g., Karatzas and Shreve, 1998, p. 103) that the optimal solution obtained from the Lagrangian is the optimal solution to the static maximization problem completes the proof. *Q.E.D.*

Proof of Proposition 3

Optimal wealth is given by

$$W_t^* = \frac{1}{M_t}\mathbb{E}_t\left[\int_t^T M_s c_s^* \, \mathrm{d}s\right]. \tag{2.9.10}$$

Substituting the optimal consumption choice into equation (2.9.10) yields

$$
\begin{aligned}
W_t^* = &\frac{1}{M_t}\mathbb{E}_t\left[\int_t^T M_s\theta_s \mathbb{1}_{[\underline{\xi}_s \le M_s \le \bar{\xi}_s]}\,\mathrm{d}s\right.\\
&+\int_t^T M_s\left\{\frac{\varphi}{\rho}\left(\frac{y\exp\{\delta s\}M_s}{\rho}\right)^{-\frac{1}{\varphi}}-\frac{\psi\varphi}{\rho}\right\}\mathbb{1}_{[M_s<\underline{\xi}_s]}\,\mathrm{d}s\\
&\left.+\int_t^T M_s\left\{\frac{\varphi}{\rho}\left(\frac{y\exp\{\delta s\}M_s}{\rho\bar{\kappa}}\right)^{-\frac{1}{\varphi}}-\frac{\psi\varphi}{\rho}\right\}\mathbb{1}_{[\bar{\xi}_s<M_s<\bar{\xi}_s^*]}\,\mathrm{d}s\right]\\
&=\mathbb{E}_t\left[\int_t^T \frac{M_s}{M_t}\theta_s\mathbb{1}_{[\underline{\xi}_s \le M_s \le \bar{\xi}_s]}\,\mathrm{d}s\right]\\
&+\frac{\varphi}{\rho}\left(\frac{y\exp\{\delta t\}}{\rho}\right)^{-\frac{1}{\varphi}}M_t^{-\frac{1}{\varphi}}\mathbb{E}_t\left[\int_t^T\left\{\frac{M_s}{M_t}\right\}^{\frac{\varphi-1}{\varphi}}\exp\left\{-\frac{\delta(s-t)}{\varphi}\right\}\mathbb{1}_{[M_s<\underline{\xi}_s]}\,\mathrm{d}s\right]\\
&-\frac{\psi\varphi}{\rho}\mathbb{E}_t\left[\int_t^T\frac{M_s}{M_t}\mathbb{1}_{[M_s<\underline{\xi}_s]}\,\mathrm{d}s\right]\\
&+\frac{\varphi}{\rho}\left(\frac{y\exp\{\delta t\}}{\rho\bar{\kappa}}\right)^{-\frac{1}{\varphi}}M_t^{-\frac{1}{\varphi}}\mathbb{E}_t\left[\int_t^T\left\{\frac{M_s}{M_t}\right\}^{\frac{\varphi-1}{\varphi}}\exp\left\{-\frac{\delta(s-t)}{\varphi}\right\}\mathbb{1}_{[\bar{\xi}_s<M_s<\bar{\xi}_s^*]}\,\mathrm{d}s\right]\\
&-\frac{\psi\varphi}{\rho}\mathbb{E}_t\left[\int_t^T\frac{M_s}{M_t}\mathbb{1}_{[\bar{\xi}_s<M_s<\bar{\xi}_s^*]}\,\mathrm{d}s\right].
\end{aligned}
\tag{2.9.11}
$$

Here, $\bar{\xi}_s^* \equiv \frac{\psi^{-\varphi}\rho\bar{\kappa}}{y\exp\{\delta t\}}$. The closed-form expression for W_t^* can be determined by computing the conditional expectations. In case the investment opportunity set is constant, we find

$$
\mathbb{E}_t\left[\left\{\frac{M_s}{M_t}\right\}^{\frac{\varphi-1}{\varphi}}\exp\left\{-\frac{\delta(s-t)}{\varphi}\right\}\mathbb{1}_{[M_s<\underline{\xi}_s]}\right]=\exp\{-C(s-t)\}\mathcal{N}\left[d_2(\underline{\xi}_s)\right],\tag{2.9.12}
$$

$$
\begin{aligned}
\mathbb{E}_t\left[\left\{\frac{M_s}{M_t}\right\}^{\frac{\varphi-1}{\varphi}}\exp\left\{-\frac{\delta(s-t)}{\varphi}\right\}\mathbb{1}_{[\bar{\xi}_s<M_s<\bar{\xi}_s^*]}\right]&=\exp\{-C(s-t)\}\\
&\times\left(\mathcal{N}\left[d_2(\bar{\xi}_s^*)\right]-\mathcal{N}\left[d_2\left(\bar{\xi}_s\right)\right]\right),
\end{aligned}
\tag{2.9.13}
$$

$$
\mathbb{E}_t\left[\frac{M_s}{M_t}\theta_s\mathbb{1}_{[\underline{\xi}_s \le M_s \le \bar{\xi}_s]}\right]=\theta_s\exp\{-r(s-t)\}\left(\mathcal{N}\left[d_1\left(\bar{\xi}_s\right)\right]-\mathcal{N}\left[d_1(\underline{\xi}_s)\right]\right),\tag{2.9.14}
$$

$$\mathbb{E}_t\left[\frac{M_s}{M_t}\mathbb{1}_{[\bar{\xi}_s<M_s<\bar{\xi}_s^*]}\right] = \exp\{-r(s-t)\}\left(\mathcal{N}\big[d_1\big(\bar{\xi}_s^*\big)\big] - \mathcal{N}\left[d_1\left(\bar{\xi}_s\right)\right]\right), \tag{2.9.15}$$

$$\mathbb{E}_t\left[\frac{M_s}{M_t}\mathbb{1}_{[M_s<\underline{\xi}_s]}\right] = \exp\{-r(s-t)\}\mathcal{N}\big[d_1\big(\underline{\xi}_s\big)\big]. \tag{2.9.16}$$

Here, $\mathcal{N}$ is the cumulative distribution function of a standard normal random variable, and C, $d_1(x)$ and $d_2(x)$ are defined as follows:

$$C = \frac{\delta + r(\varphi-1)}{\varphi} + \frac{1}{2}\frac{\varphi-1}{\varphi^2}||\lambda||^2,$$

$$d_1(x) = \frac{1}{||\lambda||\sqrt{s-t}}\left[\log(x) - \log(M_t) + \left(r - \frac{1}{2}||\lambda||^2\right)(s-t)\right],$$

$$d_2(x) = d_1(x) + \frac{1}{\varphi}||\lambda||\sqrt{s-t}.$$

Substituting the conditional expectations (2.9.12) – (2.9.16) into equation (2.9.11) yields the optimal wealth. *Q.E.D.*

2.9.3. Welfare Analysis

This appendix describes a numerical procedure for computing welfare losses. This procedure is based on the assumptions that the investment opportunity set is constant and the agent can only invest in one risky stock. We introduce the following notation:

- Δt: time step;
- $t_n \equiv n\Delta t$ for $n = 0, ..., \left\lfloor \frac{T}{\Delta t} \right\rfloor$;
- $\mathcal{S}$: total number of simulations.

The floor operator $\lfloor \cdot \rfloor$ rounds a number downward to its nearest integer.

To compute the welfare loss associated with a suboptimal consumption strategy c_t, we apply the following steps:

1. We generate $\mathcal{S}$ trajectories of the pricing kernel:

$$M^s_{t_{n+1}} = M^s_{t_n} - rM^s_{t_n}\Delta t - \lambda M^s_{t_n}\sqrt{\Delta t}\epsilon^s_{t_n}, \quad n = 0, ..., \left\lfloor \frac{T}{\Delta t} \right\rfloor, \quad s = 1, ..., \mathcal{S}.$$

   Here, $\epsilon^s_{t_n}$ is a standard normally distributed random variable.

2. We compute the optimal surplus consumption choice $\widehat{c}^{*s}_{t_n}$ for $n = 0, ..., \left\lfloor \frac{T}{\Delta t} \right\rfloor$ and $s = 1, ..., \mathcal{S}$. We note that the optimal surplus consumption choice $\widehat{c}^{*s}_{t_n}$ is a function of the dual state price density $\widehat{M}^s_{t_n} \equiv M^s_{t_n}\left(1 + \beta A_{t_n}\right)$. Expected utility can now be approximated by

$$\mathbb{E}\left[\int_0^T \exp\{-\delta t\} v\left(\widehat{c}^*_t\right) \mathrm{d}t\right] \approx \frac{1}{\mathcal{S}} \sum_{s=1}^{\mathcal{S}} \sum_{n=0}^{\left\lfloor \frac{T}{\Delta t} \right\rfloor} \exp\{-\delta t_n\} v\left(\widehat{c}^{*s}_{t_n}\right)\Delta t. \tag{2.9.17}$$

   The right-hand side of (2.9.17) approximates of $\mathbb{E}\left[\int_0^T \exp\{-\delta t\} v\left(c^*_t - \theta^*_t\right) \mathrm{d}t\right]$.

3. We solve for certainty equivalent consumption ce^*:

$$\frac{1}{\mathcal{S}} \sum_{s=1}^{\mathcal{S}} \sum_{n=0}^{\left\lfloor \frac{T}{\Delta t} \right\rfloor} \exp\{-\delta t_n\} v\left(\widehat{c}^{*s}_{t_n}\right)\Delta t = \sum_{n=0}^{\left\lfloor \frac{T}{\Delta t} \right\rfloor} \exp\{-\delta t_n\} v\left(ce^* - \theta^*_{t_n}\right)\Delta t,$$

   where

$$\theta^*_{t_n} = \theta_0 \exp\{-\alpha t_n\} + \beta \sum_{i=0}^{n-1} \exp\{-\alpha\left(t_n - t_i\right)\} ce^*\Delta t.$$

4. We compute the suboptimal consumption strategy $\widehat{c}^s_{t_n} \equiv c^s_{t_n} - \theta^s_{t_n}$ for $n = 0, ..., \left\lfloor \frac{T}{\Delta t} \right\rfloor$ and $s = 1, ..., \mathcal{S}$. Expected utility can now be approximated by

$$\mathbb{E}\left[\int_0^T \exp\{-\delta t\} v\left(\widehat{c}_t\right) \mathrm{d}t\right] \approx \frac{1}{\mathcal{S}} \sum_{s=1}^{\mathcal{S}} \sum_{n=0}^{\left\lfloor \frac{T}{\Delta t} \right\rfloor} \exp\{-\delta t_n\} v\left(\widehat{c}^s_{t_n}\right)\Delta t.$$

5. We solve for certainty equivalent consumption ce:

$$\frac{1}{\mathcal{S}} \sum_{s=1}^{\mathcal{S}} \sum_{n=0}^{\left\lfloor \frac{T}{\Delta t} \right\rfloor} \exp\{-\delta t_n\} v\left(\widehat{c}^s_{t_n}\right)\Delta t = \sum_{n=0}^{\left\lfloor \frac{T}{\Delta t} \right\rfloor} \exp\{-\delta t_n\} v\left(ce - \theta_{t_n}\right)\Delta t,$$

where

$$\theta_{t_n} = \theta_0 \exp\{-\alpha t_n\} + \beta \sum_{i=0}^{n-1} \exp\{-\alpha (t_n - t_i)\} ce\Delta t.$$

6. Finally, we compute the welfare loss WL:

$$WL = \frac{ce^* - ce}{ce^*}.$$

Chapter 3

Dynamic Consumption and Portfolio Choice under Cumulative Prospect Theory[32]

This chapter explicitly derives the optimal dynamic consumption and portfolio choice of an agent with cumulative prospect theory preferences. Specifically, the agent is loss averse, distorts probabilities, and endogenously updates his reference level over time. The optimal strategy seeks to mitigate large year-on-year fluctuations in consumption and aims to provide protection against downside risk. The first effect is due to endogenous updating of the reference level while the second effect is due to loss aversion and probability weighting. We show that if small probabilities are sufficiently overweighted, our model generates an endogenous floor on consumption.

3.1. Introduction

Since the seminal papers of Merton (1969) and Samuelson (1969), optimal consumption and portfolio choice over the life cycle has been extensively studied in the economics and finance literature. Most authors assume that preferences over consumption choices are represented by CRRA utility (see, e.g., Wachter, 2002; Liu, 2007), by Epstein-Zin utility (see, e.g., Chacko and Viceira, 2005; Gomes and Michaelides, 2008) or by habit formation utility (see, e.g., Gomes and Michaelides, 2003; Munk, 2008). However, an extensive body of literature in behavioral economics and finance documents experimentally as well as empirically departures from the key assumptions underlying these preference models, in a wide variety of risky choice situations. In response, the literature has developed several

[32]This chapter is co-authored with Roger Laeven.

alternative theories of decision making under risk. Cumulative prospect theory (CPT for short), introduced by Tversky and Kahneman (1992), is currently perhaps the most promising descriptive theory of decision making under risk. This chapter derives and analyzes the optimal dynamic consumption and portfolio choice of an agent with CPT preferences.

Specifically, we consider an agent that derives value from the difference between consumption and a so-called reference level. If consumption exceeds the reference level, the agent experiences a gain, while if consumption falls short of the reference level, the agent experiences a loss. CPT is silent on how to update the reference level over time. We follow the (internal) habit formation literature (see, e.g., Constantinides, 1990) and assume that the reference level depends on the agent's own past consumption choices.[33] As a direct consequence, consumption responds gradually to financial shocks (see Chapter 2). Our agent has a two-part power utility function and two inverse S-shaped probability weighting (or distortion) functions (one for gains and one for losses). The utility function incorporates loss aversion (i.e., losses hurt more than gains satisfy), and the probability weighting functions overweight small probabilities and underweight large probabilities.[34]

The literature on optimal consumption and portfolio choice under CPT preferences is still immature.[35] Gomes (2005) explores the optimal portfolio choice of a loss averse agent in an economy with two states of nature, and analyzes the impact of loss aversion on trading volume. Berkelaar et al. (2004) examine the optimal portfolio choice of a loss averse agent in a setting with terminal wealth and a continuum of states of nature. Chapter 2 includes intertemporal consumption choice in this setting and allows the agent to endogenously update his reference level over time. The model of Chapter 2 does, however, not accommodate probability weighting. Jin and Zhou (2008) and He and Zhou (2011) study the optimal portfolio choice of an agent that maximizes CPT value of terminal wealth. Although these authors take probability weighting into consideration, they do not consider intertemporal consumption choice and an endogenous reference level. The present chapter considers a preference model that encompasses intertemporal

[33] The reference level is backward-looking and not forward-looking as in Kőszegi and Rabin (2006, 2007, 2009).

[34] An extensive body of literature shows that individuals overweight low probabilities and underweight large probabilities (see, e.g., Wu and Gonzalez, 1996; Abdellaoui, 2000; Bleichrodt and Pinto, 2000).

[35] Several authors use CPT preferences to explain interesting features observed in financial data. For example, Benartzi and Thaler (1995) find that loss aversion helps to explain the equity premium puzzle.

consumption choice, an endogenous reference level, as well as probability weighting.

Our results can be summarized as follows. First, we find that the agent divides the state of the economy into two categories: good states and bad states. In good states, consumption is larger than the reference level, while in really bad states, consumption is smaller than the reference level. The consumption profile (i.e., consumption as a function of the log state price density) displays a 90° rotated S-shaped pattern. Second, the two inverse S-shaped probability weighting functions impede the sensitivity of the optimal consumption choice to the state of the economy, inducing endogenous guarantees. In particular, both the occurrence of really bad states and of fairly good states impact the optimal consumption choice only to a limited degree. Thus, under probability weighting, optimal consumption is fairly unresponsive to a wide range of shocks to the economy. Finally, if small probabilities are sufficiently overweighted, our preference model generates an endogenous floor level. We explicitly derive the level of this floor on consumption. Probability weighting may thus explain why some individuals buy financial products with minimum guaranteed payments.

The optimal portfolio profile (i.e., the fraction of wealth invested in the risky stock as a function of the log state price density) displays a U-shaped pattern if probabilities are not distorted (see Chapter 2): the fraction of wealth invested in the risky stock is relatively low in economic scenarios where consumption is close to the reference level (i.e., intermediate economic scenarios). If the agent overweights probabilities of bad outcomes, then the fraction of assets invested in the risky stock is relatively low in not only intermediate economic scenarios but also bad economic scenarios.

We conduct a welfare analysis to investigate the impact of implementing alternative (suboptimal) consumption strategies on the agent's welfare. More specifically, all else equal, we compute welfare losses (in terms of the relative decline in certainty equivalent consumption) associated with incorrect probability weighting functions. We still assume that the (incorrectly specified) probability weighting functions are inverse S-shaped. Our results show that welfare losses can be (relatively) modest. This is so because, due to loss aversion, consumption already displays a 90° rotated S-shaped pattern, and the inverse S-shaped probability weighting functions, whether correctly or incorrectly specified, make this pattern even more pronounced.

To obtain the optimal consumption and portfolio choice, we first invoke the solution technique proposed by Schroder and Skiadas (2002). With this method, we are able to convert our consumption and portfolio choice model with endogenous updating into a dual consumption and portfolio choice model without endogenous updating. Then, we solve the dual problem by extending to our setting the quantile method introduced by Jin and Zhou (2008) and He and Zhou (2011). These authors show that in a setting with terminal wealth and no endogenous updating of the reference level, the agent's maximization problem can be transformed into a quantile formulation. As a result, conventional techniques (such as the Lagrange method) can be used to obtain the optimal solution. We adapt the quantile method to our setting with intertemporal consumption choice. By using the equivalence relationship between the dual model and the primal model, we finally obtain explicit closed-form solutions to our initial problem under consideration.

The remainder of this chapter is structured as follows. Section 3.2 describes the economy. The agent's preferences are introduced in Section 3.3. Section 3.4 formulates the agent's maximization problem. This section also outlines the dual technique and splits the dual problem into three related sub-problems. Section 3.5 solves the agent's maximization problem. An illustration of the optimal strategies is presented in Section 3.6. Section 3.7 concludes the chapter. Proofs are relegated to the Appendix.

3.2. The Economy

Let $T > 0$ be a fixed terminal time. The randomness in the economy is represented by a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. We define on this space a standard N-dimensional Brownian motion $\{Z_t\}_{t\in[0,T]}$. The filtration $\mathbb{F} \equiv \{\mathcal{F}_t\}_{t\in[0,T]}$ is the augmentation under $\mathbb{P}$ of the natural filtration generated by the standard Brownian motion $\{Z_t\}_{t\in[0,T]}$. Throughout, (in)equalities between random variables hold $\mathbb{P}$-almost surely.

We consider a financial market consisting of an instantaneously risk-free asset and N risky stocks. We assume that trading takes place continuously over $[0, T]$ and transaction costs are absent. The price of the risk-free asset, B, satisfies

$$\frac{\mathrm{d}B_t}{B_t} = r_t \,\mathrm{d}t, \qquad B_0 = 1. \tag{3.2.1}$$

The scalar-valued risk-free rate process, r, is assumed to be $\mathcal{F}_t$-progressively measurable and uniformly bounded. The N-dimensional vector of risky stock prices, S, obeys the following stochastic differential equation:

$$\frac{\mathrm{d}S_t}{S_t} = \mu_t \,\mathrm{d}t + \sigma_t \,\mathrm{d}Z_t, \qquad S_0 = \mathbf{1}_N. \tag{3.2.2}$$

Here, $\mathbf{1}_N$ represents an N-dimensional vector consisting of all ones. The N-dimensional mean rate of return process, μ, and the $(N \times N)$-matrix-valued volatility process, σ, are both assumed to be $\mathcal{F}_t$-progressively measurable and uniformly bounded.

We impose the following condition on σ_t. For some $\epsilon > 0$,

$$\zeta^\top \sigma_t \sigma_t^\top \zeta \geq \epsilon ||\zeta||^2, \qquad \text{for all } \zeta \in \mathbb{R}^N, \tag{3.2.3}$$

where $\top$ is the transpose sign. This condition implies that σ_t is invertible and bounded. The $\mathcal{F}_t$-progressively measurable market price of risk process, λ, satisfies

$$\sigma_t \lambda_t \equiv \mu_t - r_t \mathbf{1}_N. \tag{3.2.4}$$

The unique positive-valued state price density process, M, is defined as follows (see, e.g., Karatzas and Shreve, 1998):

$$M_t \equiv \exp\left\{ -\int_0^t r_s \,\mathrm{d}s - \int_0^t \lambda_s^\top \,\mathrm{d}Z_s - \frac{1}{2}\int_0^t ||\lambda_s||^2 \,\mathrm{d}s \right\}. \tag{3.2.5}$$

The economy consists of a single agent endowed with initial wealth $W_0 \geq 0$. The agent chooses an $\mathcal{F}_t$-progressively measurable N-dimensional portfolio process π (representing the dollar amounts invested in the N risky stocks) and an $\mathcal{F}_t$-progressively measurable consumption process c in order to maximize the CPT value of consumption.[36] We impose the following integrability conditions:

$$\int_0^T \pi_t^\top \sigma_t \sigma_t^\top \pi_t \,\mathrm{d}t < \infty, \qquad \int_0^T \left| \pi_t \left(\mu_t - r_t \mathbf{1}_N \right) \right| \mathrm{d}t < \infty, \qquad \mathbb{E}\left[\int_0^T |c_t|^2 \,\mathrm{d}t \right] < \infty. \tag{3.2.6}$$

[36] We introduce the agent's preferences in Section 3.3.

The wealth process, W, satisfies the following dynamic budget constraint:

$$\mathrm{d}W_t = \left(r_t W_t + \pi_t^\top \sigma_t \lambda_t - c_t\right) \mathrm{d}t + \pi_t^\top \sigma_t \, \mathrm{d}Z_t, \qquad W_0 \geq 0 \text{ given.} \tag{3.2.7}$$

A consumption-portfolio pair (c, π) is said to be admissible if the associated wealth process is uniformly bounded from below.

3.3. Preferences

This section describes the preferences. Denote by θ_t the agent's reference level at time t. The instantaneous preferences are defined over gains and losses relative to this reference level. Inspired by CPT, we assume that the instantaneous preferences over gains and losses $\widehat{c}_t \equiv c_t - \theta_t$ are given by

$$V\left(\widehat{c}_t\right) \equiv V_+\left(\widehat{c}_t^{\,+}\right) - V_-\left(\widehat{c}_t^{\,-}\right). \tag{3.3.1}$$

Here $V_+\left(\widehat{c}_t^{\,+}\right)$ denotes the CPT value derived from gains $\widehat{c}_t^{\,+} \equiv \max\{\widehat{c}_t, 0\}$ and $V_-\left(\widehat{c}_t^{\,-}\right)$ stands for the CPT value derived from losses $\widehat{c}_t^{\,-} \equiv -\min\{\widehat{c}_t, 0\}$. More specifically,

$$V_+\left(\widehat{c}_t^{\,+}\right) \equiv \int_0^\infty v_+(x) \, \mathrm{d}\left[-w_+\left(1 - F_{\widehat{c}_t^{\,+}}(x); \vartheta_t\right)\right], \tag{3.3.2}$$

$$V_-\left(\widehat{c}_t^{\,-}\right) \equiv \int_0^\infty v_-(x) \, \mathrm{d}\left[-w_-\left(1 - F_{\widehat{c}_t^{\,-}}(x); \vartheta_t\right)\right]. \tag{3.3.3}$$

Equations (3.3.2) and (3.3.3) show that the agent's preferences consist of various elements: two probability weighting functions $w_+(\cdot; \vartheta_t)$ and $w_-(\cdot; \vartheta_t)$ (one for gains and one for losses), an endogenous reference level θ_t, and two instantaneous utility functions $v_+(\cdot)$ and $v_-(\cdot)$ (one for gains and one for losses). Here ϑ_t is a vector of (time-varying) parameters affecting the shape of the probability weighting functions. Indeed, since uncertainty changes over time, the agent may want to change the shape of the probability weighting functions as time passes.

3.3.1. Probability Weighting Functions

The probability weighting function $w_+(\cdot;\vartheta_t)$ transforms the decumulative distribution function $1 - F_{\widehat{c}_t^+}(\cdot)$ of gains $\widehat{c}_t^+$ whereas the probability weighting function $w_-(\cdot;\vartheta_t)$ transforms the decumulative distribution function $1 - F_{\widehat{c}_t^-}(\cdot)$ of losses $\widehat{c}_t^-$. Throughout, we impose the following conditions on $w_+(\cdot;\vartheta_t)$ and $w_-(\cdot;\vartheta_t)$:

Assumption 1. *Let Θ be the parameter space. For all $t \in [0, T]$ and all $\vartheta_t \in \Theta$, $w_+(\cdot;\vartheta_t)$ and $w_-(\cdot;\vartheta_t) : [0,1] \mapsto [0,1]$ are strictly increasing and differentiable, with $w_+(0;\vartheta_t) = w_-(0;\vartheta_t) = 0$ and $w_+(1;\vartheta_t) = w_-(1;\vartheta_t) = 1$.*

If $w_+(\cdot;\vartheta_t)$ is equal to the identity function, then equation (3.3.2) reduces to the ordinary expectation $\mathbb{E}\left[v_+\left(\widehat{c}_t^+\right)\right]$. Hence, expected utility maximization emerges as a special case of (3.3.1).

3.3.2. Utility Functions

This section introduces the utility function for gains $v_+(\cdot)$ and the utility function for losses $v_-(\cdot)$. Following the CPT literature (see, e.g., Tversky and Kahneman, 1992), we assume that

$$v_+\left(\widehat{c}_t^+\right) = \left(\widehat{c}_t^+\right)^{\gamma_2}, \tag{3.3.4}$$

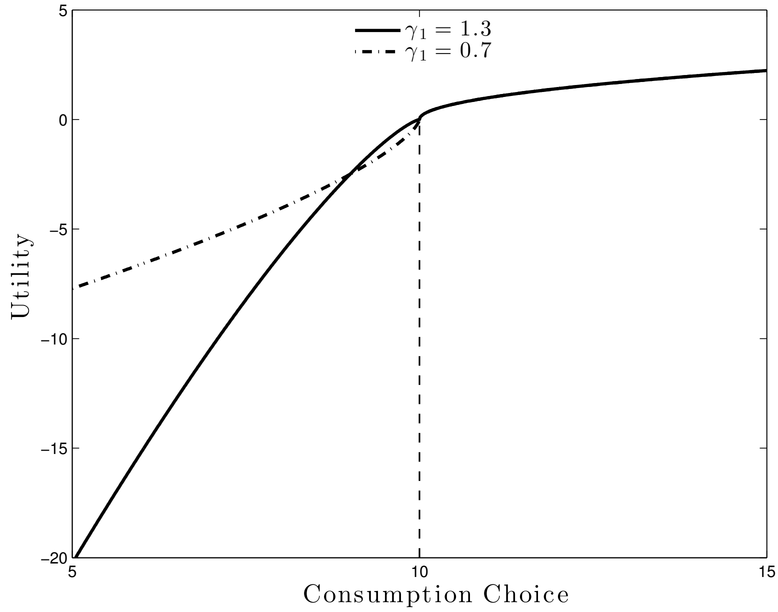
$$v_-\left(\widehat{c}_t^-\right) = \kappa\left(\widehat{c}_t^-\right)^{\gamma_1}, \tag{3.3.5}$$

where $\gamma_1 > 0$ and $\gamma_2 \in (0,1)$ are curvature parameters, and $\kappa \geq 1$ denotes the loss aversion index. Figure 3.1 shows the two-part power utility function

$$v\left(\widehat{c}_t\right) \equiv v_+\left(\widehat{c}_t^+\right) - v_-\left(\widehat{c}_t^-\right) \tag{3.3.6}$$

for $\gamma_1 = 1.3$ (solid line) and $\gamma_1 = 0.7$ (dash-dotted line). The figure shows that the two-part power utility function has a kink at the reference level, even in the case of $\kappa = 1$.

Figure 3.1.
Illustration of the two-part power utility function



The figure shows the two-part power utility function for $\gamma_1 = 1.3$ (solid line) and $\gamma_1 = 0.7$ (dash-dotted line). The agent's reference level is set equal to 10, the loss aversion index κ to 2.5 and γ_2 to 0.5.

3.3.3. Reference Level

This section describes the dynamics of the reference level θ_t. Motivated by the literature on internal habit formation (see, e.g., Constantinides, 1990; Detemple and Zapatero, 1992; Detemple and Karatzas, 2003), we assume that the agent's reference level satisfies

$$d\theta_t = (\beta c_t - \alpha\theta_t)\, dt, \qquad \theta_0 \geq 0 \text{ given.} \tag{3.3.7}$$

Here θ_0 represents the agent's initial reference level, $\alpha \geq 0$ indexes the rate at which the reference level depreciates over time, and $\beta \geq 0$ measures the sensitivity of the current reference level to current consumption. We can explicitly write the agent's reference level as follows ($s \geq t$):

$$\theta_s = \beta \int_t^s \exp\{-\alpha(s-u)\}\, c_u\, du + \exp\{-\alpha(s-t)\}\, \theta_t. \tag{3.3.8}$$

The first component on the right-hand side of equation (3.3.8) is an exponentially weighted integral of the agent's *own* past consumption choices. Hence, the agent's reference level is backward-looking and not forward-looking as in Kőszegi and Rabin (2006, 2007). The second component does not depend on the agent's past consumption choices and decreases exponentially at a rate of α.

3.4. Problem Formulation

This section formulates the agent's maximization problem. The agent aims to maximize

$$\int_0^T e^{-\delta t} V\left(c_t - \theta_t\right) \mathrm{d}t, \tag{3.4.1}$$

over all admissible consumption-portfolio pairs (c, π) subject to the dynamic budget constraint (3.2.7) and the reference level process (3.3.7). Here δ stands for the subjective rate of time preference. By virtue of the martingale approach (Pliska, 1986; Karatzas et al., 1987; Cox and Huang, 1989, 1991), we can transform the dynamic consumption and portfolio choice problem into the following equivalent static problem:

$$\begin{aligned} &\underset{c}{\text{Maximize}} && \int_0^T e^{-\delta t} V\left(c_t - \theta_t\right) \mathrm{d}t \\ &\text{subject to} && \mathbb{E}\left[\int_0^T M_t c_t \,\mathrm{d}t\right] \leq W_0, \qquad \mathrm{d}\theta_t = \left(\beta c_t - \alpha \theta_t\right) \mathrm{d}t. \end{aligned} \tag{3.4.2}$$

The optimal portfolio strategy π_t^* is determined in such a way that it finances the optimal consumption strategy c_t^*.

3.4.1. A Dual Problem

This section transforms the agent's maximization problem (3.4.2) into a *dual* (equivalent) maximization problem. Specifically, by invoking the method used in Schroder and Skiadas (2002), we can transform the agent's maximization problem (3.4.2) into the following dual problem:

$$\begin{aligned} &\underset{\widehat{c}}{\text{Maximize}} \quad \int_0^T e^{-\delta t} V\left(\widehat{c}_t\right) \mathrm{d}t \\ &\text{subject to} \quad \mathbb{E}\left[\int_0^T \widehat{M}_t \widehat{c}_t \, \mathrm{d}t\right] \leq \widehat{W}_0 \left(1+\beta A_0\right), \qquad \mathrm{d}\widehat{\theta}_t = \left(\beta\, \widehat{c}_t - (\alpha-\beta)\widehat{\theta}_t\right) \mathrm{d}t. \end{aligned} \tag{3.4.3}$$

Here the dual state price density $\widehat{M}_t$ and dual wealth $\widehat{W}_t$ are defined as follows:

$$\widehat{M}_t \equiv M_t \left(1+\beta A_t\right), \tag{3.4.4}$$

$$\widehat{W}_t \equiv \frac{W_t - A_t \widehat{\theta}_t}{1+\beta A_t}, \tag{3.4.5}$$

with

$$A_t \equiv \frac{1}{M_t} \mathbb{E}_t \left[\int_t^T M_s \exp\left\{-\left(\alpha-\beta\right)\left(s-t\right)\right\} \mathrm{d}s\right]. \tag{3.4.6}$$

The dual reference level is given by $(s \geq t)$

$$\widehat{\theta}_s = \beta \int_t^s \exp\left\{-\left(\alpha-\beta\right)\left(s-u\right)\right\} \widehat{c}_u \, \mathrm{d}u + \exp\left\{-\left(\alpha-\beta\right)\left(s-t\right)\right\} \widehat{\theta}_t. \tag{3.4.7}$$

and equals the (primal) reference level θ_s. The agent invests his dual (or surplus) wealth $\widehat{W}_t$ in a dual financial market. This dual market is characterized by the dual risk-free rate $\widehat{r}_t$, the dual volatility $\widehat{\sigma}_t$ and the dual market price of risk $\widehat{\lambda}_t$:

$$\widehat{r}_t \equiv \beta + \frac{r_t - \alpha\beta A_t}{1+\beta A_t}, \tag{3.4.8}$$

$$\widehat{\sigma}_t \equiv \sigma_t, \tag{3.4.9}$$

$$\widehat{\lambda}_t \equiv \lambda_t - \frac{\beta}{1+\beta A_t} \int_t^T \exp\left\{-(\alpha-\beta)(s-t)\right\} P_{t,s} \Psi_{t,s} \, \mathrm{d}s, \tag{3.4.10}$$

where $P_{t,s}$ represents price at time t of a zero-coupon bond that matures at time $s \geq t$, and $\Psi_{t,s}$ denotes the volatility at time t of the instantaneous return on a zero-coupon bond with maturity date $s \geq t$.

The optimal dual portfolio choice $\widehat{\pi}_t^*$ is determined such that it finances the optimal dual consumption choice $\widehat{c}_t^*$. The optimal dual reference level $\widehat{\theta}_t^*$ can be computed from substituting the optimal past consumption choices into (3.4.7). The next proposition follows from Schroder and Skiadas (2002).

Proposition 5. *Denote by $\widehat{c}_t^*$ the optimal dual consumption choice, by $\widehat{\theta}_t^*$ the optimal dual reference level, by $\widehat{W}_t^*$ optimal dual wealth, and by $\widehat{\pi}_t^*$ the optimal dual portfolio choice. Then:*

- *The optimal consumption for the agent at time $0 \le t \le T$ is given by*

$$c_t^* = \widehat{c}_t^* + \widehat{\theta}_t^*.$$

- *The optimal wealth for the agent at time $0 \le t \le T$ is given by*

$$W_t^* = \widehat{W}_t^* + \beta A_t \widehat{W}_t^* + A_t \widehat{\theta}_t^*.$$

- *The optimal portfolio choice for the agent at time $0 \le t \le T$ is given by*

$$\pi_t^* = \widehat{\pi}_t^* + \beta A_t \widehat{\pi}_t^* + \left(\beta \widehat{W}_t^* + \widehat{\theta}_t^*\right) (\widehat{\sigma}_t)^{-1} \int_t^T \exp\left\{-(\alpha - \beta)(s-t)\right\} P_{t,s} \Psi_{t,s} \, \mathrm{d}s.$$

3.4.2. Three Related Sub-Problems

Jin and Zhou (2008) explore a problem with a similar structure as our dual problem (3.4.3). However, they do not consider intertemporal consumption choice. Jin and Zhou (2008) show that their problem can be solved by splitting it into three sub-problems.

We define the first sub-problem (called the gain part problem) as follows:

$$\begin{aligned} &\underset{\widehat{c}^+}{\text{Maximize}} && \int_0^T e^{-\delta t} V_+\left(\widehat{c}_t^+\right) \mathrm{d}t \\ &\text{subject to} && \mathbb{E}\left[\int_0^T \widehat{M}_t \widehat{c}_t^+ \, \mathrm{d}t\right] \le \widehat{W}_+ \left(1 + \beta A_0\right), \qquad \widehat{c}_t^+ = 0 \text{ if } \widehat{M}_t \ge \xi_t. \end{aligned} \tag{3.4.11}$$

The gain part problem is parameterized by $\widehat{W}_+ \ge \widehat{W}_0$ (i.e., initial wealth needed to finance gains) and the process $\{\xi_t\}$. The optimal value $\mathcal{V}_+\left(\widehat{W}_+, \{\xi_t\}\right)$ is defined to be the supremum of (3.4.11).[37]

[37] If $\widehat{W}_+ > 0$ and $\xi_t = 0$ for every $t \in [0, T]$, then there is no feasible solution and $\mathcal{V}_+\left(\widehat{W}_+, \{\xi_t\}\right) = -\infty$.

The second sub-problem (called the loss part problem) is defined as follows:

$$\begin{aligned} &\underset{\widehat{c}^-}{\text{Minimize}} \quad \int_0^T e^{-\delta t} V_-\left(\widehat{c}_t^-\right) \mathrm{d}t \\ &\text{subject to} \quad \mathbb{E}\left[\int_0^T \widehat{M}_t \widehat{c}_t^- \,\mathrm{d}t\right] \le \left(\widehat{W}_+ - \widehat{W}_0\right)(1+\beta A_0), \qquad \widehat{c}_t^- = 0 \text{ if } \widehat{M}_t < \xi_t. \end{aligned} \tag{3.4.12}$$

The loss-part problem is also parameterized by $\widehat{W}_+ \ge \widehat{W}_0$ and the process $\{\xi_t\}$. The optimal value $\mathcal{V}_-\left(\widehat{W}_+, \{\xi_t\}\right)$ is now defined to be the infimum of (3.4.12).[38]

Finally, the last problem amounts to finding the 'best' $\widehat{W}_+$ and $\{\xi_t\}$:

$$\begin{aligned} &\underset{\widehat{W}_+, \{\xi_t\}}{\text{Maximize}} \quad \mathcal{V}_+\left(\widehat{W}_+, \{\xi_t\}\right) - \mathcal{V}_-\left(\widehat{W}_+, \{\xi_t\}\right) \\ &\text{subject to} \quad \widehat{W}_+ \ge \widehat{W}_0. \end{aligned} \tag{3.4.13}$$

The next proposition is adapted from Jin and Zhou (2008).

Proposition 6. *Suppose that* $\left(\widehat{W}_+^*, \{\xi_t^*\}\right)$ *is optimal for problem* (3.4.13), $\left(\widehat{c}_t^+\right)^*$ *is optimal for problem* (3.4.11) *with parameters* $\left(\widehat{W}_+^*, \{\xi_t^*\}\right)$, *and* $\left(\widehat{c}_t^-\right)^*$ *is optimal for problem* (3.4.12) *with parameters* $\left(\widehat{W}_+^*, \{\xi_t^*\}\right)$. *Then* $\widehat{c}_t^* = \left(\widehat{c}_t^+\right)^* - \left(\widehat{c}_t^-\right)^*$ *is optimal for problem* (3.4.3).

3.5. Solving the Problem

3.5.1. Quantile Method

This section demonstrates how we can convert the gain part problem (3.4.11) into a *quantile* maximization problem.[39] In the quantile formulation, the agent chooses the *quantile function* (i.e., inverse cumulative distribution function) of dual consumption. After changing the agent's decision variable from dual consumption to the quantile function of dual consumption, the agent's preference measure reduces to an ordinary linear expectation. Hence, conventional techniques (such as the Lagrange method) can

[38] If $\widehat{W}_+ \neq \widehat{W}_0$ and $\xi_t = \infty$ for every $t \in [0, T]$, then there is no feasible solution and $\mathcal{V}_-\left(\widehat{W}_+, \{\xi_t\}\right) = +\infty$.

[39] The loss part problem (3.4.12) can be converted into a quantile *minimization* problem.

be used to obtain the optimal dual consumption choice.[40] He and Zhou (2011) give a systematic account of the quantile method. The quantile method relies upon the following three crucial assumptions:

Assumption 2. *The agent's preference measure is law-invariant; that is, if $X \overset{d}{\sim} Y$, then $V(X) = V(Y)$.*

Assumption 3. *The agent is strictly better off with more initial dual wealth.*

Assumption 4. *The dual state price density admits no atoms; that is $\mathbb{P}\left\{\widehat{M}_t = a\right\} = 0$ for all $a \in \mathbb{R}^+$.*

The agent's preference measure in our setting is clearly law-invariant. Assumption 3 holds true if the probability weighting function is strictly increasing (for a proof, see He and Zhou, 2011). The last assumption is satisfied if, e.g., the investment opportunity set is deterministic. The preference measure $V_+\left(\widehat{c}_t^+\right)$ (as defined in (3.3.2)) is equivalent to

$$\mathbb{E}\left[v_+\left(F_{\widehat{c}_t^+}^{-1}(Z)\right) w_+'(1-Z;\vartheta_t)\right] = \mathbb{E}\left[v_+\left(Q_{\widehat{c}_t^+}(Z)\right) w_+'(1-Z;\vartheta_t)\right]. \tag{3.5.1}$$

Here Z is *any* uniformly distributed random variable on $[0,1]$ and $Q_{\widehat{c}_t^+}(\cdot)$ corresponds to the quantile function[41] of positive dual consumption $\widehat{c}_t^+$. The equivalence between (3.5.1) and $V_+\left(\widehat{c}_t^+\right)$ follows from Assumption 2. In a similar fashion, $V_-\left(\widehat{c}_t^-\right)$ (as defined in (3.3.3)) is equivalent to

$$\mathbb{E}\left[v_-\left(F_{\widehat{c}_t^-}^{-1}(Z)\right) w_-'(1-Z;\vartheta_t)\right] = \mathbb{E}\left[v_-\left(Q_{\widehat{c}_t^-}(Z)\right) w_-'(1-Z;\vartheta_t)\right]. \tag{3.5.2}$$

Here $Q_{\widehat{c}_t^-}(\cdot)$ corresponds to the quantile function of negative dual consumption $\widehat{c}_t^-$. Let us denote by $F_{\widehat{M}_t}(\cdot)$ the cumulative distribution function of the dual state price density $\widehat{M}_t$. By Assumptions 3 and 4, we can rewrite the left-hand sides of the static dual budget constraints as follows (the static dual budget constraints are defined in (3.4.11)

[40] The optimal solution is typically time-inconsistent. Therefore, we assume that the agent solves the maximization problem at time 0 and then commits himself to follow the optimal solution during the rest of his life.

[41] A quantile function is non-decreasing and continuous from the left.

and (3.4.12)):

$$\mathbb{E}\left[\int_0^T \widehat{M}_t \widehat{c}_t^+ \,\mathrm{d}t\right] = \mathbb{E}\left[\int_0^T Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right) Q_{\widehat{c}_t^+}\left(Z_{\widehat{M}_t}\right) \mathrm{d}t\right], \tag{3.5.3}$$

$$\mathbb{E}\left[\int_0^T \widehat{M}_t \widehat{c}_t^- \,\mathrm{d}t\right] = \mathbb{E}\left[\int_0^T Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right) Q_{\widehat{c}_t^-}\left(Z_{\widehat{M}_t}\right) \mathrm{d}t\right]. \tag{3.5.4}$$

Here $Q_{\widehat{M}_t}(\cdot)$ is the quantile function of the dual state price density $\widehat{M}_t$ and $Z_{\widehat{M}_t} \equiv 1 - F_{\widehat{M}_t}\left(\widehat{M}_t\right)$. It follows that $Z_{\widehat{M}_t}$ is a uniformly distributed random variable on $[0,1]$. Equations (3.5.1) and (3.5.2) hold true for any uniformly distributed random variable Z, whereas equations (3.5.3) and (3.5.4) are only valid for one *particular* uniformly distributed random variable $Z_{\widehat{M}_t}$.

Define $\mathbb{Q}$ to be the set of all quantile functions. The gain part problem (3.4.11) is equivalent to the following quantile maximization problem:

$$\begin{aligned}
\underset{Q_{\widehat{c}_t^+}\in\mathbb{Q}}{\text{maximize}} \quad & \mathbb{E}\left[\int_0^T e^{-\delta t} v_+\left(Q_{\widehat{c}_t^+}\right) w_+'\left(1 - Z_{\widehat{M}_t}; \vartheta_t\right) \mathrm{d}t\right] \\
\text{subject to} \quad & \mathbb{E}\left[\int_0^T Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right) Q_{\widehat{c}_t^+} \,\mathrm{d}t\right] \le \widehat{W}_+ \left(1 + \beta A_0\right) \\
& Q_{\widehat{c}_t^+} = 0 \text{ if } Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right) \ge \xi_t.
\end{aligned} \tag{3.5.5}$$

The quantile maximization problem (3.5.5) is called the quantile formulation. The agent's decision variable is the quantile function of positive dual consumption. In a similar fashion, the loss part problem (3.4.12) is equivalent to

$$\begin{aligned}
\underset{Q_{\widehat{c}_t^-}\in\mathbb{Q}}{\text{minimize}} \quad & \mathbb{E}\left[\int_0^T e^{-\delta t} v_-\left(Q_{\widehat{c}_t^-}\right) w_-'\left(1 - Z_{\widehat{M}_t}; \vartheta_t\right) \mathrm{d}t\right] \\
\text{subject to} \quad & \mathbb{E}\left[\int_0^T Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right) Q_{\widehat{c}_t^-} \,\mathrm{d}t\right] \le \left(\widehat{W}_+ - \widehat{W}_0\right) \left(1 + \beta A_0\right) \\
& Q_{\widehat{c}_t^-} = 0 \text{ if } Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right) < \xi_t.
\end{aligned} \tag{3.5.6}$$

The next proposition is adapted from He and Zhou (2011).

Proposition 7. *Suppose that $Q^*_{\widehat{c}_t^+}$ is optimal for problem (3.5.5) and $Q^*_{\widehat{c}_t^-}$ is optimal for problem (3.5.6). Then $\left(\widehat{c}_t^+\right)^* = Q^*_{\widehat{c}_t^+}$ is optimal for problem (3.4.11) and $\left(\widehat{c}_t^-\right)^* = Q^*_{\widehat{c}_t^-}$ is optimal for problem (3.4.12).*

3.5.2. Solution Procedure

This section summarizes the solution technique for solving the agent's maximization problem (3.4.2). The solution procedure consists of the following steps:

1. Solve the quantile problems (3.5.5) and (3.5.6);
2. Use Proposition 7 to obtain $\left(\widehat{c}_t^+\right)^*$ and $\left(\widehat{c}_t^-\right)^*$;
3. Solve problem (3.4.13);
4. Use Proposition 6 to obtain the optimal dual consumption choice $\widehat{c}_t^*$;
5. Use Proposition 5 to obtain the optimal consumption choice c_t^*.

The following sections explore how to solve problems (3.5.5), (3.5.6) and (3.4.13).

3.5.2.1. Solving the Gain Part Problem

This section presents the solution to problem (3.5.5). We can apply standard techniques to obtain $Q^*_{\widehat{c}_t^+}$. Let us introduce the following assumption:[42]

Assumption 5. *The quantity*

$$\frac{Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right)}{w'_+\left(1 - Z_{\widehat{M}_t}; \vartheta_t\right)} \tag{3.5.7}$$

is non-increasing in $Z_{\widehat{M}_t} \in [0, 1]$.

Section 3.6 considers a class of probability weighting functions that satisfy Assumption 5. The optimal quantile function of positive dual consumption choice is now given by (see Appendix)

$$Q^*_{\widehat{c}_t^+} = \left[\frac{e^{\delta t} y Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right)}{\gamma_2 w'_+\left(1 - Z_{\widehat{M}_t}; \vartheta_t\right)}\right]^{\frac{1}{\gamma_2 - 1}} \mathbb{1}_{\left[Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right) \leq \xi_t\right]}. \tag{3.5.8}$$

Assumption 5 ensures that the quantile function $Q^*_{\widehat{c}_t^+}$ is non-increasing in $\widehat{M}_t$. The Lagrange multiplier y is chosen such that the static budget constraint holds with equality.

[42] We note that Assumption 5 can be relaxed (see Xia and Zhou, 2014; Xu, 2014).

Straightforward computations show that the Lagrange multiplier y is given by (substitute (3.5.8) into the dual budget constraint and solve for y)

$$y = \gamma_2 \left[\frac{\widehat{W}_+ (1 + \beta A_0)}{\int_0^T \varphi(\xi_t)\, \mathrm{d}t} \right]^{\gamma_2 - 1}. \tag{3.5.9}$$

Here

$$\varphi(\xi_t) \equiv \mathbb{E}\left\{ Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right) \left[\frac{e^{\delta t} Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right)}{w'_+\left(1 - Z_{\widehat{M}_t}; \vartheta_t\right)} \right]^{\frac{1}{\gamma_2 - 1}} \mathbb{1}_{\left[Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right) \leq \xi_t\right]} \right\}. \tag{3.5.10}$$

Hence, the optimal positive dual consumption choice can be written as follows (substitute (3.5.9) into (3.5.8)):

$$Q^*_{\widehat{c}_t^+} = \frac{\widehat{W}_+ (1 + \beta A_0)}{\int_0^T \varphi(\xi_t)\, \mathrm{d}t} \left[\frac{e^{\delta t} Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right)}{w'_+\left(1 - Z_{\widehat{M}_t}; \vartheta_t\right)} \right]^{\frac{1}{\gamma_2 - 1}} \mathbb{1}_{\left[Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right) \leq \xi_t\right]}. \tag{3.5.11}$$

The supremum of the gain part problem (3.4.11) is given by

$$\mathcal{V}_+\left(\widehat{W}_+, \{\xi_t\}\right) = \left(\widehat{W}_+ (1 + \beta A_0)\right)^{\gamma_2} \left(\int_0^T \varphi(\xi_t)\, \mathrm{d}t\right)^{1 - \gamma_2}. \tag{3.5.12}$$

3.5.2.2. Solving the Loss Part Problem

This section presents the optimal solution to the quantile minimization problem (3.5.6). The two-part power utility function (3.3.6) is convex below the agent's reference level if $\gamma_1 \leq 1$, and concave otherwise. The literature is inconclusive about the shape of the utility function below the reference level (see, e.g., Etchart-Vincent, 2004; Abdellaoui et al., 2005; Booij and van de Kuilen, 2009). Etchart-Vincent (2004) found that in the case of large payoffs, the majority of subjects preferred a concave utility function below the reference level. Therefore, the present chapter considers the case of a concave utility function in the loss domain ($\gamma_1 > 1$). In future work, we intend to investigate the case where the utility function is convex below the reference level. We can apply the Lagrange method to obtain $Q^*_{\widehat{c}_t^-}$. Let us introduce the following assumption:

Assumption 6. *The quantity*

$$\frac{Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right)}{w'_-\left(1 - Z_{\widehat{M}_t}; \vartheta_t\right)} \tag{3.5.13}$$

is non-increasing in $Z_{\widehat{M}_t} \in [0,1]$.

The optimal quantile function of negative dual consumption choice is given by (see Appendix)

$$Q^*_{\widehat{c}_t^-} = \left[\frac{e^{\delta t} y Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right)}{\kappa\gamma_1 w'_-\left(1 - Z_{\widehat{M}_t}; \vartheta_t\right)}\right]^{\frac{1}{\gamma_1 - 1}} \mathbb{1}_{\left[Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right) > \xi_t\right]}. \tag{3.5.14}$$

The Lagrange multiplier y is chosen such that the static budget constraint holds with equality. Straightforward computations show that the Lagrange multiplier y is given by (substitute (3.5.14) into the dual budget constraint and solve for y)

$$y = \kappa\gamma_1 \left[\frac{\left(\widehat{W}_+ - \widehat{W}_0\right)(1 + \beta A_0)}{\int_0^T \varsigma(\xi_t)\,\mathrm{d}t}\right]^{\gamma_1 - 1}. \tag{3.5.15}$$

Here

$$\varsigma(\xi_t) \equiv \mathbb{E}\left\{Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right)\left[\frac{e^{\delta t} Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right)}{w'_-\left(1 - Z_{\widehat{M}_t}; \vartheta_t\right)}\right]^{\frac{1}{\gamma_1 - 1}} \mathbb{1}_{\left[Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right) > \xi_t\right]}\right\}. \tag{3.5.16}$$

Hence, the optimal negative dual consumption choice can be written as follows (substitute (3.5.15) into (3.5.14)):

$$Q^*_{\widehat{c}_t^-} = \frac{\left(\widehat{W}_+ - \widehat{W}_0\right)(1 + \beta A_0)}{\int_0^T \varsigma(\xi_t)\,\mathrm{d}t}\left[\frac{e^{\delta t} Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right)}{w'_-\left(1 - Z_{\widehat{M}_t}; \vartheta_t\right)}\right]^{\frac{1}{\gamma_1 - 1}} \mathbb{1}_{\left[Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right) > \xi_t\right]}. \tag{3.5.17}$$

The infimum of the loss part problem (3.4.12) is given by

$$\mathcal{V}_-\left(\widehat{W}_+, \{\xi_t\}\right) = \kappa\left(\widehat{W}_+ - \widehat{W}_0\right)^{\gamma_1}(1 + \beta A_0)^{\gamma_1}\left(\int_0^T \varsigma(\xi_t)\,\mathrm{d}t\right)^{1 - \gamma_1}. \tag{3.5.18}$$

3.5.2.3. Optimal Dual Solution

To determine the optimal dual consumption choice, the agent needs to solve problem (3.4.13). Substituting (3.5.12) and (3.5.18) into (3.4.13) yields

$$\begin{aligned} \underset{\widehat{W}_+,\{\xi_t\}}{\text{Maximize}} \quad & \left(\widehat{W}_+ (1+\beta A_0)\right)^{\gamma_2} \left(\int_0^T \varphi(\xi_t)\,\mathrm{d}t\right)^{1-\gamma_2} \\ & - \kappa \left(\widehat{W}_+ - \widehat{W}_0\right)^{\gamma_1} (1+\beta A_0)^{\gamma_1} \left(\int_0^T \varsigma(\xi_t)\,\mathrm{d}t\right)^{1-\gamma_1} \\ \text{subject to} \quad & \widehat{W}_+ \geq \widehat{W}_0. \end{aligned} \tag{3.5.19}$$

This problem can be solved numerically.

3.6. Numerical Analysis

3.6.1. Assumptions and Key Parameter Values

The agent invest his wealth in a risk-free asset and a single risky stock. We assume a constant investment opportunity set. That is, $r_t = r$, $\sigma_t = \sigma$ and $\lambda_t = \lambda$. The equity risk premium $\sigma\lambda = \mu - r$ is set at 4%, the risk-free rate r at 1%, and the volatility of innovations to the risky stock price σ at 20%. These estimates are the same as those used by Gomes et al. (2008).

The terminal time T is set equal to 4, the agent's initial wealth to 100, the curvature parameter γ_1 to 1.2, the curvature parameter γ_2 to 0.7, the subjective rate of time preference to 0.01, and the loss aversion index κ to 2.5. The literature reposts that estimates of the (median) loss aversion index range from 1 to 5 (see, e.g., Abdellaoui et al., 2008). We choose the agent's initial reference level θ_0 such that it solves the following equation:

$$W_0 = \theta_0 \int_0^T \exp\{-rt\}\,\mathrm{d}t. \tag{3.6.1}$$

We can view θ_0 as the payment from a fixed annuity with a value of W_0. The parameters α and β are set equal to zero. The reference level is thus constant over time. The impact of an endogenous reference level on the agent's optimal choice is analyzed in Chapter 2.

3.6.2. Probability Weighting Functions

Inspired by Jin and Zhou (2008) and He and Zhou (2011), we define the derivatives of the probability weighting functions $w_+(\cdot;\vartheta_t)$ and $w_-(\cdot;\vartheta_t)$ as follows:

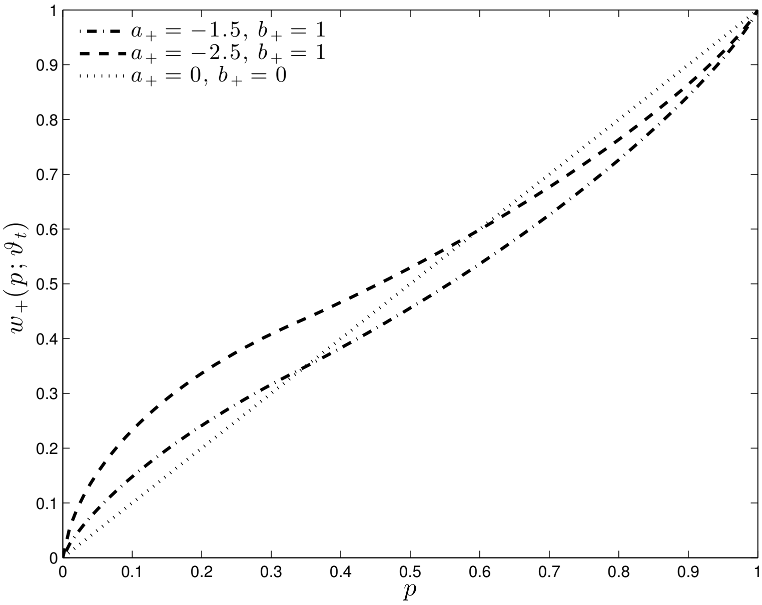
$$w'_+\left(F_{\widehat{M}_t}\left(\widehat{M}_t\right);\vartheta_t\right)=\begin{cases}k_+(\vartheta_t)\left(\widehat{M}_t\right)^{a_+}, & \text{if } \widehat{M}_t\leq Q_{\widehat{M}_t}(\bar{p}_+),\\ k_+(\vartheta_t)\left(Q_{\widehat{M}_t}(\bar{p}_+)\right)^{a_+-b_+}\left(\widehat{M}_t\right)^{b_+}, & \text{if } \widehat{M}_t> Q_{\widehat{M}_t}(\bar{p}_+),\end{cases} \tag{3.6.2}$$

$$w'_-\left(F_{\widehat{M}_t}\left(\widehat{M}_t\right);\vartheta_t\right)=\begin{cases}k_-(\vartheta_t)\left(\widehat{M}_t\right)^{a_-}, & \text{if } \widehat{M}_t\leq Q_{\widehat{M}_t}(\bar{p}_-);\\ k_-(\vartheta_t)\left(Q_{\widehat{M}_t}(\bar{p}_-)\right)^{a_--b_-}\left(\widehat{M}_t\right)^{b_-}, & \text{if } \widehat{M}_t> Q_{\widehat{M}_t}(\bar{p}_-).\end{cases} \tag{3.6.3}$$

Here $a_+\leq 0$, $a_-\leq 0$, $0\leq b_+\leq 1$, $0\leq b_-\leq 1$, $\bar{p}_+>0$ and $\bar{p}_->0$ are preference parameters. The parameter restrictions ensure that $w_+(\cdot;\vartheta_t)$ and $w_-(\cdot;\vartheta_t)$ satisfy Assumptions 1, 5 and 6. The expressions for $w_+(\cdot;\vartheta_t)$ and $w_-(\cdot;\vartheta_t)$ are given in the Appendix which also defines ϑ_t, $k_+(\vartheta_t)$ and $k_-(\vartheta_t)$. The parameter $\bar{p}_+$ is called the inflection point.[43] The probability weighting function $w_+(\cdot;\vartheta_t)$ is concave up to $\bar{p}_+$, and convex beyond $\bar{p}_+$. The parameter $a_+\leq 0$ determines the degree of concavity in the domain $0\leq p\leq\bar{p}_+$, while the parameter $0\leq b_+\leq 1$ determines the degree of convexity in the domain $\bar{p}_+\leq p\leq 1$. Figure 3.2 illustrates the probability weighting function $w_+(\cdot;\vartheta_t)$ for various sets of parameter values. The figure shows that the probability weighting function displays an inverse S-shaped pattern, consistent with CPT.

[43] It has been reported in the literature that $\bar{p}_+$ and $\bar{p}_-$ are about 1/3 (see, e.g., Wu and Gonzalez, 1996; Abdellaoui, 2000).

Figure 3.2.
Illustration of probability weighting function



The figure illustrates the probability weighting function $w_+(\cdot;\vartheta_t)$ for various sets of parameter values. We set ϑ_t equal to $(-0.28, 0.16)$, and $\bar{p}_+$ to $1/3$.

3.6.3. Optimal Consumption Choice

The optimal consumption choice for the agent at time $0 \le t \le T$ is given by (this follows from (3.6.2), (3.6.3), (3.5.8), (3.5.14), and Propositions 6 and 5)

$$c_t^* = \begin{cases} \theta_t^* + d_{1,t}\left(\widehat{M}_t\right)^{\frac{a_+-1}{1-\gamma_2}} & \text{if } \widehat{M}_t \le \min\left\{Q_{\widehat{M}_t}(\bar{p}_+), \xi_t\right\}, \\ \theta_t^* + d_{2,t}\left(\widehat{M}_t\right)^{\frac{b_+-1}{1-\gamma_2}} & \text{if } Q_{\widehat{M}_t}(\bar{p}_+) < \widehat{M}_t \le \xi_t, \\ \theta_t^* - d_{3,t}\left(\widehat{M}_t\right)^{\frac{a_--1}{1-\gamma_1}} & \text{if } \xi_t < \widehat{M}_t \le Q_{\widehat{M}_t}(\bar{p}_-), \\ \theta_t^* - d_{4,t}\left(\widehat{M}_t\right)^{\frac{b_--1}{1-\gamma_1}} & \text{if } \widehat{M}_t > \max\left\{Q_{\widehat{M}_t}(\bar{p}_-), \xi_t\right\}. \end{cases} \tag{3.6.4}$$

where

$$d_{1,t} \equiv \left[ye^{-\delta t}k_{+}\left(\vartheta_{t}\right)/\gamma_{2}\right]^{\frac{1}{1-\gamma_{2}}}, \tag{3.6.5}$$

$$d_{2,t} \equiv \left[ye^{-\delta t}k_{+}\left(\vartheta_{t}\right)\left(Q_{\widehat{M}_{t}}\left(\bar{p}_{+}\right)\right)^{a_{+}-b_{+}}/\gamma_{2}\right]^{\frac{1}{1-\gamma_{2}}}, \tag{3.6.6}$$

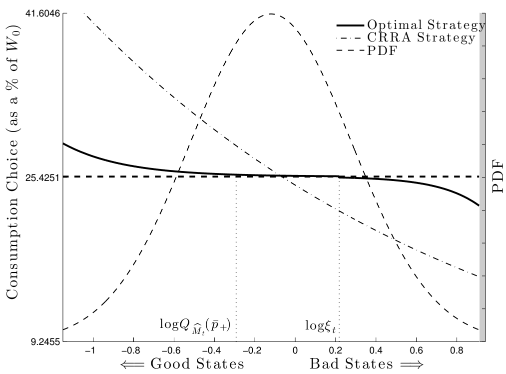
$$d_{3,t} \equiv \left[ye^{-\delta t}k_{-}\left(\vartheta_{t}\right)/\left(\kappa\gamma_{1}\right)\right]^{\frac{1}{1-\gamma_{1}}}, \tag{3.6.7}$$

$$d_{4,t} \equiv \left[ye^{-\delta t}k_{-}\left(\vartheta_{t}\right)\left(Q_{\widehat{M}_{t}}\left(\bar{p}_{-}\right)\right)^{a_{-}-b_{-}}/\left(\kappa\gamma_{1}\right)\right]^{\frac{1}{1-\gamma_{1}}}, \tag{3.6.8}$$

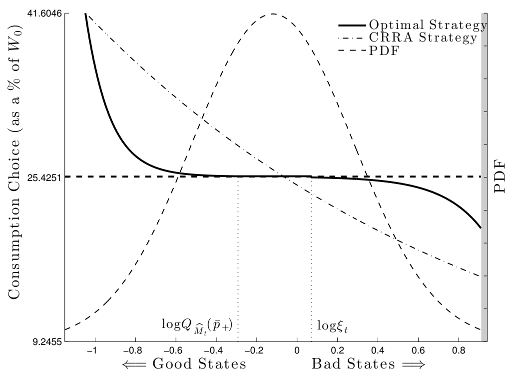
and θ_t^* denotes the optimal reference level implied by substituting the optimal past dual consumption choice into (3.4.7).

Equation (3.6.4) shows that the agent divides the states of the economy into two categories: good scenarios (low to intermediate state prices) and bad scenarios (high state prices). In good scenarios, consumption is larger than the reference level, while in bad scenarios, consumption is smaller than the reference level. The parameters a_+, b_+, a_- and b_- determine the sensitivity of consumption to the (dual) pricing kernel $\widehat{M}_t$. The sensitivity of consumption to the pricing kernel is (relatively) low in very bad scenarios, i.e., $\widehat{M}_t > \max\left\{Q_{\widehat{M}_t}\left(\bar{p}_-\right), \xi_t\right\}$, and in fairly good scenarios, i.e., $Q_{\widehat{M}_t}\left(\bar{p}_+\right) < \widehat{M}_t \leq \xi_t$. We observe that if b_- equals unity, then the agent consumes $\theta_t - d_{4,t}$ in very bad scenarios. In that case, consumption does not depend on the pricing kernel at all. Figure 3.3 shows the optimal consumption choice of an agent at time $t = 4$ as a function of the log dual pricing kernel for various sets of parameter values. The dashed-dotted lines represent the optimal 'Merton' consumption strategy (see Merton, 1969).

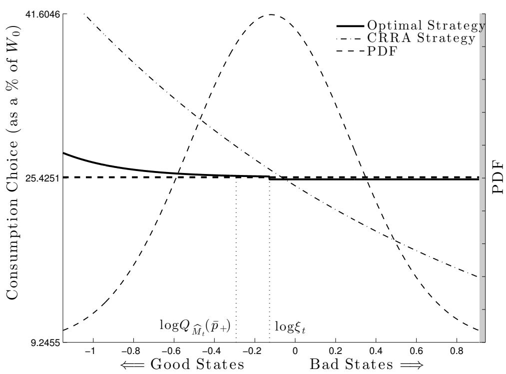
Figure 3.3.
Optimal consumption choice as function of the log dual pricing kernel



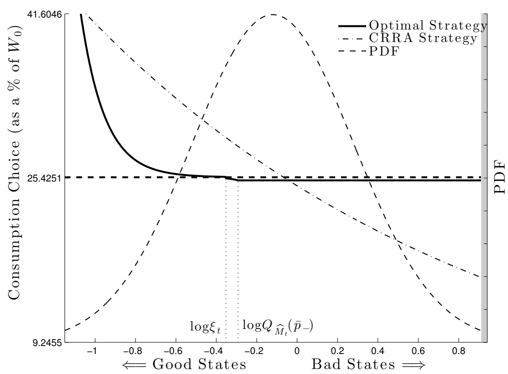
(a) No probability weighting



(b) $a_+ = -1.5$, $b_+ = 1$, $a_- = 0$ and $b_- = 0$



(c) $a_+ = 0$, $b_+ = 0$, $a_- = -2.5$ and $b_- = 1$



(d) $a_+ = -1.5$, $b_+ = 1$, $a_- = -2.5$ and $b_- = 1$

The figure illustrates the optimal consumption choice (expressed as a percentage of the agent's initial wealth W_0) of an agent at time $t = 4$ as a function of the log dual pricing kernel for various sets of parameter values. We set the inflection points $\bar{p}_+$ and $\bar{p}_-$ both equal to $1/3$. The dashed line corresponds to the reference level (expressed as a percentage of W_0). The dotted line shows the probability density function (PDF) of the current log dual pricing kernel conditional upon information available at time 0. The dash-dotted line illustrates the consumption choice (expressed as a percentage of W_0) of an agent with CRRA utility. The relative risk aversion coefficient is set equal to two.

3.6.3.1. Welfare Analysis

This section conducts a welfare analysis. We compute the welfare losses (in terms of the relative decline in certainty equivalent consumption) associated with implementing suboptimal consumption and portfolio strategies derived by solving the agent's problem on the basis of incorrect values of a_+, b_+, a_- and b_-.[44] We assume that the agent's optimal consumption and portfolio choice is characterized by the following 'true' parameter values: $a_+ = -1.5$, $b_+ = 1$, $a_- = -2.5$ and $b_- = 1$. Table 3.1 reports the welfare losses. This table shows that the welfare losses associated with incorrectly assuming incorrect parameter values are (relatively) small (all welfare losses are lower than 1%). Welfare losses are largest for the case where no probability weighting in the loss domain is applied (first column of Table 3.1).

Table 3.1.
Welfare losses

	(a_-, b_-)			
(a_+, b_+)	(0,0)	(-1,0.3)	(-1.5,0.7)	(-2.5,1)
(0,0)	0.5385	0.3173	0.2691	0.2582
(-0.5,0.3)	0.4846	0.2651	0.2086	0.1887
(-1,0.7)	0.4647	0.1970	0.1162	0.0869
(-1.5,1)	0.6571	0.1713	0.0214	0

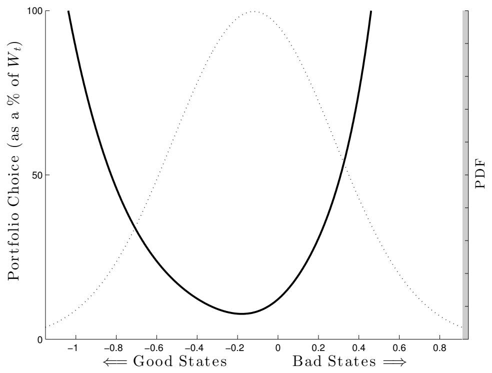
The table reports the welfare losses (in terms of the relative decline in certainty equivalent consumption) due to implementing suboptimal consumption and portfolio strategies derived on the basis of incorrect values of a_+, b_+, a_- and b_-. The numbers represent a percentage.

3.6.4. Optimal Portfolio Choice

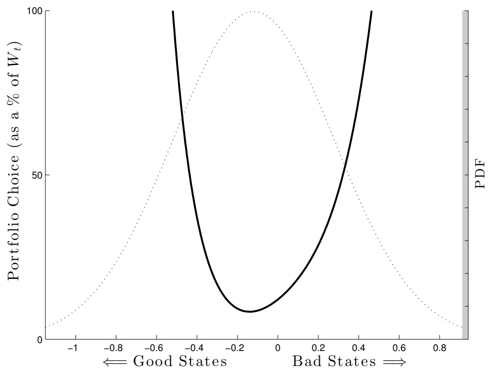
The optimal portfolio choice can be derived in closed-form by using standard hedging methods. Figure 3.4 illustrates the fraction of wealth invested in the risky stock (expressed as a percentage) of an agent at time $t = 4$ as a function of the log dual pricing kernel for various sets of parameter values. The figure shows that if consumption is close to the reference level, the fraction invested in the risky stock is relatively low. In addition, if the agent overweights unlikely extreme events (see the two panels at the bottom of the figure), the fraction of wealth invested in the risky stock tends to zero as the state of the economy worsens.

[44] We define the certainty equivalent of an uncertain consumption strategy to be the constant, certain consumption level that yields indifference to the uncertain consumption strategy.

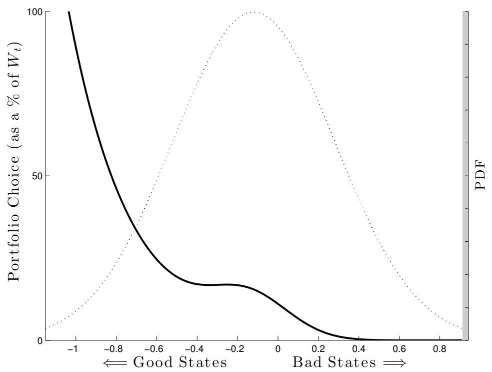
Figure 3.4.
Optimal portfolio choice as function of the log dual pricing kernel



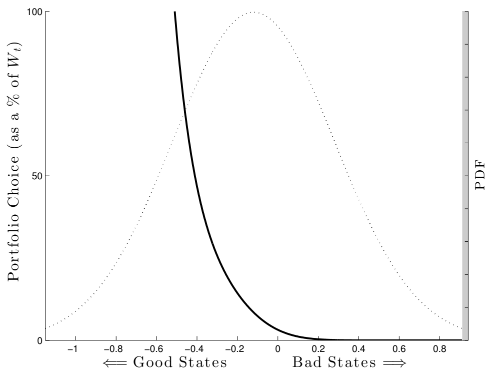
(a) No probability weighting



(b) $a_+ = -1.5$, $b_+ = 1$, $a_- = 0$ and $b_- = 0$



(c) $a_+ = 0$, $b_+ = 0$, $a_- = -2.5$ and $b_- = 1$



(d) $a_+ = -1.5$, $b_+ = 1$, $a_- = -2.5$ and $b_- = 1$

The figure illustrates the fraction of wealth invested in the risky stock (expressed as a percentage) of an agent at time $t = 4$ as a function of the log dual pricing kernel for various sets of parameter values. We set the inflection points $\bar{p}_+$ and $\bar{p}_-$ both equal to 1/3. The dotted line shows the probability density function (PDF) of the current log dual pricing kernel conditional upon information available at time 0.

3.7. Conclusion

We have explored dynamic consumption and portfolio choice of an agent with CPT preferences. Our agent is loss averse, endogenously updates his reference level over time, and distorts probabilities. We have shown that the optimal consumption profile displays a 90° rotated S-shaped pattern and that, if probabilities are sufficiently overweighted, the model generates a floor level on consumption.

3.8. Appendix

3.8.1. Proofs

Derivation of (3.5.8)

The Lagrangian $\mathcal{L}$ is given by

$$\begin{aligned}\mathcal{L} &= \mathbb{E}\left[\int_0^T e^{-\delta t} v_+\left(Q_{\widehat{c}_t^+}\right) w'_+\left(1 - Z_{\widehat{M}_t}; \vartheta_t\right) \mathrm{d}t\right] \\ &\quad + y\left\{\widehat{W}_+ (1+\beta A_0) - \mathbb{E}\left[\int_0^T Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right) Q_{\widehat{c}_t^+}\, \mathrm{d}t\right]\right\} \\ &= \int_0^T \mathbb{E}\left[e^{-\delta t} v_+\left(Q_{\widehat{c}_t^+}\right) w'_+\left(1 - Z_{\widehat{M}_t}; \vartheta_t\right) - y Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right) Q_{\widehat{c}_t^+}\right] \mathrm{d}t \\ &\quad + y\widehat{W}_+ (1+\beta A_0).\end{aligned}$$

Here $y \geq 0$ denotes the Lagrange multiplier associated with the static budget constraint. The agent maximizes $e^{-\delta t} v_+\left(Q_{\widehat{c}_t^+}\right) w'_+\left(1 - Z_{\widehat{M}_t}; \vartheta_t\right) - y Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right) Q_{\widehat{c}_t^+}$ subject to $Q_{\widehat{c}_t^+} = 0$ if $Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right) > \xi_t$. The optimal dual positive consumption choice $Q^*_{\widehat{c}_t^+}$ satisfies the following first-order optimality condition:

$$e^{-\delta t} \gamma_2 \left(Q_{\widehat{c}_t^+}\right)^{\gamma_2 - 1} w'_+\left(1 - Z_{\widehat{M}_t}; \vartheta_t\right) = y Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right).$$

After solving the first-order optimality condition, we obtain the following maximum:

$$Q^*_{\widehat{c}_t^+} = \left[\frac{e^{\delta t} y Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right)}{\gamma_2 w'_+\left(1 - Z_{\widehat{M}_t}; \vartheta_t\right)}\right]^{\frac{1}{\gamma_2 - 1}} \mathbb{1}_{\left[Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right) \leq \xi_t\right]}.$$

Derivation of (3.5.14)

The Lagrangian $\mathcal{L}$ is given by

$$
\begin{aligned}
\mathcal{L} &= \mathbb{E}\left[\int_0^T e^{-\delta t} v_-\left(Q_{\widehat{c}_t^-}\right) w_-'\left(1 - Z_{\widehat{M}_t}; \vartheta_t\right) \mathrm{d}t\right] \\
&\quad + y\left\{\left(\widehat{W}_+ - \widehat{W}_0\right)(1 + \beta A_0) - \mathbb{E}\left[\int_0^T Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right) Q_{\widehat{c}_t^-} \mathrm{d}t\right]\right\} \\
&= \int_0^T \mathbb{E}\left[e^{-\delta t} v_-\left(Q_{\widehat{c}_t^-}\right) w_-'\left(1 - Z_{\widehat{M}_t}; \vartheta_t\right) - y Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right) Q_{\widehat{c}_t^-}\right] \mathrm{d}t \\
&\quad + y\left(\widehat{W}_+ - \widehat{W}_0\right)(1 + \beta A_0).
\end{aligned}
$$

Here $y \geq 0$ denotes the Lagrange multiplier associated with the static budget constraint. The agent maximizes $e^{-\delta t} v_-\left(Q_{\widehat{c}_t^-}\right) w_-'\left(1 - Z_{\widehat{M}_t}; \vartheta_t\right) - y Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right) Q_{\widehat{c}_t^-}$ subject to $Q_{\widehat{c}_t^-} = 0$ if $Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right) \leq \xi_t$. The optimal dual negative consumption choice $Q^*_{\widehat{c}_t^-}$ satisfies the following first-order optimality condition:

$$
e^{-\delta t} \kappa \gamma_2 \left(Q_{\widehat{c}_t^-}\right)^{\gamma_1 - 1} w_-'\left(1 - Z_{\widehat{M}_t}; \vartheta_t\right) = y Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right).
$$

After solving the first-order optimality condition, we obtain the following maximum:

$$
Q^*_{\widehat{c}_t^-} = \left[\frac{e^{\delta t} y Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right)}{\kappa \gamma_2 w_-'\left(1 - Z_{\widehat{M}_t}; \vartheta_t\right)}\right]^{\frac{1}{\gamma_1 - 1}} \mathbb{1}_{\left[Q_{\widehat{M}_t}\left(1 - Z_{\widehat{M}_t}\right) > \xi_t\right]}.
$$

3.8.2. Probability Weighting Functions

This section specifies the probability weighting function for gains $w_+(\cdot; \vartheta_t)$ and the probability weighting function for losses $w_-(\cdot; \vartheta_t)$:

$$
w_+(p; \vartheta_t) = \begin{cases} k_+(\vartheta_t) f(a_+; \vartheta_t) \Phi\left(\Phi^{-1}(p) - a_+ \bar{\sigma}_t\right), & \text{if } 0 < p < \bar{p}_+; \\ \quad k_+(\vartheta_t)\left(Q_{\widehat{M}_t}(\bar{p}_+)\right)^{a_+ - b_+} f(b_+; \vartheta_t) & \\ \times \Phi\left(\Phi^{-1}(p) - b_+ \bar{\sigma}_t\right) + k_+(\vartheta_t) h_+(\vartheta_t), & \text{if } \bar{p}_+ \leq p < 1. \end{cases}
$$

$$w_-\left(p;\vartheta_t\right)=\begin{cases} k_-\left(\vartheta_t\right)f\left(a_-;\vartheta_t\right)\Phi\left(\Phi^{-1}\left(p\right)-a_-\bar{\sigma}_t\right), & \text{if } 0<p<\bar{p}_-; \\ \quad k_-\left(\vartheta_t\right)\left(Q_{\widehat{M}_t}\left(\bar{p}_-\right)\right)^{a_--b_-}f\left(b_-;\vartheta_t\right) & \\ \times\Phi\left(\Phi^{-1}\left(p\right)-b_-\bar{\sigma}_t\right)+k_-\left(\vartheta_t\right)h_-\left(\vartheta_t\right), & \text{if } \bar{p}_-\leq p<1. \end{cases}$$

Here,

$$\begin{aligned}
&\vartheta_t \equiv \left(\bar{\mu}_t,\bar{\sigma}_t^2\right) \\
&\bar{\mu}_t \equiv \log\widehat{M}_t-\left(\widehat{r}_t+\frac{1}{2}\lambda^2\right)t, \\
&\bar{\sigma}_t^2 \equiv \lambda^2 t, \\
&f\left(a_+;\vartheta_t\right) \equiv \exp\left\{a_+\bar{\mu}_t+\frac{1}{2}a_+^2\bar{\sigma}_t^2\right\}, \\
&h_+\left(\vartheta_t\right) \equiv h\left(\bar{p}_+,a_+,b_+;\vartheta_t\right) \\
&\qquad \equiv \Phi\left(\Phi^{-1}\left(\bar{p}_+\right)-a_+\bar{\sigma}_t\right)f\left(a_+;\vartheta_t\right) \\
&\qquad -\Phi\left(\Phi^{-1}\left(\bar{p}_+\right)-b_+\bar{\sigma}_t\right)\left(Q_{\widehat{M}_t}\left(\bar{p}_+\right)\right)^{a_+-b_+}f\left(b_+;\vartheta_t\right), \\
&k_+\left(\vartheta_t\right) \equiv k\left(\bar{p}_+,a_+,b_+;\vartheta_t\right) \equiv \left(h_+\left(\vartheta_t\right)+\left(Q_{\widehat{M}_t}\left(\bar{p}_+\right)\right)^{a_+-b_+}f\left(b_+;\vartheta_t\right)\right)^{-1}.
\end{aligned}$$

Chapter 4

How to Invest and Draw-Down Accumulated Wealth in Retirement? A Utility-Based Analysis[45]

This chapter explores how Baby Boomers should invest and draw-down their accumulated wealth over the rest of their lives. To answer this question we build a consumption and portfolio choice model with multiplicative internal habit formation and stochastic differential utility. We show analytically that after a wealth shock it is optimal to adjust both the level and future growth rates of consumption, implying gradual response of consumption to financial shocks. Furthermore, fostering the ability to keep catching up with the internal habit creates upward pressure on expected consumption growth. Welfare losses associated with popular alternative investment and draw-down strategies can be large.

4.1. Introduction

As the first wave of Baby Boomers moves into retirement, their future promises to be very different from that of their parents who enjoyed resilient Social Security and defined benefit pension plans. Retirees today face a very different challenge regarding retirement security, in that they are the first generation where retirement wealth was accumulated primarily in personal retirement accounts.[46] As their nest eggs will not automatically be

[45]This chapter is co-authored with Lans Bovenberg and Roger Laeven.

[46]The percentage of total U.S. retirement assets accounted for by individual retirement accounts and defined contribution pension plans rose from about 18% in 1974 to 54% in 2013 (Investment Company Institute, 2014).

annuitized, Baby Boomers thus confront the important question of how they should invest and draw-down their accumulated wealth over the rest of their lives. The objective of the present chapter is to analyze this question from the perspective of a utility maximizing individual.

Financial advisors commonly recommend to split the investment portfolio into 60% risky assets and 40% risk-free assets, and to draw-down 4 to 5% of retirement wealth per year (Polyak, 2005; Whitaker, 2005). Other popular draw-down strategies include the fixed benefit approach (i.e., the individual withdraws a specified dollar amount each year until his retirement wealth is depleted), and the remaining lifetime approach (i.e., the withdrawal fraction rises with the remaining lifetime); see, e.g., Dus, Maurer, and Mitchell (2005); Horneff, Maurer, Mitchell, and Dus (2008). However, these popular draw-down strategies are arguably ad hoc, and are typically neither founded upon nor corroborated by the individual's preferences (MacDonald, Jones, Morrison, Brown, and Hardy, 2013). Alternatively, retirees can buy annuities. But while fixed annuities are usually too expensive to be an attractive financial product (especially in low interest rate regimes), variable annuities often generate volatile fluctuations in payouts (see, e.g., Chai, Horneff, Maurer, and Mitchell, 2011; Maurer, Mitchell, Rogalla, and Kartashov, 2013b).[47]

Thus, the need for a utility-based approach to analyze the investment and draw-down strategies implemented by Baby Boomers is evident. Expected utility theory with constant relative risk aversion (CRRA) is the most commonly adopted preference model to derive an agent's optimal consumption and portfolio choice. As is well-known at least since Merton (1969, 1971) and Samuelson (1969), a CRRA agent fully absorbs a wealth shock into the level (and not future growth rates) of consumption. Under CRRA, the year-on-year volatility of consumption thus matches the year-on-year volatility of wealth. However, evidence of violations of the assumptions underlying CRRA utility – in our setting, the intertemporal independence assumption in particular – has led authors to seek for alternative models. The literature has put forward a variety of alternatives, perhaps most noticeably habit formation utility and (continuous-time) recursive utility

[47] Insurers have more recently developed variable annuities for which surpluses earned in good years support payouts in bad years (see, e.g., Guillén et al., 2006; Maurer, Rogalla, and Siegelin, 2013a; Maurer, Mitchell, Rogalla, and Siegelin, 2014). This rapidly growing form of variable annuities is, however, opaque and difficult to value.

or stochastic differential utility (SDU).[48] The present chapter proposes and analyzes a model with internal habit formation and SDU, and derives the resulting investment and draw-down strategies in closed-form.

Our contribution is three-fold. First, we build a rich consumption and portfolio choice model with multiplicative habit formation, an endogenous internal habit level, and SDU. Our general model encompasses many interesting special cases such as SDU without multiplicative internal habit formation, multiplicative internal habit formation without SDU, multiplicative habit formation with an external (deterministic) habit, and CRRA utility. Second, we develop an approximation method to accurately solve our general consumption and portfolio choice problem analytically. Third, we analyze the resulting optimal investment and draw-down strategies for a Baby Boomer, and conduct a welfare analysis. We now specify each of our contributions in more detail.

We assume that the agent derives utility from the ratio between consumption and the habit level. The ratio (or multiplicative) model of habit formation, first analyzed by Abel (1990), is the only model we know of that allows consumption to fall below the habit level while simultaneously maintaining the property of constant (i.e., state-independent) relative risk aversion. A number of authors consider an agent who derives utility from the difference – rather than the ratio – between consumption and the habit level.[49] The optimal consumption choice implied by the difference (or additive) model of habit formation (Constantinides, 1990) exceeds the habit level in each economic scenario. This addictive behavior of consumption is, however, doubtful (see, e.g., Detemple and Karatzas, 2003). Indeed, empirical evidence showing significant declines in consumption levels during recessions contradicts the addictive property. Furthermore, in the difference model of habit formation, relative risk aversion depends on (surplus) wealth. This may be undesirable from a normative point of view as it leads to very low equity holdings in bad economic scenarios.[50] Also, in our ratio model of habit formation, the portfolio

[48] The notion of SDU was introduced by Duffie and Epstein (1992) as a continuous-time limit of the preference models studied by Epstein and Zin (1989) and by Kreps and Porteus (1978). Life cycle models with Epstein-Zin preferences or internal habit formation have been widely studied in the literature. For Epstein-Zin preferences, see, e.g., Chacko and Viceira (2005); Gomes and Michaelides (2008); for the ratio model of habit formation, see, e.g., Gomes and Michaelides (2003).

[49] See, e.g., Constantinides (1990); Detemple and Zapatero (1991, 1992); Schroder and Skiadas (2002); Bodie, Detemple, Otruba, and Walter (2004); Munk (2008).

[50] Chapter 2 extends the difference model of habit formation to allow consumption to fall below the habit level. However, in the model of Chapter 2, relative risk aversion still depends on (surplus) wealth.

strategy is state-independent, and thus easy to implement. We assume that the habit level is a geometric (rather than an arithmetic) weighted average of the agent's own past consumption choices. Hence the habit level is internal to the agent and endogenously determined.[51] In the model of Abel (1990), the habit level depends only on consumption in the previous period.

Due to the endogeneity of the habit level, our consumption and portfolio choice problem cannot be solved in closed-form. By developing a linearization to the budget constraint, we are able to derive an analytical closed-form solution to the approximate optimization problem. Linearization of the agent's budget constraint is not uncommon in the economic literature; see, in a different context, e.g., Campbell and Mankiw (1991); Fuhrer (2000). Our approximation method is shown to be very accurate when consumption stays close to the habit level, and when the habit level responds slowly to consumption. Indeed, our numerical results show that the approximation error (measured in terms of the relative decline in certainty equivalent consumption) is typically of order 0.01.

Our results can be summarized as follows. First, we show analytically that after a wealth shock, it is optimal to adjust both the level and the future growth rates of consumption, implying gradual response of consumption to financial shocks. This justifies a mechanism for smoothing the change in consumption due to financial shocks. The parameters in our model (i.e., the coefficient of relative risk aversion, the strength of internal habit formation, and the deprecation rate of the habit level) have clear economic interpretations, controlling the features of the optimal strategy. The coefficient of relative risk aversion determines the effect of a wealth shock on the level of consumption. The less risk averse the agent, the larger the effect of a wealth shock on the level of consumption. The strength of internal habit formation determines the effect of a wealth shock on future growth rates of consumption. The larger the strength of internal habit formation, the larger the effect of a wealth shock on future growth rates of consumption. We also show that the lower the deprecation rate of the habit level, the longer it takes to fully absorb a wealth shock into current and future consumption.

[51] Corrado and Holly (2011) show that for the ratio model of habit formation, a geometric habit specification is more desirable than an arithmetic habit specification. In particular, they prove that under the geometric habit specification, overall utility decreases as the strength of internal habit formation increases.

Second, as the agent adjusts both the level and future growth rates of consumption after a shock, the year-on-year volatility of consumption is less than the year-on-year volatility of wealth. Thus, a risky investment portfolio does not automatically imply a high year-on-year volatility of consumption. This finding stands in sharp contrast to many popular portfolio and draw-down strategies (e.g., the remaining lifetime approach) where an increase in the risk of the investment portfolio directly translates into a higher year-on-year volatility of consumption. Furthermore, we show that the agent chooses to reduce the fraction of wealth invested in risk-bearing assets as the end of life approaches. Indeed, the agent has less time to absorb a wealth shock as he ages.

Third, we show that for a finitely-lived agent with a fixed lifetime, the expected growth rate of consumption increases with the strength of internal habit formation as well as with age.[52] If the agent were to live forever, the effects of the strength of internal habit formation and age on the expected growth rate of consumption would be absent. Indeed, in the case of an infinite horizon, the effect of a marginal change in consumption on future habit levels is independent of age, while in the more realistic case of a finite horizon, the effect of a marginal change in consumption on future habit levels decreases with age. Thus, in the finite horizon case, fostering the ability to keep catching up with the internal habit creates an upward pressure on expected consumption growth. That is, the agent prefers to postpone consumption because the utility gain of a marginal increase in consumption rises with age.

The elasticity of intertemporal substitution is in our base model, which combines CRRA utility with multiplicative internal habit formation, intimately related to the coefficient of relative risk aversion: the lower the degree of relative risk aversion, the higher the agent's willingness to engage in intertemporal substitution. In an extension of our model, we study the consumption and portfolio choice of an agent with preferences that combine SDU with multiplicative internal habit formation. This extended preference model allows us to disentangle the elasticity of intertemporal substitution from the coefficient of relative risk aversion while simultaneously maintaining the property of multiplicative internal habit formation. As a result, the change in the median growth

[52]We note that in the case of an uncertain lifetime and the absence of longevity insurance, survival probabilities create a downward pressure on expected consumption growth (see Yaari, 1965). In that case, expected consumption growth is determined by internal habit formation as well as the shape of the survival curve.

rate of consumption following a change in the interest rate is no longer related to the coefficient of relative risk aversion. Applying our linearization of the budget constraint, we are still able to derive the agent's consumption and portfolio choice under the extended model in closed-form. A model that combines SDU with multiplicative internal habit formation has, to the best of our knowledge, not yet been studied in existing literature. The closest to the current chapter in this respect is Schroder and Skiadas (1999), who analytically study SDU but do not consider multiplicative internal habit formation.

Finally, we conduct a welfare analysis in order to assess the impact of pursuing alternative suboptimal investment and draw-down strategies on the agent's welfare. More specifically, we compute welfare losses (in terms of the relative decline in certainty equivalent consumption) associated with implementing the remaining lifetime approach, the Merton approach (Merton, 1969) and the difference model of habit formation. Our results show that welfare losses can be large, especially when the agent exhibits a high degree of internal habit formation. We also show that welfare losses are typically larger for the remaining lifetime approach than for the Merton approach.

The remainder of this chapter is structured as follows. Section 4.2 describes the financial market, the preferences and the maximization problem. Section 4.3 presents the solution method used to solve the maximization problem. This solution method involves applying a linearization to the budget constraint. Section 4.4 derives *and* studies the agent's consumption and portfolio choice. This section also conducts a welfare analysis. Section 4.5 explores the consumption and portfolio choice of an agent with preferences that combine SDU with multiplicative internal habit formation. Section 4.6 studies the approximation error due to applying a linearization to the budget constraint. Finally, Section 4.7 provides concluding remarks. Proofs are relegated to the Appendix.

4.2. An Internal Habit Formation Model

4.2.1. The Financial Market

We consider a Black and Scholes financial market. Let $T > 0$ be a (possibly infinite) terminal time. The financial market consists of a money market account and a risky stock, which are traded continuously. The price of the money market account, i.e., B_t, evolves according to

$$\frac{\mathrm{d}B_t}{B_t} = r\,\mathrm{d}t. \tag{4.2.1}$$

Here r stands for the interest rate. The risky stock price S_t satisfies

$$\frac{\mathrm{d}S_t}{S_t} = (r + \lambda\sigma)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t. \tag{4.2.2}$$

Here λ denotes the equity risk premium per unit of risk (i.e., Sharpe ratio), σ stands for the diffusion parameter (i.e., stock return volatility), and W_t corresponds to a standard Brownian motion.

It is well-known that if we exclude arbitrage opportunities in this financial market, the pricing kernel (or stochastic discount factor) M_t satisfies (see, e.g., Karatzas and Shreve, 1998)

$$\frac{\mathrm{d}M_t}{M_t} = -r\,\mathrm{d}t - \lambda\,\mathrm{d}W_t. \tag{4.2.3}$$

In the numerical computations, we use the following financial market parameter values: the Sharpe ratio $\lambda = 20\%$, the risk-free rate $r = 1\%$, and the stock return volatility $\sigma = 20\%$. These parameter values are the same as those used by Gomes et al. (2008).

4.2.2. Preferences

The agent's preferences are represented by the multiplicative habit specification originally introduced by Abel (1990). More precisely, the instantaneous utility function is given by[53]

$$u\left(c_t, h_t\right) = v\left(\frac{c_t}{h_t}\right) = \frac{1}{1-\gamma}\left(\frac{c_t}{h_t}\right)^{1-\gamma}. \tag{4.2.4}$$

Here $\gamma > 0$ is the coefficient of relative risk aversion, c_t stands for consumption at time t, and h_t denotes the habit level at time t to which the agent compares consumption c_t. In the difference model of habit formation (Constantinides, 1990), relative risk aversion depends on (surplus) wealth. As a result, the optimal solution of the difference model

[53] If $\gamma = 1$, then $u\left(c_t, h_t\right) = v\left(c_t/h_t\right) = \log\left\{c_t/h_t\right\}$.

differs substantially from the optimal solution of the ratio model (see Section 4.4 for further details).

Inspired by Kozicki and Tinsley (2002) and Corrado and Holly (2011), the *log* habit level $\log h_t$ satisfies the following dynamic equation:

$$\mathrm{d}\log h_t = (\beta \log c_t - \alpha \log h_t)\,\mathrm{d}t, \qquad \log h_0 = 0. \tag{4.2.5}$$

Here $\log h_0$ denotes the *initial* log habit level. We normalize $\log h_0$ to zero (i.e., $h_0 = 1$). We thus measure initial retirement wealth, consumption and the habit level in terms of h_0.[54] The preference parameter $\alpha \geq 0$ represents the rate at which the log habit level depreciates. If α is small, then the log habit level exhibits a low degree of depreciation (or, equivalently, a high degree of persistence). The preference parameter $\beta \geq 0$ indexes the extent to which the current log habit level responds to current log consumption. If β is large, then current log consumption has a large impact on the log habit level. We impose the following restriction on the agent's preference parameters:

$$\alpha \geq \beta. \tag{4.2.6}$$

The parameter restriction (4.2.6) prevents the habit level from growing exponentially over time.

The habit level h_s conditional on information available at time t is explicitly given by[55]

$$h_s = \exp\left\{\int_t^s \beta \exp\{-\alpha(s-u)\}\log c_u\,\mathrm{d}u\right\} \times \exp\{\exp\{-\alpha(s-t)\}\log h_t\}. \tag{4.2.7}$$

Equation (4.2.7) shows that we can factor the habit level h_s into two components: one dependent upon the agent's consumption choices between time t and time s (i.e.,

[54]The utility function is *not* invariant to the unit of measurement. The agent should thus change the values of the preference parameters if the unit of measurement changes.

[55]Equation (4.2.7) is equivalent to:

$$h_s = (h_t)^{\exp\{-\alpha(s-t)\}} \times \prod_t^s (c_u)^{\beta\exp\{-\alpha(s-u)\}\,\mathrm{d}u}.$$

Here $\prod_t^s (c_u)^{\beta\exp\{-\alpha(s-u)\}\,\mathrm{d}u} = \lim_{\Delta u\to 0}\exp\left\{\sum_{i=t/\Delta u}^{s/\Delta u}\beta\exp\{-\alpha(s-i\Delta u)\}\log c_{i\Delta u}\Delta u\right\}$ denotes the geometric integral.

the stochastic component) and the other not (i.e., the deterministic component). The preference parameter β indexes the importance of the stochastic component relative to the deterministic component (given α). The stochastic component becomes less important as β decreases. Indeed, if β equals zero, then

$$\exp\{\beta \exp\{-\alpha(s-u)\} \log c_u\} = 1. \tag{4.2.8}$$

The stochastic component is an exponentially weighted *product* (and not an exponentially weighted sum) of the agent's consumption choices between time t and time s. The habit level thus depends more on consumption in the recent past than it depends on consumption in the distant past.

Corrado and Holly (2011) demonstrate that for the ratio model of habit formation, a specification in which the habit level is geometric in consumption is more desirable than a specification in which the habit level is arithmetic in consumption. In particular, they prove that under the geometric habit specification, overall utility decreases as the endogeneity parameter β increases, provided that consumption is larger than unity.[56] This property does not hold true in the arithmetic habit specification. Futhermore, the assumption of a geometric habit specification makes the agent's maximization problem analytically tractable.

4.2.3. Maximization Problem

This section formulates the agent's maximization problem. Denote by A_t the agent's retirement wealth at time t, and by π_t the fraction of wealth invested in the risky stock at time t. Wealth evolves according to

$$\mathrm{d}A_t = (r + \pi_t \lambda \sigma) A_t \,\mathrm{d}t - c_t \,\mathrm{d}t + \pi_t \sigma A_t \,\mathrm{d}W_t, \qquad A_0 \geq 0 \text{ given.} \tag{4.2.9}$$

Equation (4.2.9) is referred to as the agent's dynamic budget equation. This equation shows that the agent's retirement wealth equals the agent's initial retirement wealth, plus any gains from trading, minus cumulative consumption.

[56]Consumption is typically larger than unity because with internal habit formation, the agent has a tendency to postpone consumption (i.e., consumption tends to exhibit a positive expected growth rate); see Section 4.4.

The agent aims to maximize expected lifetime utility

$$U_0 = \mathbb{E}\left[\int_0^T e^{-\int_0^t \delta_u \,\mathrm{d}u} v\left(\frac{c_t}{h_t}\right) \mathrm{d}t\right], \tag{4.2.10}$$

over the set of all admissible consumption and portfolio strategies subject the agent's dynamic budget constraint (4.2.9) and the habit formation process (4.2.5).[57] Here $\mathbb{E}[\cdot]$ corresponds to the (unconditional) expectation operator, δ_u stands for the subjective rate of time preference at time u, and $v(c_t/h_t)$ represents the agent's instantaneous utility function (see equation (4.2.4)).[58]

We can, by virtue of the martingale approach (Pliska, 1986; Karatzas et al., 1987; Cox and Huang, 1989, 1991), transform the agent's maximization problem into the following equivalent problem:

$$\begin{aligned} &\underset{c}{\text{maximize}} \quad \mathbb{E}\left[\int_0^T e^{-\int_0^t \delta_u \,\mathrm{d}u} v\left(\frac{c_t}{h_t}\right) \mathrm{d}t\right] \\ &\text{subject to} \quad \mathbb{E}\left[\int_0^T M_t c_t \,\mathrm{d}t\right] \leq A_0, \qquad \mathrm{d}\log h_t = (\beta \log c_t - \alpha \log h_t)\,\mathrm{d}t. \end{aligned} \tag{4.2.11}$$

The optimal portfolio choice π_t^* is determined in such a way that it finances the optimal consumption choice c_t^*.

The optimal consumption choice c_t^* (if it exists) satisfies the following first-order optimality condition:

$$e^{-\int_0^t \delta_u \,\mathrm{d}u} \frac{1}{h_t}\left(\frac{c_t}{h_t}\right)^{-\gamma} - \frac{\beta}{c_t}\mathbb{E}_t\left[\int_t^T e^{-\int_0^s \delta_u \,\mathrm{d}u} e^{-\alpha(s-t)}\left(\frac{c_s}{h_s}\right)^{1-\gamma} \mathrm{d}s\right] = yM_t. \tag{4.2.12}$$

[57] For the definition of admissible consumption and portfolio strategies, see, e.g., Karatzas and Shreve (1998).

[58] Alternatively, we can view δ_u as the sum of the subjective rate of time preference at time u and the force of mortality at time u. More specifically, in the case of deterministic mortality risk, expected lifetime utility is given by (see Yaari, 1965)

$$\begin{aligned} U_0 &= \mathbb{E}\left[\int_0^T e^{-\int_0^t \rho_u \,\mathrm{d}u} \mathbb{1}_{[t\leq D]} v\left(\frac{c_t}{h_t}\right) \mathrm{d}t\right] = \mathbb{E}\left[\int_0^T e^{-\int_0^t \rho_u \,\mathrm{d}u} {}_t p_x v\left(\frac{c_t}{h_t}\right) \mathrm{d}t\right] \\ &= \mathbb{E}\left[\int_0^T e^{-\int_0^t \rho_u \,\mathrm{d}u} e^{-\int_0^t \mu_{x+u} \,\mathrm{d}u} v\left(\frac{c_t}{h_t}\right) \mathrm{d}t\right] = \mathbb{E}\left[\int_0^T e^{-\int_0^t \delta_u \,\mathrm{d}u} v\left(\frac{c_t}{h_t}\right) \mathrm{d}t\right]. \end{aligned}$$

Here ρ_u represents the subjective rate of time preference at time u, D is the stochastic date of death, ${}_t p_x$ denotes the probability that an agent aged x at time 0 will survive to time $x+t$, and μ_{x+u} is the deterministic force of mortality at age $x+u$.

Here $y \geq 0$ denotes the Lagrange multiplier. The left-hand side of equation (4.2.12) represents marginal utility, whereas the right-hand side denotes marginal cost. We can decompose marginal utility into two components: the first representing the effect of an increase in consumption on current instantaneous utility, and the second representing the effect of an increase in consumption on future instantaneous utilities. We cannot obtain the optimal consumption choice c_t^* in closed-form due to the presence of the conditional expectation operator $\mathbb{E}_t[\cdot]$ in the second component. The next section presents an approximate problem to problem (4.2.11) that can be solved analytically.

4.3. The Solution Method

4.3.1. Applying a Change of Variable

By applying a change of variable, we can transform the agent's maximization problem (4.2.11) into an equivalent *dual* problem. Denote by $\widehat{c}_t$ the ratio between the agent's consumption choice and the habit level; that is,

$$\widehat{c}_t \equiv \frac{c_t}{h_t}. \tag{4.3.1}$$

We refer to $\widehat{c}_t$ as the agent's *dual consumption choice*. We can express the dynamics of the log habit level in terms of the agent's log dual consumption choice $\log \widehat{c}_t$ (substitute $\log c_t = \log h_t + \log \widehat{c}_t$ into equation (4.2.5)):

$$\begin{aligned} \mathrm{d} \log h_t &= (\beta [\log h_t + \log \widehat{c}_t] - \alpha \log h_t) \, \mathrm{d}t \\ &= (\beta \log \widehat{c}_t - [\alpha - \beta] \log h_t) \, \mathrm{d}t. \end{aligned} \tag{4.3.2}$$

Hence the agent's habit level h_s conditional on all information available at time t is explicitly given by[59]

$$\begin{aligned} h_s = \exp &\left\{ \int_t^s \beta \exp\left\{-(\alpha-\beta)(s-u)\right\} \log \widehat{c}_u \,\mathrm{d}u \right\} \\ &\times \exp\left\{ \exp\left\{-(\alpha-\beta)(s-t)\right\} \log h_t \right\}. \end{aligned} \tag{4.3.3}$$

Equation (4.3.3) shows that due to the parameter restriction (4.2.6), the habit level does not grow exponentially over time. We define the dual *static* budget constraint as follows:

$$\mathbb{E}\left[\int_0^T M_t h_t \widehat{c}_t \,\mathrm{d}t\right] \leq A_0. \tag{4.3.4}$$

Equation (4.3.4) follows from substituting $c_t \equiv h_t \widehat{c}_t$ into the original static budget constraint. The agent's dual maximization problem is thus given by

$$\begin{aligned} &\underset{\widehat{c}_t}{\text{maximize}} \quad \mathbb{E}\left[\int_0^T e^{-\int_0^t \delta_u \,\mathrm{d}u} v\left(\widehat{c}_t\right) \mathrm{d}t\right] \\ &\text{subject to} \quad \mathbb{E}\left[\int_0^T M_t h_t \widehat{c}_t \,\mathrm{d}t\right] \leq A_0, \qquad \mathrm{d}\log h_t = \left(\beta \log \widehat{c}_t - [\alpha-\beta] \log h_t\right) \mathrm{d}t. \end{aligned} \tag{4.3.5}$$

We can obtain the optimal consumption choice c_t^* from the optimal *dual* consumption choice $\widehat{c}_t^*$ as follows:

$$c_t^* = h_t^* \widehat{c}_t^*. \tag{4.3.6}$$

Here h_t^* is the optimal habit level at time t implied by substituting the agent's optimal past dual consumption choices c_u^* $(u \leq t)$ into equation (4.3.3).

To solve the agent's maximization problem (4.2.11), we can thus restrict ourselves to finding a solution to problem (4.3.5). The agent's optimal consumption choice c_t^* then follows from applying equation (4.3.6). We can however still not solve the agent's maximization problem (4.2.11) analytically because the dual static budget constraint

[59] Equation (4.3.3) is equivalent to:

$$h_s = (h_t)^{\exp\{-(\alpha-\beta)(s-t)\}} \times \prod_t^s \left(\widehat{c}_u\right)^{\beta \exp\{-(\alpha-\beta)(s-u)\} \,\mathrm{d}u}.$$

(4.3.4) depends on the agent's dual consumption choice in a nonlinear way. Indeed, substitution of h_t into equation (4.3.4) shows that the dual static budget constraint depends on the agent's dual consumption choice $\widehat{c}_t$ nonlinearly. The next section develops a linearization to the agent's dual static budget constraint (4.3.4). After applying this linearization, we can obtain the agent's dual consumption choice in closed-form.

4.3.2. Linearizing the Budget Constraint

This section linearizes the left-hand side of the agent's dual budget constraint (4.3.4) around the consumption trajectory $\widehat{c} = c/h = 1$. We expect (and verify in Section 4.6) that the approximation error is accurate when the agent's consumption choice c_t stays close to the habit level h_t, and when the endogeneity parameter β is small or the depreciation parameter α is large. Indeed, if the habit level is completely exogenous (i.e., $\beta = 0$ or $\alpha = \infty$), the solution to problem (4.3.5) coincides with the solution to problem (4.2.11) (see also equation (4.2.12), which shows that we can solve the first-order optimality condition analytically if $\beta = 0$ or $\alpha = \infty$). We expect the approximation to be less accurate when the agent's consumption choice c_t deviates much from the habit level h_t. Section 4.6 examines the approximation error induced by applying a linearization to the agent's dual budget constraint.

By applying a first-order Taylor series approximation, we can write the left-hand side of the agent's dual budget constraint (4.3.4) as follows (see Appendix)

$$
\begin{aligned}
\mathbb{E}\left[\int_0^T M_t h_t \widehat{c}_t \,\mathrm{d}t\right] &\approx \mathbb{E}\left[\int_0^T M_t \,\mathrm{d}t\right] + \mathbb{E}\left[\int_0^T M_t \left(1+\beta P_t\right)\left(\widehat{c}_t - 1\right)\mathrm{d}t\right] \\
&= -\beta \mathbb{E}\left[\int_0^T M_t P_t \,\mathrm{d}t\right] + \mathbb{E}\left[\int_0^T \widehat{M_t} \widehat{c}_t \,\mathrm{d}t\right].
\end{aligned} \tag{4.3.7}
$$

Here $\widehat{M_t} \equiv M_t \left(1+\beta P_t\right)$ denotes the adjusted pricing kernel, and P_t stands for the price of a bond paying a continuous coupon, i.e.,[60]

$$
P_t \equiv \mathbb{E}_t\left[\int_t^T \frac{M_s}{M_t} e^{-(\alpha-\beta)(s-t)} \,\mathrm{d}s\right] = \frac{1}{r+\alpha-\beta}\left(1 - e^{-(r+\alpha-\beta)(T-t)}\right). \tag{4.3.8}
$$

[60]We can also view P_t as the amount of wealth needed to finance the consumption stream $\log c_s / \log h_t$ if $c_s = h_s$ for every $s \geq t$.

4.3.3. The Approximate Problem

This section presents an approximate problem to problem (4.3.5) based on linearizing the left-hand side of the dual budget constraint (4.3.4). The approximate problem is given by

$$\begin{aligned} &\underset{\widehat{c}_t}{\text{maximize}} \quad \mathbb{E}\left[\int_0^T e^{-\int_0^t \delta_u \,\mathrm{d}u} v\left(\widehat{c}_t\right) \mathrm{d}t\right] \\ &\text{subject to} \quad \mathbb{E}\left[\int_0^T \widehat{M_t}\widehat{c}_t \,\mathrm{d}t\right] \leq \widehat{A}_0. \end{aligned} \tag{4.3.9}$$

Here $\widehat{A}_0$ denotes the *adjusted* initial wealth. We can obtain the agent's maximization problem (4.3.9) from (4.3.5) as follows. First, we replace the left-hand side of the static dual budget constraint in (4.3.5) by equation (4.3.7). Second, we eliminate the constant term

$$-\beta\mathbb{E}\left[\int_0^T M_t P_t \,\mathrm{d}t\right] \tag{4.3.10}$$

from the *new* static dual budget constraint. We are allowed to do this because the constant term (4.3.10) does not play a role in determining the first-order optimality condition. Finally, we redefine the agent's initial wealth A_0 in such a way that the optimal solution $\widehat{c}_t^\star$ is budget-feasible. That is, we determine the initial level of the agent's optimal dual consumption choice (i.e., the Lagrange multiplier) in such a way that

$$\mathbb{E}\left[\int_0^T M_t h_t^\star \widehat{c}_t^\star \,\mathrm{d}t\right] = A_0. \tag{4.3.11}$$

Here $h_t^\star$ is the agent's habit level at time t implied by substituting the agent's optimal past dual consumption choices $\widehat{c}_u^\star$ ($u \leq t$) into (4.3.3). Straightforward computations show that the agent's adjusted initial wealth $\widehat{A}_0$ is given by

$$\begin{aligned} \widehat{A}_0 &= A_0 + \left(\mathbb{E}\left[\int_0^T \widehat{M_t}\widehat{c}_t^\star \,\mathrm{d}t\right] - \mathbb{E}\left[\int_0^T M_t h_t^\star \widehat{c}_t^\star \,\mathrm{d}t\right]\right) \\ &= \mathbb{E}\left[\int_0^T \widehat{M_t}\widehat{c}_t^\star \,\mathrm{d}t\right]. \end{aligned} \tag{4.3.12}$$

Equation (4.3.12) shows that the agent's adjusted initial wealth $\widehat{A}_0$ equals the agent's initial wealth A_0 plus the approximation error evaluated at the optimal solution. We can only compute the value of $\widehat{A}_0$ after problem (4.3.9) has been optimized.

4.4. Dynamic Consumption and Portfolio Choice

4.4.1. Consumption Choice

Theorem 4 below presents the optimal solution to the agent's maximization problem (4.3.9).

Theorem 4. *Consider an agent with the utility function (4.2.4) and habit formation process (4.2.5) who solves the consumption and portfolio choice problem (4.3.9). Denote by $h_t^\star$ the habit level implied by substituting the agent's optimal past dual consumption choices $\widehat{c}_u^\star \equiv c_u^\star / h_u^\star$ $(u \leq t)$ into equation (4.3.3), and by y the Lagrange multiplier associated with the static budget constraint in (4.3.9). Then the optimal consumption choice $c_t^\star$ is given by*

$$c_t^\star = h_t^\star \left(y e^{\int_0^t \delta_u \,\mathrm{d}u} \widehat{M}_t \right)^{-\frac{1}{\gamma}}. \tag{4.4.1}$$

The Lagrange multiplier $y \geq 0$ is determined in such a way that the agent's original budget constraint holds with equality.

4.4.1.1. Infinite Terminal Time

This section analyzes the agent's consumption choice for the case where the terminal time T equals infinity. This assumption does not necessarily imply that the agent lives forever. Indeed, if we also take into account mortality risk (see footnote 58), then T stands for the maximum age the agent can possibly reach. The Appendix shows that we can write the agent's consumption choice $c_t^\star$ (see (4.4.1) where $c_t^\star$ is expressed in terms of the state variables $h_t^\star$ and $\widehat{M}_t$) in terms of past financial shocks as follows:

$$c_t^\star = (c_0^\star)^{q_t/q_0} \exp\left\{ \int_0^t q_{t-u} \frac{1}{\gamma} \left(r + \frac{1}{2}\lambda^2 - \delta_u \right) \mathrm{d}u + \int_0^t q_{t-u} \frac{1}{\gamma} \lambda \,\mathrm{d}W_u \right\}. \tag{4.4.2}$$

The parameter q_u is defined as follows:

$$
\begin{aligned}
q_u &\equiv 1 + \frac{\beta}{\alpha - \beta}\left(1 - \exp\left\{-(\alpha-\beta)u\right\}\right) \\
&= q_0 + \left(q_\infty - q_0\right)\left(1 - \exp\left\{-\eta u\right\}\right).
\end{aligned} \tag{4.4.3}
$$

Here

$$q_0 = 1, \tag{4.4.4}$$

$$q_\infty = 1 + \frac{\beta}{\alpha - \beta}, \tag{4.4.5}$$

$$\eta \equiv \alpha - \beta. \tag{4.4.6}$$

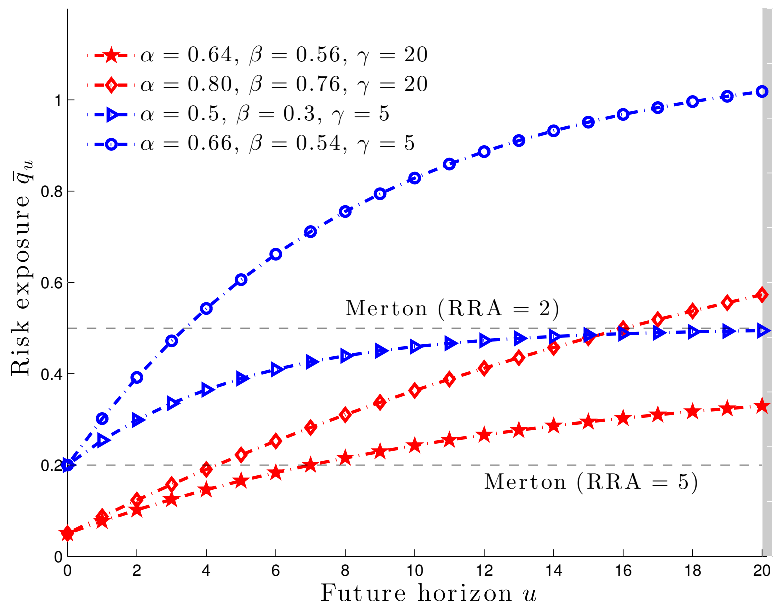
We can view

$$\bar{q}_u \equiv q_u / \gamma \tag{4.4.7}$$

as the exposure of future log consumption $\log c^\star_{t+u}$ to a current financial shock $\lambda \, \mathrm{d}W_t$. We make the following observations. First, the risk exposure $\bar{q}_u$ increases with the horizon u: a current financial shock has a larger impact on log consumption in the distant future than it has on log consumption in the near future. This implies that consumption responds gradually to financial shocks. It provides a utility-based foundation for the existence of smoothing mechanisms in drawing-down accumulated wealth by dampening the change in consumption due to financial shocks. Second, the risk exposure of current log consumption $\log c^\star_t$ to a current financial shock $\lambda \, \mathrm{d}W_t$, i.e., $\bar{q}_0$, decreases with the coefficient of relative risk aversion γ. Hence the coefficient of relative risk aversion γ determines the effect of a current financial shock on the level of log consumption (i.e., current log consumption). Third, $\beta/(\alpha - \beta)$ determines the effect of a current financial shock on future growth rates of consumption. If the endogeneity parameter β is large or the depreciation parameter η is small, then a current financial shock has a large effect on future growth rates of consumption (see also equation (4.3.3)). Fourth, we can view $\eta = \alpha - \beta$ as the rate at which $\bar{q}_u$ converges to $\bar{q}_\infty$. If η is small (i.e., the habit level depreciates at a slow pace), then it takes a long time to fully absorb a financial shock into current and future consumption. Finally, the Merton consumption strategy (see

Merton, 1969) emerges as a special case when $\bar{q}_u = 1/\gamma$ for all u. The risk exposure of an agent with CRRA utility is always smaller than the risk exposure of an agent with utility function (4.2.4), given γ. Figure 4.1 shows $\bar{q}_u$ (expressed relative to $\sigma = 20\%$) as a function of the horizon u for various sets of parameter values. We choose the parameter values such that the average risk exposure matches the risk exposure of a CRRA agent.

Figure 4.1.
Illustration of the risk exposure of future log consumption to a current financial shock



The figure illustrates the risk exposure $\bar{q}_u$ (i.e., the risk exposure of future log consumption $\log c^\star_{t+u}$ to a current financial shock $\lambda\, \mathrm{d}W_t$) as a function of the horizon u for various sets of parameter values. The figure also shows the Merton risk exposure for RRA = 2 and RRA = 5. Here RRA stands for relative risk aversion.

Equation (4.4.7) demonstrates that the parameters $\bar{q}_0$ (i.e., the exposure of current log consumption to a current shock), $\bar{q}_\infty$ (i.e., the exposure of long-term log consumption to a current financial shock), and η (i.e., the time it takes to absorb a financial shock) fully characterize the risk exposure $\bar{q}_u$. We can *uniquely* identify the agent's original preference parameters α, β and γ from $\bar{q}_0$, $\bar{q}_\infty$ and η:

$$\alpha = \frac{\bar{q}_\infty}{\bar{q}_0}\eta, \tag{4.4.8}$$

$$\beta = \frac{\bar{q}_\infty - \bar{q}_0}{\bar{q}_0}\eta, \tag{4.4.9}$$

$$\gamma = \frac{1}{\bar{q}_0}. \tag{4.4.10}$$

The Appendix shows that log consumption $\log c_t^\star$ evolves according to

$$\mathrm{d}\log c_t^\star = \log F_t^{\mathrm{d}t} + q_0 \frac{1}{\gamma}\left(r + \frac{1}{2}\lambda^2 - \delta_t\right)\mathrm{d}t + q_0 \frac{1}{\gamma}\lambda\,\mathrm{d}W_t. \tag{4.4.11}$$

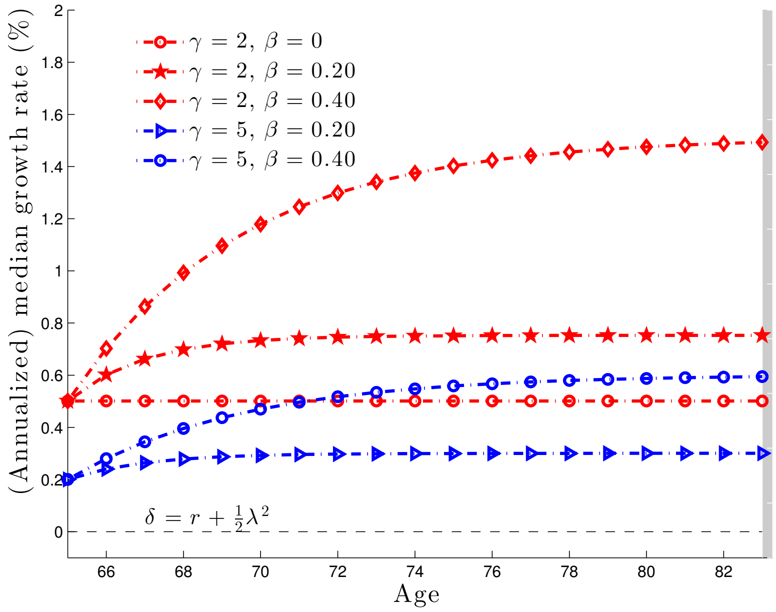
Here

$$\begin{aligned}\log F_t^v \equiv & \int_0^t (q_{t+v-u} - q_{t-u})\frac{1}{\gamma}\left(r + \frac{1}{2}\lambda^2 - \delta_u\right)\mathrm{d}u + \left(\frac{q_{t+v}}{q_0} - \frac{q_t}{q_0}\right)\frac{1}{\gamma}\log c_0^\star \\ & + \int_0^t (q_{t+v-u} - q_{t-u})\frac{1}{\gamma}\lambda\,\mathrm{d}W_u.\end{aligned} \tag{4.4.12}$$

The left-hand side of equation (4.4.11) consists of three terms. The first two terms represent the median (or expected) growth rate of log consumption. The term $\log F_t^{\mathrm{d}t}$ represents past financial shocks that are reflected into the current median growth rate of log consumption. This term disappears if $\beta = 0$ or $\bar{q}_u = 1/\gamma$ for all u. The second term represents the desired growth rate of consumption. The median value of log consumption stays constant over time if $\beta = 0$, $\delta_u = \delta$ and $r = \delta - \frac{1}{2}\lambda^2$ for all u. Finally, the last term corresponds to current financial shocks that are absorbed into the level of log consumption. The (annualized) volatility of $\mathrm{d}\log c_t^\star$ equals $q_0/\gamma \cdot \lambda$.

Figure 4.2 shows the median growth rate of consumption as a function of age for various sets of parameter values. The black dashed line corresponds the case where $\delta_u = r + \frac{1}{2}\lambda^2 = 3\%$ for all u. In that case, the median growth rate of consumption is zero (i.e., median consumption stays constant over time). The other lines illustrate the median growth rate of consumption if δ changes at the age of retirement (65 years) from 3% to 2%. The parameter δ can change because of a (discretionary) change in the force of mortality (see also footnote 58). We observe that the agent reallocates consumption from the short-run to the long-run. Indeed, the agent expects to live longer so that he postpones consumption and saves more for the future. The effect of a decrease in δ on the median growth rate of consumption is more pronounced for large values of β and small values of γ. Hence, if β is large or γ is small, then the agent is more willing to substitute consumption over time. Section 4.5 considers a utility specification in which the coefficient of relative risk aversion does not affect the agent's willingness to substitute consumption over time.

Figure 4.2.
Illustration of the median growth rate of consumption ($T = \infty$)



The figure illustrates the median growth rate of consumption as a function of age for various sets of parameter values. The depreciation parameter α is set equal to 0.6, the subjective rate of time preference δ to 0.02, and the Lagrange multiplier y to unity. The black dashed line represents the median growth of consumption in the case of $\delta = r + \frac{1}{2}\lambda^2 = 0.03$.

4.4.1.2. Finite Terminal Time

This section analyzes the agent's consumption choice for the case where the terminal time T is finite (i.e., $T < \infty$). The Appendix shows that we can write the agent's consumption choice $c_t^\star$ in terms of past financial shocks as follows:

$$c_t^\star = (c_0^\star)^{q_t/q_0} \exp\left\{ \int_0^t q_{t-u} \frac{1}{\gamma} \left(\widehat{r}_u + \frac{1}{2}\lambda^2 - \delta_u \right) \mathrm{d}u + \int_0^t q_{t-u} \frac{1}{\gamma} \lambda \, \mathrm{d}W_u \right\}. \qquad (4.4.13)$$

Here

$$\widehat{r}_u \equiv \beta + \frac{r - \alpha\beta P_u}{1 + \beta P_u} \tag{4.4.14}$$

We can obtain equation (4.4.13) from equation (4.4.2) by replacing the interest rate r with the adjusted interest rate $\widehat{r}_u$. The Appendix proofs the following theorem.

Theorem 5. *Let the adjusted interest rate $\widehat{r}_u$ be defined by equation (4.4.14). Then:*

1. *The adjusted interest rate $\widehat{r}_u$ increases as the endogeneity parameter β increases, given $\eta = \alpha - \beta$.*

2. *The adjusted interest rate $\widehat{r}_u$ decreases as the terminal time T increases. In particular, $\widehat{r}_u \to r$ if $T \to \infty$.*

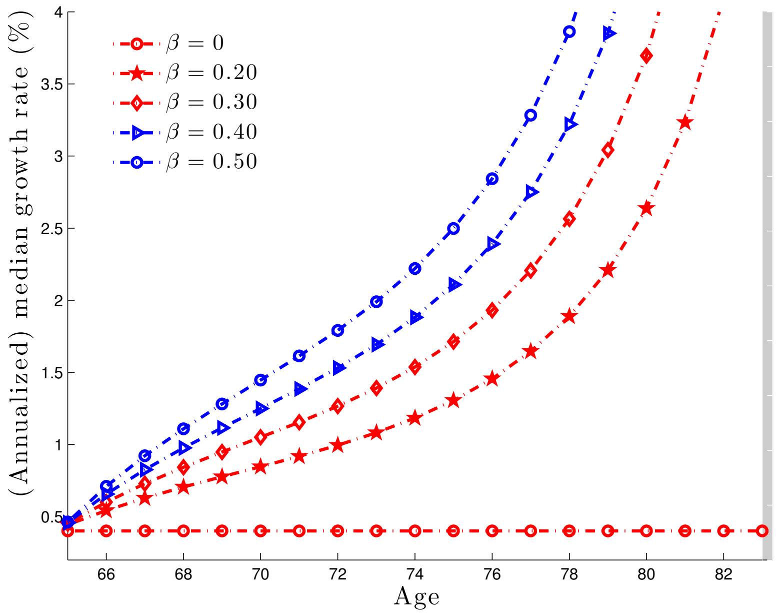
Current consumption has a large impact on future habit levels if β is large. Furthermore, the utility gain of an infinitesimal increase in consumption is smaller when the agent is (relatively) young (i.e., small t) than when the agent is (relatively) old (i.e., large t). These two facts together explain why the median consumption choice tends to go up with age if the endogeneity parameter β is large. Indeed, as already pointed out by Deaton (1992), the agent derives utility not only from consumption levels but also from consumption growth. If T equals infinity, the utility gain of an infinitesimal increase in consumption is age-independent. Hence, the agent does no longer have a desire to postpone consumption. In our model, four factors thus affect the median consumption choice (see also equation (4.4.13)): the financial market (i.e., r, λ and σ), the subjective rate of time preference, the survival curve, and the strength of internal habit formation. Figure 4.3 illustrates the median growth rate of consumption for various values of the endogeneity parameter β.

4.4.2. Portfolio Choice

This section analyzes the agent's portfolio choice $\pi_t^\star$. The Appendix shows that the replicating portfolio strategy $\pi_t^\star$ is given by

$$\pi_t^\star = \widehat{q}_t \frac{\lambda}{\sigma}. \tag{4.4.15}$$

Figure 4.3.
Illustration of median growth rate of consumption ($T < \infty$)



The figure illustrates the median growth rate of consumption as a function of age for various values of the endogeneity parameter β. The parameter $\eta \equiv \alpha - \beta$ is set equal to 0.2, the coefficient of relative risk aversion γ to 5, the subjective rate of time preference δ to 0.01, the terminal time T to 20, and the Lagrange multiplier y to unity.

Here $0 \leq \widehat{q}_t \leq 1$ denotes the (weighted) average risk exposure. That is,

$$\widehat{q}_t = \int_t^T \bar{q}_u \frac{V_t^u}{V_t} \, \mathrm{d}u, \tag{4.4.16}$$

where $V_t \equiv \int_t^T V_t^u \, \mathrm{d}u$ and V_t^u denotes the (market) value at time t of $c_{t+u}^{\star}$:

$$V_t^u \equiv c_t^{\star} F_t^u C_t^u. \tag{4.4.17}$$

Equation (4.4.17) shows that the market value of future consumption, i.e., $V_t^u / c_t^{\star}$, consists of two factors. The first factor, i.e., F_t^u, represents past financial shocks that are absorbed into future growth rates of consumption. This factor equals unity if the agent directly adjusts consumption after unexpected financial shocks (i.e., $\bar{q}_u = 1/\gamma$ for all u). The horizon-dependent *annuity factor* C_t^u summarizes the impacts of desired consumption streams and future rates of return on the market value of future consumption. The Appendix provides an explicit analytical expression for the horizon-dependent annuity factor C_t^u (see equation (4.8.3) in the Appendix).

Table 4.1 shows the median portfolio choice as a function of age for various sets of parameter values. The agent implements a life cycle investment strategy (that is, the fraction of wealth invested in the risky stock decreases on average as the agent ages), even without taking human capital into account. Indeed, the agent has less time to absorb a wealth shock as he grows older.

Table 4.1.
The agent's median portfolio choice

Age	(1)	(2)	(3)	(4)	Merton (RRA = 2)	Merton (RRA = 5)
65	0.20	0.27	0.41	0.65	0.50	0.20
70	0.18	0.24	0.39	0.64	0.50	0.20
75	0.15	0.19	0.36	0.56	0.50	0.20
80	0.10	0.12	0.29	0.40	0.50	0.20
85	0.05	0.05	0.20	0.20	0.50	0.20

The table reports the agent's median portfolio choice (i.e., the median fraction of assets invested in the risky stock) as a function of age for various sets of parameter values. The table also reports the Merton portfolio strategy. (1) corresponds to $\alpha = 0.64$, $\beta = 0.56$, $\gamma = 20$; (2) to $\alpha = 0.80$, $\beta = 0.76$, $\gamma = 20$; (3) to $\alpha = 0.5$, $\beta = 0.3$, $\gamma = 5$; and (4) to $\alpha = 0.66$, $\beta = 0.54$, $\gamma = 5$.

Table 4.2 shows the (annualized) volatility of the relative change in consumption and the (annualized) volatility of the relative change in wealth. With internal habit formation, the volatility of the relative change is consumption is smaller than the volatility of the relative change in wealth. Hence the agent can take substantial stock market risk without affecting the year-on-year volatility of consumption. Indeed, the degree of internal habit formation largely determines the fraction of wealth invested in the risky stock, while the coefficient of relative risk aversion largely determines the year-on-year fluctuations in consumption.

4.4.3. Welfare Analysis

This section conducts a welfare analysis. More specifically, we compare a number of alternative popular draw-down and investment strategies to the draw-down and investment strategy implied by the agent's maximization problem (4.3.9). The welfare loss associated with implementing an alternative draw-down and investment strategy is computed *relative* to the agent's optimal draw-down and investment strategy. More precisely, the performance

Table 4.2.
Volatility of the change in consumption and the volatility of the change in wealth

Age	(1)		(2)		Merton (RRA = 2)		Merton (RRA = 5)	
	σ_c	σ_A	σ_c	σ_A	σ_c	σ_A	σ_c	σ_A
65	1.00	3.95	4.00	12.99	10.00	10.00	4.00	4.00
70	1.00	3.56	4.00	12.81	10.00	10.00	4.00	4.00
75	1.00	2.92	4.00	11.12	10.00	10.00	4.00	4.00
80	1.00	1.99	4.00	7.90	10.00	10.00	4.00	4.00
85	1.00	1.00	4.00	4.00	10.00	10.00	4.00	4.00

The table reports the volatility of the change in consumption σ_c and the volatility of the change in wealth σ_A as a function of age for various sets of parameter values. (1) corresponds to $\alpha = 0.64$, $\beta = 0.56$, $\gamma = 20$; and (2) to $\alpha = 0.66$, $\beta = 0.54$, $\gamma = 5$. The numbers represent a percentage.

of an alternative strategy is evaluated by measuring the relative decline in certainty equivalent consumption. We define the certainty equivalent of an uncertain consumption strategy to be the constant, certain consumption level that yields indifference to the uncertain consumption strategy. Certainty equivalents are computed using the lifetime utility function (4.2.10). Due to the presence of internal habit formation, the computation of certainty equivalents is non-standard; see also Chapter 2. In the welfare analysis, we consider the following alternative draw-down and investment strategies:

- The remaining lifetime approach (i.e., the $1/T$-rule): the proportion of wealth withdrawn from the agent's retirement wealth is given by

$$\frac{c_t}{A_t} = \frac{1}{T-t}. \tag{4.4.18}$$

  Here $T-t$ denotes the agent's remaining lifetime which is assumed to be non-random. We assume that the agent invests a fixed percentage (0, 20, 40, 60 or 80 percent) of wealth into the risky stock. Equation (4.4.18) shows that consumption responds directly to a financial shock.

- The Merton approach: the proportion of wealth withdrawn from the agent's retirement wealth is given by

$$\frac{c_t}{A_t} = \frac{x_1 - x_2}{\exp\{(x_1 - x_2)(T-t)\} - 1}, \tag{4.4.19}$$

where

$$x_1 \equiv \frac{1-\gamma}{\gamma}\left(r + \frac{1}{2}\lambda^2\right) + \frac{1}{2}\left(\frac{1-\gamma}{\gamma}\right)^2 \lambda^2 \tag{4.4.20}$$

$$x_2 \equiv \frac{1}{\gamma}\delta. \tag{4.4.21}$$

We assume that the coefficient of relative risk aversion equals 2, 5 or 20. Like the remaining lifetime approach, consumption is directly adjusted after a wealth shock.

- Difference model of habit formation. We assume that the agent maximizes

  $$U_0 = \mathbb{E}\left[\int_0^T e^{-\int_0^t \delta_u \,\mathrm{d}u} \frac{1}{1-\gamma} (c_t - h_t)^{1-\gamma} \,\mathrm{d}t\right]$$

  subject to the dynamic budget constraint (4.2.9) and the habit formation process

  $$\mathrm{d}h_t = (\beta c_t - \alpha h_t)\,\mathrm{d}t.$$

  The consumption strategy is given by

  $$c_t = h_t + \left(\widehat{M}_t y e^{\int_0^t \delta_u \,\mathrm{d}u}\right)^{-\frac{1}{\gamma}}. \tag{4.4.22}$$

  Here y is a Lagrange multiplier. Consumption (4.4.22) responds gradually to a financial shock. The investment strategy follows from replicating the consumption strategy (4.4.22). Unlike the investment strategy implied by the ratio model (see equation (4.4.15)), the investment strategy implied by the difference model depends on the habit level. Indeed, if the habit level approaches consumption, the agent reduces the fraction of wealth invested in the risky stock.

Table 4.3 reports welfare losses associated with implementing the remaining lifetime approach. The welfare losses are relatively large if the agent exhibits a significant degree of internal habit formation; see the first two rows of Table 4.3. Table 4.4 reports the welfare losses due to implementing the Merton strategy. The welfare losses are smaller compared to the remaining lifetime approach. Indeed, the Merton strategy emerges as a special case of our model when $\beta = 0$. However, welfare losses associated with implementing the Merton strategy may still be significant when the strength of internal habit formation is large. Finally, Table 4.5 reports the welfare losses due to implementing the difference model of habit formation. Again, the welfare loss increases as the strength of internal habit formation (i.e., β) increases. Also, the welfare losses are larger compared to the Merton approach.

Table 4.3.
Welfare losses due to implementing the remaining lifetime approach

Optimal Strategy	Fraction of Assets Invested in the Risky Stock				
	0%	20%	40%	60%	80%
$\alpha = 0.64$, $\beta = 0.56$, $\gamma = 20$	35.16	34.07	36.97	42.76	50.37
$\alpha = 0.80$, $\beta = 0.76$, $\gamma = 20$	42.93	42.35	43.84	46.97	51.45
$\alpha = 0.50$, $\beta = 0.30$, $\gamma = 5$	7.45	2.90	2.79	7.11	15.44
$\alpha = 0.66$, $\beta = 0.54$, $\gamma = 5$	8.85	3.16	1.71	4.79	11.87

The table reports the welfare losses (in terms of the relative decline in certainty equivalent consumption) due to implementing the remaining lifetime approach. The table reports the welfare losses for various values of the fraction of wealth invested in the risky stock. The numbers represent a percentage. We assume that $\delta_u = \delta = 3\%$ for all u and $T = 20$.

Table 4.4.
Welfare losses due to implementing the Merton approach

Optimal Strategy	Relative Risk Aversion Coefficient		
	2	5	20
$\alpha = 0.64$, $\beta = 0.56$, $\gamma = 20$	23.85	5.19	2.44
$\alpha = 0.80$, $\beta = 0.76$, $\gamma = 20$	29.55	13.13	8.50
$\alpha = 0.50$, $\beta = 0.30$, $\gamma = 5$	2.95	0.63	3.93
$\alpha = 0.66$, $\beta = 0.54$, $\gamma = 5$	2.85	1.86	5.31

The table reports the welfare losses (in terms of the relative decline in certainty equivalent consumption) due to the consumption and portfolio strategy of an agent with CRRA utility (i.e., the Merton strategy). The table reports the welfare losses for various values of the coefficient of relative risk aversion γ underlying the Merton strategy. The numbers represent a percentage. We assume that $\delta_u = \delta = 3\%$ for all u and $T = 20$.

Table 4.5.
Welfare losses due to implementing the difference model of habit formation

Optimal Strategy	Welfare Loss
$\alpha = 0.64,\ \beta = 0.56,\ \gamma = 20$	5.63
$\alpha = 0.80,\ \beta = 0.76,\ \gamma = 20$	9.10
$\alpha = 0.50,\ \beta = 0.30,\ \gamma = 5$	4.33
$\alpha = 0.66,\ \beta = 0.54,\ \gamma = 5$	4.31

The table reports the welfare losses (in terms of the relative decline in certainty equivalent consumption) due to implementing the difference model of habit formation. The numbers represent a percentage. We assume that $\delta_u = \delta = 3\%$ for all u and $T = 20$.

4.5. Stochastic Differential Utility

4.5.1. Preferences and Maximization Problem

This section considers consumption and portfolio choice of an agent with preferences that combine SDU with multiplicative internal habit formation. We specify the agent's utility process $\{V_t\}_{t\in[0,T]}$ in terms of the intertemporal aggregator f. More specifically, $\{V_t\}_{t\in[0,T]}$ satisfies the following integral equation ($t \in [0,T]$):

$$V_t = \mathbb{E}_t\left[\int_t^T f\left(\frac{c_s}{h_s}, V_s, s\right) \mathrm{d}s\right]. \tag{4.5.1}$$

As in the previous sections, the log habit level $\log h_t$ evolves according to equation (4.2.5). The intertemporal aggregator f is characterized by[61]

$$f\left(\frac{c_t}{h_t}, V_t, t\right) = (1+\zeta)\left[\frac{\left(\frac{c_t}{h_t}\right)^{\varphi}}{\varphi}|V_t|^{\frac{\zeta}{1+\zeta}} - \delta V_t\right]. \tag{4.5.2}$$

Here $\zeta > -1$ and $\varphi < \min\{1, 1/(1+\zeta)\}$ are preference parameters. Equation (4.5.2) is usually referred to as the Kreps-Porteus aggregator. If $\zeta = 0$ and the habit level h_t equals unity (i.e., $\alpha = \beta = 0$), then $f(c_t/h_t, V_t, t)$ reduces to

$$f\left(\frac{c_t}{h_t}, V_t, t\right) = \frac{1}{\varphi}c_t^{\varphi} - \delta V_t. \tag{4.5.3}$$

[61] If $\zeta = 0$, then $f(c_t/h_t, V_t, t) = (1+\zeta V_t)\left[\log\{c_t/h_t\} - (\delta/\zeta)\log\{1+\zeta V_t\}\right]$.

Equation (4.5.1) is then equivalent to the additive utility specification

$$V_t = \mathbb{E}_t \left[\int_t^T e^{-\delta(s-t)} \frac{1}{\varphi} c_s^{\varphi} \mathrm{d}s \right]. \tag{4.5.4}$$

The agent aims to maximize V_0 in equation (4.5.1) (at $t = 0$) with $f\left(\frac{c_t}{h_t}, V_t, t\right)$ given by equation (4.5.2) subject to the habit formation process (4.2.5) and the dynamic budget constraint (4.2.9).

We can transform this dynamic consumption and portfolio choice problem into an equivalent static consumption and portfolio choice problem (similar to what we did in Section 4.2.3). After transforming this static problem into a dual problem and applying a linearization to the static dual budget constraint (which takes the same form as in our base model; see Section 4.3.2 for further details), we obtain the following maximization problem:

$$\begin{aligned} &\underset{\widehat{c}_t}{\text{maximize}} \quad \mathbb{E}\left[\int_0^T f\left(\widehat{c}_t, V_t, t\right) \mathrm{d}t\right] \\ &\text{subject to} \quad \mathbb{E}\left[\int_0^T \widehat{M}_t \widehat{c}_t \, \mathrm{d}t\right] \leq \widehat{A}_0. \end{aligned} \tag{4.5.5}$$

The next section presents the optimal solution to problem (4.5.5).

4.5.2. Dynamic Consumption and Portfolio Choice

The agent's maximization problem (4.5.5) obtained in the dual model upon linearizing the dual budget constraint can be solved by invoking the approach of Schroder and Skiadas (1999). The next theorem presents the optimal consumption choice.

Theorem 6. *Consider an agent with utility process* (4.5.1)*, intertemporal aggregator* (4.5.2) *and habit formation process* (4.2.5) *who solves the consumption and portfolio choice problem* (4.5.5)*. Let $h_t^\star$ be the agent's habit level implied by substituting the agent's optimal past dual consumption choices $\widehat{c}_u^\star \equiv c_u^\star / h_u^\star$ $(u \leq t)$ into equation* (4.3.3) *and let y be the Lagrange multiplier associated with the static budget constraint in* (4.5.5)*. Then the agent's optimal consumption choice $c_t^\star$ is given by*

$$c_t^\star = h_t^\star \exp\left\{ \int_0^t \left(\psi \left[\widehat{r}_u + \frac{1}{2}\frac{\lambda^2}{\gamma} - \delta \right] + \frac{1}{2}\frac{\lambda^2}{\gamma^2}[\gamma - 1] \right) \mathrm{d}u + \psi y + \frac{\lambda}{\gamma} \int_0^t \mathrm{d}W_u \right\}, \tag{4.5.6}$$

where

$$\psi = \frac{1}{1-\varphi},$$
$$\zeta = \frac{1-\gamma}{\varphi} - 1.$$

We can write the agent's consumption choice $c_t^\star$ in terms of past financial shocks as follows:

$$c_t^\star = (c_0^\star)^{\frac{q_t}{q_0}} \exp\left\{ \int_0^t q_{t-u} \left(\psi \left[\widehat{r}_u + \frac{1}{2}\frac{\lambda^2}{\gamma} - \delta \right] + \frac{1}{2}\frac{\lambda^2}{\gamma^2}\left[\gamma - 1\right] \right) \mathrm{d}u + \right.$$
$$\left. + q_{t-u}\frac{1}{\gamma}\lambda \int_0^t \mathrm{d}W_u \right\}. \tag{4.5.7}$$

The optimal portfolio choice $\pi_t^\star$ is the same as in Section 4.4.2. Equation (4.5.7) shows that with SDU, the parameter ψ determines the willingness to substitute consumption over time. Relative risk aversion is thus decoupled from the elasticity of intertemporal substitution. Figure 4.4 shows the median growth rate of consumption as a function of age for various values of ψ and δ. We observe that the change in the median growth rate of consumption following a (permanent) change in the subjective rate of time preference δ at the age of retirement (65 years) is small if ψ is small in absolute terms.

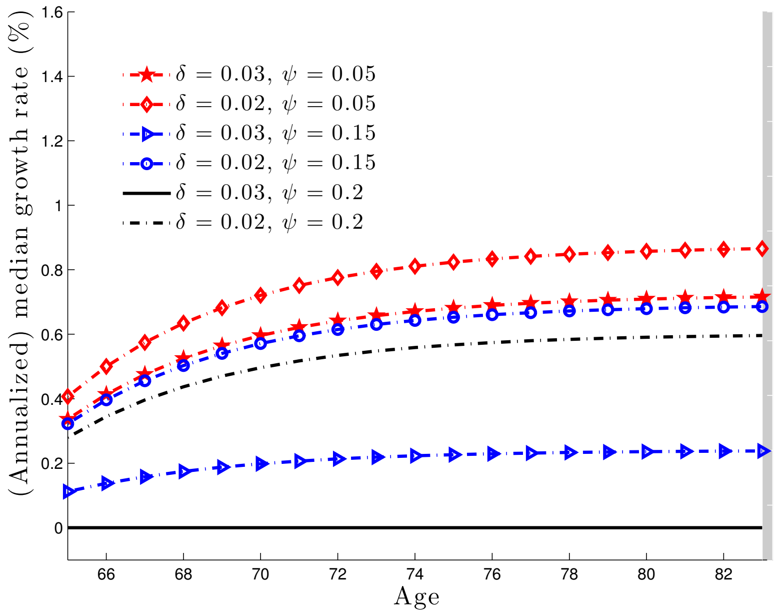
4.6. The Accuracy of the Approximation Method

The consumption and portfolio strategies presented in Section 4.4 are exact only in the case of $\beta = 0$ and/or $\alpha = \infty$. In all other cases the consumption and portfolio strategies are approximate based upon linearizing the left-hand side of the agent's dual budget constraint (4.3.4) around the consumption trajectory $c/h = 1$. This section analyzes the approximation error due to applying a linearization to the dual budget constraint.[62]

We determine the optimal consumption choice c_t^* by using the method of backward induction. That is, first, we determine the optimal consumption choice at the terminal time T. Then, the optimal consumption choice at time $T-1$ is determined taking the

[62] The Appendix linearizes the left-hand side of the dual budget constraint (4.3.4) in a multi-period, discrete-time setting.

Figure 4.4.
Illustration of median growth rate of consumption (SDU utility)



The figure illustrates the median growth rate of consumption as a function of age for various values of ψ and δ. The endogeneity parameter β is set equal to 0.4, the depreciation parameter α to 0.6, the coefficient of relative risk aversion γ to 5, the terminal time T to infinity, and the Lagrange multiplier y to unity. The black solid line corresponds to the case where median consumption growth is zero.

optimal consumption choice at time T as given. We continue this process backwards in time until all optimal consumption choices have been determined. The terminal time T is set equal to three (we also consider the case where the terminal time T equals four), the time interval Δt equals unity and the underlying uncertainty is described by a binomial tree.[63] The computation of the optimal consumption choice c_t^* rapidly becomes infeasible as the number of time steps increases.

We evaluate the performance of the (sub-optimal) consumption choice $c_t^\star$ by measuring the relative decline in certainty equivalent consumption (see Section 4.4.3 for the definition of certainty equivalent consumption).[64] Tables 4.6 – 4.9 report our results. The first three tables show the welfare losses (in terms of the relative decline in certainty equivalent

[63]By considering a binomial tree, we can exactly compute the conditional expectations involved in the optimization technique.

[64]The certainty equivalent consumption choice $\bar{c}$ always exists if $\alpha \geq \beta$. In particular, $\frac{\partial U_0}{\partial \bar{c}} \geq 0$ if $\beta \int_0^T e^{-\alpha t}\,\mathrm{d}t \leq 1$. If T is large, then $\int_0^T e^{-\alpha t}\,\mathrm{d}t \approx \frac{1}{\alpha}$. Hence we can always compute (for any T) the certainty equivalent consumption choice $\bar{c}$ if $\frac{\beta}{\alpha} \leq 1$.

consumption) for the case where the terminal time equals three. We find that the approximation error is an increasing function of β, and an decreasing function of α and γ. Indeed, if α is large and/or β is small, the impact of an increase in consumption on future habit levels is limited. Also, if γ is large, consumption stays close to the habit level. In all cases, the approximation error is smaller than 1%. Table 4.9 reports the approximation error for case where the terminal time T equals four. The approximation error is still very small.

Table 4.6.
Welfare losses ($\gamma = 2$ and $T = 3$)

	β				
α	0	0.15	0.2	0.3	0.6
0	0	–	–	–	–
0.15	0	0.0229	–	–	–
0.2	0	0.0178	0.0516	–	–
0.3	0	0.0149	0.03053	0.1263	–
0.6	0	0.0153	0.0293	0.0689	0.6840

The table reports the welfare losses (in terms of the relative decline in certainty equivalent consumption) associated with implementing the consumption choice $c_t^\star$. The numbers represent a percentage. We only report welfare losses for the case $\alpha \geq \beta$. The terminal time T is set equal to 3, initial wealth to 3, and the subjective rate of time preference δ to 3%.

Table 4.7.
Welfare losses ($\gamma = 5$ and $T = 3$)

	β				
α	0	0.15	0.2	0.3	0.6
0	0	–	–	–	–
0.15	0	0.0078	–	–	–
0.2	0	0.0012	0.0101	–	–
0.3	0	0.0164	0.0021	0.0197	–
0.6	0	0.0017	0.0285	0.0096	0.0479

The table reports the welfare losses (in terms of the relative decline in certainty equivalent consumption) associated with implementing the consumption choice $c_t^\star$. The numbers represent a percentage. We only report welfare losses for the case $\alpha \geq \beta$. The terminal time T is set equal to 3, initial wealth to 3, and the subjective rate of time preference δ to 3%.

Table 4.8.
Welfare losses ($\gamma = 20$ and $T = 3$)

	β				
α	0	0.15	0.2	0.3	0.6
0	0	–	–	–	–
0.15	0	0.0000	–	–	–
0.2	0	0.0002	0.0003	–	–
0.3	0	0.0019	0.0008	0.0024	–
0.6	0	0.0000	0.0008	0.0012	0.0017

The table reports the welfare losses (in terms of the relative decline in certainty equivalent consumption) associated with implementing the consumption choice $c_t^\star$. The numbers represent a percentage. We only report welfare losses for the case $\alpha \geq \beta$. The terminal time T is set equal to 3, initial wealth to 3, and the subjective rate of time preference δ to 3%.

Table 4.9.
Welfare losses ($\gamma = 2$ and $T = 4$)

	β				
α	0	0.15	0.2	0.3	0.6
0	0	–	–	–	–
0.15	0	0.0486	–	–	–
0.2	0	0.0561	0.0989	–	–
0.3	0	0.0777	0.1025	0.2309	–
0.6	0	0.3997	0.2779	0.2081	1.3957

The table reports the welfare losses (in terms of the relative decline in certainty equivalent consumption) associated with implementing the consumption choice $c_t^\star$. The numbers represent a percentage. We only report welfare losses for the case $\alpha \geq \beta$. The terminal time T is set equal to 4, initial wealth to 4, and the subjective rate of time preference δ to 3%.

4.7. Concluding Remarks

In this chapter, we have built a rich consumption-portfolio choice model with preferences that combine both multiplicative internal habit formation and stochastic differential utility. To solve our preference model, we have developed an approximation method based upon linearizing the agent's (dual) budget constraint. For reasonable values of the preference parameters, the approximation error induced by our method is very small. We have shown that after a wealth shock, the agent optimally chooses to adjust both the level and future growth rates of consumption, giving rise to gradual response of consumption to financial shocks. Furthermore, expected consumption tends to grow with age, and relative risk aversion does not affect the willingness to substitute consumption over time. A possible venue for future research would be to confront our preference model with actual consumption data.

4.8. Appendix

4.8.1. Proofs

Derivation of (4.3.7)

This appendix linearizes the left-hand side of the agent's dual static budget constraint (4.3.4) around the consumption trajectory $\widehat{c} = c/h = 1$. The partial derivative of h_s with respect to $\widehat{c}_t\,\mathrm{d}t$ is given by

$$\frac{\partial h_s}{\partial \widehat{c}_t}\frac{1}{\mathrm{d}t} = \beta \exp\left\{-(\alpha-\beta)(s-t)\right\}\frac{h_s}{\widehat{c}_t}. \tag{4.8.1}$$

Equation (4.8.1) follows from differentiating (4.3.3) with respect to $\widehat{c}_t\,\mathrm{d}t$. The partial derivative (4.8.1) evaluated along the consumption trajectory $\widehat{c} = 1$ yields

$$\left.\frac{\partial h_s}{\partial \widehat{c}_t}\frac{1}{\mathrm{d}t}\right|_{\widehat{c}=1} = \beta \exp\left\{-(\alpha-\beta)(s-t)\right\}.$$

Define the function $f\left(\widehat{c}\right)$ as follows:

$$f\left(\widehat{c}\right) \equiv \mathbb{E}\left[\int_0^T M_t h_t \widehat{c}_t\,\mathrm{d}t\right].$$

Straightforward computations show

$$\begin{aligned} f(1) &= \mathbb{E}\left[\int_0^T M_t\,\mathrm{d}t\right], \\ \left.\frac{\partial f\left(\widehat{c}\right)}{\partial \widehat{c}_t}\frac{1}{\mathrm{d}t}\right|_{\widehat{c}=1} &= M_t + \beta \mathbb{E}_t\left[\int_t^T M_s \exp\left\{-(\alpha-\beta)(s-t)\right\}\mathrm{d}s\right] \\ &= M_t\left(1+\beta P_t\right). \end{aligned}$$

We can, by virtue of Taylor series expansion, approximate the dual budget constraint $f(\widehat{c})$ by

$$f(\widehat{c}) \approx f(1) + \mathbb{E}\left[\int_0^T M_t (1 + \beta P_t)(\widehat{c}_t - 1)\,\mathrm{d}t\right].$$

Proof of Theorem 4

The Lagrangian $\mathcal{L}$ is given by

$$\begin{aligned}\mathcal{L} &= \mathbb{E}\left[\int_0^T e^{-\int_0^t \delta_u\,\mathrm{d}u} v(\widehat{c}_t)\,\mathrm{d}t\right] + y\left(\widehat{A}_0 - \mathbb{E}\left[\int_0^T \widehat{M}_t \widehat{c}_t\,\mathrm{d}t\right]\right)\\ &= \int_0^T \mathbb{E}\left[e^{-\int_0^t \delta_u\,\mathrm{d}u} v(\widehat{c}_t) - y\widehat{M}_t\widehat{c}_t\right]\mathrm{d}t + y\widehat{A}_0.\end{aligned}$$

Here $y \geq 0$ denotes the Lagrange multiplier associated with the static budget constraint. The agent aims to maximize $e^{-\int_0^t \delta_u\,\mathrm{d}u} v(\widehat{c}_t) - y\widehat{M}_t\widehat{c}_t$. The optimal dual consumption choice $\widehat{c}_t$ satisfies the following first-order optimality condition:

$$e^{-\int_0^t \delta_u\,\mathrm{d}u}\widehat{c}_t^{\,-\gamma} = y\widehat{M}_t.$$

After solving the first-order optimality condition, we obtain the following maximum:

$$\widehat{c}_t^{\star} = \left(e^{\int_0^t \delta_u\,\mathrm{d}u} y\widehat{M}_t\right)^{-\frac{1}{\gamma}}.$$

Hence (use equation (4.3.6))

$$c_t^{\star} = h_t^{\star}\left(ye^{\int_0^t \delta_u\,\mathrm{d}u}\widehat{M}_t\right)^{-\frac{1}{\gamma}}.$$

A standard verification (see, e.g., Karatzas and Shreve, 1998, p. 103) stating that the optimal solution to the Lagrangian equals the optimal solution to the dual problem completes the proof.

Derivation of (4.4.2), (4.4.11) and (4.4.13)

This appendix explicitly writes the agent's consumption choice $c_t^{\star}$ in terms of past financial shocks. We can write the adjusted pricing kernel $\widehat{M}_t \equiv M_t(1 + \beta P_t)$ as follows (this follows from applying Itô's Lemma to $\widehat{M}_t = f(M_t, P_t) = M_t(1+\beta P_t)$):

$$\widehat{M}_t = \widehat{M}_0 \exp\left\{-\int_0^t \left(\widehat{r}_u + \frac{1}{2}\lambda^2\right) \mathrm{d}u\right\} \exp\left\{\lambda \int_0^t \mathrm{d}W_u\right\}. \tag{4.8.2}$$

Substituting equation (4.8.2) into equation (4.4.1) yields

$$\widehat{c}_t^\star \equiv \frac{c_t^\star}{h_t^\star} = \exp\left\{\frac{1}{\gamma}\int_0^t \left(\widehat{r}_u + \frac{1}{2}\lambda^2 - \delta_u\right) \mathrm{d}u + \frac{\bar{y}}{\gamma}\right\} \exp\left\{\frac{\lambda}{\gamma}\int_0^t \mathrm{d}W_u\right\}.$$

Here $\bar{y} \equiv -\left(\log y + \log \widehat{M}_0\right)$.

We can write the habit level as follows:

$$\begin{aligned}
h_t^\star &= \exp\left\{\int_0^t \beta \exp\{-(\alpha-\beta)(t-u)\} \log \widehat{c}_u^\star \mathrm{d}u\right\} \\
&= \exp\left\{\int_0^t \beta \exp\{-(\alpha-\beta)(t-u)\}\right. \\
&\left.\left[\frac{1}{\gamma}\int_0^u \left(\widehat{r}_v + \frac{1}{2}\lambda^2 - \delta_v\right) \mathrm{d}v + \frac{\bar{y}}{\gamma} + \frac{\lambda}{\gamma}\int_0^u \mathrm{d}W_v\right] \mathrm{d}u\right\} \\
&= \exp\left\{\int_0^t \left(\bar{q}_{t-u} - \frac{1}{\gamma}\right)\left(\widehat{r}_u + \frac{1}{2}\lambda^2 - \delta_u\right) \mathrm{d}u\right\} \\
&\exp\left\{\left(\bar{q}_t - \frac{1}{\gamma}\right)\bar{y} + \int_0^t \left(\bar{q}_{t-u} - \frac{1}{\gamma}\right)\lambda \, \mathrm{d}W_u\right\}.
\end{aligned}$$

Here

$$\begin{aligned}
\bar{q}_{t-u} &\equiv \frac{1}{\gamma}\left[1 + \beta\int_u^t \exp\{-(\alpha-\beta)(t-v)\} \mathrm{d}v\right] \\
&= \frac{1}{\gamma}\left[1 + \frac{\beta}{\alpha-\beta}\left(1 - \exp\{-(\alpha-\beta)(t-u)\}\right)\right].
\end{aligned}$$

Hence,

$$\begin{aligned}
c_t^\star &= h_t^\star \exp\left\{\frac{1}{\gamma}\int_0^t \left(\widehat{r}_u + \frac{1}{2}\lambda^2 - \delta_u\right) \mathrm{d}u + \bar{y}\right\} \exp\left\{\frac{\lambda}{\gamma}\int_0^t \mathrm{d}W_u\right\} \\
&= \exp\left\{\int_0^t \bar{q}_{t-u}\left(\widehat{r}_u + \frac{1}{2}\lambda^2 - \delta_u\right) \mathrm{d}u + \bar{q}_t\bar{y} + \int_0^t \bar{q}_{t-u}\lambda \, \mathrm{d}W_u\right\} \\
&= (c_0^\star)^{\bar{q}_t/\bar{q}_0} \exp\left\{\int_0^t \bar{q}_{t-u}\left(\widehat{r}_u + \frac{1}{2}\lambda^2 - \delta_u\right) \mathrm{d}u + \int_0^s \bar{q}_{t-u}\lambda \, \mathrm{d}W_u\right\}.
\end{aligned}$$

Equation (4.4.2) follows from equation (4.4.13) and Theorem 5.

Dividing $\log c_{t+\Delta t}^\star$ by $\log c_t^\star$ and taking the limit $\Delta t \Rightarrow 0$ yields equation (4.4.11).

Proof of Theorem 5

We first proof that the (partial) derivate of $\widehat{r}_u$ with respect to β is positive given $\eta = \alpha - \beta$. Substituting $\alpha = \eta + \beta$ into equation (4.4.14) yields

$$\widehat{r}_u = \beta + \frac{r - (\eta + \beta)\beta P_u}{1 + \beta P_u}.$$

The derivate of $\widehat{r}_u$ with respect to β is given by

$$\begin{aligned}
\frac{\partial \widehat{r}_u}{\partial \beta} &= 1 + \frac{-\left(1 + \beta P_u\right)\left(\eta + 2\beta\right) P_u - \left(r - (\eta + \beta)\beta P_u\right) P_u}{\left(1 + \beta P_u\right)^2} \\
&= 1 + \frac{-\eta P_u - 2\beta P_u - \eta\beta P_u^2 - 2\left(\beta P_u\right)^2 - r P_u + \eta\beta P_u^2 + \left(\beta P_s\right)^2}{1 + 2\beta P_u + \left(\beta P_u\right)^2} \\
&= 1 + \frac{-\eta P_u - 2\beta P_u - \left(\beta P_u\right)^2 - r P_u}{1 + 2\beta P_u + \left(\beta P_u\right)^2}.
\end{aligned}$$

Hence

$$\begin{aligned}
\frac{\partial \widehat{r}_u}{\partial \beta} \geq 0 \Leftrightarrow & \frac{-\eta P_u - 2\beta P_u - \left(\beta P_u\right)^2 - r P_u}{1 + 2\beta P_u + \left(\beta P_u\right)^2} \geq -1 \\
& \Leftrightarrow \eta P_u + 2\beta P_u + \left(\beta P_u\right)^2 + r P_u \leq 1 + 2\beta P_u + \left(\beta P_u\right)^2 \\
& \Leftrightarrow \left(r + \eta\right) P_u \leq 1 \Leftrightarrow 1 - \exp\left\{-(r + \eta)(T - u)\right\} \leq 1.
\end{aligned}$$

This last inequality is obviously true. Hence $\partial \widehat{r}_u / \partial \beta$ is positive (given η).

Finally, we proof that the (partial) derivate of $\widehat{r}_u$ with respect to T is negative. The derivate of $\widehat{r}_u$ with respect to T is given by

$$\frac{\partial \widehat{r}_u}{\partial T} = -r\left(1 + \beta P_u\right)^{-2} \frac{\partial P_u}{\partial T} - \alpha\beta\left(1 + \beta P_u\right)^{-2} \frac{\partial P_u}{\partial T}.$$

Using the fact that $\partial P_u / \partial T$ is positive, we find that $\partial \widehat{r}_u / \partial T$ is negative. Furthermore, simple algebra yields that $\widehat{r}_u = r$ if $T = \infty$ (here we use the fact that $P_u \Rightarrow 1/(r + \alpha - \beta)$ as $T \Rightarrow \infty$).

Derivation of (4.4.15)

Straightforward computations show that

$$
\begin{aligned}
V_t^u &= \mathbb{E}_t\left[\frac{M_{t+u}}{M_t} c_{t+u}^{\star}\right] \\
&= c_t^{\star} F_t^u \mathbb{E}_t\left[\exp\left\{-\int_0^u \left(r+\frac{1}{2}\lambda^2\right) \mathrm{d}v - \int_0^u \lambda \,\mathrm{d}W_{t+u-v}\right\}\right. \\
&\left.\exp\left\{\int_0^u q_v \frac{1}{\gamma}\left(r+\frac{1}{2}\lambda^2-\delta_{t+u-v}\right) \mathrm{d}v + \int_0^u q_v \frac{1}{\gamma}\lambda \,\mathrm{d}W_{t+u-v}\right\}\right] \\
&= c_t^{\star} F_t^u C_t^u,
\end{aligned}
$$

where

$$
C_t^u \equiv \exp\left\{-\int_0^u \left([1-\bar{q}_v]\, r + \frac{1}{2}\bar{q}_v\lambda^2 + \bar{q}_v\delta_{t+u-v} - \frac{1}{2}\bar{q}_v^2\lambda^2\right) \mathrm{d}v\right\}. \tag{4.8.3}
$$

It follows from Itô's Lemma that

$$
\frac{\partial \log V_t}{\partial W_t} = \frac{1}{V_t}\int_t^T \frac{\partial V_t^u}{\partial W_t}\,\mathrm{d}u = \int_t^T \bar{q}_u \frac{V_t^u}{V_t}\lambda \,\mathrm{d}u. \tag{4.8.4}
$$

We also have (this follows from applying Itô's Lemma to the dynamic budget constraint (4.2.9))

$$
\frac{\partial \log A_t}{\partial W_t} = \pi_t \sigma. \tag{4.8.5}
$$

Setting equation (4.8.5) equal to equation (4.8.4) and solving for π_t yields (4.4.15).

Proof of Theorem 6

Schroder and Skiadas (1999) derive the optimal dual consumption choice $\widehat{c}_t^{\star}$. The optimal consumption choice $c_t^{\star}$ follows from equation (4.3.6).

4.8.2. Multi-Period Discrete-Time Model

This section linearizes the left-hand side of the agent's dual budget constraint (4.3.4) in a multi-period, discrete-time setting. Let us denote by Δt the time step (the magnitude of Δt is usually taken to be small). The habit level is given by

$$
\begin{aligned}
h_{n\Delta t} &= \exp\left\{\beta\left[\sum_{i=1}^{n}(1-[\alpha-\beta]\Delta t)^{n-i}\log\widehat{c}_{i\Delta t}\right]\Delta t\right\} \\
&= \prod_{i=1}^{n}\widehat{c}_{i\Delta t}^{\beta(1-[\alpha-\beta]\Delta t)^{n-i}\Delta t}.
\end{aligned}
$$

Here, $n \in \{0, ..., \lfloor T/\Delta t\rfloor - 1\}$. The agent's dual budget constraint can be written as follows:

$$
\begin{aligned}
\mathbb{E}\left[\sum_{n=0}^{\lfloor T/\Delta t\rfloor-1} M_{(n+1)\Delta t}h_{n\Delta t}\widehat{c}_{(n+1)\Delta t}\Delta t\right] = \mathbb{E}\Bigg[\sum_{n=0}^{\lfloor T/\Delta t\rfloor-1} M_{(n+1)\Delta t} \\
\left(\times\prod_{i=1}^{n}\widehat{c}_{i\Delta t}^{\beta(1-[\alpha-\beta]\Delta t)^{n-i}\Delta t}\right)\widehat{c}_{(n+1)\Delta t}\Delta t\Bigg].
\end{aligned}
$$

Let us define the following function:

$$
\begin{aligned}
f(\widehat{c}) &\equiv f\left(\widehat{c}_{\Delta t}, ..., \widehat{c}_{\lfloor T/\Delta t\rfloor\Delta t}\right) \\
&\equiv \mathbb{E}\left[\sum_{n=0}^{\lfloor T/\Delta t\rfloor-1} M_{(n+1)\Delta t}\left(\prod_{i=1}^{n}\widehat{c}_{i\Delta t}^{\beta(1-[\alpha-\beta]\Delta t)^{n-i}\Delta t}\right)\widehat{c}_{(n+1)\Delta t}\Delta t\right].
\end{aligned}
$$

By Taylor series expansion,

$$
\begin{aligned}
f(\widehat{c}) &\approx f(1) + \mathbb{E}\left[\sum_{n=0}^{\lfloor T/\Delta t\rfloor-1}\left.\frac{\partial f(\widehat{c})}{\partial\widehat{c}_{(n+1)\Delta t}}\right|_{\widehat{c}=1}\left(\widehat{c}_{(n+1)\Delta t}-1\right)\right] \\
&= f(1) + \mathbb{E}\left[\sum_{n=0}^{\lfloor T/\Delta t\rfloor-1}\left.\frac{\partial f(\widehat{c})}{\partial\widehat{c}_{(n+1)\Delta t}}\right|_{\widehat{c}=1}\left(\widehat{c}_{(n+1)\Delta t}-1\right)\right].
\end{aligned}
$$

Straightforward computations show that

$$
\begin{aligned}
\left.\frac{\partial f(\widehat{c})}{\partial\widehat{c}_{(n+1)\Delta t}}\right|_{\widehat{c}=1} &= M_{(n+1)\Delta t}\Delta t \\
&\quad + \beta\mathbb{E}_{(n+1)\Delta t}\left[\sum_{i=n+2}^{\lfloor T/\Delta t\rfloor-1} M_{i\Delta t}(1-[\alpha-\beta]\Delta t)^{i-(n+2)}\Delta t\right]\Delta t \\
&= M_{(n+1)\Delta t}\left(1+\beta P_{(n+1)\Delta t}\right)\Delta t,
\end{aligned}
$$

where

$$P_{(n+1)\Delta t} \equiv \mathbb{E}_{(n+1)\Delta t}\left[\sum_{i=n+2}^{\lfloor T/\Delta t\rfloor-1} \frac{M_{i\Delta t}}{M_{(n+1)\Delta t}}(1-[\alpha-\beta]\Delta t)^{i-(n+2)}\Delta t\right].$$

Hence,

$$f\left(\widehat{c}\right) \approx f\left(1\right) + \mathbb{E}\left[\sum_{n=0}^{\lfloor T/\Delta t\rfloor-1} M_{(n+1)\Delta t}\left(1+\beta P_{(n+1)\Delta t}\right)\left(\widehat{c}_{(n+1)\Delta t}-1\right)\Delta t\right].$$

Part II:

Modelling Pension Plans

Chapter 5

Personal Pension Plans with Risk Pooling: Investment Approach versus Consumption Approach[65]

Personal pension plans with risk pooling (PPRs) promise to play a new role in the provision of retirement income. These plans unbundle the main functions of variable annuities. In particular, a PPR individualizes the savings, investment and withdrawal functions of variable annuities, and organizes the insurance function collectively. A policyholder can adopt an investment approach or a consumption approach to a PPR. This chapter explores these two approaches in detail. We show that in the investment approach, a policyholder can freely adjust the investment policy without affecting the intertemporal allocation of the market value of the consumption stream. This is not the case for the consumption approach. We also explore the collective versions of the investment approach and the consumption approach.

5.1. Introduction

Private pension provision is in transition, moving from defined benefit (DB) pension plans towards defined contribution (DC) pension plans (Investment Company Institute, 2014). The global financial crisis and its aftermath have accelerated this trend. The move towards DC pension plans, however, may be problematic as these pension plans primarily focus on wealth accumulation rather than providing stable lifelong income streams. Indeed, according to The Melbourne Mercer Global Pension Index (2013), "there is an urgent need to find a better balance between the individual orientation of

[65]This chapter is co-authored with Lans Bovenberg.

a DC plan and a collective (or pooled) approach where there is some sharing of risks within and between generations."

The pension plans being proposed by Bovenberg and Nijman (2015) promise to play a new role in the provision of retirement income. These pension plans, which are called personal pension plans with risk pooling (PPRs), unbundle the main functions of variable annuities. In particular, a PPR individualizes the savings, investment and withdrawal functions of variable annuities[66] and arranges the insurance function (i.e., pooling of idiosyncratic longevity risk) collectively. By pooling idiosyncratic longevity risk and taking systematic risks on behalf of the policyholders, pension funds can provide lifelong income streams during retirement at relatively low costs.[67] A policyholder can adopt an investment approach or a consumption approach to a PPR. This chapter examines these two approaches in detail.

In the investment approach, the policyholder exogenously specifies the contribution level, the investment policy and the assumed interest rate. The annuity units (i.e., the lifelong consumption streams in retirement) are derived endogenously. We show that in this approach, the policyholder can freely adjust the investment policy without affecting the intertemporal allocation of the market value of the consumption stream. This property facilitates pooling of idiosyncratic longevity risk. Indeed, a myopic policyholder cannot reallocate market value from the long-run to the short-run by changing the investment policy.

In the consumption approach, the policyholder exogenously specifies the volatility of annuity units, the expected growth rate of annuity units and the initial annuity units. The contribution level, the investment policy and the assumed interest rate are derived endogenously. Brown, Kling, Mullainathan, and Wrobel (2008, 2013) find that people value annuities more when framed in terms of consumption than when framed in terms of investment. In the consumption approach, policyholders can reallocate value from the long-run to the short-run by raising the volatility of annuity units. Hence, myopic policyholders face an incentive to raise the risk of annuity units to consume more today.

[66]Individualization of these functions is possible without any welfare losses. Indeed, pooling of systematic risks does not generate any welfare gain. In fact, individualization of these functions may even lead to an improvement in welfare because these functions can now be tailored to the specific needs of the policyholders (see Mehlkopf, Boelaars, Bovenberg, and van Bilsen, 2015).

[67]We can view a PPR as a middle ground between a DB and a DC pension plan. It aims to provide a stable lifelong income stream during retirement (as in DB pension plans), while property rights are defined in terms of a personal investment account (as in DC pension plans).

We also examine a collective defined contribution (CDC) pension system (which is the collective analogue of the investment approach) and a collective defined ambition (CDA) pension system (which is the collective analogue of the consumption approach). These pension plans allow policyholders to share non-traded risk factors. However, collective pension systems have several disadvantages. In particular, they allow for less tailor-made solutions, require the valuation of annuity units, and may well lead to intergenerational conflicts about the (unobservable) parameters.

The remainder of this chapter is structured as follows. Section 5.2 describes the underlying assumptions. Section 5.3 formalizes the investment approach and Section 5.4 the consumption approach. Section 5.5 explores a CDC pension system and Section 5.6 a CDA pension system. Section 5.7 concludes the chapter. Proofs are relegated to the Appendix.

5.2. Assumptions

5.2.1. Financial Market

The financial market consists of a money market account and a risky stock. The price of the money market account, i.e., P_t, satisfies

$$\frac{\mathrm{d}P_t}{P_t} = r\mathrm{d}t. \tag{5.2.1}$$

Here r can be viewed as the risk-free interest rate. The price of the risky stock, i.e., S_t, evolves according to

$$\frac{\mathrm{d}S_t}{S_t} = (r + \lambda\sigma)\,\mathrm{d}t + \sigma\mathrm{d}W_t. \tag{5.2.2}$$

Here λ denotes the equity risk premium per unit of risk (i.e., Sharpe ratio), σ stands for the diffusion parameter (i.e., stock return volatility), and W_t corresponds to a standard Brownian motion. The pricing kernel (or stochastic discount factor) m_t is subject to (see, e.g., Karatzas and Shreve, 1998):

$$\frac{\mathrm{d}m_t}{m_t} = -r\mathrm{d}t - \lambda\mathrm{d}W_t. \tag{5.2.3}$$

5.2.2. Longevity Insurance

The insurer pools idiosyncratic longevity risk, so that policyholders are protected from outliving their retirement wealth. Denote by y the date of birth, by x_r the age at which policyholders retire, and by $x_{\max}$ the maximum age policyholders can reach. The policyholder receives a pension payment at time t if he is alive and his birthdate y falls between time $t-x_r$ and time $t-x_{\max}$. The probability that a policyholder aged $x = t-y$ will survive to age $x+h$ is given by

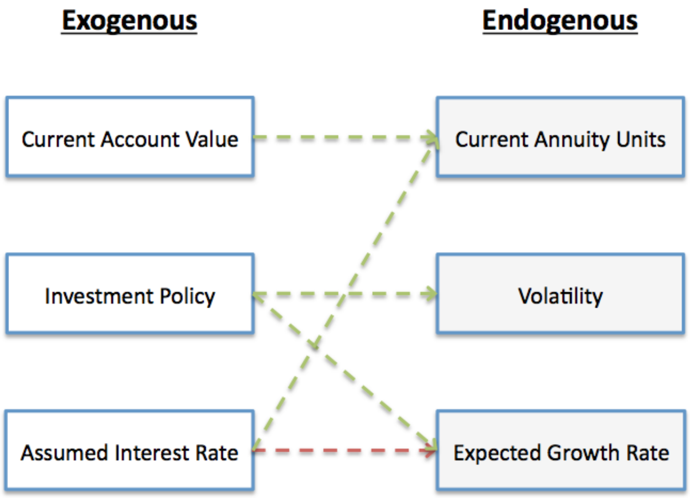
$$ {}_h p_x \equiv \exp\left\{-\int_0^h \mu_{x+v} \mathrm{d}v\right\}, \tag{5.2.4}$$

where μ_{x+v} corresponds to the force of mortality at age $x+v$.

5.3. The Investment Approach

The investment approach defines pension entitlements in terms of a personal investment account. The policyholder exogenously specifies the (current) value of his account, the investment policy (i.e., the fraction of assets invested in the risky stock) and the assumed interest rate. The assumed interest rate determines how fast retirement wealth is depleted. Current annuity units, the volatility of annuity units and the expected growth rate of annuity units are determined endogenously. Figure 5.1 summarizes the investment approach. The left-hand side of the figure shows the exogenous parameters of the pension contract. These exogenous parameters determine the endogenous variables on the right-hand side of the figure. A green line describes a positive relationship between an exogenous parameter and an endogenous variable whereas a red line describes a negative relationship. Section 5.3.1 formalizes the pension contract. This section also derives the rate at which retirement wealth is depleted. Section 5.3.2 determines the endogenous variables of the pension contract. Finally, Section 5.3.3 explores the impact of (discretionary) changes in the exogenous parameters on the endogenous variables.

Figure 5.1.
Illustration of the investment approach



The figure illustrates the investment approach. The left-hand side of the figure shows the exogenous parameters of the pension contract. These exogenous parameters determine the variables on the right-hand side of the figure.

5.3.1. The Pension Contract

Denote by $X^i_{y,t}$ the value of the personal investment account at time t of policyholder i born at time y, and by ϕ^i the fraction of wealth invested in the risky stock by policyholder i. The value of the personal investment account evolves according to[68]

$$\frac{\mathrm{d}X^i_{y,t}}{X^i_{y,t}} = \left(\mu_{t-y} + r + \phi^i\lambda\sigma\right)\mathrm{d}t + \phi^i\sigma\mathrm{d}W_t - d^i_{y,t}\mathrm{d}t, \qquad X^i_{y,t} \geq 0 \text{ given.} \tag{5.3.1}$$

Here $d^i_{y,t}$ is the decumulation rate (or withdrawal fraction) at time t of policyholder i born at time y. Equation (5.3.1) shows that the expected rate of return on the assets $X^i_{y,t}$ is equal to the sum of the biometric rate of return μ_{t-y} and the expected financial rate of return $r + \phi^i\lambda\sigma$. The assumed interest rate determines the decumulation rate $d^i_{y,t}$. Denote by δ^i_{t+v} the (forward) assumed interest rate at time t for horizon $v \geq 0$. The

[68] The value of policyholder i's personal investment account accrues to the insurer if policyholder i passes away.

decumulation rate is now given by

$$d^i_{y,t} \equiv \begin{cases} 1/A^i_{y,t}, & \text{if } t > y + x_r; \\ 0, & \text{if } t \leq y + x_r. \end{cases} \tag{5.3.2}$$

Here $A^i_{y,t}$ denotes the annuity (or conversion) factor:

$$A^i_{y,t} \equiv \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} \exp\left\{ -\int_0^h \left(\mu_{t-y+v} + \delta^i_{t+v}\right) \mathrm{d}v \right\} \mathrm{d}h. \tag{5.3.3}$$

Equations (5.3.2) and (5.3.3) show that the decumulation rate increases with the assumed interest rate. Hence an increase in the assumed interest rate reallocates consumption from the long-run to the short-run.

5.3.2. The Endogenous Variables

Denote by $B^i_{y,t}$ the annuity units at time t of policyholder i born at time y. The value of the personal investment account $X^i_{y,t}$ together with the annuity factor $A^i_{y,t}$ determine current annuity units $B^i_{y,t}$ as follows:

$$B^i_{y,t} = \frac{X^i_{y,t}}{A^i_{y,t}}. \tag{5.3.4}$$

Applying Itô's Lemma to equation (5.3.4) shows

$$\frac{\mathrm{d}B^i_{y,t}}{B^i_{y,t}} = \frac{1}{A^i_{y,t}B^i_{y,t}}\mathrm{d}X^i_{y,t} - \frac{X^i_{y,t}}{\left(A^i_{y,t}\right)^2 B^i_{y,t}}\mathrm{d}A^i_{y,t} = \frac{\mathrm{d}X^i_{y,t}}{X^i_{y,t}} - \frac{\mathrm{d}A^i_{y,t}}{A^i_{y,t}}. \tag{5.3.5}$$

Hence current annuity units $B^i_{y,t}$ fully absorb the mismatch between assets and the annuity factor. The dynamic equation of the annuity factor is given by (this follows from applying Itô's Lemma to equation (5.3.3))

$$\frac{\mathrm{d}A^i_{y,t}}{A^i_{y,t}} = \left(\mu_{t-y} + \delta^i_t\right) \mathrm{d}t - d^i_{y,t}\mathrm{d}t. \tag{5.3.6}$$

Accordingly, current annuity units $B_{y,t}^i$ satisfy (this follows from equations (5.3.1), (5.3.5), and (5.3.6))

$$\frac{\mathrm{d}B_{y,t}^i}{B_{y,t}^i} = \left(r + \phi^i\lambda\sigma - \delta_t^i\right)\mathrm{d}t + \phi^i\sigma\mathrm{d}W_t. \tag{5.3.7}$$

The (annualized) volatility of annuity units ω^i is thus given by

$$\omega^i = \phi^i\sigma. \tag{5.3.8}$$

Furthermore, equation (5.3.7) shows that the current expected growth of annuity units π_t^i is equal to the difference between the expected financial rate of return $r + \phi^i\lambda\sigma$ and the assumed interest rate δ_t^i:

$$\pi_t^i = r + \phi^i\lambda\sigma - \delta_t^i. \tag{5.3.9}$$

A high assumed interest rate thus implies not only a high decumulation rate but also a low expected growth rate of annuity units.

5.3.3. Changes in Parameters

This section explores the impact of (discretionary) changes in the exogenous parameters (i.e., the investment policy and the assumed interest rate) on the endogenous variables. We also consider changes in the Sharpe ratio and the force of mortality. Whereas the investment policy and the assumed interest rate are parameters specified by the policyholder, the Sharpe ratio and the force of mortality are parameters describing the external environment (i.e., non-traded risk factors). Changes in parameters are not traded on financial markets, and are thus not valued ex ante.

5.3.3.1. Investment Policy

Policyholder i can freely adjust its investment policy ϕ^i before or during the retirement period. A change in the investment policy of $\Delta\phi^i$ causes – ceteris paribus – the expected growth rate of annuity units π_t^i to change by $\lambda\sigma\Delta\phi^i$ (see equation (5.3.9)). The expected growth rate of annuity units thus changes as a result of a change in the investment policy. The assumed interest rate is not affected by a change in the investment policy. A change in the investment policy thus impact neither the speed with which retirement wealth is depleted nor the horizon-dependent market value $V_{y,t}^{i,h}$:

$$V_{y,t}^{i,h} \equiv B_{y,t}^{i} \exp\left\{-\int_0^h \left(\mu_{t-y+v} + \delta_{t+v}^{i}\right) \mathrm{d}v\right\}. \tag{5.3.10}$$

Hence the investment policy is not mixed up with the intertemporal allocation of the market value of the consumption stream. Perverse incentives in the investment policy are thus absent: a myopic policyholder (whose discount rate exceeds the assumed interest rate because of hyperbolic discounting or a short life expectancy) does not face incentives to change the investment policy in order to reallocate consumption away from the future to the present. This property facilitates pooling of idiosyncratic longevity risk. Indeed, policyholders whose life expectancy declines as a result of adverse health shocks cannot reallocate market value from the long-run to the short-run by increasing the fraction of wealth invested in the risky stock. A change in investment policy does, however, impact expected annuity units. More specifically, we find (see Appendix)

$$\frac{\Delta\mathbb{E}_t\left[B_{y,t+h}^{i}\right]}{\mathbb{E}_t\left[B_{y,t+h}^{i}\right]} \approx h\lambda\sigma\Delta\phi^{i}. \tag{5.3.11}$$

Hence an increase in ϕ^i leads to an increase in not only the volatility of annuity units but also expected annuity units. Indeed, a change in the investment policy causes a different position on the risk-return frontier.

5.3.3.2. Assumed Interest Rate

The policyholder can also change the assumed interest rate δ_{t+v}^{i}.[69] We assume that a change in the assumed interest rate δ_{t+v}^{i} is a weighted average of a change in the short-term assumed interest rate δ_t^i and a change in the long-term assumed interest rate δ_∞^i:

$$\Delta\delta_{t+v}^{i} = e^{-\kappa v}\Delta\delta_t^{i} + \left(1 - e^{-\kappa v}\right)\Delta\delta_\infty^{i}. \tag{5.3.12}$$

[69] In a model with interest rate risk and/or (expected) inflation risk, also changes in these risk factors may affect the assumed interest rate.

Here $\kappa \geq 0$ is a weight parameter. If κ is small, then $\Delta\delta^i_{t+v}$ is largely determined by the short-term shock $\Delta\delta^i_t$. Policyholder i thus only needs to specify $\Delta\delta^i_t$, $\Delta\delta^i_\infty$ and κ.

A change in the assumed interest rate of $\Delta\delta^i_{t+v}$ causes – ceteris paribus – current annuity units to change by (see Appendix)

$$\frac{\Delta B^i_{y,t}}{B^i_{y,t}} \approx \widehat{D}^{i,0}_{y,t}\Delta\delta^i_\infty + \widehat{D}^{i,\kappa}_{y,t}\left(\Delta\delta^i_t - \Delta\delta^i_\infty\right). \tag{5.3.13}$$

Here the κ-adjusted duration $\widehat{D}^{i,\kappa}_{y,t}$ is defined as follows:[70]

$$\widehat{D}^{i,\kappa}_{y,t} \equiv \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} \gamma^{i,h}_{y,t} D^{\kappa,h}\mathrm{d}h, \tag{5.3.14}$$

where

$$D^{\kappa,h} \equiv \frac{1}{\kappa}\left(1-e^{-\kappa h}\right), \tag{5.3.15}$$

$$\gamma^{i,h}_{y,t} \equiv \frac{V^{i,h}_{y,t}}{V^i_{y,t}}. \tag{5.3.16}$$

An increase in the assumed interest rate causes not only an increase in current annuity units (see equation (5.3.13)) but also a decrease in the expected growth rate of annuity units π^i_{t+v} ($v \geq 0$) (see equation (5.3.9)). We note that a change in the investment policy leaves current annuity units unaffected (see Section 5.3.3.1).

A change in the assumed interest rate also affects expected annuity units and the intertemporal allocation of the market value of the consumption stream. We find (see Appendix)

$$\frac{\Delta\mathbb{E}_t\left[B^i_{y,t+h}\right]}{\mathbb{E}_t\left[B^i_{y,t+h}\right]} \approx \left(\widehat{D}^{i,0}_{y,t} - h\right)\Delta\delta^i_\infty + \left(\widehat{D}^{i,\kappa}_{y,t} - D^{\kappa,h}\right)\left(\Delta\delta^i_t - \Delta\delta^i_\infty\right), \tag{5.3.17}$$

$$\frac{\Delta V^{i,h}_{y,t}}{V^{i,h}_{y,t}} \approx \left(\widehat{D}^{i,0}_{y,t} - h\right)\Delta\delta^i_\infty + \left(\widehat{D}^{i,\kappa}_{y,t} - D^{\kappa,h}\right)\left(\Delta\delta^i_t - \Delta\delta^i_\infty\right). \tag{5.3.18}$$

[70] We note that

$$\widehat{D}^{i,0}_{y,t} \equiv \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} \gamma^{i,h}_{y,t} h\mathrm{d}h \geq \widehat{D}^{i,\kappa}_{y,t}.$$

Indeed, whereas $\widehat{D}^{i,0}_{y,t}$ denotes the impact of a permanent shock, $\widehat{D}^{i,\kappa}_{y,t}$ represents the impact of a temporary shock. The impact of the temporary shock declines with the speed κ with which a temporary shock dies out.

An increase in the assumed interest rate thus reallocates market value from the long-run to the short-run. In particular, a permanent shock $\Delta\delta_\infty^i > 0$ raises the market value at horizon h as long as h is smaller than the average duration, i.e., $h < \widehat{D}_{y,t}^{i,0}$. Similarly, a temporary shock $\Delta\delta_t^i - \Delta\delta_\infty^i > 0$ raises the market value at horizon h as long as the κ-adjusted duration is smaller than the average κ-adjusted duration, i.e., $D^{\kappa,h} < \widehat{D}_{y,t}^{i,\kappa}$. Intuitively, a higher assumed interest rate yields two effects. First, the direct effect is to reduce the annuity factor. Second, the indirect effect is through an immediate increase in current annuity units on account of a lower annuity factor (see equation (5.3.4)). Whereas the first effect (i.e., the reduction in the annuity factor) increases with the horizon, the second effect (i.e., the increase in current annuity units) is uniform for all horizons and amounts (in absolute value) to the weighted average of the first effects. Hence the sum of the two effects declines with the horizon and adds up to zero (weighted with the market shares).

5.3.3.3. Sharpe Ratio

A change in the Sharpe ratio of $\Delta\lambda$ leads to similar effects on the endogenous variables as a change in the investment policy. That is, the intertemporal allocation of the market value of the consumption stream is unaffected, while expected annuity units change as follows:

$$\frac{\Delta\mathbb{E}_t\left[B_{y,t+h}^i\right]}{\mathbb{E}_t\left[B_{y,t+h}^i\right]} \approx h\phi^i\sigma\Delta\lambda. \tag{5.3.19}$$

5.3.3.4. Force of Mortality

This section considers a change in the force of mortality μ_{t-y+v}. Inspired by Lee and Carter (1992), we assume that a change in the force of mortality μ_{t-y+v}, is driven by a change in a common risk factor f:

$$\Delta\mu_{t-y+v} = g_{t-y+v}\Delta f. \tag{5.3.20}$$

Here $g_{t-y+v} \geq 0$ represents the exposure of μ_{t-y+v} to the risk factor f. We can view $g_{t-y+v}\Delta f$ as the change in the assumed biometric rate of return.

The Appendix shows that current annuity units change as follows:

$$\frac{\Delta B^i_{y,t}}{B^i_{y,t}} \approx \widehat{G}^i_{y,t} \Delta f, \tag{5.3.21}$$

where $\widehat{G}^i_{y,t}$ can be viewed as the g-adjusted duration:

$$\widehat{G}^i_{y,t} \equiv \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} \gamma^{i,h}_{y,t} G^h_{y,t} \mathrm{d}h \tag{5.3.22}$$

with

$$G^h_{y,t} \equiv \int_0^h g_{t-y+v}\, \mathrm{d}v. \tag{5.3.23}$$

The intertemporal allocation of the market value of the consumption stream also changes following a change in the force of mortality. We find (see Appendix)

$$\frac{\Delta V^{i,h}_{y,t}}{V^{i,h}_{y,t}} \approx \left(\widehat{G}^i_{y,t} - G^h_{y,t}\right) \Delta f. \tag{5.3.24}$$

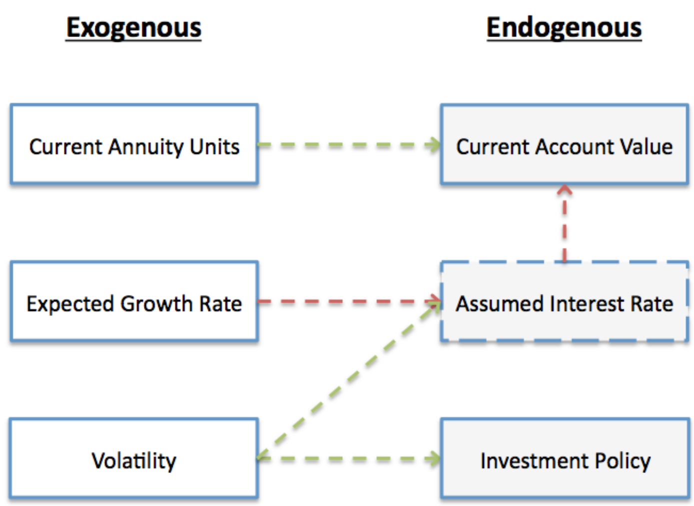
Improved longevity (i.e., $\Delta f < 0$) thus reallocates market value from the short-run to the long-run. Intuitively, after a positive longevity shock, policyholders on average live longer so that it is optimal to save more for the future. We note that the expected growth rate of annuity units is unaffected by a change in the force of mortality.

5.4. The Consumption Approach

The consumption approach defines pension entitlements in terms of a personal investment account, just like the investment approach. However, the policyholder now exogenously specifies current annuity units, the expected growth rate of annuity units and the volatility of annuity units. These parameters characterize the entire consumption stream in retirement under the assumption that future annuity units are log-normally distributed. The current value of the personal investment account, the assumed interest rate (i..e, the discount rate used to determine the current account value) and the investment policy (i.e., the fraction of assets invested in the risky stock) are determined endogenously.

Figure 5.2 summarizes the consumption approach, where the assumed interest rate is an intermediate variable that links the current account value with the expected growth rate of annuity units and the volatility of annuity units. The left-hand side of the figure shows the exogenous parameters of the pension contract. These exogenous parameters determine the endogenous variables on the right-hand side of the figure. A green line describes a positive relationship between an exogenous parameter and an endogenous variable whereas a red line describes a negative relationship. Section 5.4.1 formalizes the pension contract. This section also derives the entire consumption stream in retirement. Section 5.4.2 determines the endogenous variables of the pension contract. Finally, Section 5.4.3 explores the impact of (discretionary) changes in the exogenous parameters on the endogenous variables.

Figure 5.2.
Illustration of the consumption approach



The figure illustrates the consumption approach. The left-hand side of the figure shows the exogenous parameters of the pension contract. These exogenous parameters determine the variables on the right-hand side of the figure.

5.4.1. The Pension Contract

Policyholder i exogenously specifies current annuity units $B^i_{y,t}$, the expected growth rate of annuity units π^i_t and the volatility of annuity units ω^i. We can view the consumption approach as a generalization of defined benefit to stochastic retirement consumption streams. Indeed, the consumption approach in a stochastic setting is sometimes called defined ambition. The annuity units $B^i_{y,t}$, which can be viewed as the liabilities of the pension contract, evolve according to:

$$\frac{\mathrm{d}B^i_{y,t}}{B^i_{y,t}} = \pi^i_t \mathrm{d}t + \omega^i \mathrm{d}W_t, \qquad B^i_{y,t} \geq 0 \text{ given.} \tag{5.4.1}$$

Straightforward application of Itô's Lemma to (5.4.1) shows that the annuity units at time $t+h$ ($h \geq 0$) are given by

$$B^i_{y,t+h} = B^i_{y,t} \exp\left\{ \int_t^{t+h} \left(\pi^i_s - \frac{1}{2}\omega^i\omega^i \right) \mathrm{d}s + \omega^i \int_t^{t+h} \mathrm{d}W_s \right\}. \tag{5.4.2}$$

The exogenous parameters of the pension contract (i.e., $B^i_{y,t}$, π^i_t and ω^i) thus characterize the entire consumption stream in retirement.

5.4.2. The Endogenous Variables

The Appendix shows that the market value of the consumption stream in retirement $V^i_{y,t}$ is equal to

$$V^i_{y,t} = X^i_{y,t} = B^i_{y,t} A^i_{y,t}, \tag{5.4.3}$$

where the annuity factor $A^i_{y,t}$ is given by (5.3.3). The assumed interest rate can be computed from the expected growth rate of annuity units and the volatility of annuity units as follows:

$$\delta^i_t = r + \omega^i \lambda - \pi^i_t. \tag{5.4.4}$$

With a non-stochastic consumption stream (i.e., $\omega^i = 0$), equation (5.4.4) boils down to the assumed interest rate of a DB pension plan. The fraction of assets invested in the risky stock is given by (see equation (5.3.8))

$$\phi^i = \frac{\omega^i}{\sigma}. \tag{5.4.5}$$

Equation (5.4.5) shows that the volatility of annuity units determines the investment policy (see also Figure 5.2): the more volatile annuity units are, the higher the fraction of assets invested in the risky stock. Equation (5.4.5) also implies that financial shocks

$\phi^i \sigma \mathrm{d}W_t = \omega^i \mathrm{d}W_t$ are directly absorbed into current annuity units $B^i_{y,t}$ (see also equation (5.4.1)). Chapter 6 considers a pension contract in which annuity units respond gradually, rather than directly, to financial shocks.

5.4.3. Changes in Parameters

This section explores the impact of (discretionary) changes in the exogenous parameters (i.e., the volatility of annuity units and the expected growth rate of annuity units) on the endogenous variables. We also investigate changes in the Sharpe ratio and the force of mortality. Whereas the volatility of annuity units and the expected growth rate of annuity units are parameters specified by the policyholder, the Sharpe ratio and the force of mortality are parameters describing the external environment (i.e., non-traded risk factors).

5.4.3.1. Ex Ante Changes in Parameters

This section considers changes in parameters before or at the beginning of the retirement period (i.e., ex ante). We assume that a change in the expected growth rate of annuity units π^i_{t+v} is a weighted average of a change in the short-term expected growth rate of annuity units π^i_t and a change in the long-term expected growth rate of annuity units π^i_∞:

$$\Delta\pi^i_{t+v} = e^{-\theta v}\Delta\pi^i_t + \left(1 - e^{-\theta v}\right)\Delta\pi^i_\infty. \tag{5.4.6}$$

Here $\theta \geq 0$ is a weight parameter. As in Section 5.3.3.4, we assume that a change in the force of mortality μ_{t-y+v} is driven by a change in a common risk factor f:

$$\Delta\mu_{t-y+v} = g_{t-y+v}\Delta f. \tag{5.4.7}$$

The market value of future annuity units (i.e., the current account value which is an endogenous variable in the consumption approach) now changes as follows (see Appendix):

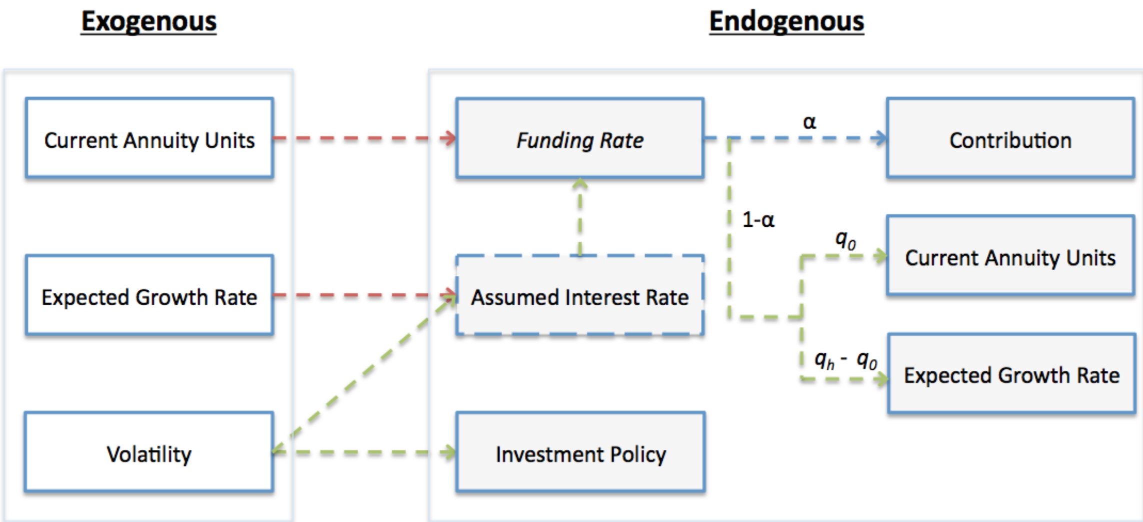
$$\frac{\Delta V^i_{y,t}}{V^i_{y,t}} \approx \widehat{D}^{i,0}_{y,t}\left(\Delta\pi^i_\infty - \omega^i\Delta\lambda - \Delta\omega^i\lambda\right) + \widehat{D}^{i,\theta}_{y,t}\left(\Delta\pi^i_t - \Delta\pi^i_\infty\right) - \widehat{G}^i_{y,t}\Delta f. \tag{5.4.8}$$

Equation (5.4.8) shows that the market value of future annuity units is an increasing function of the expected growth rate of annuity units, and a decreasing function of the volatility of annuity units, the Sharpe ratio and the force of mortality. The market value of the future consumption stream decreases if the policyholder increases the volatility of annuity units. This tempts myopic policyholders to raise the volatility of annuity units in order to consume more today. We note that in the investment approach, an increase in the investment policy (which implies an increase in the volatility of annuity units) does *not* affect the market value of the future consumption stream.

5.4.3.2. Ex Post Changes in Parameters

Parameters may also change during the retirement period (i.e., ex post). Ex post we allow for a more general closure rule than ex ante. In particular, the policyholder can adjust not only the current account value (which was endogenous ex ante) but also current annuity units and/or the expected growth rate of annuity units (which were exogenous ex ante). Figure 5.3 summarizes the consumption approach (ex post).

Figure 5.3.
Illustration of the consumption approach (ex post)



The figure illustrates the consumption approach (ex post). The left-hand side of the figure shows the exogenous parameters of the pension contract. These exogenous parameters determine the variables on the right-hand side of the figure.

Changes in parameters lead to mismatch between assets and liabilities (see equation (5.4.8)). Mismatch causes the funding rate, which is defined as the ratio between the value of the assets and the value of the liabilities, to deviate from unity. We assume that the policyholder contributes a fraction $0 \leq \alpha \leq 1$ of the mismatch (or funding gap) between assets and liabilities $\Delta V_{y,t}^{i}$ into his personal investment account. As a consequence, a fraction $(1-\alpha)$ of $\Delta V_{y,t}^{i}$ is absorbed into current annuity units and the expected growth rate of annuity units. The new balance of the investment account $X_{y,t}^{i} + \alpha \Delta V_{y,t}^{i}$ must match the new market value of future annuity units:

$$\begin{aligned} X_{y,t}^{i} + \alpha \Delta V_{y,t}^{i} &= V_{y,t}^{i} + \alpha \Delta V_{y,t}^{i} \\ &= \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} \left(V_{y,t}^{i,h} + \Delta V_{y,t}^{i,h} \right) \mathrm{d}h, \end{aligned} \tag{5.4.9}$$

where $V_{y,t}^{i,h} + \Delta V_{y,t}^{i,h}$ stands for the horizon-dependent market value of future annuity units after annuity units are adjusted:

$$\begin{aligned} V_{y,t}^{i,h} + \Delta V_{y,t}^{i,h} = B_{y,t}^{i} \exp\left\{ \frac{q_h}{\widehat{q}_{y,t}^{i}} M_{y,t}^{i} \right\} \exp\Bigg\{ - \int_0^h \big[\mu_{t-y+v} + \Delta\mu_{t-y+v} + r \\ - \left(\pi_{t+v}^{i} + \Delta\pi_{t+v}^{i} \right) + \left(\omega^{i} + \Delta\omega^{i} \right) (\lambda + \Delta\lambda) \big] \mathrm{d}v \Bigg\}. \end{aligned} \tag{5.4.10}$$

Here $M_{y,t}^{i}$ is given by (see Appendix)

$$M_{y,t}^{i} \approx -(1-\alpha) \left[\widehat{D}_{y,t}^{i,0} \left(\pi_{\infty}^{i} - \omega^{i}\Delta\lambda - \Delta\omega^{i}\lambda \right) + \widehat{D}_{y,t}^{i,\theta} \left(\Delta\pi_{t}^{i} - \pi_{\infty}^{i} \right) - \widehat{G}_{y,t}^{i} \Delta f \right], \tag{5.4.11}$$

and q_h denotes the exposure of $\log B_{y,t+h}^{i}$ to $M_{y,t}^{i} / \widehat{q}_{y,t}^{i}$. If q_h strictly increases with the horizon h, then annuity units in the distant future are more exposed to mismatch risk as compared to annuity units in the near future. We have

$$\widehat{q}_{y,t}^{i} \equiv \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} \gamma_{y,t}^{i,h} q_h \mathrm{d}h \tag{5.4.12}$$

so that

$$\frac{1}{\widehat{q}_{y,t}^{i}} \int_{\max\{x_r+y-t,0)}^{x_{\max}+y-t} \gamma_{y,t}^{i,h} q_h \mathrm{d}h = 1. \tag{5.4.13}$$

We note that the ex post closure rule coincides with the ex ante closure rule if $\alpha = 1$ (mismatch risk is not absorbed into current and future annuity units at all). We can consider the following special case for q_h (see also Chapters 4 and 6).

$$q_h = q_0 + \frac{1}{\eta}\left(1 - e^{-\eta h}\right)(1 - q_0). \tag{5.4.14}$$

The so-called growth rate method (i.e., $q_h = h$) is given by $q_0 = 0$ and $\eta \Downarrow 0$. With $q_0 = 0$, we model exponential decay, i.e., $q_h = \frac{1}{\eta}\left(1 - e^{-\eta h}\right)$.

The relative change in $V_{y,t}^{i,h}$ is given by (see Appendix)

$$\begin{aligned}\frac{\Delta V_{y,t}^{i,h}}{V_{y,t}^{i,h}} \approx{}& \left(h - \frac{q_h}{\widehat{q}_{y,t}^{i}}(1-\alpha)\widehat{D}_{y,t}^{i,0}\right)\left(\Delta\pi_\infty^i - \omega^i\Delta\lambda - \Delta\omega^i\lambda\right)\\ &+\left(D^{\theta,h} - \frac{q_h}{\widehat{q}_{y,t}^{i}}(1-\alpha)\widehat{D}_{y,t}^{i,\theta}\right)\left(\Delta\pi_t^i - \Delta\pi_\infty^i\right)\\ &-\left(G_{y,t}^{h} - \frac{q_h}{\widehat{q}_{y,t}^{i}}(1-\alpha)\widehat{G}_{y,t}^{i}\right)\Delta f.\end{aligned} \tag{5.4.15}$$

Permanent shocks do not affect $V_{y,t}^{i,h}$ if the growth rate method is adopted (i.e., $\alpha = 0$ and $q_h = h$). A temporary shock in the expected growth rate of annuity units leaves the intertemporal allocation of the market value of the consumption stream unaffected if $q_0 = \alpha = 0$ and η equals the speed θ with which the temporary shock dies out. Endogenous changes in the assumed interest rate (through endogenous changes in the expected growth rate of annuity units) exactly offset the exogenous changes in the assumed interest through the exogenous changes in the parameters. Finally, we note that a myopic policyholder has an incentive to raise the volatility of annuity units if current annuity units (rather than the expected growth rate of annuity units) is the endogenous closure variable.

5.5. Collective Defined Contribution

5.5.1. The Pension Contract

This section considers a collective defined contribution (CDC) pension system. In such a pension system, the insurer has one general pooled account. Property rights are defined in terms of annuity units, rather than in terms of a personal investment account. The current value of the collective investment account, the investment policy and the assumed interest rate are specified exogenously. In addition, the insurer specifies the annuity units at time t for each generation y. The value of the collective investment account X_t evolves according to:

$$\frac{\mathrm{d}X_t}{X_t} = (\widehat{\mu}_t + r + \phi\lambda\sigma)\,\mathrm{d}t + \phi\sigma\mathrm{d}W_t - \widehat{d}_t\mathrm{d}t, \qquad X_t \geq 0 \text{ given.} \tag{5.5.1}$$

Here

$$\widehat{\mu}_t \equiv \int_{t-x_{\max}}^{t-x_s} \mu_{t-y}\bar{c}_{y,t}\,\mathrm{d}y, \tag{5.5.2}$$

$$\widehat{d}_t \equiv \int_{t-x_{\max}}^{t-x_s} d_{y,t}\bar{c}_{y,t}\,\mathrm{d}y, \tag{5.5.3}$$

$$\bar{c}_{y,t} \equiv c_y V_{y,t}/V_t, \tag{5.5.4}$$

where c_y denotes the number of policyholders born at time y, x_s stands for the age at which policyholders enter the pension fund, $V_{y,t}$ represents the price at time t of the consumption stream in retirement for a policyholder born at time y:

$$V_{y,t} \equiv B_{y,t} \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} \exp\left\{-\int_0^h \left(\mu_{t-y+v} + \delta_{t+v}\right)\mathrm{d}v\right\}\mathrm{d}h, \tag{5.5.5}$$

V_t stands for the market value of the total liabilities:

$$V_t \equiv \int_{t-x_{\max}}^{t-x_s} c_y V_{y,t}\,\mathrm{d}y, \tag{5.5.6}$$

and $d_{y,t}$ denotes the decumulation rate at time t for a policyholder born at time y:

$$d_{y,t} = \begin{cases} B_{y,t}/V_{y,t}, & \text{if } t > y + x_r; \\ 0, & \text{if } t \leq y + x_r. \end{cases} \tag{5.5.7}$$

The assumed interest rate and the investment policy do no longer depend on i (i.e., all policyholders face the same risk profile). Hence, a pension system with a single collective asset pool allows for less tailor-made solutions. It follows from equations (5.5.1), (5.5.5) and (5.5.6) that the current annuity units $B_{y,t}$ evolve according to:

$$\frac{\mathrm{d}B_{y,t}}{B_{y,t}} = (r + \phi\lambda\sigma - \delta_t)\,\mathrm{d}t + \phi\sigma\,\mathrm{d}W_t, \qquad B_{y,t} \geq 0 \text{ given.} \tag{5.5.8}$$

5.5.2. Changes in Parameters

5.5.2.1. Investment Policy and Sharpe Ratio

Discretionary changes in the investment policy and/or the Sharpe ratio do not affect the horizon-dependent market value

$$V_{y,t}^{h} \equiv B_{y,t} \exp\left\{-\int_0^h \left(\mu_{t-y+v} + \delta_{t+v}\right)\mathrm{d}v\right\}. \tag{5.5.9}$$

Hence, in a CDC pension system, the insurer can adjust the investment policy and/or the Sharpe ratio without causing value transfers between generations.

5.5.2.2. Assumed Interest Rate and Force of Mortality

The market value of the total liabilities V_t changes as a result of changes in the assumed interest rate δ_{t+v} and/or the force of mortality μ_{t+v}. We assume that the change in the assumed interest rate and the change in the force of mortality are given by equations (5.3.12) and (5.3.20), respectively. The policyholders of the pension system contribute a fraction $0 \leq \alpha \leq 1$ of the mismatch between assets and liabilities ΔV_t into the collective asset pool. As a consequence, a fraction $(1-\alpha)$ of ΔV_t is absorbed into current and future annuity units. Figure 5.4 illustrates a CDC pension system. If we allow for value transfers between generations, then $V_{y,t}$ changes as follows (see Appendix):

$$\begin{aligned}\frac{\Delta V_{y,t}}{V_{y,t}} \approx{}& \left(\frac{\widehat{q}_{y,t}}{\widehat{q}_t}(1-\alpha)\widehat{D}_t^0 - \widehat{D}_{y,t}^0\right)\Delta\delta_\infty \\ &+ \left(\frac{\widehat{q}_{y,t}}{\widehat{q}_t}(1-\alpha)\widehat{D}_t^\kappa - \widehat{D}_{y,t}^\kappa\right)(\Delta\delta_t - \Delta\delta_\infty) \\ &- \left(\widehat{G}_{y,t} - \frac{\widehat{q}_{y,t}}{\widehat{q}_t}(1-\alpha)\widehat{G}_t\right)\Delta f,\end{aligned} \tag{5.5.10}$$

where

$$\widehat{q}_{y,t} \equiv \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} \gamma_{y,t}^{h} q_h \mathrm{d}h, \tag{5.5.11}$$

$$\widehat{q}_t \equiv \int_{t-x_{\max}}^{t-x_s} \bar{c}_{y,t} \widehat{q}_{y,t} dy, \tag{5.5.12}$$

$$\widehat{D}_t^{\kappa} \equiv \int_{t-x_{\max}}^{t-x_s} \bar{c}_{y,t} \widehat{D}_{y,t}^{\kappa} \, \mathrm{d}y, \tag{5.5.13}$$

$$\widehat{G}_t \equiv \int_{t-x_{\max}}^{t-x_s} \bar{c}_{y,t} \widehat{G}_{y,t} \, \mathrm{d}y, \tag{5.5.14}$$

with

$$\gamma_{y,t}^{h} \equiv \frac{V_{y,t}^{h}}{V_{y,t}} \tag{5.5.15}$$

$$\widehat{D}_{y,t}^{\kappa} \equiv \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} \gamma_{y,t}^{h} D^{\kappa,h} \mathrm{d}h, \tag{5.5.16}$$

$$\widehat{G}_{y,t} \equiv \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} \gamma_{y,t}^{h} G_{y,t}^{h} \mathrm{d}h. \tag{5.5.17}$$

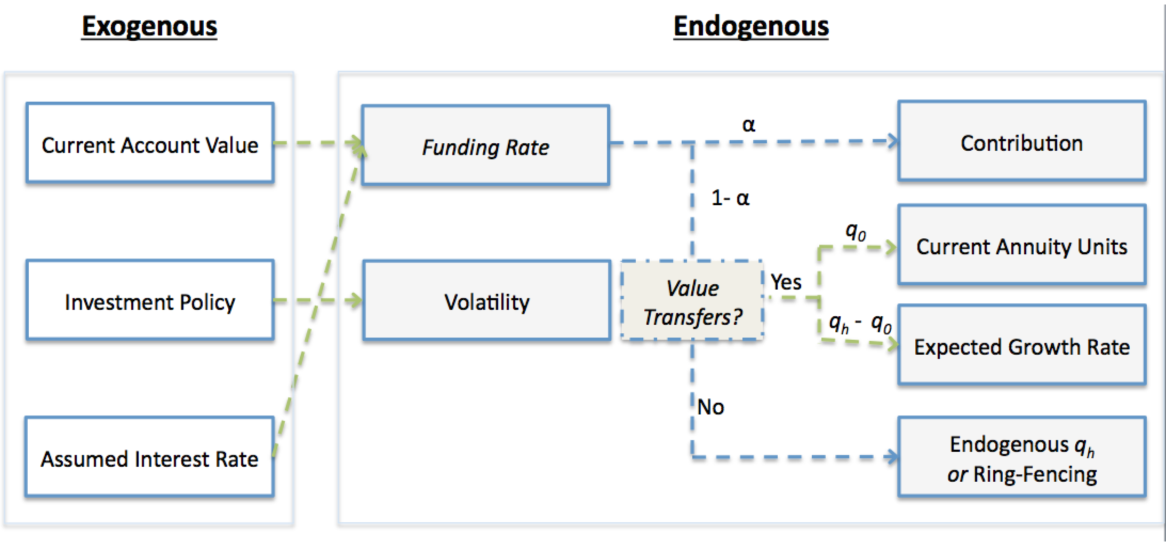
Permanent shocks do not affect $V_{y,t}$ if the growth rate method is adopted (i.e., $\alpha = 0$ and $q_h = h$). If, however, the level method is adopted (i.e., $\alpha = 0$ and $q_h = 1$), then a change in the assumed interest rate redistributes value from long horizons (i.e., young policyholders) to short horizons (i.e., old policyholders). Intuitively, a higher assumed interest rate raises the funding rate, thereby increasing the scope to pay out today. There are no value transfers between generations if the insurer chooses q_h such that $\Delta V_{y,t}/V_{y,t} = 0$. Alternatively, the insurer can ring-fence the assets of each generation. In that case, adjustments in current annuity units typically depends on the age of the policyholder.

5.6. Collective Defined Ambition

5.6.1. The Pension Contract

This section considers a collective defined ambition (CDA) pension system. The pension fund has one general pooled account and defines property rights in terms of annuity units. As in the consumption approach, current annuity units, the expected growth

Figure 5.4.
Illustration of a CDC pension contract



The figure illustrates a CDC pension contract. The left-hand side of the figure shows the exogenous parameters of the pension contract. These exogenous parameters determine the variables on the right-hand side of the figure.

rate of annuity units, and the volatility of annuity units are specified exogenously. The annuity units are adjusted according to a collective version of (5.4.1):

$$\frac{\mathrm{d}B_{y,t}}{B_{y,t}} = \pi_t \mathrm{d}t + \omega \mathrm{d}W_t, \qquad B_{y,t} \geq 0 \text{ given.} \tag{5.6.1}$$

The volatility of annuity units and the expected growth rate of annuity units do no longer depend on i.

In a collective pension system without personal investment accounts, proper pricing of future annuity units is essential. Indeed, if annuity units are not priced properly, some generations may sponsor other generations, thereby giving rise to intergenerational value transfers. The price at time t of the consumption stream in retirement for a policyholder born at time y is given by

$$V_{y,t} = \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} V_{y,t}^h \,\mathrm{d}h. \tag{5.6.2}$$

Here $V_{y,t}^h$ represents the price at time t of annuity units at time $t+h$ for a policyholder born at time y:

$$V_{y,t}^{h} = B_{y,t} \exp\left\{ -\int_{0}^{h} \left(\mu_{t-y+v} + r - \pi_{t+v} + \omega\lambda \right) \mathrm{d}v \right\}. \tag{5.6.3}$$

The total market value is now defined as follows:

$$V_t \equiv \int_{t-x_{\max}}^{t-x_s} c_y \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} V_{y,t}^{h} \, \mathrm{d}h \, \mathrm{d}y, \tag{5.6.4}$$

5.6.2. Changes in Parameters

5.6.2.1. Ex Ante Changes in Parameters

In a CDA pension system, changes in parameters before or at the beginning of retirement (i.e., ex ante) typically lead to a change in the purchase price of annuity units. As in Section 5.4.3, we consider changes in the volatility of annuity units, the expected growth rate of annuity units, the Sharpe ratio and the force of mortality. The change in the expected growth of annuity units and the change in the force of mortality are given by equations (5.4.6) and (5.4.7), respectively. As a result, the purchase price of annuity units $V_{y,t}$ changes as follows (see Appendix):[71]

$$\frac{\Delta V_{y,t}}{V_{y,t}} \approx \widehat{D}_{y,t}^{0} \left(\Delta\pi_\infty - \omega\Delta\lambda - \Delta\omega\lambda \right) + \widehat{D}_{y,t}^{\theta} \left(\Delta\pi_t - \Delta\pi_\infty \right) - \widehat{G}_{y,t} \Delta f. \tag{5.6.5}$$

Younger policyholders (i.e., larger h) are more affected by a change in parameters than older policyholders (i.e., smaller h).

5.6.2.2. Ex Post Changes in Parameters

In a CDA pension system, a change in parameters during retirement (i.e., ex post) typically leads to intergenerational redistribution of market value. The policyholders of the pension system contribute a fraction $0 \leq \alpha \leq 1$ of the mismatch between assets and liabilities ΔV_t into the collective asset pool. A fraction $(1 - \alpha)$ of ΔV_t is thus absorbed

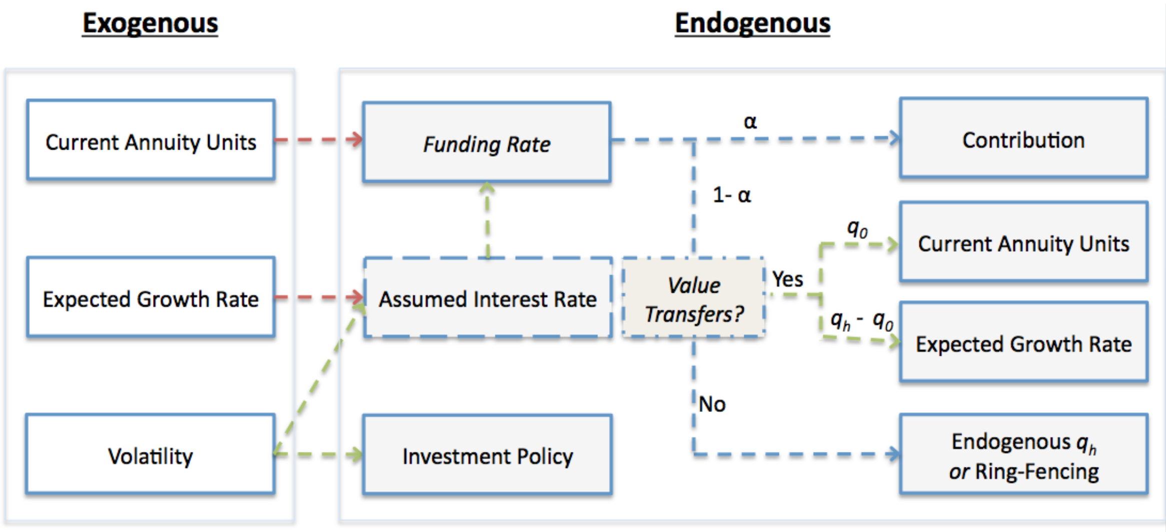
[71] Rauh (2008) addresses a change in discounting as a result of a change in the volatility of annuity units in the context of DB pension plans with corporate risk sponsors. He shows that if corporations can employ the expected return on their investment portfolio to discount annuity units, corporate sponsors face an incentive to raise the risk of future annuity units. Intuitively, by raising the risk of future annuity units, they can reduce the cost to the corporation of sponsoring these annuity units at the expense of the policyholders of the pension fund.

into current and future annuity units. Figure 5.5 illustrates a CDA pension system. If we allow for value transfers between generations, then $V_{y,t}$ changes as follows (see Appendix):

$$\begin{aligned}\frac{\Delta V_{y,t}}{V_{y,t}} \approx & \left(\widehat{D}^0_{y,t} - \frac{\widehat{q}_{y,t}}{\widehat{q}_t}(1-\alpha)\widehat{D}^0_t\right)(\Delta\pi_\infty - \omega\Delta\lambda - \lambda\Delta\omega) \\ & + \left(\widehat{D}^\theta_{y,t} - \frac{\widehat{q}_{y,t}}{\widehat{q}_t}(1-\alpha)\widehat{D}^\theta_t\right)(\Delta\pi_t - \Delta\pi_\infty) \\ & - \left(\widehat{G}_{y,t} - \frac{\widehat{q}_{y,t}}{\widehat{q}_t}(1-\alpha)\widehat{G}_t\right)\Delta f.\end{aligned} \tag{5.6.6}$$

Whether one wants to ensure that discretionary changes in the Sharpe ratio do not lead to intergenerational distribution is a matter of debate. On the one hand, one could argue that changes in the Sharpe ratio should lead to similar intergenerational risk sharing as with changes in the interest rate. In particular, a higher Sharpe ratio (at a given risk) raises the expected future rates of return and thus reduces the current price of funding an uncertain future pension. The pension contract thus allows generations to share risk factors that are not traded on financial markets. This approach views the pension contract as a social contract that allows for trade in risk factors (such as the Sharpe ratio) that are not traded on financial markets. On the other hand, allowing changes in the Sharpe ratio to redistribute market value across horizons and therefore across generations may well lead to intergenerational conflicts about the unobservable Sharpe ratio. Moreover, pension funds cannot hedge discretionary changes in the Sharpe ratio. Hence, to avoid these problems, one could argue that the horizon-dependent market value $V_{y,t}$ should remain constant if λ changes. This approach views the pension contract as a pure financial contract that includes only risk factors that are traded on financial markets. A value neutral transfer can be accomplished by ring-fencing the assets of each generation or choosing q_h such that $\Delta V_{y,t}/V_{y,t} = 0$.

Figure 5.5.
Illustration of a CDA pension contract



The figure illustrates a CDA pension contract. The left-hand side of the figure shows the exogenous parameters of the pension contract. These exogenous parameters determine the variables on the right-hand side of the figure.

5.7. Concluding Remarks

Private pension provision faces the challenge of providing adequate retirement income. PPRs promise to play a new role in the provision of retirement income. These pension plans individualize the savings, investment and withdrawal functions of variable annuities and arrange the insurance function collectively. We have explored two approaches to a PPR: the investment approach and the consumption approach. In the first approach, the contribution level, the investment policy and the assumed interest rate are specified exogenously, while in the second approach, current annuity units, the expected growth rate of annuity and the volatility of annuity units are specified exogenously. We have demonstrated that in the investment approach, the policyholder can freely adjust the investment policy without affecting the intertemporal allocation of the market value of the consumption stream. This property does not hold true in the consumption approach.

5.8. Appendix

Proofs

Derivation of (5.3.11), (5.3.13), (5.3.17), (5.3.18), (5.3.21) and (5.3.24)

Expected annuity units are given by (this follows from equation (5.3.7))

$$\mathbb{E}_t\left[B^i_{y,t+h}\right] = B^i_{y,t}\exp\left\{\int_0^h \pi^i_{t+v}\mathrm{d}v\right\} = B^i_{y,t}\exp\left\{\int_0^h \left(r+\phi^i\lambda\sigma-\delta^i_{t+v}\right)\mathrm{d}v\right\}.$$

A change in the investment policy causes expected annuity units to change by

$$\Delta\mathbb{E}_t\left[B^i_{y,t+h}\right] = B^i_{y,t}\exp\left\{\int_0^h \left(r+\phi^i\lambda\sigma-\delta^i_{t+v}\right)\mathrm{d}v\right\}\left(\exp\left\{\int_0^h \lambda\sigma\Delta\phi^i\mathrm{d}v\right\}-1\right).$$

Hence, by Taylor series expansion,

$$\frac{\Delta\mathbb{E}_t\left[B^i_{y,t+h}\right]}{\mathbb{E}_t\left[B^i_{y,t+h}\right]} = \exp\left\{\int_0^h \lambda\sigma\Delta\phi^i\mathrm{d}v\right\}-1 \approx h\lambda\sigma\Delta\phi^i.$$

A change in the assumed interest rate causes the annuity factor to change by

$$\begin{aligned}
\frac{\Delta A^i_{y,t}}{A^i_{y,t}} &= \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} \gamma^{i,h}_{y,t}\exp\left\{-\int_0^h \left(e^{-\kappa v}\Delta\delta^i_t+\left(1-e^{-\kappa v}\right)\Delta\delta^i_\infty\right)\mathrm{d}v\right\}\mathrm{d}h-1\\
&\approx \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} \gamma^{i,h}_{y,t}\left(1-\int_0^h \left(e^{-\kappa v}\Delta\delta^i_t+\left(1-e^{-\kappa v}\right)\Delta\delta^i_\infty\right)\mathrm{d}v\right)\mathrm{d}h-1\\
&= -\int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} \gamma^{i,h}_{y,t}\int_0^h \left(e^{-\kappa v}\Delta\delta^i_t+\left(1-e^{-\kappa v}\right)\Delta\delta^i_\infty\right)\mathrm{d}v\mathrm{d}h\\
&= -\widehat{D}^{i,\kappa}_{y,t}\left(\Delta\delta^i_t-\Delta\delta^i_\infty\right)-\widehat{D}^{i,0}_{y,t}\Delta\delta^i_\infty.
\end{aligned}$$

It follows from equation (5.3.5) that

$$\frac{\Delta B^i_{y,t}}{B^i_{y,t}} \approx -\frac{\Delta A^i_{y,t}}{A^i_{y,t}} \approx \widehat{D}^{i,0}_{y,t}\Delta\delta^i_\infty+\widehat{D}^{i,\kappa}_{y,t}\left(\Delta\delta^i_t-\Delta\delta^i_\infty\right).$$

The relative change in $V^{i,h}_{y,t}$ is given by

$$
\begin{aligned}
\frac{\Delta V^{i,h}_{y,t}}{V^{i,h}_{y,t}} &\approx \frac{\Delta B^{i}_{y,t}}{B^{i}_{y,t}} + \exp\left\{-\int_0^h \left(e^{-\kappa v}\Delta\delta^i_t + \left(1-e^{-\kappa v}\right)\Delta\delta^i_\infty\right)\mathrm{d}v\right\} - 1 \\
&\approx \frac{\Delta B^{i}_{y,t}}{B^{i}_{y,t}} - \int_0^h \left(e^{-\kappa v}\Delta\delta^i_t + \left(1-e^{-\kappa v}\right)\Delta\delta^i_\infty\right)\mathrm{d}v \\
&\approx \left(\widehat{D}^{i,0}_{y,t} - h\right)\Delta\delta^i_\infty + \left(\widehat{D}^{i,\kappa}_{y,t} - D^{\kappa,h}\right)\left(\Delta\delta^i_t - \Delta\delta^i_\infty\right).
\end{aligned}
$$

The relative change in $\mathbb{E}_t\left[B^i_{y,t+h}\right]$ can be computed in a similar fashion. A change in the force of mortality causes the annuity factor to change by

$$
\begin{aligned}
\frac{\Delta A^{i}_{y,t}}{A^{i}_{y,t}} &= \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} \gamma^{i,h}_{y,t}\exp\left\{-G^h_{y,t}\Delta f\right\}\mathrm{d}h - 1 \\
&\approx \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} \gamma^{i,h}_{y,t}\left(1 - G^h_{y,t}\Delta f\right)\mathrm{d}h - 1 \\
&= -\int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} \gamma^{i,h}_{y,t}G^h_{y,t}\mathrm{d}h\Delta f \\
&= -\widehat{G}^i_{y,t}\Delta f.
\end{aligned}
$$

It follows from equation (5.3.5) that

$$
\frac{\Delta B^{i}_{y,t}}{B^{i}_{y,t}} \approx -\frac{\Delta A^{i}_{y,t}}{A^{i}_{y,t}} \approx \widehat{G}^i_{y,t}\Delta f.
$$

The relative change in $V^{i,h}_{y,t}$ is given by

$$
\frac{\Delta V^{i,h}_{y,t}}{V^{i,h}_{y,t}} \approx \frac{\Delta B^{i}_{y,t}}{B^{i}_{y,t}} + \exp\left\{-G^h_{y,t}\Delta f\right\} - 1 \approx \frac{\Delta B^{i}_{y,t}}{B^{i}_{y,t}} - G^h_{y,t}\Delta f \approx \left(\widehat{G}^i_{y,t} - G^h_{y,t}\right)\Delta f.
$$

Derivation of (5.4.3), (5.4.8), (5.4.11) and (5.4.15)

The market value is given by

$$
V^i_{y,t} = \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} \mathbb{E}_t\left[\frac{m_{t+h}}{m_t}B^i_{y,t+h}\right]\mathrm{d}h.
$$

Straightforward computations show that

$$\begin{aligned} V^i_{y,t} &= \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} \mathbb{E}_t \left[\exp\left\{ \int_0^h -\left(r+\frac{1}{2}\lambda^2\right) \mathrm{d}v - \lambda \int_0^h \mathrm{d}W_{t+v} \right\} \right. \\ &\quad \left. B^i_{y,t} \exp\left\{ \int_0^h \left(\pi_{t+v} - \frac{1}{2}\omega^i\omega^i\right) \mathrm{d}v + \omega^i \int_0^h \mathrm{d}W_{t+v} \right\} \right] \mathrm{d}h \\ &= \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} B^i_{y,t} \mathbb{E}_t \left[\exp\left\{ -\int_0^h \left(r+\frac{1}{2}\lambda^2 - \pi_{t+v} + \frac{1}{2}\omega^i\omega^i\right) \mathrm{d}v \right\} \right. \\ &\quad \left. \exp\left\{ \left(\omega^i - \lambda\right) \int_0^h \mathrm{d}W_{t+v} \right\} \right] \mathrm{d}h \\ &= B^i_{y,t} \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} \exp\left\{ -\int_0^h \left(r + \lambda\omega^i - \pi_{t+v}\right) \mathrm{d}v \right\} \mathrm{d}h = B^i_{y,t} A^i_{y,t}. \end{aligned}$$

The relative change in the market value of future annuity units is given by (the second equality follows from equations (5.4.6) and (5.4.7))

$$\begin{aligned} \frac{\Delta V^i_{y,t}}{V^i_{y,t}} &\approx \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} \gamma^{i,h}_{y,t} \exp\left\{ -\int_0^h \left(\Delta\mu_{x+v} - \Delta\pi^i_{t+v}\right.\right. \\ &\quad \left.\left. + \omega^i\Delta\lambda + \lambda\Delta\omega^i\right) \mathrm{d}v \right\} \mathrm{d}h - 1 \\ &\approx \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} \gamma^{i,h}_{y,t} \exp\left\{ -G^h_{y,t}\Delta f + h\left(\Delta\pi^i_\infty - \omega^i\Delta\lambda - \lambda\Delta\omega^i\right) \right. \\ &\quad \left. + D^{\theta,h}\left(\Delta\pi^i_t - \Delta\pi^i_\infty\right) \right\} \mathrm{d}h - 1 \end{aligned} \tag{5.8.1}$$

Here we assume that $\Delta\omega^i\Delta\lambda \approx 0$. By Taylor series expansion, we can write

$$\frac{\Delta V^i_{y,t}}{V^i_{y,t}} \approx \widehat{D}^{i,0}_{y,t}\left(\Delta\pi^i_\infty - \omega^i\Delta\lambda - \lambda\Delta\omega^i\right) + \widehat{D}^{i,\theta}_{y,t}\left(\Delta\pi^i_t - \Delta\pi^i_\infty\right) - \widehat{G}^i_{y,t}\Delta f. \tag{5.8.2}$$

We also have (this follows from equation (5.4.9))

$$\begin{aligned} \alpha\frac{\Delta V^i_{y,t}}{V^i_{y,t}} &= \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} \gamma^{i,h}_{y,t} \exp\left\{ \frac{q_h}{\widehat{q}^i_{y,t}} M^i_{y,t} \right\} \times \\ &\quad \exp\left\{ -G^h_{y,t}\Delta f + h\left(\Delta\pi^i_\infty - \omega^i\Delta\lambda - \Delta\omega^i\lambda\right) + D^{\theta,h}\left(\Delta\pi^i_t - \Delta\pi^i_\infty\right) \right\} \mathrm{d}h - 1. \end{aligned}$$

By Taylor series expansion, we can write

$$\alpha\frac{\Delta V^i_{y,t}}{V^i_{y,t}} \approx M^i_{y,t} + \widehat{D}^{i,0}_{y,t}\left(\pi^i_\infty - \omega^i\Delta\lambda - \Delta\omega^i\lambda\right) + \widehat{D}^{i,\theta}_{y,t}\left(\Delta\pi^i_t - \pi^i_\infty\right) - \widehat{G}^i_{y,t}\Delta f. \tag{5.8.3}$$

It follows from (5.8.2) that

$$\alpha\frac{\Delta V^{i}_{y,t}}{V^{i}_{y,t}} \approx \alpha\left[\widehat{D}^{i,0}_{y,t}\left(\Delta\pi^{i}_{\infty}-\omega^{i}\Delta\lambda-\lambda\Delta\omega^{i}\right)+\widehat{D}^{i,\theta}_{y,t}\left(\Delta\pi^{i}_{t}-\Delta\pi^{i}_{\infty}\right)-\widehat{G}^{i}_{y,t}\Delta f\right]. \quad (5.8.4)$$

Setting (5.8.3) equal to (5.8.4) and solving for $M^{i}_{y,t}$ yields

$$\begin{aligned} M^{i}_{y,t} \approx -(1-\alpha)&\left[\widehat{D}^{i,0}_{y,t}\left(\Delta\pi^{i}_{\infty}-\omega^{i}\Delta\lambda-\lambda\Delta\omega^{i}\right)\right.\\ &\left.+\widehat{D}^{i,\theta}_{y,t}\left(\Delta\pi^{i}_{t}-\Delta\pi^{i}_{\infty}\right)-\widehat{G}^{i}_{y,t}\Delta f\right]. \end{aligned} \quad (5.8.5)$$

We also have

$$\begin{aligned} \frac{\Delta V^{i,h}_{y,t}}{V^{i,h}_{y,t}} = \exp&\left\{\frac{q_h}{\widehat{q}^{i}_{y,t}}M^{i}_{y,t}+h\left(\Delta\pi^{i}_{\infty}-\omega^{i}\Delta\lambda-\lambda\Delta\omega^{i}\right)+D^{\theta,h}\left(\Delta\pi^{i}_{t}-\Delta\pi^{i}_{\infty}\right)\right.\\ &\left.-G^{h}_{y,t}\Delta f\right\}-1. \end{aligned}$$

By Taylor series expansion, we can write

$$\begin{aligned} \frac{\Delta V^{i,h}_{y,t}}{V^{i,h}_{y,t}} \approx &\frac{q_h}{\widehat{q}^{i}_{y,t}}M^{i}_{y,t}+h\left(\Delta\pi^{i}_{\infty}-\omega^{i}\Delta\lambda-\lambda\Delta\omega^{i}\right)\\ &+D^{\theta,h}\left(\Delta\pi^{i}_{t}-\Delta\pi^{i}_{\infty}\right)-G^{h}_{y,t}\Delta f. \end{aligned} \quad (5.8.6)$$

Substituting (5.8.5) into (5.8.6) yields (5.4.15).

Derivation of (5.5.10)

We have (a fraction α of $\Delta V_t/V_t$ is absorbed into current and future annuity units)

$$\begin{aligned} \alpha\frac{\Delta V_t}{V_t} = &\int_{t-x_{\max}}^{t-x_s}\int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t}\frac{c_y V^{h}_{y,t}}{V_t}\exp\left\{\frac{q_h}{\widehat{q}_t}M_t\right\}\times\\ &\exp\left\{-G^{h}_{y,t}\Delta f-h\Delta\delta_{\infty}-D^{\kappa,h}\left(\Delta\delta_t-\Delta\delta_{\infty}\right)\right\}\mathrm{d}h\,\mathrm{d}y-1. \end{aligned}$$

By Taylor series expansion, we can write

$$\alpha\frac{\Delta V_t}{V_t} \approx M_t-\widehat{D}^{0}_{t}\Delta\delta_{\infty}-\widehat{D}^{\kappa}_{t}\left(\Delta\delta_t-\Delta\delta_{\infty}\right)-\widehat{G}_t\Delta f. \quad (5.8.7)$$

It also follows that

$$\alpha\frac{\Delta V_t}{V_t} \approx \int_{t-x_{\max}}^{t-x_s}\int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t}\frac{c_y V_{y,t}^h}{V_t}\exp\Bigg\{-G_{y,t}^h\Delta f - h\Delta\delta_\infty \\ - D^{\kappa,h}\left(\Delta\delta_t - \Delta\delta_\infty\right)\Bigg\}\mathrm{d}h\,\mathrm{d}y - 1 \\ \approx \alpha\left[-\widehat{D}_t^0\Delta\delta_\infty - \widehat{D}_t^\kappa\left(\Delta\delta_t - \Delta\delta_\infty\right) - \widehat{G}_t\Delta f\right]. \tag{5.8.8}$$

Setting (5.8.7) equal to (5.8.8) and solving for M_t yields

$$M_t \approx (1-\alpha)\left[-\widehat{D}_t^0\Delta\delta_\infty - \widehat{D}_t^\kappa\left(\Delta\delta_t - \Delta\delta_\infty\right) - \widehat{G}_t\Delta f\right]. \tag{5.8.9}$$

We also have

$$\frac{\Delta V_{y,t}}{V_{y,t}} = \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t}\gamma_{y,t}^h\exp\Bigg\{\frac{q_h}{\widehat{q}_t}M_t - h\Delta\delta_\infty^i - D^{\kappa,h}\left(\Delta\delta_t - \Delta\delta_\infty^i\right) \\ - G_{y,t}^h\Delta f\Bigg\}\,\mathrm{d}h - 1.$$

By Taylor series expansion, we can write

$$\frac{\Delta V_{y,t}}{V_{y,t}} \approx \frac{q_h}{\widehat{q}_t}M_t - \widehat{D}_{y,t}^0\Delta\delta_\infty - \widehat{D}_{y,t}^\kappa\left(\Delta\delta_t - \Delta\delta_\infty^i\right) - \widehat{G}_{y,t}\Delta f. \tag{5.8.10}$$

Substituting (5.8.9) into (5.8.10) yields (5.5.10).

Derivation of (5.6.5) and (5.6.6)

The relative change in the market value of future annuity units is given by

$$\frac{\Delta V_{y,t}}{V_{y,t}} \approx \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t}\gamma_{y,t}^h\exp\left\{-\int_0^h\left(\Delta\mu_{x+v} - \Delta\pi_{t+v} + \omega\Delta\lambda + \lambda\Delta\omega\right)\mathrm{d}v\right\}\mathrm{d}h - 1 \\ \approx \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t}\gamma_{y,t}^h\exp\Bigg\{-G_{y,t}^h\Delta f + h\left(\Delta\pi_\infty - \omega\Delta\lambda - \lambda\Delta\omega\right) \\ + D^{\theta,h}\left(\Delta\pi_t - \Delta\pi_\infty\right)\Bigg\}\mathrm{d}h - 1 \tag{5.8.11}$$

Here we assume that $\Delta\omega\Delta\lambda \approx 0$. By Taylor series expansion, we can write

$$\frac{\Delta V_{y,t}}{V_{y,t}} \approx \widehat{D}^{0}_{y,t}\left(\Delta\pi_\infty - \omega\Delta\lambda - \lambda\Delta\omega\right) + \widehat{D}^{\theta}_{y,t}\left(\Delta\pi_t - \Delta\pi_\infty\right) - \widehat{G}_{y,t}\Delta f. \tag{5.8.12}$$

We also have (a fraction α of $\Delta V_t/V_t$ is absorbed into current and future annuity units)

$$\alpha\frac{\Delta V_t}{V_t} = \int_{t-x_{\max}}^{t-x_s}\int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} \frac{c_y V^h_{y,t}}{V_t}\exp\left\{\frac{q_h}{\widehat{q}_t}M_t\right\}\times$$
$$\exp\left\{-G^h_{y,t}\Delta f + h\left(\Delta\pi_\infty - \omega\Delta\lambda - \Delta\omega\lambda\right) + D^{\theta,h}\left(\Delta\pi_t - \Delta\pi_\infty\right)\right\}\mathrm{d}h\,\mathrm{d}y - 1.$$

By Taylor series expansion, we can write

$$\alpha\frac{\Delta V_t}{V_t} \approx M_t + \widehat{D}^{0}_{t}\left(\pi_\infty - \omega\Delta\lambda - \Delta\omega\lambda\right) + \widehat{D}^{\theta}_{t}\left(\Delta\pi_t - \pi_\infty\right) - \widehat{G}_t\Delta f. \tag{5.8.13}$$

It also follows that

$$\begin{aligned}\alpha\frac{\Delta V_t}{V_t} &\approx \int_{t-x_{\max}}^{t-x_s}\int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} \frac{c_y V^h_{y,t}}{V_t}\exp\Bigg\{-G^h_{y,t}\Delta f \\ &\quad + h\left(\Delta\pi_\infty - \omega\Delta\lambda - \lambda\Delta\omega\right) + D^{\theta,h}\left(\Delta\pi_t - \Delta\pi_\infty\right)\Bigg\}\mathrm{d}h\,\mathrm{d}y - 1 \\ &\approx \alpha\left[\widehat{D}^{0}_{t}\left(\Delta\pi_\infty - \omega\Delta\lambda - \lambda\Delta\omega\right) + \widehat{D}^{\theta}_{t}\left(\Delta\pi_t - \Delta\pi_\infty\right) - \widehat{G}_t\Delta f\right]\end{aligned} \tag{5.8.14}$$

Setting (5.8.13) equal to (5.8.14) and solving for M_t yields

$$M_t \approx -(1-\alpha)\left[\widehat{D}^{0}_{t}\left(\Delta\pi_\infty - \omega\Delta\lambda - \lambda\Delta\omega\right) + \widehat{D}^{\theta}_{t}\left(\Delta\pi_t - \Delta\pi_\infty\right) - \widehat{G}_t\Delta f\right]. \tag{5.8.15}$$

We also have

$$\frac{\Delta V_{y,t}}{V_{y,t}} = \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t}\gamma^h_{y,t}\exp\left\{\frac{q_h}{\widehat{q}_t}M_t + h\left(\Delta\pi^i_\infty - \omega\Delta\lambda - \lambda\Delta\omega\right)\right.$$
$$\left. + D^{\theta,h}\left(\Delta\pi_t - \Delta\pi^i_\infty\right) - G^h_{y,t}\Delta f\right\}\mathrm{d}h - 1.$$

By Taylor series expansion, we can write

$$\begin{aligned}\frac{\Delta V_{y,t}}{V_{y,t}} \approx\ & \frac{q_h}{\widehat{q}_t} M_t + \widehat{D}^{0}_{y,t} \left(\Delta\pi_\infty - \omega\Delta\lambda - \lambda\Delta\omega\right) \\ & + \widehat{D}^{\theta}_{y,t} \left(\Delta\pi_t - \Delta\pi^{i}_\infty\right) - \widehat{G}_{y,t}\Delta f.\end{aligned} \tag{5.8.16}$$

Substituting (5.8.15) into (5.8.16) yields (5.6.6).

Chapter 6

Buffering Shocks in Variable Annuities: Valuation, Investment and Communication[72]

This chapter explores defined ambition pension plans that allocate stock market risk among policyholders on the basis of complete pension contracts while pooling idiosyncratic longevity risk. Entitlements respond gradually to financial shocks. We show how pension entitlements can be valued in a market-consistent fashion. The market-consistent discount rate includes a risk premium that rises with the horizon. Proper valuation ensures efficient intertemporal consumption smoothing and protects the value of property rights of existing policyholders. In the tradition of liability-driven investment, we determine the investment policy from the stochastic pension promises. We show that gradual absorption of financial shocks leads to life cycle investment.

6.1. Introduction

Corporations are increasingly withdrawing as sponsors from defined benefit (DB) pension plans (Investment Company Institute, 2014). As a consequence, pension funds become mutual insurers in which policyholders, rather than corporations, bear stock market risk. These so-called defined ambition (DA) pension plans retain several advantages of DB pension plans.[73] In particular, by pooling idiosyncratic longevity risk, pension funds can still provide lifelong benefits at relatively low costs. Also, by specifying (dis)saving and investment decisions, DA pension plans protect financially illiterate policyholders

[72]This chapter is co-authored with Lans Bovenberg and Roel Mehlkopf.

[73]DA pension plans share some characteristics with non-financial defined contribution (NDC) pension plans (Holzmann, Palmer, and Robalino, 2012). In particular, both NDC and DA pension plans lack outside sponsors.

against behavioral biases. To illustrate, risk management and payout policies seek to provide policyholders with stable income streams after they retire.

Pension entitlements are adjusted gradually after an unexpected financial shock that causes a mismatch between assets and the value of liabilities. Hence retirees take investment risk but can take some time to adjust their standard of living after an unexpected financial shock. Gradual absorption of financial shocks is consistent with internal habit formation (see, e.g., Fuhrer, 2000); see also Chapter 4. It reconciles defined benefit thinking in which current consumption is stable with defined contribution thinking in which policyholders bear investment risk. Indeed, future rather than current consumption bears most of current investment risk. As a consequence, the year-on-year volatility of current consumption is smaller than the year-on-year volatility of wealth. This property stands in sharp contrast to many (popular) variable annuity products. Indeed, traditional variable annuities usually assume that payouts respond directly, rather than gradually, to an unexpected financial shock (see, e.g., Chai et al., 2011; Maurer et al., 2013b).[74] Other important parameters of the pension contract are the assumed expected rates of return on risky securities. If actual expected rates of return exceed assumed expected rates of return, then the actual growth rate of annuity units – ceteris paribus – exceeds the desired growth rate of annuity units, and vice versa.

We show how pension entitlements can be valued in a market-consistent fashion. Proper valuation is relevant for determining the prices at which variable annuities can be bought and sold. Indeed, we show how pension contributions can be derived endogenously from the stochastic pension promises, which are the liabilities of the pension contract. This is reminiscent of DB pension plans in which pension contributions are determined by the costs of the desired consumption stream in retirement. The discount rate is equal to the sum of the interest rate and a risk premium that rises with the investment horizon. We show that the discount rate for valuing variable annuities is typically in between the expected rate of return on the actual investment portfolio and the risk-free interest rate. Market-consistent valuation ensures efficient intertemporal consumption smoothing and intergenerational fairness. It also helps to protect the value of property rights of existing policyholders. The contract can thus be changed without giving rise to value transfers.

[74] Insurers have also developed variable annuities for which payouts respond sluggishly to an unexpected financial shock (see, e.g., Guillén et al., 2006; Maurer et al., 2013a, 2014). However, these variable annuity products are often based on complex profit-sharing rules and hence difficult to value.

We demonstrate how the investment policy can be determined endogenously from the stochastic pension promises. The current chapter thus extends the principle of liability-driven investment from DB pension plans to DA pension plans. We in fact generalize asset-liability management (ALM) to stochastic liabilities. Indeed, the pension contract is complete not only in terms of the allocation of stock market risk but also in terms of the investment policy so that policyholders obtain the exposures that have been communicated to them. Furthermore, we show that gradual absorption of financial shocks leads to life cycle investment in which the stock market exposure declines with age. This is because retired agents have less time to absorb financial shocks when their remaining expected lifetime declines. Stock market exposures are thus tailored to the investment horizon.

DA pension plans are based on proposed risk-sharing systems in the Netherlands, and evolved from traditional DB pension plans with (nominally) guaranteed pension entitlements. Also in public-sector pension plans in the United States, risk sharing is being considered as a way to reduce the costs of these plans (see, e.g., Novy-Marx and Rauh, 2014). This chapter contributes to the emerging literature on the implications of moving from a DB design towards a DA design. The DA plans considered in the current chapter encompass both the accumulation and the decumulation phases; however, a DA plan can be limited to the decumulation phase or the accumulation phase only. Indeed, one can view a DA pension plan as a particular way to draw-down accumulated retirement wealth.

The remainder of this chapter is structured as follows. Section 6.2 describes the financial market. Section 6.3 specifies the pension contract. This section also investigates how pension funds can calibrate and communicate the risk of future pension entitlements. Section 6.4 explores the pricing of future pension entitlements, and derives the replicating portfolio strategy. Section 6.5 shows how the pension contract can be expressed in terms of mismatch between assets and liabilities. This section also investigates the impact of using an incorrect discount rate on intertemporal consumption smoothing. Section 6.6 explores a subclass of risk profiles in which adjustments in pension entitlements are determined by a single state variable. Section 6.7 concludes the chapter and explores the roles of public supervision. Proofs are relegated to the Appendix.

6.2. The Financial Market

We assume a simple continuous-time financial market with a single risk factor, which we interpret as an aggregate stock market index. Let us denote by S_t the value of the stock market index at time t. Throughout, boldface type is used to denote uncertain variables at time t. The stock market return in any discrete-time period $t+j$ $(j > 1)$ is given by

$$\frac{\mathbf{S}_{t+j}}{\mathbf{S}_{t+j-1}} = \exp\left\{\left(r + \lambda\sigma - \frac{1}{2}\sigma^2\right) + \sigma \int_{t+j-1}^{t+j} \mathrm{d}\mathbf{W}_s\right\}. \tag{6.2.1}$$

Here r is the nominal risk-free interest rate, λ denotes the equity risk premium per unit of risk (i.e., Sharpe ratio), σ represents the stock return volatility, and W_t corresponds to a standard Brownian motion. Empirically, the value of $\lambda\sigma - \sigma^2/2$ is found to be positive (see, e.g., Brennan and Xia, 2002).

6.3. The Pension Contract

6.3.1. Specification

In this section, we specify the pension contract. Entitlements are defined in terms of annuity units. Denote by $B_{y,t}$ the annuity units at time t of a policyholder born at time y, by x_r the age at which a policyholder retires, and by $x_{\max}$ the maximum age a policyholder can reach. The insurer adjusts annuity units at discrete points in time $t = t_0{+}1, t_0{+}2, ..., y{+}x_{\max}$. Here t_0 represents the time at which the pension contribution is paid. If the birth date y falls between $t - x_r$ and $t - x_{\max}$ and the policyholder is still alive at time t, then this policyholder receives a pension payment at time t. We denote the probability that a policyholder currently aged $x = t - y$ will survive to age $x + h$ by

$${}_hp_x \equiv \exp\left\{-\sum_{j=1}^{h} \mu_{x+j}\right\}. \tag{6.3.1}$$

Here μ_{x+j} represents the force of mortality between age $x + j - 1$ and age $x + j$.[75] The force of mortality μ_{x+j} is assumed to not change over time.

[75] The force of mortality μ_{x+j} can be written as follows: $\mu_{x+j} = -\log\{1 - d_{x+j}\}$ where d_{x+j} stands for the one-year death probability between age $x + j - 1$ and age $x + j$.

The annuity units at time $t+h$ ($h \in \mathbb{N}$) of a policyholder born at time y, i.e., $\mathbf{B}_{y,t+h}$, are specified in terms of past and future stock market shocks as follows:

$$\begin{aligned}
\mathbf{B}_{y,t+h} &= B_{y,t_0} \exp\left\{\pi_{t_0+1} + \ldots + \pi_{t+h}\right\} \\
&\quad \exp\left\{q_{t+h-t_0}\omega \int_{t_0}^{t_0+1} \mathrm{d}W_s^* + \ldots + q_1\omega \int_{t+h-1}^{t+h} \mathrm{d}\mathbf{W}_s^*\right\} \\
&= B_{y,t_0} \prod_{j=t_0+1}^{t} \exp\left\{\pi_j + q_{j+h-t_0}\omega \int_{t+t_0-j}^{t+t_0-(j-1)} \mathrm{d}W_s^*\right\} \\
&\quad \prod_{j=1}^{h} \exp\left\{\pi_{t+j} + q_j\omega \int_{t+h-j}^{t+h-(j-1)} \mathrm{d}\mathbf{W}_s^*\right\}.
\end{aligned} \tag{6.3.2}$$

Here $\mathrm{d}\mathbf{W}_s^* \equiv \mathrm{d}\mathbf{W}_s + (\lambda - \lambda^*)\,\mathrm{d}t$ where λ^* denotes the assumed Sharpe ratio. This assumed Sharpe ratio may differ from the actual Sharpe ratio λ.[76] The parameter q_h represents the exposure of future log annuity units $\log \mathbf{B}_{y,t+h}$ to current stock market shocks $\omega \int_t^{t+1} \mathrm{d}\mathbf{W}_s^*$. We define q_h at discrete points in time. Indeed, h is an integer because adjustments of pension entitlements occur only once per period. We require that the marginal risk exposure $q_h \Rightarrow 1$ as $h \Rightarrow \infty$. Hence the parameter ω can be viewed as the exposure of long-term annuity units $\log \mathbf{B}_{y,\infty}$ to current stock market shocks. Finally, the parameter π_{t+1} represents the desired (or targeted) growth rate of annuity units between time t and time $t+1$. The desired growth rate of annuity units coincides with the median growth rate of annuity units (conditional upon information available at time t_0) if the assumed Sharpe ratio λ^* equals the actual Sharpe ratio λ (see also Section 6.3.2).

6.3.2. Horizon Differentiation

Specification (6.3.2) allows the marginal risk exposure q_h to depend on the investment horizon h.[77] Indeed, if q_h *strictly* increases with the investment horizon h, then long investment horizons exhibit a larger exposure to current stock market shocks than shorter investment horizons. We impose that the marginal risk exposure q_h is non-decreasing with the investment horizon h, i.e., $q_h \geq q_{h-1}$ for all $h > 1$. Specification (6.3.2) allows

[76] Merton (1980) shows that estimates of expected returns are less accurate than estimates of variances. Therefore, we distinguish only between the assumed expected return and the actual expected return. In particular, we assume that the actual variance matches the assumed variance.

[77] The present chapter assumes that the marginal risk exposure does not depend on age nor time. Specification (6.3.2) can however be extended by allowing the marginal risk exposure to depend on age and time.

the marginal risk exposure to depend only on the investment horizon h and thus only indirectly on age $x = t - y$. As a direct consequence, life cycle investment in which the stock market exposure declines with age continues during the decumulation phase. Internal habit formation can explain this type of horizon differentiation in marginal risk exposures (see, e.g., Fuhrer, 2000); see also Chapter 4.[78]

Figure 6.1 shows horizon differentiation in marginal risk exposures.[79] The dash-dotted line displays the case where $q_h = h/N$ for $h < N$ and $q_h = 1$ for $h \geq N$. In that case, short investment horizons $h < N = 10$ exhibit a smaller risk exposure than longer investment horizons $h \geq N = 10$. In particular, the marginal risk exposure at a one-year investment horizon ($h = 1$) is only one-tenth of the marginal risk exposure at a ten-year investment horizon ($h = 10$). We can view the parameter N as the smoothing period. If N increases, then it takes longer to fully absorb current stock market shocks into annuity units. The solid line corresponds to the case of exponential smoothing. That is, $q_h = 1 - \exp\{-\eta h\}$. The parameter η can be regarded as a smoothing parameter. If $\eta \Rightarrow 0$, horizon differentiation in marginal risk exposures is maximal, whereas horizon differentiation in marginal risk exposures is absent if $\eta \Rightarrow \infty$. The dotted line illustrates the case where horizon differentiation in marginal risk exposures is absent (i.e., $q_h = 1$ for all h). Finally, the dashed line displays a linear combination of the last two rules, i.e.,

$$q_h = (1 - q_1) \times 1 + q_1 \times (1 - \exp\{-\eta h\}) = 1 - (1 - q_1)\exp\{-\eta h\} \tag{6.3.3}$$

with $q_1 \geq 0$. The rule (6.3.3) characterizes horizon differentiation in terms of two parameters: $q_1 \leq 1$ (i.e., the part of current stock market shocks that is absorbed into the level of annuity units) and η (i.e., the speed at which the remaining part of current stock market shocks is absorbed into future growth rates of annuity units).

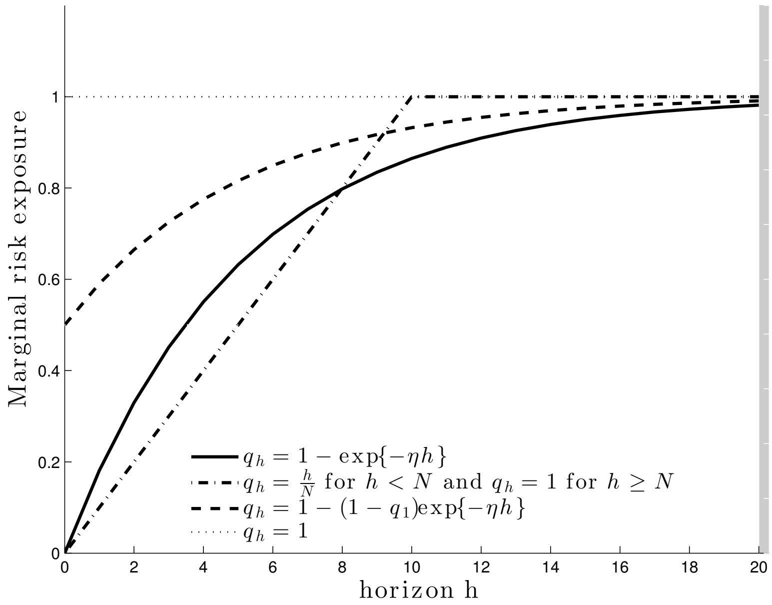
Dividing $\mathbf{B}_{y,t+h}$ by $B_{y,t}$ yields

$$\frac{\mathbf{B}_{y,t+h}}{B_{y,t}} = F_t^h \times \prod_{j=1}^{h} \exp\left\{\pi_{t+j} + q_j \omega \int_{t+h-j}^{t+h-(j-1)} \mathrm{d}\mathbf{W}_s^*\right\}, \tag{6.3.4}$$

[78] In the case where risk differentiation is based on human capital risk rather than internal habit formation, marginal risk exposures depend on not only the investment horizon h but also age $x = t - y$ (see Bodie et al., 1992).

[79] Although adjustments in pension entitlements occur only once per period, the figures in this chapter show continuous (rather than discrete) graphs.

Figure 6.1.
Illustration of horizon differentiation in marginal risk exposures



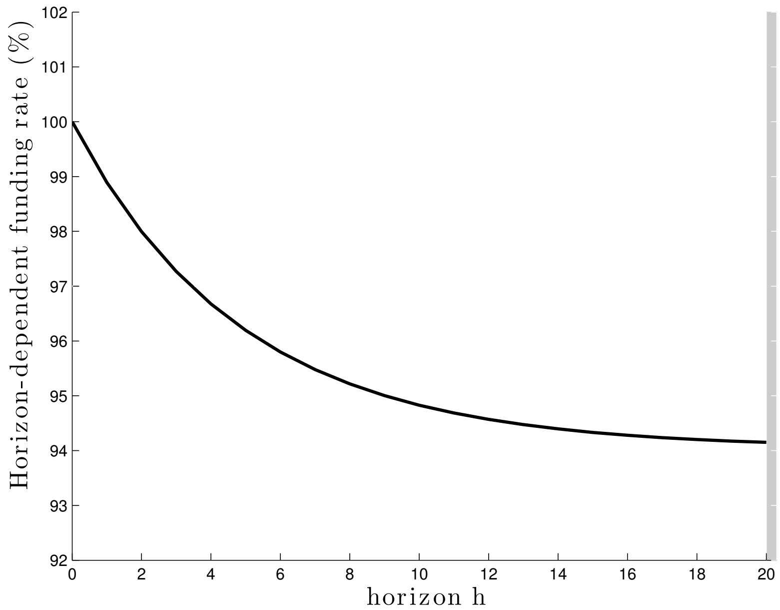
The figure illustrates horizon differentiation in marginal risk exposures. The dashed-dotted line displays the case where $q_h = h/N$ for $h < N$ and $q_h = 1$ for $h \geq N$ (with $N = 10$). The solid line corresponds to the case where $q_h = 1 - \exp\{-\eta h\}$ (with $\eta = 0.2$). The dotted line illustrates the case where $q_h = 1$ for all h. The dashed line represents the case where $q_h = 1 - (1 - q_1)\exp\{-\eta h\}$ (with $q_1 = 0.5$ and $\eta = 0.2$).

where

$$
\begin{aligned}
F_t^h &\equiv \frac{\prod_{j=t_0+1}^{t} \exp\left\{\pi_j + q_{j+h-t_0}\omega \int_{t+t_0-j}^{t+t_0-(j-1)} \mathrm{d}W_s^*\right\}}{\prod_{j=t_0+1}^{t} \exp\left\{\pi_j + q_{j-t_0}\omega \int_{t+t_0-j}^{t+t_0-(j-1)} \mathrm{d}W_s^*\right\}} \\
&= \prod_{j=t_0+1}^{t} \exp\left\{\left(q_{j+h-t_0} - q_{j-t_0}\right)\omega \int_{t+t_0-j}^{t+t_0-(j-1)} \mathrm{d}W_s^*\right\}
\end{aligned}
\tag{6.3.5}
$$

captures how past stock market shocks affect future annuity units. In the case of no horizon differentiation in marginal risk exposures, F_t^h is equal to unity (substitute $q_h = 1$ for all h in (6.3.5)). Indeed, in the absence of horizon differentiation in marginal risk exposures, stock market shocks are absorbed immediately into current annuity units. The horizon-dependent funding ratio (6.3.5) is thus the direct consequence of the gradual adjustment of annuity units to stock market shocks. Intuitively, sluggish adjustment of annuity units to stock market shocks gives rise to funding imbalances that must be absorbed in the future. As a direct consequence, future adjustments of pension entitlements become predictable. The horizon-dependent funding ratio F_t^h summarizes the predictable changes of future annuity units $\mathbf{B}_{y,t+h}$ as a result of past stock market shocks that have not yet been fully absorbed into current annuity units. Figure 6.2 illustrates the horizon-dependent funding ratio F_t^h as a function of the investment horizon h, with a single unexpected stock-market shock that occurred one year ago.

Figure 6.2.
Illustration of the horizon-dependent funding ratio



The figure illustrates the horizon-dependent funding rate F_t^h as a function of the investment horizon h, with a single unexpected stock-market shock that occurred one year ago. That is, $\log\{S_t/S_{t-1}\} - \left(r + \lambda\sigma - \frac{1}{2}\sigma^2\right) = -0.3$ while $\log\left\{S_{t-j}/S_{t-j-1}\right\} - \left(r + \lambda\sigma - \frac{1}{2}\sigma^2\right) = 0$ for all $j \geq 1$. The figure is based on $\omega = 0.5$ and $q_h = 1 - (1 - q_1)\exp\{-\eta h\}$ (with $q_1 = 0.5$ and $\eta = 0.2$).

The median value of future annuity units is given by (see Appendix)

$$\frac{\mathbb{M}_t\left[\mathbf{B}_{y,t+h}\right]}{B_{y,t}} = F_t^h \times \exp\left\{\sum_{j=1}^{h} \pi_{t+j}\right\} \times \exp\left\{(\lambda - \lambda^*)\,\omega \sum_{j=1}^{h} q_j\right\}, \tag{6.3.6}$$

where $\mathbb{M}_t[\cdot]$ stands for the median operator conditional upon all information available at time t. Equation (6.3.6) shows that the median value of future annuity units differs from the desired benefit (or pension ambition) $B_{y,t}\exp\left\{\sum_{j=1}^{h} \pi_{t+j}\right\}$ due to two factors:

one factor capturing the past and the other factor corresponding to future stock market returns. In particular, the first factor F_t^h represents past stock market shocks that have not yet been fully absorbed into annuity units. The second factor $(\lambda - \lambda^*)\,\omega \sum_{j=1}^{h} q_j$ is due to the gap between the *actual* Sharpe ratio λ and the *assumed* Sharpe ratio λ^*. The desired growth rate of median pension entitlements can thus be increased by not only increasing π_h but also reducing the assumed Sharpe ratio λ^*.

6.3.3. Bonus Policy

This section explores how annuity units (and thus pensions in payment) change as time proceeds. At the begin of each time period, pension entitlements are adjusted according to (see Appendix)

$$\frac{B_{y,t+1}}{B_{y,t}} = \exp\left\{\pi_{t+1}\right\} \times \exp\left\{q_1\omega \int_t^{t+1} \mathrm{d}W_s^*\right\} \times F_t^1. \tag{6.3.7}$$

Equation (6.3.7) can be viewed as the bonus (or dividend) policy of the pension scheme, showing how annuity units develop as time proceeds. The first term at the right-hand side of equation (6.3.7) represents the desired growth rate. The second term denotes the impact of current stock market shocks on current annuity units while the last term reflects the impact of past stock market shocks on current annuity units.

The bonus policy (6.3.7) can be rewritten as follows:

$$\frac{B_{y,t+1}}{B_{y,t}} = \exp\left\{\pi_{t+1}\right\} \times \bar{F}_{t+1}^0. \tag{6.3.8}$$

Here $\bar{F}_{t+1}^{h-1}$ denotes the horizon-dependent funding ratio *before* annuity units are adjusted but *after* stock market shocks between time t and time $t+1$, i.e., $\int_t^{t+1} \mathrm{d}W_s^*$, have been realized:

$$\bar{F}_{t+1}^{h-1} \equiv F_t^h \times \exp\left\{q_h\omega \int_t^{t+1} \mathrm{d}W_s^*\right\}. \tag{6.3.9}$$

The horizon-dependent funding ratio *after* annuity units are adjusted is given by (see Appendix)

$$F_{t+1}^{h-1} = F_t^h \times \exp\left\{q_h\omega \int_t^{t+1} \mathrm{d}W_s^*\right\} \times \frac{1}{\bar{F}_{t+1}^0}. \tag{6.3.10}$$

The second term at the right-hand side of (6.3.10) denotes current stock market shocks that result in the new funding ratio $\bar{F}_{t+1}^{h-1}$. The last term represents past stock market shocks that are gradually being absorbed into annuity units so that they are no longer included in the funding ratio.

6.3.4. Calibrating the Risk of Future Annuity Units

Future annuity units exhibit a lognormal distribution. The pth quantile of future annuity units $\mathbf{B}_{y,t+h}$ conditional upon all information available at time t is given by $(0 \leq p \leq 1)$

$$\mathbb{Q}_t^p \left[\mathbf{B}_{y,t+h}\right] = \mathbb{M}_t \left[\mathbf{B}_{y,t+h}\right] \exp\left\{\Phi^{-1}(p)\omega\sqrt{\sum_{j=1}^{h} q_j^2}\right\}. \tag{6.3.11}$$

Here $\Phi(\cdot)$ is the cumulative distribution function of a standard normal random variable. We can calibrate the marginal risk exposure q_h and the long-term risk exposure ω from the desired risk of the median value of future annuity units. For example, the insurer can set the desired difference between the log median value of future annuity units and the log 2.5% quantile of future annuity units for each horizon h:

$$\log \mathbb{Q}_t^{2.5\%} \left[\mathbf{B}_{y,t+h}\right] - \log \mathbb{M}_t \left[\mathbf{B}_{y,t+h}\right] = \Phi^{-1}\left(2.5\%\right)\omega\sqrt{\sum_{j=1}^{h} q_j^2}. \tag{6.3.12}$$

The parameters q_h and ω can then be endogenously derived from (6.3.12).[80]

6.3.5. Communicating the Risk of Future Annuity Units

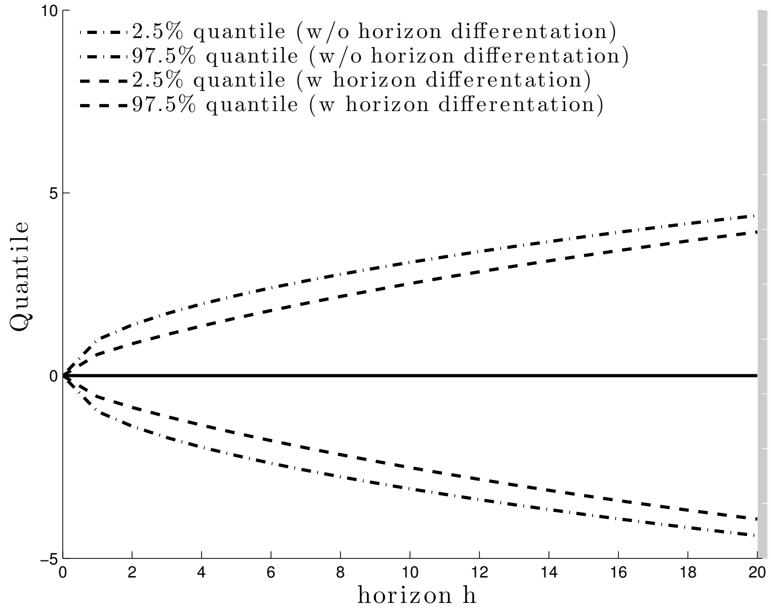
Given current annuity units $B_{y,t}$, insurers should communicate ex ante both the median value of future annuity units (6.3.6) and the 2.5% quantile of future annuity units (6.3.11), and ex post the difference between the realized pension outcome and the pension ambition. The median $\mathbb{M}_t\left[\mathbf{B}_{y,t+h}\right]$ and the 2.5% quantile $\mathbb{Q}_t^{2.5\%}\left[\mathbf{B}_{y,t+h}\right]$ depend on the stochastic model used by the insurer. Supervisory authorities may have to prescribe parameter values to prevent insurers from providing excessively optimistic projections.

Figure 6.3 illustrates the 2.5% and 97.5% quantiles of future annuity units (the quantiles are expressed relative to the median value of future annuity units). The

[80]Certain restrictions need to be imposed on (6.3.12) to ensure that q_h increases with the investment horizon h.

dashed lines represent the case of horizon differentiation in marginal risk exposures (i.e., $q_h = 1 - (1 - q_1)\exp\{-\eta h\}$ with $q_1 = 0.5$ and $\eta = 0.2$), while the dash-dotted lines correspond to the case of no horizon differentiation in marginal risk exposures (i.e., $q_h = 1$ for all h).

Figure 6.3.
Illustration of the 2.5% and 97.5% quantiles of future annuity units



The figure illustrates the 2.5% and 97.5% (log) quantiles of future annuity units (the quantiles are expressed relative to the median value of future annuity units). The solid lines represent the case of horizon differentiation in marginal risk exposures (i.e., $q_h = 1 - (1 - q_1)\exp\{-\eta h\}$ with $q_1 = 0.5$ and $\eta = 0.2$), while the dash-dotted lines correspond to the case of no horizon differentiation in marginal risk exposures (i.e., $q_h = 1$ for all h). The figure is based on $\omega = 0.5$.

6.4. Market-Consistent Valuation

6.4.1. Useful Decomposition

Denote by $V_{y,t}^h$ the market-consistent value at time t of annuity units $\mathbf{B}_{y,t+h}$. We can compute $V_{y,t}^h$ by solving the following conditional expectation (see, e.g., Cochrane, 2001):

$$V^h_{y,t} = {}_hp_{t-y}\mathbb{E}_t\left[\frac{\mathbf{m}_{t+h}}{m_t}\mathbf{B}_{y,t+h}\right], \tag{6.4.1}$$

where m_t stands for the pricing kernel (or stochastic discount factor) at time t. The Appendix provides an exact analytical expression for m_t and shows that

$$C^h_{y,t} \equiv V^h_{y,t}/B_{y,t} = F^h_t A^h_{y,t}. \tag{6.4.2}$$

Here

$$A^h_{y,t} \equiv \exp\left\{-\sum_{j=1}^{h}\delta^j_{y,t}\right\}. \tag{6.4.3}$$

Equation (6.4.2) shows that the market price of future annuity units in terms of current annuity units, i.e., $C^h_{y,t}$, consists of two factors. The horizon-dependent funding ratio F^h_t corresponds to past stock market shocks that have not yet been fully absorbed into current annuity units. This factor equals unity if current stock market shocks are fully absorbed into current annuity units (i.e., $q_h = 1$ for all h). The horizon-dependent annuity factor $A^h_{y,t}$ summarizes the impacts of desired growth and risk of annuity units, and future assumed expected rates of return on the market price of future annuity units. The next section gives an exact analytical expression for the forward (market-consistent) discount rate $\delta^j_{y,t}$.

6.4.2. The Forward Discount Rate

The forward discount rate $\delta^j_{y,t}$ is given by (see Appendix)

$$\begin{aligned}\delta^j_{y,t} &= \mu_{t-y+j} + r - \pi_{t+j} + q_j\omega\lambda^* + \xi_j \\ &= \mu_{t-y+j} + r - \pi_{t+j} + q_j\omega\left(\lambda^* - \frac{1}{2}q_j\omega\right)\end{aligned} \tag{6.4.4}$$

where $\xi_j \equiv -\frac{1}{2}q_j^2\omega^2$ denotes a second-order term. Equation (6.4.4) collapses to the survival premium μ_{t-y+j} plus the (forward) interest rate r if annuity units are fixed and guaranteed (i.e., $\pi_{t+j} = \omega = 0$). More generally, the discount rate (6.4.4) is a decreasing function of the desired growth rate π_{t+j}, and an increasing function of future biometric and (assumed) expected financial rates of return provided that $\lambda^* > \frac{1}{2} q_j \omega$. The survival credit μ_{t-y+j} represents the future biometric rate of return. The risk premium $q_j \omega \lambda^*$ is due to the impact of stock market shocks on annuity units: riskier annuity units yield higher expected returns, thereby raising discount rates and reducing the costs of the consumption stream. It depends on the exposure of annuity units to stock market shocks (determined by both the marginal risk exposure q_j and the long-term risk exposure ω) and the assumed Sharpe ratio λ^*.

6.4.3. The Discount Curve

The average discount rate $\bar{\delta}^h_{y,t}$ is given by

$$\bar{\delta}^h_{y,t} \equiv \frac{1}{h} \sum_{j=1}^{h} \delta^j_{y,t} = \bar{\mu}^h_{t-y} + r - \bar{\pi}^h_t + \bar{q}_h \omega \lambda^* + \bar{\xi}_h. \tag{6.4.5}$$

Here $0 \leq \bar{q}_h \leq 1$ represents the average risk exposure (or term structure of risk):

$$\bar{q}_h \equiv \frac{1}{h} \sum_{j=1}^{h} q_j. \tag{6.4.6}$$

The quantities $\bar{\mu}^h_{t-y}$, $\bar{\pi}^h_t$ and $\bar{\xi}_h$ are defined as follows:

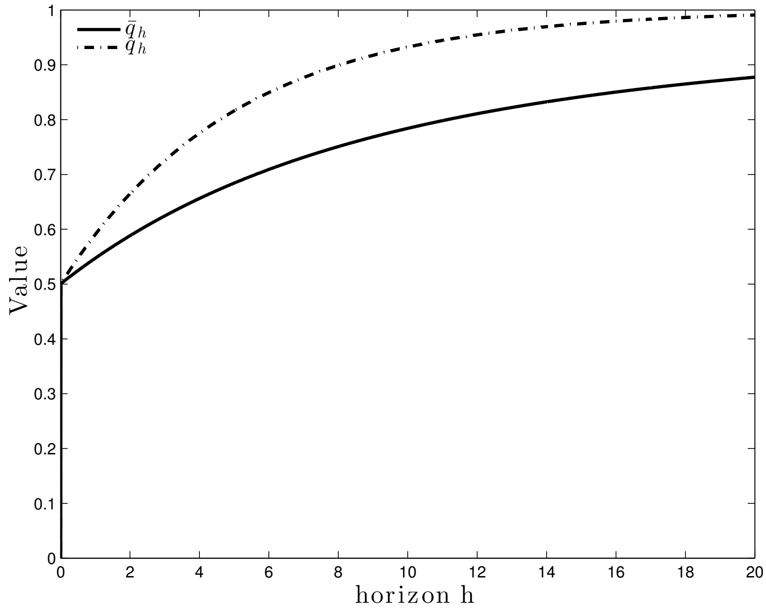
$$\bar{\mu}^h_{t-y} \equiv \frac{1}{h} \sum_{j=1}^{h} \mu_{t-y+j}, \tag{6.4.7}$$

$$\bar{\pi}^h_t \equiv \frac{1}{h} \sum_{j=1}^{h} \pi_{t+j}, \tag{6.4.8}$$

$$\bar{\xi}_h \equiv \frac{1}{h} \sum_{j=1}^{h} \xi_j. \tag{6.4.9}$$

The average and marginal risk exposure, i.e., $\bar{q}_h$ and q_h, relate to each other in an analogous way as the YTM relates to the forward interest rate. The non-decreasing nature of the marginal risk exposure q_h (i.e., $q_h \geq q_{h-1}$ for all $h > 1$) implies that the average risk exposure $\bar{q}_h$ is non-decreasing as well (i.e., $\bar{q}_h \geq \bar{q}_{h-1}$ for $h > 1$). Also, it follows that the average risk exposure does not exceed the marginal risk exposure (i.e., $\bar{q}_h \leq q_h$ for all $h \geq 1$). Figure 6.4 illustrates both the average risk exposure $\bar{q}_h$ and the marginal risk exposure q_h with horizon differentiation in marginal risk exposures.

Figure 6.4.
Illustration of the average and marginal risk exposure



The figure illustrates the average risk exposure $\bar{q}_h$ as well as the marginal risk exposure q_h with horizon differentiation in marginal risk exposures. The marginal risk exposure is given by $q_h = 1 - (1 - q_1) \exp\{-\eta h\}$ (with $q_1 = 0.5$ and $\eta = 0.2$).

In the case of gradual adjustment of annuity units to stock market shocks, the risk premium $\bar{q}_h \omega \lambda^*$ increases with the investment horizon h. Figure 6.5 shows the risk premium $\bar{q}_h \omega \lambda^*$ with horizon differentiation in marginal risk exposures.

6.4.4. Discounting the Median Value of Future Annuity Units

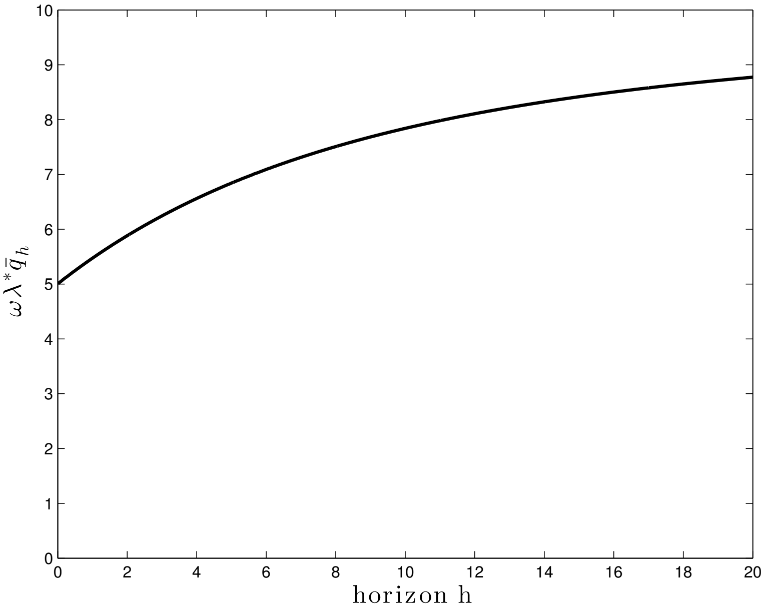
We can also discount the median value of future annuity units to find the market-consistent value $V_{y,t}^h$. We find

$$V_{y,t}^h = \mathbb{M}_t \left[\mathbf{B}_{y,t+h} \right] \exp \left\{ - \sum_{j=1}^{h} \widetilde{\delta}_{y,t}^j \right\}, \tag{6.4.10}$$

where the forward discount rate $\widetilde{\delta}_{y,t}^j$ is given by (this follows from equations (6.3.6) and (6.4.2))

$$\widetilde{\delta}_{y,t}^j = \delta_{y,t}^j + \pi_{t+j} + q_j \omega \left(\lambda - \lambda^* \right). \tag{6.4.11}$$

Figure 6.5.
Illustration of the horizon-dependent risk premium



The figure illustrates the horizon-dependent risk premium $\bar{q}_h\omega\lambda^*$ with horizon differentiation in marginal risk exposures. The marginal risk exposure is given by $q_h = 1-(1-q_1)\exp\{-\eta h\}$ (with $q_1 = 0.5$ and $\eta = 0.2$). The figure is based $\omega = 0.5$ and $\lambda^* = 0.2$.

Whereas the *actual* risk premium λ features in discounting the median value of future annuity units (see equation (6.4.11)), only the *assumed* risk premium λ^* features in discounting current annuity units $B_{y,t}$ (see equation (6.4.4)). This implies that the market-consistent value of annuity units does not depend on the *subjective* parameter λ of the stochastic model but depends only on the parameters of the contract π_{t+j}, λ^* and q_j.[81] Whereas model risk (i.e., the value of λ) does not affect the market-consistent value of future annuity units, it does impact the median value of future annuity units (6.3.6) and the quantiles of future annuity units (6.3.11).

[81]Here we implicitly assume that the actual variance coincides with the assumed variance.

6.4.5. Comparison with Traditional Annuities

6.4.5.1. Nominal Fixed Annuities

The market-consistent value at time t of a guaranteed nominal annuity payment (i.e., $\pi_{t+j} = 0$ for every j and $\omega = 0$) at time $t + h$ is given by

$$V_{y,t}^{h} = B_{y,t} \exp\left\{ -\sum_{j=1}^{h} \left(\mu_{t-y+j} + r \right) \right\}. \tag{6.4.12}$$

We identify three differences between the market price of variable annuities (see equation (6.4.2)) and the market price of traditional DB pension plans (see equation (6.4.12)). Whereas traditional DB pension plans keep current annuity units $B_{y,t}$ constant in nominal terms, variable annuities vary with stock market returns. First, a horizon-dependent risk premium $\bar{q}_h \omega \lambda^* + \bar{\xi}_h$, which typically rises with the investment horizon, is added in (6.4.5) to account for the conditional, risky nature of future annuity units. Second, desired indexation $\bar{\pi}_t^h$ is included in (6.4.5) to measure the costs of the desired bonus payments. Third, a factor F_t^h is included in (6.4.2) representing past stock market shocks that have not yet been fully absorbed into current annuity units.

6.4.5.2. Traditional Variable Annuities

In the case of a traditional variable annuity in which shocks are absorbed immediately into current annuity units (i.e., $q_h = \bar{q}_h = 1$ for all $h \geq 1$) and the median annuity payments are constant in nominal terms (i.e., $\pi_{t+j} = 0$ for every j), equation (6.4.2) boils down to

$$V_{y,t}^{h} = B_{y,t} \exp\left\{ -\left(\bar{\mu}_{t-y}^{h} + r + \omega \lambda^* - \frac{1}{2} \omega^2 \right) h \right\}. \tag{6.4.13}$$

Comparing equation (6.4.13) with equation (6.4.2), we observe that the term representing past stock market shocks F_t^h is not present in this case because stock market shocks are absorbed immediately into current annuity units $B_{y,t}$. Hence, predictable future changes in pension entitlements resulting from past stock market shocks are absent. Moreover, the risk premium in the discount rate does not depend on the investment horizon because there is no horizon differentiation in marginal risk exposures.

6.4.6. Conversion Factor versus Annuity Factor

Assuming that newly bought pension entitlements share in current funding gaps, we can calculate the market price at time t of an annuity unit for a policyholder born at time y as follows:

$$C_{y,t} \equiv \frac{V_{y,t}}{B_{y,t}} = \sum_{h=\max\{x_r+y-t,1\}}^{x_{\max}+y-t} F_t^h A_{y,t}^h. \tag{6.4.14}$$

We employ the formula (6.4.14) to calculate the price of newly bought entitlements. This is in line with the DB tradition in which the pension premium is determined by the costs of the desired consumption stream in retirement. The conversion factor $C_{y,t}$ represents the economically fair price that a policyholder with birth year y should pay for each annuity unit to ensure that the newly bought annuity units do not affect the value of existing pension entitlements.

In the case of gradual adjustment of annuity units to stock market shocks, the conversion factor differs from the aggregate annuity factor

$$A_{y,t} = \sum_{h=\max\{x_r+y-t,1\}}^{x_{\max}+y-t} A_{y,t}^h, \tag{6.4.15}$$

because newly bought annuity units share in current funding gaps. As a consequence, the conversion factor is smaller than the annuity factor in the presence of underfunding (i.e., $F_t^h < 1$). Intuitively, underfunding reduces the growth of future consumption streams, thereby reducing the costs of future annuity units. The conversion factor $C_{y,t}$ would coincide with the aggregate annuity factor $A_{y,t}$ if newly bought annuity units do not share in past stock market shocks.

The valuation of annuity units (6.4.2) assumes that policyholders pay a fair economic price for new annuity units so that the value of existing annuity units is not affected by the purchase of new annuity units. Indeed, valuation of existing annuity units (6.4.2) relies on the premium rule (6.4.14) for newly bought annuity units. With this premium rule, the aggregate value of the variable annuities matches the current value of aggregate assets X_t:

$$X_t = V_t. \tag{6.4.16}$$

The pension contract thus exhibits a defined contribution character in the sense that stock market shocks are absorbed into pension entitlements through adjustments in current annuity units (i.e., the so-called bonus payments) and future predictable changes F_t^h. Outside sponsors (such as companies, future policyholders, insurance companies, tax payers, shareholders) are absent: risks are borne by the current policyholders. The pension promises are backed by financial assets so that the system is always fully funded on a so-called discontinuity basis.[82] Indeed, the funding ratio is unity if we measure liabilities in terms of the market value of promised cash flows V_t.

If newly bought annuity units share in current funding gaps but pricing is based on the annuity factor (6.4.15), the pension contract involves risk sharing between current policyholders and new policyholders. In particular, new policyholders subsidize (tax) current policyholders in case of underfunding, i.e., $F_t^h < 1$ (overfunding, i.e., $F_t^h > 1$).

6.4.7. Replicating Portfolio Strategy

This section derives the portfolio strategy that replicates the pension contract (6.3.2). We allow the insurer to invest in a risky stock and a nominal money market account. The portfolio strategy is determined in such a way that the value of the assets matches the value of the liabilities in each state of the world. We thus apply the principle of liability-driven investment familiar from DB pension plans to arrive at the replicating portfolio strategy.

Replication of the pension contract requires a fraction $\widehat{q}_{y,t}$ of the assets to be invested in the risky stock (see Appendix)

$$\widehat{q}_{y,t} = \frac{\omega}{\sigma} \sum_{h=\max\{x_r+y-t,1\}}^{x_{\max}+y-t} \alpha_{y,t}^h q_h, \tag{6.4.17}$$

where $\alpha_{y,t}^h$ is defined as follows:

$$\alpha_{y,t}^h \equiv V_{y,t}^h / V_{y,t} \tag{6.4.18}$$

Equation (6.4.17) shows that in the case of gradual adjustment of annuity units to stock market shocks, the portfolio weight decreases as the policyholder ages.

[82] This means that the policyholders receive their promised benefits even if the insurer is wound up. Hence, the policyholders are not exposed to the credit risk of the insurer.

6.5. Mismatch Risk and an Incorrect Discount Rate

6.5.1. Mismatch Risk

The expression for the bonus rate (6.3.7) can be written in terms of mismatch. In particular, we can define mismatch risk M_{t+1} as follows:

$$M_{t+1} \equiv \sigma \widehat{q}_{y,t} \int_t^{t+1} \mathrm{dW}_s. \tag{6.5.1}$$

The bonus rate expression (6.3.7) can now be written as

$$\frac{B_{y,t+1}}{B_{y,t}} = \exp\left\{\pi_{t+1}\right\} \times \exp\left\{\frac{q_1\omega}{\widehat{q}_{y,t}\sigma} M_{t+1}\right\} \times F_t^1. \tag{6.5.2}$$

The bonus rate $\log\left\{B_{y,t+1}/B_{y,t}\right\}$ is thus equal to the ambition π_{t+1} adjusted for past shocks $\log\left\{F_t^1\right\}$ and current mismatch risk $q_1\omega/\left(\widehat{q}_{y,t}\sigma\right) M_{t+1}$. We can rewrite the funding ratio (6.3.5) in terms of mismatch in past (substitute (6.5.1) into (6.3.5)):

$$F_t^h = \exp\left\{\sum_{j=t_0+1}^{t} \omega \frac{q_{h+j-t_0} - q_{j-t_0}}{\widehat{q}_{y,t+t_0-j}\sigma} M_{t+t_0-j+1}\right\}. \tag{6.5.3}$$

We can also write (6.3.4) in terms of mismatch in the past:

$$\frac{\mathbf{B}_{y,t+h}}{B_{y,t}} = F_t^h \times \exp\left\{\sum_{j=1}^{h} \pi_{t+j}\right\} \times \exp\left\{\sum_{j=1}^{h} \frac{q_j\omega}{\widehat{q}_{y,t+t_0+h-j}\sigma} \mathbf{M}_{t+t_0+h-j+1}\right\}. \tag{6.5.4}$$

6.5.2. Discounting with Expected Returns

A popular way to compute the value of annuity units is to employ the expected return on the investment portfolio $\widehat{q}_{y,t_0}$ as the discount rate:

$$\widehat{V}_{y,t_0} = \sum_{h=\max\{x_r+y-t_0,1\}}^{x_{\max}+y-t_0} \widehat{V}_{y,t_0}^h, \tag{6.5.5}$$

where

$$\widehat{V}_{y,t_0}^h \equiv B_{y,t_0} \exp\left\{-\left(\bar{\mu}_{t-y}^h + r - \bar{\pi}_t^h + \lambda^* \sigma \widehat{q}_{y,t_0} + \bar{\xi}_h\right) h\right\}. \tag{6.5.6}$$

Linearization of $\widehat{V}^h_{y,t_0}$ around $\widehat{q}_{y,t_0} = \omega/\sigma\bar{q}_h$ yields

$$\frac{\widehat{V}^h_{y,t_0} - V^h_{y,t_0}}{V^h_{y,t_0}} \approx -\sigma\lambda^* h \left(\widehat{q}_{y,t_0} - \frac{\omega}{\sigma}\bar{q}_h\right), \tag{6.5.7}$$

so that

$$\frac{\widehat{V}_{y,t_0} - V_{y,t_0}}{V_{y,t_0}} \approx \sigma\lambda^* \widehat{D}_{y,t_0} \left(\frac{\omega}{\sigma}\theta_{y,t_0} - \widehat{q}_{y,t_0}\right), \tag{6.5.8}$$

where the duration $\widehat{D}_{y,t_0}$ is defined as follows:

$$\widehat{D}_{y,t_0} \equiv \sum_{h=\max\{x_r+y-t_0,1\}}^{x_{\max}+y-t_0} \alpha^h_{y,t_0} h \tag{6.5.9}$$

and

$$\theta_{y,t_0} \equiv \frac{1}{\widehat{D}_{y,t_0}} \times \sum_{h=\max\{x_r+y-t_0,1\}}^{x_{\max}+y-t_0} \alpha^h_{y,t_0} h\bar{q}_h = \sum_{h=\max\{x_r+y-t_0,1\}}^{x_{\max}+y-t_0} \beta^h_{y,t_0}\bar{q}_h \tag{6.5.10}$$

with $\beta^h_{y,t_0} \equiv \alpha^h_{y,t_0} \times \frac{h}{\widehat{D}_{y,t_0}}$.

For traditional variable annuities without gradual absorption of shocks (i.e., $q_h = \bar{q}_h = 1$), the traditional method of using expected returns on the current portfolio yields the correct result (since in equation (6.5.8), $\omega/\sigma\theta_{y,t_0} = \widehat{q}_{y,t_0} = 1$ if $q_h = \bar{q}_h = 1$ for all horizons $h \geq 1$). With horizon differentiation, in contrast, the method of using current expected returns tends to understate the value of actual liabilities since $\omega/\sigma\theta_{y,t_0} < \widehat{q}_{y,t_0}$.[83] Hence the overall discount curve for valuing variable annuities is typically in between the expected return on the actual investment portfolio and the risk-free term structure. Intuitively, current expected returns exceed future returns because with life cycle investment, risk is taken back when policyholders age. Also, using an incorrect discount rate leads to inefficient intertemporal consumption smoothing. Indeed, a higher discount rate raises the funding ratio (i.e., $V_{y,t_0}/\widehat{V}_{y,t_0} > 1$), thereby increasing the scope to pay out today (i.e., consumption is reallocated from the long-run to the short-run).

[83] This is always the case if liabilities are concentrated around a certain horizon because horizon differentiation (i.e., $q_h > q_1$) implies $q^h_h > \bar{q}_h$ and thus $\widehat{q}_{y,t_0} > \omega/\sigma\theta_{y,t_0}$ if $\beta^h_{y,t_0} \approx \alpha^h_{y,t_0}$. If liabilities are dispersed over various horizons and $\beta^h_{y,t_0} > \alpha^h_{y,t_0}$ for long horizons h, we may theoretically have $\widehat{q}_{y,t_0} < \theta_{y,t_0}$ because longer horizons with larger $\bar{q}^h$ and q^h receive a larger weight in the calculation of θ_{y,t_0} than in the calculation of $\widehat{q}_{y,t_0}$.

6.6. Exponential Decay and the Cash-Flow Funding Rate

6.6.1. Exponential Decay

Equation (6.3.5) implies that for each horizon, we need a separate state variable to summarize past stock market shocks. For a specific specification of the marginal risk exposure q_h, we can, however, summarize past stock market shocks in one state variable. Specifically, we assume that

$$q_h = 1 - \rho^h. \tag{6.6.1}$$

The coefficient ρ in equation (6.6.1) governs horizon differentiation in marginal risk exposures. With $\rho = 0$, horizon differentiation in marginal risk exposures is absent and $q_h = 1$. In that case, shocks are absorbed immediately so that $F_t^h = 1$ (see equation (6.3.5) with $q_h = 1$ for all h). With $\rho \uparrow 1$, horizon differentiation in marginal risk exposures is maximal and $q_h/q_1 \Rightarrow h$. Specification (6.6.1) thus implies that the risk of future annuity units is parameterized by the long-term risk exposure ω and the parameter $0 \leq \rho < 1$.

With (6.6.1), we can write the horizon-dependent funding ratios in terms of one state variable (use $q_{h+j-t_0} - q_{j-t_0} = q_h \left(1 - q_{j-t_0}\right) = q_h \rho^{j-t_0}$ in equation (6.3.5))

$$F_t^h = \left(\widehat{F}_t\right)^{q_h} = \left(\widehat{F}_t\right)^{1-\rho^h}. \tag{6.6.2}$$

Here

$$\widehat{F}_t \equiv \exp\left\{\omega \sum_{j=t_0+1}^{t} \rho^{j-t_0} \int_{t+t_0-j}^{t+t_0-(j-1)} \mathrm{d}W_s^*\right\}. \tag{6.6.3}$$

6.6.2. The Cash-Flow Funding Rate

We can write the single state variable $\widehat{F}_t$ in terms of a so-called cash-flow funding ratio. This funding ratio is computed on the basis of an alternative definition of the liabilities. With this alternative definition, the aggregate value of the liabilities is no longer necessarily equal to the value of the assets so that the so-called cash-flow funding ratio can deviate from unity. The calculation of a funding ratio unequal to one makes the pension system reminiscent of DB pension systems. The difference with traditional DB pension systems is that funding disequilibria are absorbed by the policyholders themselves rather than an outside sponsor such as a corporation.

The alternative definition of the liabilities is based on the ambition to increase current pension entitlements $B_{y,t}$ in line with desired indexation (in median and with the desired risk of future annuity units). In particular, the value of the liability at horizon h of an policyholder born at time y is given by

$$L^h_{y,t} \equiv \frac{B_{y,t}}{\exp\left\{\sum_{j=1}^{h} \delta^j_{y,t}\right\}} = \frac{V^h_{y,t}}{F^h_t}. \tag{6.6.4}$$

We can view F^h_t as the horizon-dependent funding ratio because it represents the ratio between the actual value of annuity payments ('assets') at horizon h and the value of the defined ambition ('liabilities') at horizon h, i.e., $F^h_t = V^h_{y,t}/L^h_{y,t}$.

The overall liability $L_{y,t}$ corresponding to a policyholder born at time y is given by

$$L_{y,t} \equiv B_{y,t} \sum_{h=\max\{x_r+y-t,1\}}^{x_{\max}+y-t} \frac{1}{\exp\left\{\sum_{j=1}^{h} \delta^j_{y,t}\right\}}. \tag{6.6.5}$$

These liabilities are the resources that are currently needed to consistently increase current pension entitlements $B_{y,t}$ in line with desired indexation (in median and with the desired risk of future annuity units parameterized by ω and ρ). This definition of liabilities thus abstracts from predictable changes in future annuity payments that are the result of gradual adjustment of annuity units to past stock market shocks.

The 'cash-flow' funding ratio $F_{y,t}$ is equal to the weighted average of horizon-specific funding ratios (the second equality follows from equation (6.6.4) to eliminate $V^h_{y,t}$)

$$F_{y,t} \equiv \frac{V_{y,t}}{L_{y,t}} = \sum_{h=\max\{x_r+y-t,1\}}^{x_{\max}+y-t} \frac{V^h_{y,t}}{L_{y,t}} = \sum_{h=\max\{x_r+y-t,1\}}^{x_{\max}+y-t} \gamma^h_{y,t} F^h_t = \sum_{h=\max\{x_r+y-t,1\}}^{x_{\max}+y-t} \gamma^h_{y,t} F^h_t, \tag{6.6.6}$$

where the aggregate value of liabilities $L_{y,t}$ is defined by

$$L_{y,t} \equiv \sum_{h=\max\{x_r+y-t,1\}}^{x_{\max}+y-t} L_{y,t}^h = \sum_{h=\max\{x_r+y-t,1\}}^{x_{\max}+y-t} \frac{B_{y,t}}{\exp\left\{\sum_{j=1}^{h} \delta_{y,t}^j\right\}}, \tag{6.6.7}$$

and

$$\gamma_{y,t}^h \equiv \frac{L_{y,t}^h}{L_{y,t}}. \tag{6.6.8}$$

The Appendix employs a linear approximation to write the state variable $\widehat{F}_t$ (see (6.6.3)) in terms of the cash-flow funding ratio $F_{y,t} \equiv \frac{V_{y,t}}{L_{y,t}}$ (see (6.6.6))

$$\left(\widehat{F}_t\right)^{\check{q}_{y,t}} \approx F_{y,t}, \tag{6.6.9}$$

where

$$\check{q}_{y,t} \equiv \sum_{h=\max\{x_r+y-t,1\}}^{x_{\max}+y-t} \gamma_{y,t}^h q_h.$$

Equation (6.6.9) represents the relationship between the economy-wide state variable $\widehat{F}_t$, which summarizes realized economy-wide risk in the past, and the cash-flow funding ratio $F_{y,t}$, which depends on age determining the exposure of the fund to this aggregate risk. The older the policyholder, the less he is exposed to past macro-economic shocks if horizon differentiation implies smoothing of adjustment to shocks. With horizon differentiation, q_h rises with h so that larger weights $\gamma_{y,t}^h$ of the shorter horizons reduces the exposure $\check{q}_{y,t}$ of the fund to macro-economic shocks $\widehat{F}_t$.

Substitution of equation (6.6.9) into equation (6.6.2) to eliminate $\widehat{F}_t$ yields

$$F_t^h \approx \left(F_{y,t}\right)^{\frac{q_h}{\check{q}_{y,t}}}. \tag{6.6.10}$$

We can write $V_{y,t}$ as (use equation (6.6.10) to eliminate F_t^h)

$$V_{y,t} = B_{y,t} \sum_{h=\max\{x_r+y-t,1\}}^{x_{\max}+y-t} \frac{\left(F_{y,t}\right)^{\frac{q_h}{\check{q}_{y,t}}}}{\exp\left\{\sum_{j=1}^{h} \delta_{y,t}^j\right\}}. \tag{6.6.11}$$

Equation (6.6.11) shows how the assets $V_{y,t}$ are distributed across the various horizons. The funding ratio $F_{y,t} = V_{y,t}/L_{y,t}$ is computed by using the observed actual assets $V_{y,t}$ and the liabilities $L_{y,t}$ from (6.6.5). We do not need information on how past shocks have occurred over time.

6.7. Concluding Remarks

This chapter has explored DA plans that provide variable annuities to policyholders. The pension fund exogenously specifies the entire stochastic income stream in retirement. In line with internal habit formation, pension payments respond gradually to financial shocks. The specification of the pension contract endogenously determines contribution levels and the investment policy. We have shown that the discount rate includes a risk premium that rises with the horizon and that the fraction of assets invested in the risky stock decreases as the policyholder ages. Also, the gradual absorption of financial shocks leads to predictable changes in pension payments. The effects of past stock market shocks on future adjustments in annuity units can be captured in one state variable (i.e., the funding ratio) if financial shocks are smoothed out in an exponentially declining manner.

Public supervision plays four important roles. The first two of these four roles involve the risk of future annuity units and the last two roles are associated with proper valuation and ensuring intergenerational fairness. First, the supervisory authorities should induce the funds to communicate the expected income streams and the risks involved (e.g., based on a ‘bad weather’ scenario) on the basis of standardized stochastic models. Second, they should monitor that the investment policy of the fund is consistent with the desired risk of future annuity units. Third, public supervision should ensure that the annuities are priced fairly, especially when participation is compulsory and competition does not discipline funds. Fourth, if funds change the pension contract (e.g., the way that annuity units are discounted), the authorities should check whether the exchange of annuities occurs at fair prices. By preventing intergenerational transfers, public supervision protects individual property rights.

6.8. Appendix

Proofs

Derivation of (6.3.6), (6.3.7) and (6.3.10)

We can write (6.3.4) as

$$\begin{aligned}\frac{\mathbf{B}_{y,t+h}}{B_{y,t}} &= F_t^h \exp\left\{\sum_{j=1}^{h} \pi_{t+j}\right\} \exp\left\{\omega\left(\lambda-\lambda^*\right)\sum_{j=1}^{h} q_j\right\} \\ &\exp\left\{\omega \sum_{j=1}^{h} q_j \int_{t+h-j}^{t+h-(j-1)} \mathrm{d}\mathbf{W}_s\right\}.\end{aligned} \tag{6.8.1}$$

Taking the median of (6.8.1) yields (6.3.6).

Equation (6.3.7) follows from (6.3.4) with $h=1$ where $\int_t^{t+1} \mathrm{d}W_s$ is now known:

$$\begin{aligned}\frac{\mathbf{B}_{y,t+1}}{B_{y,t}} &= F_t^1 \exp\left\{\pi_{t+1} + q_1\omega \int_t^{t+1} \mathrm{d}W_s^*\right\} \\ &= \frac{\prod_{j=t_0+1}^{t} \exp\left\{q_{j+1-t_0}\omega \int_{t+t_0-j}^{t+t_0-(j-1)} \mathrm{d}W_s^*\right\}}{\prod_{j=t_0+1}^{t} \exp\left\{q_{j-t_0}\omega \int_{t+t_0-j}^{t+t_0-(j-1)} \mathrm{d}W_s^*\right\}} \\ &\exp\left\{\pi_{t+1} + q_1\omega \int_t^{t+1} \mathrm{d}W_s^*\right\} \\ &= \exp\left\{\pi_{t+1}\right\} \exp\left\{\sum_{j=t_0}^{t} \left(q_{j+1-t_0} - q_{j-t_0}\right)\omega \int_{t+t_0-j}^{t+t_0-(j-1)} \mathrm{d}W_s^*\right\} \\ &= \exp\left\{\pi_{t+1}\right\} \times \bar{F}_{t+1}^{0},\end{aligned}$$

where $q_0 = 0$ by convention and $\bar{F}_{t+1}^{h-1}$ is the horizon-dependent funding ratio *before* annuity units are adjusted:

$$\begin{aligned}\bar{F}_{t+1}^{h-1} &\equiv F_t^h \times \exp\left\{q_h\omega \int_t^{t+1} \mathrm{d}W_s^*\right\} \\ &= \exp\left\{\sum_{j=t_0}^{t} \left(q_{j+1-t_0} - q_{j-t_0}\right)\omega \int_{t+t_0-j}^{t+t_0-(j-1)} \mathrm{d}W_s^*\right\}.\end{aligned} \tag{6.8.2}$$

The horizon-dependent funding ratio *after* annuity units are adjusted is given by (see equation (6.3.5))

$$F_{t+1}^{h-1} = \frac{F_t^h \times \exp\left\{q_h \omega \int_t^{t+1} \mathrm{d}W_s^*\right\}}{\bar{F}_{t+1}^0} = \frac{F_t^h \times \exp\left\{q_h \omega \int_t^{t+1} \mathrm{d}W_s^*\right\}}{F_t^1 \times \exp\left\{q_1 \omega \int_t^{t+1} \mathrm{d}W_s^*\right\}},$$

so that $F_{t+1}^0 = 1$.

Derivation of (6.4.2) and (6.4.17)

The pricing kernel m_t is subject to the following dynamic equation:

$$\frac{\mathrm{d}\mathbf{m}_t}{m_t} = -r\mathrm{d}t - \lambda\,\mathrm{d}\mathbf{W}_t.$$

Application of Ito's lemma yields

$$\mathrm{d}\log\mathbf{m}_t = -\left(r + \frac{1}{2}\lambda^2\right)\mathrm{d}t - \lambda\,\mathrm{d}\mathbf{W}_t.$$

Hence,

$$\frac{\mathbf{m}_{t+h}}{m_t} = \exp\left\{-\left(r + \frac{1}{2}\lambda^2\right)h - \lambda\int_t^{t+h}\mathrm{d}\mathbf{W}_s\right\} = \exp\left\{-\left(r + \frac{1}{2}\lambda^2\right)h - \lambda\sum_{j=1}^{h}\int_{t+h-j}^{t+h-(j-1)}\mathrm{d}\mathbf{W}_s\right\}. \tag{6.8.3}$$

Substituting equation (6.8.3) and equation (6.3.4) into equation (6.4.1), we arrive at

$$\begin{aligned}
V_{y,t}^{h} =& {}_h p_{t-y} F_t^h B_{y,t} \exp\left\{\sum_{j=1}^{h} \pi_{t+j}\right\} \\
& \exp\left\{\omega\left(\lambda-\lambda^*\right)\sum_{j=1}^{h} q_j\right\} \exp\left\{-\left(r+\frac{1}{2}\lambda^2\right)h\right\} \\
& \mathbb{E}_t\left[\exp\left\{\sum_{j=1}^{h}\left(q_j\omega-\lambda\right)\int_{t+h-j}^{t+h-(j-1)} \mathrm{d}\mathbf{W}_s\right\}\right] \\
=& {}_h p_{t-y} F_t^h B_{y,t} \exp\left\{\sum_{j=1}^{h} \pi_{t+j}\right\} \exp\left\{\omega\left(\lambda-\lambda^*\right)\sum_{j=1}^{h} q_j\right\} \\
& \exp\left\{-\left(r+\frac{1}{2}\lambda^2\right)h\right\} \exp\left\{\frac{1}{2}\sum_{j=1}^{h}\left(q_j\omega-\lambda\right)^2\right\} \\
=& {}_h p_{t-y} F_t^h B_{y,t} \exp\left\{-\sum_{j=1}^{h}\left(r-\pi_{t+j}+q_j\omega\lambda-\frac{1}{2}q_j^2\omega^2\right)\right\}.
\end{aligned} \tag{6.8.4}$$

We can rewrite $V_{y,t}^{h}$ as follows:

$$\begin{aligned}
V_{y,t}^{h} &= B_{y,t} F_t^h \exp\left\{-\left(\bar{\mu}_{t-y}^h + r - \bar{\pi}_t^h + \bar{q}_h\omega\lambda^* - \bar{\xi}_h\right)h\right\} \\
&= \prod_{j=t_0+1}^{t} \exp\left\{\pi_j + q_{j+h-t_0}\omega\int_{t+t_0-j}^{t+t_0-(j-1)} \mathrm{d}W_s^*\right\} \\
&\quad \exp\left\{-\left(\bar{\mu}_{t-y}^h + r - \bar{\pi}_t^h + \bar{q}_h\omega\lambda^* - \bar{\xi}_h\right)h\right\} B_{y,t_0}.
\end{aligned}$$

Hence,

$$\begin{aligned}
\frac{V_{y,t+1}^{h-1}}{V_{y,t}^{h}} &= \exp\left\{\pi_{t+1} + q_h\omega\int_t^{t+1} \mathrm{d}W_s^*\right\} \\
&\quad \exp\left\{\mu_{t-y+1} + r - \pi_{t+1} + q_h\omega\lambda^* - \xi_h\right\} \\
&= \exp\left\{\mu_{t-y+1} + r + q_h\omega\lambda^* - \xi_h + q_h\omega\int_t^{t+1} \mathrm{d}W_s^*\right\}.
\end{aligned} \tag{6.8.5}$$

Let $X_{y,t}^{h}$ be the value of the assets at time t that finances $\mathbf{B}_{y,t+h}$, and let $\widehat{q}_{y,t}^{h}$ be the corresponding investment policy (i.e., $\widehat{q}_{y,t}^{h}$ denotes the fraction of $X_{y,t}^{h}$ invested in the risky stock). We can write $X_{y,t}^{h}$ as follows:

$$X_{y,t}^{h} = X_{y,t_0}^{t-t_0+h} \exp\left\{\sum_{j=t_0}^{t-1} \mu_{j+1-y} + r\left(t-t_0\right) + \sum_{j=t_0}^{t-1} \widehat{q}_{y,j}^{t+h-j}\lambda\sigma \right.$$
$$\left. -\frac{1}{2}\sigma^2 \sum_{j=t_0}^{t-1}\left(\widehat{q}_{y,j}^{t+h-j}\right)^2 + \sum_{j=t_0}^{t-1}\widehat{q}_{y,j}^{t+h-j}\int_j^{j+1} \mathrm{d}W_s\right\},$$

where

$$X_{y,t_0}^{t-t_0+h} = V_{y,t_0}^{t-t_0+h}.$$

Hence,

$$\frac{X_{y,t+1}^{h-1}}{X_{y,t}^{h}} = \exp\left\{\mu_{t+1-y} + r + \widehat{q}_{y,t}^{h}\lambda\sigma - \frac{1}{2}\sigma^2\left(\widehat{q}_{y,t}^{h}\right)^2 + \widehat{q}_{y,t}^{h}\sigma\int_t^{t+1}\mathrm{d}W_s\right\}. \tag{6.8.6}$$

Comparing equation (6.8.5) with equation (6.8.6) and using $X_{y,t}^{h} = V_{y,t}^{h}$, we find

$$\widehat{q}_{y,t}^{h} = q_h\frac{\omega}{\sigma}.$$

We have that $\widehat{q}_{y,t}X_{y,t} = \sum_{h=\max\{x_r+y-t,1\}}^{x_{\max}} \widehat{q}_{y,t}^{h}V_{y,t}^{h} = \sum_{h=\max\{x_r+y-t,1\}}^{x_{\max}} \widehat{q}_{y,t}^{h}\alpha_{y,t}^{h}X_{y,t}$. Hence, $\widehat{q}_{y,t} = \sum_{h=\max\{x_r+y-t,1\}}^{x_{\max}} \widehat{q}_{y,t}^{h}\alpha_{y,t}^{h} = \omega\sum_{h=\max\{x_r+y-t,1\}}^{x_{\max}} q_h\alpha_{y,t}^{h}/\sigma$.

Derivation of (6.6.9)

Linearizing $\left(\widehat{F}_t\right)^{q_h}$ around $\widehat{F}_t = 1$, we arrive at $\left(\widehat{F}_t\right)^{q_h} \approx 1 + q_h\left(\widehat{F}_t - 1\right)$ so that (using (6.6.2))

$$F_t^{h} \approx 1 + q_h\left(\widehat{F}_t - 1\right).$$

Substitution of this approximation in (6.6.6) yields

$$F_{y,t} - 1 \approx \check{q}_{y,t}\left(\widehat{F}_t - 1\right), \tag{6.8.7}$$

where

$$\check{q}_{y,t} \equiv \sum_{h=\max\{x_r+y-t,1\}}^{x_{\max}+y-t} \gamma_{y,t}^{h} q_h.$$

We can write (6.8.7) as

$$\exp\left\{F_{y,t}-1\right\} \approx \exp\left\{\check{q}_{y,t}\left(\widehat{F}_t-1\right)\right\}.$$

Subsequently using the linear approximation $\exp\left\{F_{y,t}-1\right\} \approx F_{y,t}$, we arrive at (6.6.9).

Chapter 7

Pricing and Risk Management of Variable Annuities in Defined Ambition Pension Plans[84]

This chapter explores defined ambition pension plans, which are pension plans that allocate various risks (i.e., real interest rate risk, expected inflation risk and stock market risk) among policyholders on the basis of complete pension contracts while simultaneously pooling idiosyncratic longevity risk. We demonstrate how to value these pension plans in a market-consistent fashion. Market-consistent valuation of entitlements is important for avoiding conflicts between the insurer's policyholders and ensuring efficient intertemporal consumption smoothing. We also show that the costs of variable real annuities may be less sensitive to the nominal interest rate as compared to the costs of fixed nominal annuities, thereby reducing the nominal interest rate duration of the intertemporal hedging portfolio.

7.1. Introduction

Around the world, many firms are declining to continue to sponsor employer-sponsored defined benefit (DB) pension plans, due to the high risks these pension plans are now seen to impose on corporate sponsors (Investment Company Institute, 2014). In response, the defined ambition (DA) pension plans being designed in the Netherlands promise to play a new role, serving as mutual insurers in which policyholders, rather than corporate sponsors, bear investment risk. These DA pension plans aim to retain several advantages of traditional DB pension plans. In particular, by pooling idiosyncratic longevity risk, lifelong benefits can be provided at relatively low costs. Furthermore, risk management

[84] This chapter is co-authored with Lans Bovenberg.

seeks to provide retirees with stable income streams after they leave employment. To this end, real interest rate risk and expected inflation risk are actively managed during both the accumulation and the decumulation phase.

This chapter investigates the pricing and risk management of DA pension plans that provide variable annuities to policyholders.[85] Property rights of individual policyholders (i.e., pension entitlements) are defined in terms of annuity units (i.e., payouts) that vary with financial shocks. The pension contract specifies not only how annuity units respond to financial shocks but also how the desired (or targeted) growth rate of annuity units develops over time. In particular, the desired growth rate depends on the expected rate of inflation to protect the purchasing power of consumption, and on the interest rate to account for intertemporal substitution in consumption. Traditional variable annuities typically assume a constant (i.e., state-independent) desired growth rate (see, e.g., Horneff, Maurer, Mitchell, and Stamos, 2009, 2010; Maurer et al., 2013b).

We show how annuity units can be valued in a market-consistent fashion. Proper valuation of annuity units is relevant for determining the prices at which variable annuities can be bought and sold. It ensures that buying and selling of variable annuities does not impose externalities on other policyholders. Furthermore, market-consistent valuation helps protect the value of property rights if the pension contract is changed. Accordingly, the pension contract can be adapted to new circumstances without giving rise to conflicts between the insurer's policyholders. Also, proper pricing of annuity units ensures efficient intertemporal consumption smoothing, and allows policyholders to endogenously set their saving levels in order to realize a particular pension ambition in terms of a lifelong income stream during retirement.

We show that the costs of annuity units are an increasing function of the desired growth rate, and a decreasing function of *assumed* expected financial and biometric rates of return.[86] To account for the uncertain nature of future annuity units, the discount rate includes a risk premium that depends on assumed expected financial rates of return. Indeed, the annuity factor formalizes how the costs of annuity units depend not only on

[85] We define a variable annuity as an insurance contract in which annuity payments depend on the performance of the investment portfolio. A variable annuity does *not* include a guarantee.

[86] The annuity factor depends on assumed expected returns rather than actual expected returns. Selling and buying of annuities on the basis of prices that depend on assumed, rather than actual, expected returns does not impose externalities on other policyholders. In that case, however, actual expected consumption growth deviates from assumed expected consumption growth. Hence, intertemporal consumption smoothing is inefficient.

the median level (and desired growth rate) of future annuity units, but also on the risk of future annuity units. The more uncertain future annuity units are, the higher – ceteris paribus – this risk premium can be and thus the lower the costs of annuity units become. Biometric rates of return are the survival premia that depend on mortality rates. Higher mortality rates result in higher biometric rates of return reducing the costs of annuity units. This chapter focuses on stochastic financial rates of return and assumes that mortality rates are non-stochastic.

In a complete financial market, the portfolio strategy can be derived in closed-form by generalizing the principle of liability-driven investment from DB pension plans to variable annuities. In particular, the so-called replicating portfolio strategy can be decomposed into two components: a speculative component and an intertemporal hedging component. This decomposition is familiar from the literature on optimal consumption and portfolio choice under a stochastic investment opportunity set (see, e.g., Brennan and Xia, 2002; Wachter, 2002; Chacko and Viceira, 2005; Liu, 2007). The speculative portfolio allows policyholders to take advantage of risk premia. We show how the exposures to the various risk factors should be chosen if the policyholder aims to maximize the expected rate of return on the assets subject to a given amount of consumption risk. Unlike the replicating portfolio strategy with exogenous risk exposures, the *efficient* portfolio strategy depends on actual risk premia and thus suffers from model risk. The intertemporal hedging portfolio hedges changes in the future investment opportunity set that affect the costs of annuity units (see Merton, 1971). These changes in the future investment opportunity set are due to shocks in the real interest rate and the expected rate of inflation. The intertemporal hedging portfolio depends on the extent to which the interest rate affects the desired growth rate of annuity units. In the special case where a one percent point increase in the interest rate leads to a one percent point increase in the desired growth rate, the intertemporal hedging portfolio fully disappears.

We allow the actual portfolio strategy to differ from the replicating portfolio strategy. In that case, the actual portfolio strategy determines how annuity units develop over time (i.e., assets determine liabilities instead of the other way around). We show how annuity units should be adjusted such that the actual portfolio strategy does not affect – ex ante – the intertemporal allocation of the market value of annuity units. A mutual insurer can thus change its portfolio strategy without causing value transfers between generational cohorts. Alternatively, a mutual insurer can ring-fence the assets of each generation so that he needs not to worry about value transfers between generational cohorts. An advantage of ring-fenced accounts over one general pooled account is that the pension plan can be tailored to the needs of each generation (see also Bovenberg and Nijman, 2015). Furthermore, we show how an incorrect interest rate sensitivity of the intertemporal hedging portfolio gives rise to inefficient intertemporal consumption smoothing. In a collective pension fund without ring-fenced accounts, an incorrect interest rate sensitivity of the intertemporal hedging portfolio results in not only inefficient intertemporal consumption smoothing but also inefficient intergenerational risk sharing and intergenerational conflicts about the choice of the intertemporal hedging portfolio for the pension fund as a whole.

We allow the equity risk premium to be stochastic through a negative relationship with the nominal interest rate. Our specification of the equity risk premium causes the intertemporal hedging portfolio to depend on the speculative portfolio: a larger speculative portfolio renders the annuity factor less sensitive to the nominal interest rate, thereby reducing the nominal interest rate sensitivity of the intertemporal hedging portfolio. The stochastic model of the equity risk premium can be explained by stochastic variations in risk aversion that cause the rates of return on safe securities to move in an opposite direction from the rates of return on risky securities. Empirically, whereas nominal interest rates tend to vary pro-cyclically over the business cycle, the equity risk premium typically varies in a countercyclical fashion.[87] Campbell and Cochrane (1999) attribute this time variation in equity risk premia to the countercyclical behavior of risk aversion. In addition, countercyclical monetary policy causes nominal interest rates to behave pro-cyclically. These two stylized facts motivate our stochastic model of the equity risk premium.

We show how in an incomplete financial market, a mutual insurer can determine the intertemporal hedging portfolio so as to minimize the mismatch between the fund's aggregate assets and liabilities. In particular, if only a single nominal bond is available to hedge both real interest rate risk and expected inflation risk, then the duration of the best hedging portfolio (i.e., the portfolio that minimizes the mismatch between

[87] The literature has shown that equity risk premia tend to be larger in economic troughs than in booms (see, e.g., Fama and French, 1989; Harvey, 1989; Ferson and Harvey, 1991; Li, 2001; Lettau and Ludvigson, 2009).

the fund's aggregate assets and liabilities) trades-off hedging real interest risk against hedging inflation risk. If real interest rate risk dominates expected inflation risk, then the duration of the best hedging portfolio is close to the duration of the liabilities. The duration of the best hedging portfolio becomes shorter, however, if expected inflation risk dominates real interest rate risk. In an incomplete financial market, the intertemporal hedging portfolio depends on the financial model and thus becomes subject to model risk. The same holds true for the valuation of a variable *real* annuity. Indeed, in the absence of real securities that hedge expected inflation risk, real annuities cannot be priced objectively. We thus face a trade-off between optimal *risk sharing* on the one hand and objective market-consistent pricing of annuities on the other hand. To avoid conflicts with policyholders about the pricing of annuities, the insurer may want to provide variable annuities that can be valued objectively. In that case, non-traded expected inflation risk is also not traded between the insurer and its policyholders. Alternatively, the insurer can ring-fence the assets of a each generation so that the valuation of a variable real annuity cannot give rise to conflicts between generations. A disadvantage of ring-fenced accounts over one general pooled account is that non-traded risks (e.g., systematic longevity risk) can no longer be shared between generations.[88]

This chapter extends Chapter 6 in a number of ways. First, we consider continuous rather than discrete adjustments of entitlements. Second, and most importantly, we extend the number of risk factors by considering not only stock market risk but also real interest rate risk and expected inflation risk. These additional risk factors affect future investment opportunities so that the annuity factor becomes stochastic. With a stochastic annuity factor, the costs of annuity units (and hence contribution levels) depend on the macro-economic environment (i.e., the real interest rate and the expected rate of inflation). Indeed, the nominal interest rate sensitivity of the annuity factor yields conversion risk and thus results in intertemporal hedging demands aimed at hedging this risk. In fact, compared to Chapter 6, we include a number of extensions that affect how sensitive the annuity factor is with respect to changes in the nominal interest rate.

The remainder of this chapter is structured as follows. Section 7.2 describes the economy. Section 7.3 specifies the DA pension contract. Section 7.4 values the variable

[88]One could conclude separate swap contracts on these risks. These swap contracts, however, cannot be priced objectively (see also Bovenberg and Nijman, 2015).

annuities in a market-consistent fashion and explores the sensitivity of the annuity factor to the nominal interest rate. Section 7.5 determines the replicating portfolio strategy as well as the efficient portfolio strategy. Section 7.6 considers the case where the actual portfolio strategy deviates from the replicating portfolio strategy. This section also investigates the case of ring-fenced individual accounts while idiosyncratic longevity risk is still being pooled. Section 7.7 extends our results to a financial model with a stochastic equity risk premium. Section 7.8 considers an incomplete financial market in which real interest rate risk and expected inflation risk cannot be hedged simultaneously. Section 7.9 concludes the chapter. Proofs are relegated to the Appendix.

7.2. The Economy

This section outlines the economy. Section 7.2.1 describes the dynamics of the state variables. The price of a zero-coupon bond is derived in Section 7.2.2. Throughout, boldface type is used to denote uncertain variables at time t.

7.2.1. Dynamics of the State Variables

We consider a continuous-time economy with three state variables: the short-term real interest rate r_t, the short-term expected rate of inflation π_t and the nominal stock price S_t. The real interest rate and the expected rate of inflation follow mean reverting processes of the Ornstein-Uhlenbeck type. The nominal stock price is driven by a geometric Brownian motion. More specifically,

$$\mathrm{d}\mathbf{r}_t = \kappa\left(\bar{r} - r_t\right)\mathrm{d}t + \sigma_r \mathrm{d}\mathbf{W}_t^r, \tag{7.2.1}$$

$$\mathrm{d}\boldsymbol{\pi}_t = \theta\left(\bar{\pi} - \pi_t\right)\mathrm{d}t + \sigma_\pi \mathrm{d}\mathbf{W}_t^\pi, \tag{7.2.2}$$

$$\frac{\mathrm{d}\mathbf{S}_t}{S_t} = \left(R_t + \lambda_S \sigma_S\right)\mathrm{d}t + \sigma_S \mathrm{d}\mathbf{W}_t^S. \tag{7.2.3}$$

Here $\kappa > 0$ and $\theta > 0$ are mean reversion coefficients, $\bar{r}$ and $\bar{\pi}$ denote long-term means, R_t stands for the short-term nominal interest rate at time t, λ_S is the constant equity risk premium per unit of risk (i.e., the Sharpe ratio of the risky stock), $\sigma_r > 0$, $\sigma_\pi > 0$ and $\sigma_S > 0$ correspond to diffusion coefficients, and W_t^r, W_t^π and W_t^S represent standard Brownian motions. The real interest rate, the expected rate of inflation and the nominal stock price follow the same dynamics as in Brennan and Xia (2002).[89]

The correlation coefficients between the Brownian increments are summarized in the correlation matrix ρ:

$$\rho \equiv \begin{pmatrix} 1 & \rho_{r\pi} & \rho_{rS} \\ \rho_{r\pi} & 1 & \rho_{\pi S} \\ \rho_{rS} & \rho_{\pi S} & 1 \end{pmatrix}, \tag{7.2.4}$$

where ρ_{ij} ($i, j \in \{r, \pi, S\}$ and $i \neq j$) denotes the correlation coefficient between $\mathrm{d}\mathbf{W}_t^i$ and $\mathrm{d}\mathbf{W}_t^j$.

The real pricing kernel m_t evolves according to (see, e.g., Brennan and Xia, 2002)

$$\begin{aligned} \frac{\mathrm{d}\mathbf{m}_t}{m_t} &= -r_t \mathrm{d}t + \phi_r \mathrm{d}\mathbf{W}_t^r + \phi_\pi \mathrm{d}\mathbf{W}_t^\pi + \phi_S \mathrm{d}\mathbf{W}_t^S \\ &= -r_t \mathrm{d}t + \phi^\top \mathrm{d}\mathbf{W}_t. \end{aligned} \tag{7.2.5}$$

Here $\top$ denotes the transpose sign, $\phi \equiv (\phi_r, \phi_\pi, \phi_S)^\top$ and $W_t \equiv \left(W_t^r, W_t^\pi, W_t^S\right)^\top$.[90] The constant coefficients ϕ_r, ϕ_π and ϕ_S determine the market prices of risk associated with the state variables. More specifically, the vector of market prices of risk $\lambda \equiv (\lambda_r, \lambda_\pi, \lambda_S)$ can be computed from ϕ as follows:

$$\lambda = -\rho\phi. \tag{7.2.6}$$

7.2.2. Price of a Zero-Coupon Bond

Denote by $P_{\alpha,t}^h$ the price at time t of a zero-coupon bond with *fixed* maturity date $t + h$. Here $h \geq 0$ represents the time to maturity, and $\alpha \in [0, 1]$ is a parameter indicating the extent to which the payoff of the zero-coupon bond is linked to the price index

$$\Pi_t \equiv \exp\left\{\int_0^t \pi_s \mathrm{d}s\right\}. \tag{7.2.7}$$

[89] In contrast to Brennan and Xia (2002), we assume that the expected rate of inflation coincides with the realized rate of inflation. The results that follow can, however, be extended to the case where the expected rate of inflation differs from the realized rate of inflation.

[90] For notational convenience, we often write a *column* vector in the form $z = (z_1, z_2, ..., z_n)^\top$.

If $\alpha = 1$, the payoff of the bond is fully linked to the price index, while the payoff of the bond is not linked to the price index at all if $\alpha = 0$. For values of α in between zero and one, the payoff of the bond is only partially linked to the price index. We can view Π_t as the consumer price index or the wage price index.

The price of the bond can be obtained by computing the following conditional expectation:

$$\begin{aligned} P_{\alpha,t}^h &= \mathbb{E}_t \left[\frac{\mathbf{m}_{t+h}}{m_t} \frac{\Pi_t}{\mathbf{\Pi}_{t+h}} \left(\frac{\mathbf{\Pi}_{t+h}}{\Pi_t} \right)^{\alpha} \right] \\ &= \mathbb{E}_t \left[\exp \left\{ - \int_0^h \left(\mathbf{r}_{t+v} + (1-\alpha) \boldsymbol{\pi}_{t+v} + \frac{1}{2} \phi^\top \rho \phi \right) \mathrm{d}v \right. \right. \\ &\quad \left. \left. + \int_0^h \phi^\top \mathrm{d}\mathbf{W}_{t+v} \right\} \right], \end{aligned} \tag{7.2.8}$$

where $\mathbb{E}_t [\cdot]$ denotes the expectation operator conditional on all information available at time t. We find (see Appendix)

$$P_{\alpha,t}^h = \exp \left\{ - \int_0^h r_{\alpha,t}^v \mathrm{d}v \right\}. \tag{7.2.9}$$

Here $r_{\alpha,t}^v$ stands for the instantaneous forward interest rate at time t for horizon $v \geq 0$. The exact expression for $r_{\alpha,t}^v$ can be found in the Appendix (see (7.10.9)).

The yield to maturity (YTM) at time t for horizon $h \geq 0$ is given by (see Appendix)

$$\bar{r}_{\alpha,t}^h \equiv - \frac{\log P_{\alpha,t}^h}{h} = \frac{D^{\kappa,h}}{h} r_t + (1-\alpha) \frac{D^{\theta,h}}{h} \pi_t + \frac{E_\alpha^h}{h}. \tag{7.2.10}$$

Here $D^{x,h} \equiv \left(1 - e^{-xh} \right) / x$ with $x = \kappa$ or $x = \theta$. We note that $D^{x,h}/h$ decreases with the horizon h and $D^{x,h} \Rightarrow 0$ as $x \Rightarrow \infty$. Long-term YTMs are thus less variable as compared to short-term YTMs, especially when the mean reversion coefficients κ and θ are large. This property is consistent with empirical data (see Ang, Bekaert, and Wei, 2008). The exact expression for the horizon-dependent constant E_α^h can be found in the Appendix (see (7.10.11)).

The bond price $P^h_{\alpha,t}$ evolves according to (see Appendix)

$$\begin{aligned}\frac{\mathrm{d}\mathbf{P}^h_{\alpha,t}}{P^h_{\alpha,t}} &= \left(r_t + (1-\alpha)\pi_t - \lambda_r \sigma_r D^{\kappa,h} - (1-\alpha)\lambda_\pi \sigma_\pi D^{\theta,h}\right)\mathrm{d}t \\ &\quad - \sigma_r D^{\kappa,h}\mathrm{d}\mathbf{W}^r_t - (1-\alpha)\sigma_\pi D^{\theta,h}\mathrm{d}\mathbf{W}^\pi_t.\end{aligned} \tag{7.2.11}$$

We make the following observations. First, the expected return on the bond in excess of $r_t + (1-\alpha)\pi_t$ (i.e., the bond risk premium) is given by

$$-\lambda_r \sigma_r D^{\kappa,h} - (1-\alpha)\lambda_\pi \sigma_\pi D^{\theta,h}. \tag{7.2.12}$$

Estimates of λ_r and λ_π are typically negative (see, e.g., Brennan and Xia, 2002), so that bond risk premia are usually positive. Second, $D^{\kappa,h}$ and $D^{\theta,h}$ increase at a declining rate with the horizon h, implying that long-term bond risk premia exceed short-term bond risk premia. Third, the short-term nominal interest rate can be obtained from (7.2.11) by taking the limit $h \Rightarrow 0$. We find that the short-term nominal interest rate equals the short-term real interest rate plus the short-term expected rate of inflation. The Fisher equation thus holds true in this economy. Fourth, $D^{x,h}$ decreases with x. Hence bond risk premia are small if the mean reversion coefficients κ and θ are large. Finally, $D^{\kappa,h}$ and $(1-\alpha)D^{\theta,h}$ measure the sensitivity of the bond price with respect to (unexpected) changes in the real interest rate and the expected rate of inflation, respectively. Hence we can view $D^{\kappa,h}$ and $(1-\alpha)D^{\theta,h}$ as the real interest rate duration and the expected inflation duration of the bond, respectively.

The numerical illustrations in the chapter use the parameter values contained in Table 7.1 (see Appendix).

7.3. Specification of the Pension Contract

This section specifies the pension contract. Pension entitlements are framed in terms of (deferred) variable annuity units (i.e., payouts).[91] Let us denote by $B_{y,t}$ the annuity units at time t of a policyholder born at time y, by x_r the age at which a policyholder

[91] Brown et al. (2008, 2013) show that agents value annuities more when presented in a consumption frame than when presented in an investment frame.

retires (e.g., $x_r = 65$ years of age), and by $x_{\max}$ the maximum age a policyholder can reach (e.g., $x_{\max} = 120$ years of age). If the birth date y of a policyholder falls between time $t - x_r$ and time $t - x_{\max}$ and the policyholder has survived up to time t, then this policyholder receives a pension payment at time t. The probability that a policyholder currently aged $x = t - y$ will survive to age $x + h$ is denoted by

$$
{}_hp_x \equiv \exp\left\{-\int_0^h \mu_{x+v}\mathrm{d}v\right\}. \tag{7.3.1}
$$

Here μ_{x+v} denotes the force of mortality at age $x + v$. We assume that the force of mortality μ_{x+v} does not change over time. Systematic longevity risk is thus absent. Furthermore, in view of the law of large numbers, the insurer pools idiosyncratic longevity risk so that policyholders are insured against outliving their retirement assets.

The annuity units at time $t+h$ ($h \geq 0$) of a policyholder born at time y, i.e., $\mathbf{B}_{y,t+h}$, are specified in terms of past and future financial shocks as follows:[92]

$$
\begin{aligned}
\mathbf{B}_{y,t+h} = {} & B_{y,t_0} (\Pi_t)^{\beta} \exp\left\{\psi \int_{t_0}^{t} (r_s + (1-\beta)\pi_s)\,\mathrm{d}s + g\cdot(t - t_0)\right. \\
& \left. + \int_{t_0}^{t} \omega^{*\top}\mathrm{d}W_s^*\right\} \left(\frac{\mathbf{\Pi}_{t+h}}{\Pi_t}\right)^{\beta} \\
& \exp\left\{\psi \int_{t}^{t+h} (\mathbf{r}_s + (1-\beta)\boldsymbol{\pi}_s)\,\mathrm{d}s + g\cdot h + \int_t^{t+h} \omega^{*\top}\mathrm{d}\mathbf{W}_s^*\right\} \\
= {} & B_{y,t} \left(\frac{\mathbf{\Pi}_{t+h}}{\Pi_t}\right)^{\beta} \exp\left\{\psi \int_0^h (\mathbf{r}_{t+v} + (1-\beta)\boldsymbol{\pi}_{t+v})\,\mathrm{d}v + g\cdot h\right. \\
& \left. + \int_0^h \omega^{*\top}\mathrm{d}\mathbf{W}_{t+v}^*\right\}.
\end{aligned} \tag{7.3.2}
$$

Here t_0 is the time at which the (single) contribution is paid, $\omega^* \equiv (\omega_r^*, \omega_\pi^*, \omega_S^*)$, and $\mathrm{d}\mathbf{W}_t^* \equiv \mathrm{d}\mathbf{W}_t + (\lambda - \lambda^*)\,\mathrm{d}t$ with $\lambda^* \equiv (\lambda_r^*, \lambda_\pi^*, \lambda_S^*)$.[93] The parameter ω_i^* is the exposure of current annuity units $B_{y,t}$ to the (observed) financial shock $\mathrm{d}\mathbf{W}_t^{i*}$ ($i \in \{r, \pi, S\}$). The coefficients λ_r^*, λ_π^* and λ_S^* are the assumed market prices of risk. The assumed market prices of risk are allowed to differ from the actual market prices of risk λ_r, λ_π and λ_S.[94]

[92] Specification (7.3.2) assumes that no pension premia are paid after time t_0. We thus adopt a discontinuity perspective in which we only consider future annuity units on account of annuity units that have been accumulated up to time t_0.

[93] The financial shocks $\mathrm{d}\mathbf{W}_t^{r*}$ and $\mathrm{d}\mathbf{W}_t^{\pi*}$ can be determined from the observed price dynamics of two nominal zero-coupon bonds (with different times of maturity).

[94] As shown by Merton (1980), estimates of expected returns are less accurate than estimates of (co)variances. Therefore, we distinguish only between actual risk premia and assumed risk premia. In

The desired growth rate of annuity units (i.e., the growth rate of annuity units if $\omega^* = 0$) is affected by three factors. First, the parameter β represents the sensitivity of annuity units to the price index. If $\beta = 1$, then annuity units aim to keep up with price inflation, while annuity units are not linked to the price index at all if $\beta = 0$. Second, the parameter ψ measures how the desired growth rate varies with the interest rate.[95] If ψ is positive, then the desired growth rate increases as the interest rate (i.e., the return on savings) rises. The parameter ψ thus models intertemporal substitution in consumption.[96] Finally, the parameter g denotes a constant growth rate.

Log annuity units are adjusted according to (this follows from (7.3.2))

$$\mathrm{d}\log \mathbf{B}_{y,t} = \left(\beta\pi_t + \psi\left(r_t + (1-\beta)\pi_t\right) + g\right)\mathrm{d}t + {\omega^*}^\top \mathrm{d}\mathbf{W}_t^*. \tag{7.3.3}$$

Equation (7.3.3) can be viewed as the bonus (or dividend) policy of the pension plan, showing how annuity units develop as time proceeds. The right-hand side of equation (7.3.3) does not depend on age. Hence annuity units are adjusted uniformly across policyholders: each policyholder faces the same uniform adjustment of annuity units. The first term at the right-hand side of equation (7.3.3) represents the desired growth rate of annuity units. The second term denotes the impact of current financial shocks on current annuity units. We observe that current financial shocks are fully absorbed into current annuity units. Chapter 6 considers a pension plan for which annuity units respond gradually, rather than directly, to financial shocks. Our results that follow can be extended to the case of gradual absorption of financial shocks.

particular, we assume that actual (co)variances coincide with assumed (co)variances.

[95] If annuity units are fully linked to the price index (i.e., $\beta = 1$), then the desired growth rate depends on the real interest rate, while the desired growth rate depends on the nominal interest rate if annuity units are not linked to the price index at all (i.e., $\beta = 0$).

[96] Schroder and Skiadas (1999) show that if the investment opportunity set is constant, ψ can be viewed as the elasticity of intertemporal substitution (see also Chapter 4).

7.4. Pricing of Future Annuity Units

7.4.1. Market-Consistent Valuation

This section computes the market-consistent value of future annuity units. Denote by $V^h_{y,t}$ the market-consistent value at time t of future annuity units $\mathbf{B}_{y,t+h}$. We can compute $V^h_{y,t}$ by solving the following conditional expectation (see, e.g., Cochrane, 2001):

$$V^h_{y,t} = {}_hp_{t-y}\mathbb{E}_t\left[\frac{\mathbf{m}_{t+h}}{m_t}\frac{\Pi_t}{\mathbf{\Pi}_{t+h}}\mathbf{B}_{y,t+h}\right]. \tag{7.4.1}$$

Straightforward computations show that (see Appendix)

$$V^h_{y,t} = B_{y,t}A^h_{y,t}, \tag{7.4.2}$$

where the horizon-dependent annuity factor $A^h_{y,t}$ is defined as follows:

$$A^h_{y,t} \equiv \exp\left\{-\int_0^h \delta^v_{y,t}\mathrm{d}v\right\}. \tag{7.4.3}$$

Here $\delta^v_{y,t}$ denotes the forward discount rate at time t for maturity $v \geq 0$ for a policyholder born at time y:

$$\delta^v_{y,t} = \mu_{t-y+v} + (1-\psi)r^v_{\beta,t} + \omega^{*\top}\lambda^* + \xi_v - g. \tag{7.4.4}$$

The horizon-dependent annuity factor $A^h_{y,t}$ summarizes the impacts of the desired growth rate and the risk of annuity units on the costs of future annuity units. The value of the annuity factor $A^h_{y,t}$ is determined by the forward discount rate $\delta^v_{y,t}$ which depends on the forward biometric rate of return, the expected rate of return on the investment portfolio and the desired growth rate of annuity units. The term ξ_v includes second-order and interaction terms and represents the impact of the correlation between the underlying speculative and intertemporal hedging portfolio on the rate of return of the investment portfolio as a whole (see equation (7.10.12) in the Appendix).[97] If annuity units are constant over time (i.e., $\psi = g = \omega^* = 0$), then the forward discount rate $\delta^v_{y,t}$ is equal to

[97] The forward interest rate $r^v_{\beta,t}$ also includes second-order and interaction terms (see equation (7.10.9) in the Appendix).

the sum of the survival premium μ_{t-y+v} and the forward interest rate $r_{\beta,t}^{v}$. More generally, the forward discount rate $\delta_{y,t}^{v}$ is a decreasing function of the desired growth rate of annuity units (i.e., β, ψ and g), and an increasing function of future (assumed) expected biometric and financial rates of return. The biometric rate of return is represented by the force of mortality μ_{t-y+v}, whereas future (assumed) expected financial rates of return are represented by the other terms (except g). The risk premium $\omega^{*\top}\lambda^{*}$ is due to the impact of financial shocks on future annuity units. It depends on the exposure of future annuity units to financial shocks ω^{*} and the vector of assumed market prices of risk λ^{*}. The 'speculative' risk premium $\omega^{*\top}\lambda^{*}$ reflects the expected excess rate of return on the underlying (liability-driven) speculative investment portfolio.

7.4.2. Interest Rate Sensitivity of the Annuity Factor

The horizon-dependent annuity factor $A_{y,t}^{h}$ is stochastic and depends on the real interest rate r_t and the expected rate of inflation π_t. The sensitivity of the log annuity factor $\log A_{y,t}^{h}$ with respect to unexpected changes in the real interest rate is given by (see Appendix)

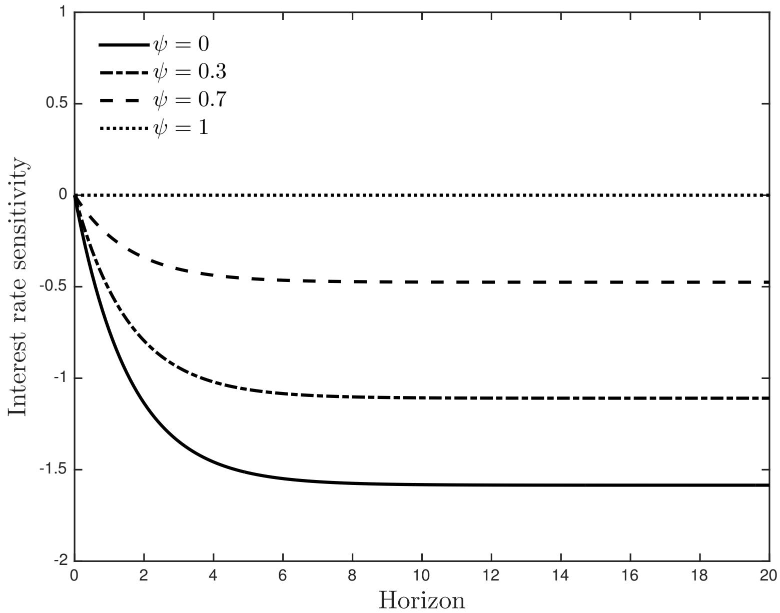
$$\frac{\partial \log \mathbf{A}_{y,t}^{h}}{\partial \mathbf{W}_{t}^{r}}\frac{1}{\sigma_r} = -(1-\psi)D^{\kappa,h}. \tag{7.4.5}$$

Equation (7.4.5) is usually referred to as the real interest rate duration of the annuity factor. This equation shows that changes in the real interest rate do not affect the annuity factor $A_{y,t}^{h}$ if $\psi = 1$. Intuitively, by raising future returns, a higher real interest rate reduces the *price* of a given consumption stream. With intertemporal substitution in consumption (i.e., $\psi > 0$), a higher real interest rate also raises the desired growth rate of future annuity units, thereby increasing the *magnitude* of future consumption streams. In the special case of $\psi = 1$, the price and volume effects of movements in the real interest rate cancel each other out. Figure 7.1 illustrates the interest rate sensitivity of $\log A_{y,t}^{h}$ for various values of ψ.

We can also compute the sensitivity of the log annuity factor $\log A_{y,t}^{h}$ with respect to unexpected changes in the expected rate of inflation (see Appendix):

$$\frac{\partial \log \mathbf{A}_{y,t}^{h}}{\partial \mathbf{W}_{t}^{\pi}}\frac{1}{\sigma_\pi} = -(1-\psi)(1-\beta)D^{\theta,h}. \tag{7.4.6}$$

Figure 7.1.
Illustration of the interest rate sensitivity of the market price



The figure illustrates the interest rate sensitivity of log $A^h_{y,t}$ for various values of ψ. The financial market parameter values are given in Table 7.1.

This equation shows that if annuity units are fully linked to the price index Π_t (i.e., $\beta = 1$), the annuity factor $A^h_{y,t}$ is affected only by changes in the real interest rate while an expected inflation shock leaves the annuity factor unaffected. Intuitively, in that case, the pension contract is defined in real terms so that pure nominal variables do not impact the costs of future annuity units. If annuity units are only partially linked to the price index (i.e., $\beta < 1$), a higher expected rate of inflation – ceteris paribus – reduces the costs of future annuity units.

7.4.3. The Conversion Factor

The market-consistent value at time t of an annuity unit for a policyholder born at time y is given by

$$A_{y,t} \equiv \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} A^h_{y,t} \mathrm{d}h = \frac{V_{y,t}}{B_{y,t}}, \tag{7.4.7}$$

where

$$V_{y,t} \equiv \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} V_{y,t}^h \mathrm{d}h. \tag{7.4.8}$$

The sensitivity of the log conversion factor $\log A_{y,t}$ with respect to unexpected changes in the real interest rate is given by (see Appendix)

$$\frac{\partial \log \mathbf{A}_{y,t}}{\partial \mathbf{W}_t^r} \frac{1}{\sigma_r} = -(1-\psi)\widehat{D}_{y,t}^{\kappa}, \tag{7.4.9}$$

where $\widehat{D}_{y,t}^{\kappa}$ is the κ-adjusted duration:

$$\widehat{D}_{y,t}^{\kappa} \equiv \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} \gamma_{y,t}^h D^{\kappa,h} \mathrm{d}h, \tag{7.4.10}$$

with $\gamma_{y,t}^h \equiv V_{y,t}^h / V_{y,t} = \left(B_{y,t}A_{y,t}^h\right) / \left(B_{y,t}A_{y,t}\right) = A_{y,t}^h / A_{y,t}$.

In an analogous way, we find (see Appendix)

$$\frac{\partial \log \mathbf{A}_{y,t}}{\partial \mathbf{W}_t^{\pi}} \frac{1}{\sigma_{\pi}} = -(1-\psi)(1-\beta)\widehat{D}_{y,t}^{\theta}. \tag{7.4.11}$$

Equation (7.4.7) is relevant for computing pension contributions and the portfolio strategy aimed at hedging conversion risk. We can distinguish between two alternative methods for determining pension premia, depending on what is assumed to be exogenous. A first method assumes that the newly bought annuity units $B_{y,t}$ are exogenously set and that the contribution $V_{y,t}$ varies endogenously with the real interest rate and the expected rate of inflation affecting the aggregate annuity factor $A_{y,t}$ according to (7.4.9) and (7.4.11), i.e., $V_{y,t} = B_{y,t}A_{y,t}$. This is consistent with defined ambition thinking in which annuity units (or pension ambitions) $B_{y,t}$ endogenously determine pension contributions. The second method assumes that the contribution $V_{y,t}$ is exogenously set and that the newly bought annuity units $B_{y,t}$ depend on the macro-economic environment determining $A_{y,t}$, i.e., $B_{y,t} = V_{y,t}/A_{y,t}$. This is consistent with defined contribution thinking in which contributions are fixed.

The method in which $V_{y,t}$ is exogenous and $B_{y,t}$ is endogenous is also relevant for determining the annuity units that can be bought at retirement from a capital sum. Indeed, $A_{y,t}$ can be viewed as the conversion factor at which a given amount of capital can be transformed into a consumption stream, i.e., $B_{y,t} = V_{y,t}/A_{y,t}$. More generally, during the decumulation phase, we can view $V_{y,t} = B_{y,t}A_{y,t}$ as the value of an individual account that corresponds to a certain number of annuity units $B_{y,t}$ of a policyholder born at time y. If property rights are defined in terms of (variable) annuity units (as in defined ambition thinking), $V_{y,t}$ corresponds to the market value of the annuity units $B_{y,t}$.

The intertemporal hedging portfolio aims at hedging the impact of macro-economic shocks (i.e., real interest rate shocks and expected inflation shocks) on the conversion factor and thus the consumption stream. That is why (7.4.7) is also important for the portfolio strategy of a policyholder who plans to buy an annuity at or during retirement. Hedging conversion risk ensures that an individual account can buy a fixed amount of annuity units without putting in more capital if the real interest rate and the expected rate of inflation change. Hedging the costs of future annuity units is also essential for an insurer providing (deferred) variable annuities. The next section explores the portfolio strategy in more detail.

7.5. Liability-Driven Investment

7.5.1. The Replicating Portfolio Strategy

This section derives the portfolio strategy that replicates the contract (7.3.2) for a policyholder born at time y. We allow the insurer to invest in three *risky* securities: two nominal zero-coupon bonds (with different times of maturity) and a risky stock. The number of risky securities thus equals the number of sources of risk. Let $X_{y,t}$ be the assets at time t of a policyholder born at time y, $\varpi^i_{y,t}$ be the fraction of assets invested in a nominal bond with time to maturity n_i $(i = 1, 2)$, and $\varpi^3_{y,t}$ be the fraction of assets invested in the risky stock. The fraction of assets invested in the nominal money market account is given by $1-\sum_{i=1}^{3} \varpi^i_{y,t}$. The *replicating* portfolio weights $\varpi^{1*}_{y,t}$, $\varpi^{2*}_{y,t}$ and $\varpi^{3*}_{y,t}$ are determined such that the value of the assets matches the value of the liabilities in each state of the world. We thus apply the principle of liability-driven investment familiar from DB pension plans to arrive at the replicating portfolio strategy. Specifically, the replicating portfolio weights solve the following system of equations (see Appendix)

$$-\left(\varpi^{1*}_{y,t}D^{\kappa,n_1}+\varpi^{2*}_{y,t}D^{\kappa,n_2}\right)=\frac{\omega^*_r}{\sigma_r}-(1-\psi)\widehat{D}^{\kappa}_{y,t}, \tag{7.5.1}$$

$$-\left(\varpi^{1*}_{y,t}D^{\theta,n_1}+\varpi^{2*}_{y,t}D^{\theta,n_2}\right)=\frac{\omega^*_\pi}{\sigma_\pi}-(1-\psi)(1-\beta)\widehat{D}^{\theta}_{y,t}, \tag{7.5.2}$$

$$\varpi^{3*}_{y,t}=\frac{\omega^*_S}{\sigma_S}. \tag{7.5.3}$$

The exact expressions for $\varpi^{1*}_{y,t}$, $\varpi^{2*}_{y,t}$ and $\varpi^{3*}_{y,t}$ are given in the Appendix (see (7.10.14), (7.10.15) and (7.10.16)). The right-hand side of equation (7.5.1) denotes the real interest rate sensitivity of the (log) value of the liabilities, i.e.,

$$\frac{\partial \log \mathbf{V}_{y,t}}{\partial \mathbf{W}^r_t}\frac{1}{\sigma_r}=\frac{\omega^*_r}{\sigma_r}-(1-\psi)\widehat{D}^{\kappa}_{y,t}, \tag{7.5.4}$$

while the left-hand side of equation (7.5.1) corresponds to the real interest rate sensitivity of the (log) value of the assets, i.e.,

$$\frac{\partial \log \mathbf{X}_{y,t}}{\partial \mathbf{W}^r_t}\frac{1}{\sigma_r}=-\left(\varpi^{1*}_{y,t}D^{\kappa,n_1}+\varpi^{2*}_{y,t}D^{\kappa,n_2}\right). \tag{7.5.5}$$

In an analogous way, we find

$$\frac{\partial \log \mathbf{V}_{y,t}}{\partial \mathbf{W}^\pi_t}\frac{1}{\sigma_\pi}=\frac{\omega^*_\pi}{\sigma_\pi}-(1-\psi)(1-\beta)\widehat{D}^{\theta}_{y,t}, \tag{7.5.6}$$

$$\frac{\partial \log \mathbf{X}_{y,t}}{\partial \mathbf{W}^\pi_t}\frac{1}{\sigma_\pi}=-\left(\varpi^{1*}_{y,t}D^{\theta,n_1}+\varpi^{2*}_{y,t}D^{\theta,n_2}\right). \tag{7.5.7}$$

The replicating portfolio strategy can be decomposed into two terms. The first terms at the right-hand sides of equations (7.5.1), (7.5.2) and (7.5.3) denote the speculative demands, whereas the second terms at the right-hand sides of (7.5.1) and (7.5.2) represent the intertemporal hedging demands. The intertemporal hedging demands depend on the sensitivity of the annuity factor with respect to unexpected shocks in the real interest rate and the expected rate of inflation. The intertemporal hedging portfolio is thus determined by the impact of financial shocks on the aggregate annuity factor. Indeed, the intertemporal hedging portfolio hedges the impact of these shocks on the aggregate annuity factor.

7.5.2. Mismatch Risk

The bonus rule (7.3.3) can be rewritten as follows:

$$\mathrm{d}\log \mathbf{B}_{y,t} = \left(\beta\pi_t + \psi\left(r_t + (1-\beta)\pi_t\right) + g\right)\mathrm{d}t + \mathrm{d}\log \mathbf{M}_{y,t}, \tag{7.5.8}$$

where

$$\begin{aligned}\mathrm{d}\log \mathbf{M}_{y,t} \equiv &\Big[-(1-\psi)\widehat{D}^{\kappa}_{y,t}\sigma_r\left(\mathrm{d}\mathbf{W}^{r*}_t + \lambda^*_r \mathrm{d}t\right)\\ &-(1-\psi)(1-\beta)\widehat{D}^{\theta}_{y,t}\sigma_\pi\left(\mathrm{d}\mathbf{W}^{\pi*}_t + \lambda^*_\pi \mathrm{d}t\right) + \omega^{*\top}\mathrm{d}\mathbf{W}^*_t\Big]\\ &-\Big[-(1-\psi)\widehat{D}^{\kappa}_{y,t}\sigma_r\left(\mathrm{d}\mathbf{W}^{r*}_t + \lambda^*_r \mathrm{d}t\right)\\ &-(1-\psi)(1-\beta)\widehat{D}^{\theta}_{y,t}\sigma_\pi\left(\mathrm{d}\mathbf{W}^{\pi*}_t + \lambda^*_\pi \mathrm{d}t\right)\Big] = \omega^{*\top}\mathrm{d}\mathbf{W}^*_t\end{aligned} \tag{7.5.9}$$

can be viewed as the *mismatch* between the replicating portfolio strategy (i.e., the portfolio strategy that replicates the pension contract (7.3.2)) and the intertemporal hedging portfolio strategy (i.e., the portfolio strategy that hedges stochastic variations in the aggregate annuity factor). Equation (7.5.8) shows that the speculative portfolio strategy determines how annuity units develop over time.

Mismatch is the difference between the development of assets and the development of 'norm' liabilities. The 'norm' liabilities are defined excluding unexpected shocks (i.e., under the assumption that expectations are met). The intertemporal hedging portfolio represents the value of these 'norm' liabilities. Mismatch (7.5.9) is absorbed by the policyholders themselves. This causes liabilities (*including* unexpected shocks) to continue to match assets such that the funding ratio remains unity.

7.5.3. The Efficient Portfolio Strategy

This section shows how the risk exposures ω^*_r, ω^*_π and ω^*_S should be chosen if the policyholder aims to maximize the expected rate of return on the assets subject to a given amount of consumption risk $\left(\omega^*_\rho\right)^2 = \omega^{*\top}\rho\omega^*$. Here ω^*_ρ is exogenously given. The expected return on the assets $X_{y,t}$ is given by

$$\begin{aligned}\mathbb{E}_t\left[\frac{\mathrm{d}\mathbf{X}_{y,t}}{X_{y,t}}\right] &= \mathbb{E}_t\left[\frac{\mathrm{d}\mathbf{V}_{y,t}}{V_{y,t}}\right]\\ &= r_t + \pi_t - (1-\psi)\lambda_r\sigma_r\widehat{D}^{\kappa}_{y,t} - (1-\psi)(1-\beta)\sigma_\pi\lambda_\pi\widehat{D}^{\theta}_{y,t} + \omega^{*\top}\lambda.\end{aligned} \tag{7.5.10}$$

The policyholder maximizes (7.5.10) over ω^* subject to $\left(\omega^*_\rho\right)^2 = \omega^{*\top}\rho\omega^*$. This yields

$$\omega^\star = -\phi \cdot \frac{\omega^*_\rho}{\phi_\rho} = \rho^{-1}\lambda\frac{\omega^*_\rho}{\phi_\rho}. \tag{7.5.11}$$

The efficient portfolio strategy is obtained by substituting (7.5.11) for ω^* in the replicating portfolio weight vector. Equation (7.5.11) shows that the vector of optimal risk exposures $\omega^\star$ depends on the actual risk premia λ_r, λ_π and λ_S. The efficient portfolio strategy is thus vulnerable to model risk.

7.6. Asset-Driven Liabilities

This section allows the actual portfolio strategy to differ from the replicating portfolio strategy. As a result, assets determine liabilities instead of the other way around. We thus speak of *asset-driven liabilities* instead of liability-driven investment.

7.6.1. Collective Defined Contribution

This section assumes that the mutual insurer has one general pooled account. Mismatch risk is shared between policyholders. We call this plan collective defined contribution (CDC): an external sponsor is absent and the annuity units are determined endogenously by the investment policy and by how mismatch risk is measured and allocated among policyholders with different ages.

We define total mismatch risk $\mathrm{d}\log\mathbf{M}_t$ as follows:

$$\mathrm{d}\log\mathbf{M}_t \equiv \mathrm{d}\log\mathbf{M}^1_t + \mathrm{d}\log\mathbf{M}^2_t + \mathrm{d}\log\mathbf{M}^3_t, \tag{7.6.1}$$

where

$$\mathrm{d}\log\mathbf{M}^1_t \equiv \omega^{*\top}\mathrm{d}\mathbf{W}^*_t, \tag{7.6.2}$$

$$\mathrm{d}\log\mathbf{M}^2_t \equiv \left(\varphi_t - \omega^*\right)^\top\left(\mathrm{d}\mathbf{W}^*_t + \lambda^*\mathrm{d}t\right), \tag{7.6.3}$$

$$\begin{aligned}\mathrm{d}\log\mathbf{M}^3_t \equiv &-\left(H^r_t - (1-\psi)\widehat{D}^\kappa_t\right)\sigma_r\left(\mathrm{d}\mathbf{W}^{r*}_t + \lambda^*_r\mathrm{d}t\right)\\ &-\left(H^\pi_t - (1-\psi)(1-\beta)\widehat{D}^\theta_t\right)\sigma_\pi\left(\mathrm{d}\mathbf{W}^{\pi*}_t + \lambda^*_\pi\mathrm{d}t\right).\end{aligned} \tag{7.6.4}$$

Here

$$\widehat{D}_t^{\kappa} \equiv \int_{t-x_{\max}}^{t-x_s} \gamma_{y,t} \widehat{D}_{y,t}^{\kappa} \mathrm{d}y, \tag{7.6.5}$$

where x_s denotes the age at which policyholders start their working career and $\gamma_{y,t} \equiv V_{y,t}/V_t$ with $V_t \equiv \int_{t-x_{\max}}^{t-x_s} V_{y,t}\mathrm{d}y$.[98] The vector $\varphi_t \equiv \left(\varphi_t^r, \varphi_t^{\pi}, \varphi_t^S\right)$ consists of the actual speculative demands, while $-H_t^r$ and $-H_t^{\pi}$ denote the actual hedging demands. Equation (7.6.3) can be viewed as the mismatch between the actual speculative portfolio and the 'desired' speculative portfolio (i.e., the speculative portfolio that finances (7.3.2)), whereas equation (7.6.4) represents the mismatch between the actual intertemporal hedging portfolio and the 'desired' intertemporal hedging portfolio (i.e., the intertemporal hedging portfolio that finances (7.3.2)). Total mismatch risk is defined as the mismatch between the actual portfolio and the 'desired' hedging portfolio (as determined by the discount rate that is used to compute the value of the 'norm' liabilities). Indeed, we measure total mismatch as the difference between the development of assets and the development of the 'norm' liabilities (as measured by the 'desired' intertemporal hedging portfolio).

The bonus rule is determined in such a way that the actual portfolio strategy does not affect – ex ante – the value of the liabilities $V_{y,t}^h$. This is important for avoiding conflicts between the insurer's policyholders. Using the requirement that the aggregate portfolio strategy does not redistribute market value among policyholders, we find that the bonus rule is given by (see Appendix)

$$\mathrm{d}\log \mathbf{B}_{y,t} = \left(\beta\pi_t + \psi\left(r_t + (1-\beta)\pi_t\right) + g + \widetilde{\xi}_{y,t} - \widehat{\xi}_{y,t}\right) dt + \mathrm{d}\log \mathbf{M}_t. \tag{7.6.6}$$

Here $\widetilde{\xi}_{y,t}$ and $\widehat{\xi}_{y,t}$ are second-order terms defined in the Appendix. Equation (7.6.6) shows that although the market value does not change as a result of a change in the actual portfolio strategy, the median value and the risk of future annuity units do change.

The ratio between the actual annuity units at time $t+h$ and the 'desired' annuity units at time $t+h$ (defined in (7.3.2)) is given by

$$\exp\left\{\int_0^h \left(\mathrm{d}\log \mathbf{M}_{t+v}^2 + \mathrm{d}\log \mathbf{M}_{t+v}^3\right) + \int_0^h \left(\widetilde{\xi}_{y,t+v} - \widehat{\xi}_{y,t+v}\right) \mathrm{d}v\right\}. \tag{7.6.7}$$

[98] $V_{y,t}$ now represents the market value at time t of *all* policyholders born at time y.

Equations (7.6.6) and (7.6.7) show that the actual portfolio strategy determines how future annuity units develop over time.

7.6.2. A Special Case

This section assumes that the insurer promises a *real* fixed annuity to its policyholders. However, supervisory authorities force the insurer to employ the nominal term structure to discount future annuity units (i.e., $\beta = 0$, $\omega^* = 0$ and $\psi = 0$).

7.6.2.1. Inefficient Intertemporal Consumption Smoothing

In case the actual portfolio strategy aims to mimic a real fixed annuity (i.e., $\varphi_t = 0$, $H_t^r = \widehat{D}_t^\kappa$ and $H_t^\pi = 0$), equation (7.6.1) collapses to

$$\mathrm{d}\log \mathbf{M}_t = \mathrm{d}\log \mathbf{M}_t^3 = \widehat{D}_t^\theta \sigma_\pi \left(\mathrm{d}\mathbf{W}_t^{\pi *} + \lambda_\pi^* \mathrm{d}t\right). \tag{7.6.8}$$

The actual sensitivity of log future annuity units $\log \mathbf{B}_{y,t+h}$ with respect to unexpected changes in the expected rate of inflation $\sigma_\pi \left(\mathrm{d}\mathbf{W}_t^{\pi *} + \lambda_\pi^* \mathrm{d}t\right)$ is not $D^{\theta,h}$ (implied by (7.3.2) with $\beta = 1$, $\omega^* = 0$ and $\psi = 0$) but rather $\widehat{D}_t^\theta$ (implied by (7.6.7) and (7.6.8) with $\beta = 0$, $\omega^* = 0$ and $\psi = 0$). Accordingly, the difference between the actual and the 'desired' sensitivity of log future annuity units $\log \mathbf{B}_{y,t+h}$ is given by

$$\widehat{D}_t^\theta - D^{\theta,h} \neq 0. \tag{7.6.9}$$

An expected inflation shock leads to a shock in real consumption, even though the actual portfolio strategy aims to mimic a real fixed annuity. Intuitively, by using a nominal discount rate rather than a real discount rate for calculating liabilities, intertemporal consumption smoothing is not efficient. In particular, a *positive* expected inflation shock typically causes a decline in real long-term consumption (i.e., $\widehat{D}_t^\theta < D^{\theta,h}$ for large h) and an increase in real short-term consumption (i.e., $\widehat{D}_t^\theta > D^{\theta,h}$ for small h). Indeed, a higher nominal interest rate on account of a higher expected rate of inflation depresses the value of the 'norm' intertemporal hedging portfolio, thereby understating the value of the 'true' intertemporal hedging portfolio (which takes into account the impact of a higher expected rate of inflation). The mismatch on account of an understatement of real liabilities raises consumption in the short run. The gain of consumption at short horizons is at the expense of long-term consumption, which receives inadequate compensation for a higher expected rate of inflation (the cost of a higher expected rate of inflation at horizon h is measured by $D^{\theta,h}$ which exceeds $\widehat{D}_t^{\theta}$ for large h).

7.6.2.2. Inefficient Intergenerational Risk Sharing

Equation (7.6.9) shows that an incorrect discount rate produces not only inefficient intertemporal consumption smoothing but also inefficient intergenerational risk sharing. In particular, in the case of a positive expected inflation shock, old generations gain at the expense of young generations, thereby making real consumption more risky than necessary. These inefficiencies in the allocation of consumption across generations become larger in more heterogeneous pension funds with large discrepancies in horizons (which causes $D^{\theta,h}$ to differ substantially from the average duration $\widehat{D}_t^{\theta}$ for large and small horizons h).

7.6.2.3. Intergenerational Conflict about the Portfolio Strategy

Inefficient intergenerational risk sharing of inflation shocks leads to intergenerational conflicts about the investment policy. In particular, to hedge against expected inflation shocks, young policyholders would prefer to invest in real bonds with a long duration such that the expected inflation duration of these bonds matches the expected inflation duration of their own consumption stream. Older policyholders, in contrast, would prefer to invest in nominal bonds with no or small expected inflation duration. Intuitively, each generation would like to distort the aggregate investment policy so as to offset the distortions of intergenerational risk sharing. The changes in the investment policy desired by old generations worsen expected inflation risk for young generations further, thereby causing an intergenerational conflict about the aggregate intertemporal hedging portfolio.

7.6.2.4. Inefficient Portfolio Strategy

If the insurer matches the prescribed nominal liabilities to avoid conflicts with the supervisor, then the difference between the actual and the 'desired' sensitivity of log future annuity units $\log \mathbf{B}_{y,t+h}$ with respect to unexpected changes in the expected rate of inflation is given by

$$-D^{\theta,h}. \tag{7.6.10}$$

Hence inefficiencies on account of a shock in the expected rate of inflation would on average be larger compared to (7.6.9), even though expected inflation risk for older generations with small horizons h would be smaller. Intuitively, the mutual insurer engages in not only inefficient intertemporal consumption smoothing and inefficient intergenerational risk sharing but also inefficient portfolio strategy: the incorrect discount rate introduces departures from the efficient portfolio, thereby worsening the risk-return trade-off further.

7.6.3. Ring-Fenced Accounts

This section assumes that each generation bears its own mismatch risk. That is, the assets belonging to cohort y are *ring-fenced* from the other assets in the fund. An advantage of ring-fenced accounts over one general pooled account is that the portfolio strategy (and hence the payout profile) can be tailored to the needs of each generation. Hence intergenerational conflicts about the investment policy are absent. At the same time, longevity risk is still being shared within a generation. Moreover, ring-fencing eliminates intergenerational conflicts about the valuation of financial risks.

Allowing for the actual portfolio to differ from the replicating portfolio, we can define total mismatch risk of each generation y in analogy of (7.6.1) as follows:

$$\mathrm{d}\log \mathbf{M}_{y,t} \equiv \mathrm{d}\log \mathbf{M}^1_{y,t} + \mathrm{d}\log \mathbf{M}^2_{y,t} + \mathrm{d}\log \mathbf{M}^3_{y,t}, \tag{7.6.11}$$

where

$$\mathrm{d}\log \mathbf{M}^1_{y,t} \equiv \omega^{*\top} \mathrm{d}\mathbf{W}^*_t \tag{7.6.12}$$

$$\mathrm{d}\log \mathbf{M}^2_{y,t} \equiv \left(\varphi_{y,t} - \omega^*\right)^\top \left(\mathrm{d}\mathbf{W}^*_t + \lambda^* \mathrm{d}t\right), \tag{7.6.13}$$

$$\begin{aligned} \mathrm{d}\log \mathbf{M}^3_{y,t} \equiv &- \left(H^r_{y,t} - (1-\psi)\widehat{D}^{\kappa}_{y,t}\right) \sigma_r \left(\mathrm{d}\mathbf{W}^{r*}_t + \lambda^*_r \mathrm{d}t\right) \\ &- \left(H^\pi_{y,t} - (1-\psi)(1-\beta)\widehat{D}^{\theta}_{y,t}\right) \sigma_\pi \left(\mathrm{d}\mathbf{W}^{\pi*}_t + \lambda^*_\pi \mathrm{d}t\right). \end{aligned} \tag{7.6.14}$$

Here the vector $\varphi_{y,t} \equiv \left(\varphi^r_{y,t}, \varphi^\pi_{y,t}, \varphi^S_{y,t}\right)$ consists of the actual speculative demands, while

$-H^r_{y,t}$ and $-H^\pi_{y,t}$ denote the actual hedging demands. The bonus rule is determined in such a way that $V^h_{y,t}$ is not affected by changes in the actual portfolio strategy so that a generation cannot affect the intertemporal allocation of market value. This facilitates the pooling of idiosyncratic longevity risk (see Chapter 5). We find (see Appendix)

$$\mathrm{d}\log \mathbf{B}_{y,t} = \left(\beta\pi_t + \psi\left(r_t + (1-\beta)\pi_t\right) + g + \widetilde{\xi}_{y,t} - \widehat{\xi}_{y,t}\right)\mathrm{d}t + \mathrm{d}\log \mathbf{M}_{y,t}. \tag{7.6.15}$$

Here $\widetilde{\xi}_{y,t}$ and $\widehat{\xi}_{y,t}$ are second-order terms defined in the Appendix. We note that equation (7.6.15) coincides with equation (7.5.8) if the actual portfolio strategy matches the replicating portfolio strategy. The ratio between the actual annuity units at time $t+h$ and the 'desired' annuity units at time $t+h$ (defined in (7.3.2)) is given by

$$\exp\left\{\int_0^h \left(\mathrm{d}\log \mathbf{M}^2_{y,t+v} + \mathrm{d}\log \mathbf{M}^3_{y,t+v}\right) + \int_0^h \left(\widetilde{\xi}_{y,t+v} - \widehat{\xi}_{y,t+v}\right)\mathrm{d}v\right\}. \tag{7.6.16}$$

Equations (7.6.15) and (7.6.16) show that the actual portfolio strategy determines how future annuity units develop over time.

The special case in which the mutual insurer aims to mimic a real fixed annuity but supervisory authorities force the insurer to employ the nominal term structure to discount future annuity units still produces inefficient intertemporal consumption smoothing. With ring-fenced accounts, however, inefficient intergenerational risk sharing and intergenerational conflicts about the portfolio strategy are no longer present. We also note that supervisory authorities may grant more discretion to mutual insurers to select and modify their own discount rates because these discount rates no longer affect the distribution of value among policyholders.

7.7. Stochastic Equity Risk Premium

The previous sections have assumed that the equity risk premium is constant over time. This section assumes that the equity risk premium varies stochastically over time.

7.7.1. Specification of the Equity Risk Premium

The real interest rate r_t and the expected rate of inflation π_t follow the same dynamic equations as in Section 7.2.1. The nominal stock price satisfies the following dynamic equation:

$$\frac{\mathrm{d}\mathbf{S}_t}{S_t} = (R_t + e_t)\,\mathrm{d}t + \sigma_S \mathrm{d}\mathbf{W}_t^S, \tag{7.7.1}$$

where the equity risk premium $e_t \equiv \lambda_t^S \sigma_S$ is a *linear* function of the real interest rate and the expected rate of inflation. That is,

$$e_t = \nu - a r_t - b\pi_t \tag{7.7.2}$$

for some positive constants ν, a and b. The equity risk premium e_t is subject to the following dynamic equation:

$$\mathrm{d}\mathbf{e}_t = -a\kappa\,(\bar{r} - r_t)\,\mathrm{d}t - b\theta\,(\bar{\pi} - \pi_t)\,\mathrm{d}t - a\sigma_r \mathrm{d}\mathbf{W}_t^r - b\sigma_\pi \mathrm{d}\mathbf{W}_t^\pi. \tag{7.7.3}$$

If $a = 1$ and $b = 1$, then the expected nominal rate of return on the risky stock is equal to ν, while if $a = 0$ and $b = 0$, then the equity risk premium is constant. The real pricing kernel m_t satisfies the following dynamic equation:

$$\begin{aligned}\frac{\mathrm{d}\mathbf{m}_t}{m_t} &= -r_t \mathrm{d}t + \phi_t^r \mathrm{d}\mathbf{W}_t^r + \phi_t^\pi \mathrm{d}\mathbf{W}_t^\pi + \phi_t^S \mathrm{d}\mathbf{W}_t^S \\ &= -r_t \mathrm{d}t + \phi_t^\top \mathrm{d}\mathbf{W}_t.\end{aligned} \tag{7.7.4}$$

Here $\phi_t \equiv \left(\phi_t^r, \phi_t^\pi, \phi_t^S\right)$. The coefficients ϕ_t^r, ϕ_t^π and ϕ_t^S determine the market prices of risk λ_r, λ_π and $\lambda_t^S = (\nu - a r_t - b\pi_t)\,/\sigma_S$ that are associated with the state variables. The vector of market prices of risk $\lambda_t \equiv \left(\lambda_r, \lambda_\pi, \lambda_t^S\right)$ can be computed from ϕ_t as follows:

$$\lambda_t = -\rho\phi_t. \tag{7.7.5}$$

The YTM is given by equation (7.2.10).[99]

[99]See also Vasicek (1977).

7.7.2. The Pension Contract

The annuity units at time $t+h$ ($h \geq 0$) of a policyholder born at time y, $\mathbf{B}_{y,t+h}$, are given by specification (7.3.2). Financial shocks are now defined as follows:

$$\mathrm{d}\mathbf{W}_t^* \equiv \mathrm{d}\mathbf{W}_t + (\lambda_t - \lambda_t^*)\,\mathrm{d}t \tag{7.7.6}$$

with $\lambda_t^* \equiv \left(\lambda_r^*, \lambda_\pi^*, \lambda_t^{S*}\right)$. The *assumed* equity risk premium is denoted by $e_t^* \equiv \nu^* - a^* r_t - b^* \pi_t$. We assume that the vector of long-term risk exposures ω^* is constant.

7.7.3. Pricing of Future Annuity Units

7.7.3.1. Market-Consistent Valuation

Let $V_{y,t}^h$ be the market-consistent value at time t of $\mathbf{B}_{y,t+h}$. Straightforward computations show that (see Appendix)

$$V_{y,t}^h = B_{y,t} A_{y,t}^h, \tag{7.7.7}$$

where

$$A_{y,t}^h \equiv \exp\left\{-\int_0^h \delta_{y,t}^{v,h} \mathrm{d}v\right\}, \tag{7.7.8}$$

$$\begin{aligned}\delta_{y,t}^{v,h} &= \mu_{t-y+v} + (1-\psi) r_{\beta,t}^v + \lambda_r \sigma_r \frac{\omega_S^* a^*}{\sigma_S} D^{\kappa,v} + \lambda_\pi \sigma_\pi \frac{\omega_S^* b^*}{\sigma_S} D^{\theta,v} \\ &\quad + \omega^{*\top} \mathbb{E}_t \left[\boldsymbol{\lambda}_{t+h-v}^*\right] + \xi_v - g.\end{aligned} \tag{7.7.9}$$

The second and third term at the right-hand side of equation (7.7.9) arise because the equity risk premium is negatively linearly related to the real interest rate and the expected rate of inflation. The risk premium $\omega^{*\top}\mathbb{E}_t\left[\boldsymbol{\lambda}_{t+h-v}^*\right]$ arises because the insurer takes speculative risk. This risk premium is not constant but time-dependent. Indeed, the equity risk premium varies stochastically over time. The risk premium ξ_v includes second-order and interaction terms. The exact expression for ξ_v can be found in the Appendix.

7.7.3.2. Interest Rate Sensitivity of the Annuity Factor

The sensitivity of $\log A^h_{y,t}$ with respect to unexpected changes in the real interest rate is given by (see Appendix)

$$\frac{\partial \log \mathbf{A}^h_{y,t}}{\partial \mathbf{W}^r_t}\frac{1}{\sigma_r} = \frac{\omega^*_S a^*}{\sigma_S} D^{\kappa,h} - (1-\psi)D^{\kappa,h}. \tag{7.7.10}$$

The sensitivity of $\log A^h_{y,t}$ with respect to unexpected changes in the expected rate of inflation is given by (see Appendix)

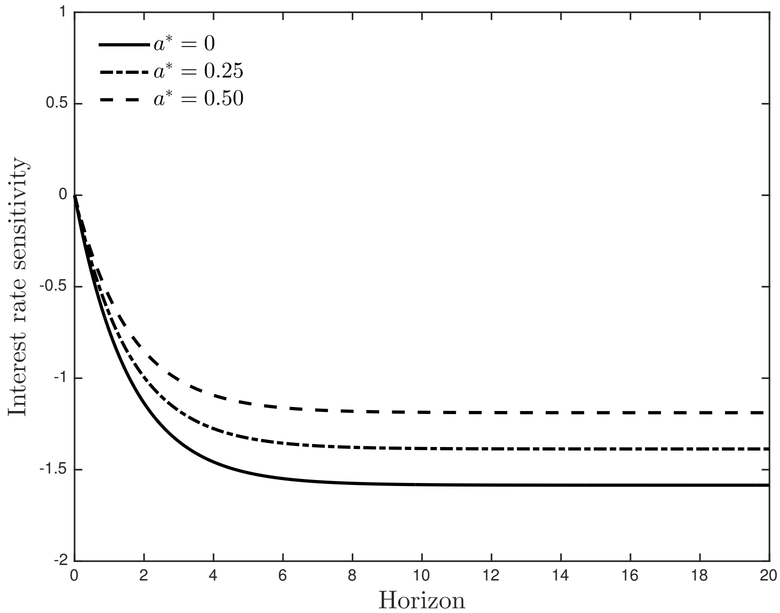
$$\frac{\partial \log \mathbf{A}^h_{y,t}}{\partial \mathbf{W}^r_\pi}\frac{1}{\sigma_\pi} = \frac{\omega^*_S b^*}{\sigma_S} D^{\theta,h} - (1-\psi)(1-\beta)D^{\theta,h}. \tag{7.7.11}$$

The horizon-dependent annuity factor $A^h_{y,t}$ may become less sensitive to unexpected changes in the real interest rate and the expected rate of inflation if the equity risk premium varies stochastically over time. Intuitively, a *low* nominal interest rate implies a *high* equity risk premium, so that the costs of future annuity units may become less sensitive to unexpected changes in the real interest rate and the expected rate of inflation if the insurer takes stock market risk. Figure 7.2 shows the real interest rate sensitivity of $\log A^h_{y,t}$ for various values of a^* ($\psi = 0$). We assume that the insurer invests 50% of wealth into the risky stock. The figure shows that the real interest rate duration of the liabilities decreases by 25% if a^* goes up from 0 to 0.5.

7.7.4. Liability-Driven Investment

This section derives the replicating portfolio strategy for a policyholder born at time y. We allow the insurer to invest in three *risky* securities: two nominal zero-coupon bonds (with different times of maturity) and a risky stock. Let $\varpi^i_{y,t}$ be the fraction of assets invested in a nominal bond with time to maturity n_i ($i = 1, 2$) and $\varpi^3_{y,t}$ be the fraction of assets invested in the risky stock. The fraction of assets invested in the nominal money market account is given by $1 - \sum_{i=1}^3 \varpi^i_{y,t}$. The *replicating* portfolio weights $\varpi^{1*}_{y,t}$, $\varpi^{2*}_{y,t}$ and $\varpi^{3*}_{y,t}$ are determined in such a way that the value of the assets matches the value of the liabilities in each state of the world. The replicating portfolio weights solve the

Figure 7.2.
Illustration of the interest rate sensitivity of the annuity factor



The figure shows the interest rate sensitivity of $\log A^h_{y,t}$ for various values of a^* ($\psi = 0$). The insurer invests 50% of wealth into the risky stock. The financial market parameter values are given in Table 7.1.

following system of equations (see Appendix)

$$-\left(\varpi^{1*}_{y,t} D^{\kappa,n_1} + \varpi^{2*}_{y,t} D^{\kappa,n_2}\right) = \frac{\omega^*_r}{\sigma_r} - (1-\psi)\widehat{D}^{\kappa}_{y,t} + \widehat{D}^{\kappa}_{y,t}\frac{\omega^*_S a^*}{\sigma_S}, \tag{7.7.12}$$

$$-\left(\varpi^{1*}_{y,t} D^{\theta,n_1} + \varpi^{2*}_{y,t} D^{\theta,n_2}\right) = \frac{\omega^*_\pi}{\sigma_\pi} - (1-\psi)(1-\beta)\widehat{D}^{\theta}_{y,t} + \widehat{D}^{\theta}_{y,t}\frac{\omega^*_S b^*}{\sigma_S}, \tag{7.7.13}$$

$$\varpi^{3*}_{y,t} = \frac{\omega^*_S}{\sigma_S}. \tag{7.7.14}$$

The replicating portfolio strategy can be decomposed into two terms. The first terms at the right-hand sides of equations (7.7.12), (7.7.13) and (7.7.14) denote the speculative demands, while the second and third terms at the right-hand sides of (7.7.12) and (7.7.13) correspond to the intertemporal hedging demands. The intertemporal hedging portfolio now depends on ω^*_S because the equity risk premium is stochastic. Indeed, a larger ω^*_S typically renders the annuity factor less sensitive to the nominal interest rate, thereby reducing the nominal interest rate sensitivity of the intertemporal hedging portfolio. Hence the intertemporal hedging portfolio is now affected by the speculative portfolio.

7.7.5. An Incorrect Discount Rate

This section assumes that the actual portfolio strategy takes into account the fact that the equity risk premium varies stochastically over time (i.e., $\varphi = \omega^*$, $H_t^r = (1-\psi)\widehat{D}_t^\kappa - \widehat{D}_t^\kappa \omega_S^* a^*/\sigma_S$ and $H_t^\pi = (1-\psi)(1-\beta)\widehat{D}_t^\theta - \widehat{D}_t^\theta \omega_S^* b^*/\sigma_S$). However, supervisory authorities force the insurer to discount liabilities by (7.4.4). The mutual insurer has one general pooled account. Mismatch risk is thus shared between generations. We can define total mismatch risk as follows (note that $\mathrm{d}\log \mathbf{M}_t^2 = 0$):

$$\begin{aligned}
\mathrm{d}\log \mathbf{M}_t &= \mathrm{d}\log \mathbf{M}_t^1 + \mathrm{d}\log \mathbf{M}_t^3 \\
&= \omega^{*\top} \mathrm{d}W_t^* + \widehat{D}_t^\kappa \frac{\omega_S^* a^* \sigma_r}{\sigma_S} \left(\mathrm{d}\mathbf{W}_t^{r*} + \lambda_r^* \, \mathrm{d}t\right) \\
&\quad + \widehat{D}_t^\theta \frac{\omega_S^* b^* \sigma_\pi}{\sigma_S} \left(\mathrm{d}\mathbf{W}_t^{\pi*} + \lambda_\pi^* \, \mathrm{d}t\right).
\end{aligned} \tag{7.7.15}$$

The bonus rule is determined in such a way that $V_{y,t}^h$ is not affected by a change in the actual portfolio strategy. We find (see Appendix)

$$\mathrm{d}\log \mathbf{B}_{y,t} = \left(\beta\pi_t + \psi\left(r_t + (1-\beta)\pi_t\right) + g + \widetilde{\xi}_{y,t} - \widehat{\xi}_{y,t}\right) \mathrm{d}t + \mathrm{d}\log \mathbf{M}_t. \tag{7.7.16}$$

Here $\widetilde{\xi}_{y,t}$ and $\widehat{\xi}_{y,t}$ are second-order terms (see Appendix). The ratio between the actual annuity units at time $t+h$ and the 'desired' annuity units at time $t+h$ is given by

$$\exp\left\{\int_0^h \mathrm{d}\log \mathbf{M}_{t+v}^3 + \int_0^h \left(\widetilde{\xi}_{y,t+v} - \widehat{\xi}_{y,t+v}\right) \mathrm{d}v\right\}. \tag{7.7.17}$$

The actual sensitivity of log future annuity units $\log \mathbf{B}_{y,t+h}$ with respect to unexpected changes in the real interest rate $\sigma_r\left(\mathrm{d}\mathbf{W}_t^{r*} + \lambda_r^* \mathrm{d}t\right)$ is not $\psi D^{r,h} + \omega_r^*/\sigma_r$ but rather $\psi D^{r,h} + \omega_r^*/\sigma_r + \widehat{D}_t^\kappa \frac{\omega_S^* a^*}{\sigma_S}$. Accordingly, the difference between the actual and the desired sensitivity of log future annuity units $\log \mathbf{B}_{y,t+h}$ is given by

$$\widehat{D}_t^\kappa \frac{\omega_S^* a^*}{\sigma_S} \neq 0. \tag{7.7.18}$$

An incorrect discount rate thus produces inefficient intertemporal consumption smoothing and inefficient intergenerational risk sharing. In the same way as in Section 7.6.2.4, this leads to intergenerational conflicts about the investment policy.

7.8. Incomplete Financial Market

This section assumes that annuity units at time $t+h$ ($h \geq 0$) are specified as follows:[100]

$$\mathbf{B}_{y,t+h} = B_{y,t} \left(\frac{\mathbf{\Pi}_{t+h}}{\Pi_t} \right)^{\beta}. \tag{7.8.1}$$

In addition, we assume that the investment opportunity set consists of a nominal money market account and a single zero-coupon nominal bond. The pension contract (7.8.1) can thus not be replicated unless $\beta = 0$. That is, the actual portfolio strategy is *forced* to differ from the replicating portfolio strategy. As in Section 7.6, we assume that the insurer shifts the mismatch between the actual portfolio strategy and the replicating portfolio strategy back to its policyholders. The insurer can, to some degree, control mismatch risk by appropriately choosing the duration of the actual portfolio strategy. We show that the duration of the best hedging portfolio strategy (i.e., the portfolio strategy that minimizes mismatch risk) is small (large) if fluctuations in the nominal interest rate are largely driven by fluctuations in the expected rate of inflation (real interest rate).[101] Intuitively, if changes in the nominal interest rate are primarily driven by changes in the expected rate of inflation, then investing in short-term financial instruments (such as a nominal money market account) provides a 'good' hedge against expected inflation risk. Section 7.8.1 assumes one general account for all policyholders (i.e., mismatch risk is shared between generations), while Section 7.8.2 considers the case of ring-fenced accounts (i.e., mismatch risk is not shared between generations).

7.8.1. Collective Defined Contribution

In what follows, we assume that $\kappa = \theta$. This assumption implies that $D^{\kappa,h} = D^{\theta,h}$. The *best hedging portfolio* is defined as the one that minimizes the variance of the mismatch between the actual portfolio strategy and the replicating portfolio strategy (note that the replicating portfolio strategy does not exist because real bonds are not available). The mutual insurer chooses to shift the mismatch between the actual portfolio and the

[100] Specification (7.8.1) arises as a special case of specification (7.3.2) if $\psi = g = \omega^* = 0$.

[101] This assumes that the insurer aims to hedge a *real* annuity. However, if the insurer aims to hedge a nominal annuity, then the duration of the best hedging portfolio strategy *exactly* matches the duration of the liabilities.

replicating portfolio back to *all* its policyholders (i.e., mismatch risk is shared between generations). Let $D^{\kappa,n}$ be the duration of the single zero-coupon nominal bond with time to maturity n, and let ϖ_t be the fraction of assets invested in the zero-coupon nominal bond (with the remaining fraction of assets being invested in the nominal money market account). The insurer determines ϖ_t such that the variance of mismatch risk is minimized. Mismatch risk is defined as follows:

$$\begin{aligned} \mathrm{d}\log \mathbf{M}_t = &\left(\varpi_t D^{\kappa,n} - \widehat{D}_t^{\kappa}\right) \sigma_r \left(\mathrm{d}\mathbf{W}_t^{r*} + \lambda_r^* \,\mathrm{d}t\right) \\ &+ \left(\varpi_t D^{\kappa,n} - (1-\beta)\widehat{D}_t^{\kappa}\right) \sigma_\pi \left(\mathrm{d}\mathbf{W}_t^{\pi*} + \lambda_\pi^* \,\mathrm{d}t\right). \end{aligned} \tag{7.8.2}$$

The insurer faces the following minimization problem:

$$\underset{\varpi_t}{\mathrm{Minimize}} \mathbb{V}\left[\mathrm{d}\log \mathbf{M}_t\right]. \tag{7.8.3}$$

Solving (7.8.3) yields

$$\varpi_t^* = \frac{\sigma_r^2 + (1-\beta)\sigma_\pi^2 + (2-\beta)\rho_{r\pi}\sigma_r\sigma_\pi}{\sigma_r^2 + \sigma_\pi^2 + 2\rho_{r\pi}\sigma_r\sigma_\pi} \frac{\widehat{D}_t^{\kappa}}{D^{\kappa,n}}. \tag{7.8.4}$$

The ratio between the duration of the best hedging portfolio (i.e., $\varpi_t^* D^{\kappa,n}$) and the duration of the liabilities (i.e., $\widehat{D}_t^{\kappa}$) is given by

$$\frac{\varpi_t^* D^{\kappa,n}}{\widehat{D}_t^{\kappa}} = \frac{\sigma_r^2 + (1-\beta)\sigma_\pi^2 + (2-\beta)\rho_{r\pi}\sigma_r\sigma_\pi}{\sigma_r^2 + \sigma_\pi^2 + 2\rho_{r\pi}\sigma_r\sigma_\pi} \leq 1. \tag{7.8.5}$$

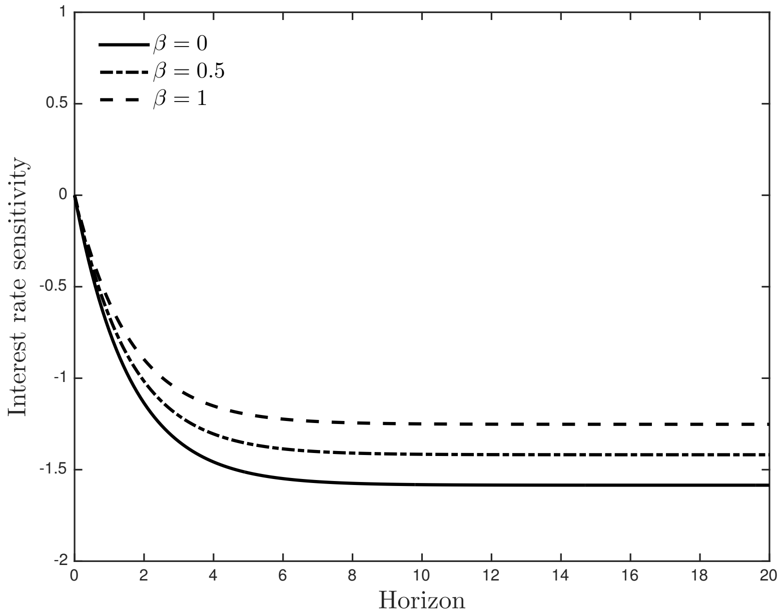
We observe that $\varpi_t^* D^{\kappa,n} \Rightarrow \widehat{D}_t^{\kappa}$ if $\sigma_\pi \Rightarrow 0$. On the other hand, if $\sigma_\pi \Rightarrow \infty$, then $\varpi_t^* D^{\kappa,n} \Rightarrow 0$. The duration of the best hedging portfolio is thus small (large) if fluctuations in the nominal interest rate are largely driven by fluctuations in the expected rate of inflation (real interest rate). Figure 7.3 shows the function

$$\frac{\sigma_r^2 + (1-\beta)\sigma_\pi^2 + (2-\beta)\rho_{r\pi}\sigma_r\sigma_\pi}{\sigma_r^2 + \sigma_\pi^2 + 2\rho_{r\pi}\sigma_r\sigma_\pi} D^{\kappa,h} \tag{7.8.6}$$

for various values of h and β. We observe that for $\beta = 1$, the duration of the best hedging portfolio is approximately 80% of $D^{\kappa,h}$.

Valuation of the pension contract (7.8.1) is relevant for determining the duration $\widehat{D}_t^{\kappa}$. Since the financial market is incomplete (i.e., real bonds are not traded), V_t^h and $\widehat{D}_t^{\kappa}$

Figure 7.3.
Illustration of the interest rate sensitivity of the best hedging portfolio



The figure illustrates the interest rate sensitivity of the best hedging portfolio for various values of β. The financial market parameter values are given in Table 7.1.

cannot be objectively determined. The insurer thus faces a trade-off between optimal risk sharing on the one hand and objective market-consistent pricing of annuities on the other hand. Indeed, in order to avoid conflicts with policyholders about the pricing of annuities, the insurer may want to provide variable annuities that can be valued objectively. In that case, non-traded expected inflation risk is also not traded between the insurer and its policyholders.

7.8.2. Ring-fenced Accounts

This section assumes that mismatch risk is not shared between generations. Let $\varpi_{y,t}$ be the fraction of assets invested in the zero-coupon nominal bond (with the remaining fraction of assets being invested in the nominal money market account). The insurer determines $\varpi_{y,t}$ such that the variance of mismatch risk is minimized. Mismatch risk is now defined as follows:

$$\begin{aligned} \mathrm{d}\log \mathbf{M}_{y,t} &= \left(\varpi_{y,t} D^{\kappa,n} - \widehat{D}^{\kappa}_{y,t}\right) \sigma_r \left(\mathrm{d}\mathbf{W}^{r*}_t + \lambda^*_r \,\mathrm{d}t\right) \\ &\quad + \left(\varpi_{y,t} D^{\kappa,n} - (1-\beta)\widehat{D}^{\kappa}_{y,t}\right) \sigma_\pi \left(\mathrm{d}\mathbf{W}^{\pi*}_t + \lambda^*_\pi \,\mathrm{d}t\right). \end{aligned} \tag{7.8.7}$$

The insurer faces the following minimization problem:

$$\underset{\varpi_{y,t}}{\text{Minimize}} \mathbb{V}\left[\mathrm{d}\log \mathbf{M}_{y,t}\right]. \tag{7.8.8}$$

We find

$$\varpi^{*}_{y,t} = \frac{\sigma_r^2 + (1-\beta)\sigma_\pi^2 + (2-\beta)\rho_{r\pi}\sigma_r\sigma_\pi}{\sigma_r^2 + \sigma_\pi^2 + 2\rho_{r\pi}\sigma_r\sigma_\pi} \frac{\widehat{D}^{\kappa}_{y,t}}{D^{\kappa,n}}. \tag{7.8.9}$$

As in Section 7.8.1, the insurer faces a trade-off between objective market-consistent pricing of annuities on the one hand (i.e., the value of the generational account $V_{y,t}$ can be objectively determined because the assets of cohort y are ring-fenced from the other assets in the fund) and sharing of systemic risks (e.g., expected inflation risk) between generations on the other hand.

7.9. Concluding Remarks

This chapter has explored pricing and risk management of variable annuities in DA pension plans. We have shown that the costs of variable real annuities may be less sensitive to the nominal interest rate as compared to the costs of fixed nominal annuities. This is so because of three reasons. First of all, the desired growth rate of annuity units may increase with the interest rate due to intertemporal substitution in consumption. Second, the desired growth rate rises with the expected rate of inflation so that the costs of these annuities depend on the real rather than the nominal interest rate. Hence, changes in nominal interest rates impact the cost of an annuity only if these changes in nominal interest rates reflect changes in real interest rates. The costs of real annuities tend to be more stable than the costs of fixed nominal annuities because the real interest rate is less volatile than the nominal interest rate: fluctuating inflation expectations affect mainly the nominal rather than the real interest rate. In an incomplete financial market in which real interest rate risk and expected inflation risk cannot be hedged simultaneously, insurers must trade-off hedging real interest rate risk against hedging expected inflation risk. This reduces the nominal interest sensitivity of the annuity factor, especially when fluctuations in the nominal interest rate are driven by changes in the expected rate of inflation rather than by changes in the real interest rate. A third factor reducing the nominal interest sensitivity of the annuity factor is that the expected rates of return on risky securities tend to be less sensitive to the nominal interest rate when compared to the rates of return on safe securities. Overall then, the cost of real variable annuities tend to be more stable than the costs of nominal fixed annuities because real expected rates of return on risky securities are less volatile than nominal rates of return on safe securities. Indeed, TIAA-CREF has fixed the assumed real expected rate of return on its variable annuities at 4% since it started to provide this retirement product in 1952.

7.10. Appendix

7.10.1. Parameter Values

Table 7.1.
Parameter values

	Parameter	Value
Real Interest Rate Process	κ	0.631
	$\bar{r}$	0.012
	σ_r	0.026
	λ_r	-0.209
Expected Inflation Process	θ	0.027
	$\bar{\pi}$	0.054
	σ_π	0.014
	λ_π	-0.105
Stock Return Process	σ_S	0.343
	λ_S	0.158
Correlation Matrix	$\rho_{r\pi}$	-0.061
	ρ_{rS}	-0.129
	$\rho_{\pi S}$	-0.024

The table reports the parameter values employed in the numerical illustrations. The parameter values are taken from Brennan and Xia (2002).

7.10.2. Proofs

Derivation of (7.2.9)**,** (7.2.10) **and** (7.2.11)

We start by deriving the analytical solution to the stochastic differential equation (SDE) for the Ornstein-Uhlenbeck process. After applying Itô's Lemma to the function $f(t, r_t) \equiv e^{\kappa t}(r_t - \bar{r})$, we find (where the second equality follows from equation (7.2.1))

$$\begin{aligned} \mathrm{d}f(t, r_t) &= \kappa e^{\kappa t}(r_t - \bar{r})\,\mathrm{d}t + e^{\kappa t}\mathrm{d}r_t \\ &= \kappa e^{\kappa t}(r_t - \bar{r})\,\mathrm{d}t - e^{\kappa t}\kappa(r_t - \bar{r})\,\mathrm{d}t + e^{\kappa t}\sigma_r \mathrm{d}\mathbf{W}_t^r = \sigma_r e^{\kappa t}\mathrm{d}\mathbf{W}_t^r. \end{aligned} \tag{7.10.1}$$

The solution to the SDE (7.10.1) is given by

$$f(t, \mathbf{r}_{t+v}) = f(t, r_t) + \sigma_r \int_t^{t+v} e^{\kappa u}\mathrm{d}\mathbf{W}_u^r. \tag{7.10.2}$$

The real interest rate at time $t + v > t$ is given by (where the first and third equality follow from the definition of $f(t, r_t)$, and the second equality follows from (7.10.2))

$$\begin{aligned} \mathbf{r}_{t+v} &= \bar{r} + e^{-\kappa(t+v)} f(t, \mathbf{r}_{t+v}) \\ &= \bar{r} + e^{-\kappa(t+v)} f(t, r_t) + \sigma_r \int_t^{t+v} e^{-\kappa(t+v-u)}\mathrm{d}\mathbf{W}_u^r \\ &= \bar{r} + e^{-\kappa v}(r_t - \bar{r}) + \sigma_r \int_0^{v} e^{-\kappa(v-u)}\mathrm{d}\mathbf{W}_{t+u}^r \\ &= r_t + \left(1 - e^{-\kappa v}\right)(\bar{r} - r_t) + \sigma_r \int_0^{v} e^{-\kappa(v-u)}\mathrm{d}\mathbf{W}_{t+u}^r. \end{aligned} \tag{7.10.3}$$

In a similar fashion, we find

$$\boldsymbol{\pi}_{t+v} = \pi_t + \left(1 - e^{-\theta v}\right)(\bar{\pi} - \pi_t) + \sigma_\pi \int_0^{v} e^{-\theta(v-u)}\mathrm{d}\mathbf{W}_{t+u}^\pi.$$

The (conditional) expectation of the real interest rate $\mathbb{E}_t[\mathbf{r}_{t+v}]$ and the (conditional) expectation of the expected rate of inflation $\mathbb{E}_t[\boldsymbol{\pi}_{t+v}]$ are given by

$$\mathbb{E}_t[\mathbf{r}_{t+v}] = r_t + \kappa(\bar{r} - r_t) D^{\kappa,v}, \tag{7.10.4}$$

$$\mathbb{E}_t[\boldsymbol{\pi}_{t+v}] = \pi_t + \theta(\bar{\pi} - \pi_t) D^{\theta,v}. \tag{7.10.5}$$

The average real interest rate $\check{\mathbf{r}}_t^h \equiv \frac{1}{h}\int_0^h \mathbf{r}_{t+v}\mathrm{d}v$ and the average expected rate of inflation $\check{\boldsymbol{\pi}}_t^h \equiv \frac{1}{h}\int_0^h \boldsymbol{\pi}_{t+v}\mathrm{d}v$ play a key role in determining the yield to maturity. We find (where the first equality follows from substituting (7.10.3) to eliminate $\mathbf{r}_{t+v}$)

$$\begin{aligned}
\check{\mathbf{r}}_t^h &\equiv \frac{1}{h}\int_0^h \mathbf{r}_{t+v}\mathrm{d}v \\
&= \frac{1}{h}\int_0^h \left(r_t + (\bar{r} - r_t)\left(1 - e^{-\kappa v}\right)\right)\mathrm{d}v + \frac{\sigma_r}{h}\int_0^h \int_0^v e^{-\kappa(v-u)}\mathrm{d}\mathbf{W}_{t+u}^r \mathrm{d}v \\
&= \frac{1}{h}\int_0^h \left(r_t + (\bar{r} - r_t)\left(1 - e^{-\kappa v}\right)\right)\mathrm{d}v + \frac{\sigma_r}{h}\int_0^h \int_v^h e^{-\kappa(h-u)}\mathrm{d}u\mathrm{d}\mathbf{W}_{t+v}^r \\
&= \frac{1}{h}\int_0^h \left(r_t + (\bar{r} - r_t)\kappa D^{\kappa,v}\right)\mathrm{d}v + \frac{\sigma_r}{\kappa h}\int_0^h \left(1 - e^{-\kappa(h-v)}\right)\mathrm{d}\mathbf{W}_{t+v}^r \\
&= \frac{1}{h}\int_0^h \mathbb{E}_t\left[\mathbf{r}_{t+v}\right]\mathrm{d}v + \frac{\sigma_r}{h}\int_0^h D^{\kappa,h-v}\mathrm{d}\mathbf{W}_{t+v}^r.
\end{aligned} \tag{7.10.6}$$

In a similar fashion, we find that the average expected rate of inflation $\check{\boldsymbol{\pi}}_t^h$ is given by

$$\check{\boldsymbol{\pi}}_t^h \equiv \frac{1}{h}\int_0^h \boldsymbol{\pi}_{t+v}\mathrm{d}v = \frac{1}{h}\int_0^h \mathbb{E}_t\left[\boldsymbol{\pi}_{t+v}\right]\mathrm{d}v + \frac{\sigma_\pi}{h}\int_0^h D^{\theta,h-v}\mathrm{d}\mathbf{W}_{t+v}^\pi. \tag{7.10.7}$$

Substituting (7.10.6) and (7.10.7) into (7.2.8) to eliminate $\int_0^h \mathbf{r}_{t+v}\mathrm{d}v$ and $\int_0^h \boldsymbol{\pi}_{t+v}\mathrm{d}v$ yields

$$\begin{aligned}
P_{\alpha,t}^h = \exp&\left\{-\int_0^h \left(\mathbb{E}_t\left[\mathbf{r}_{t+v} + \bar{\alpha}\boldsymbol{\pi}_{t+v}\right] + \frac{1}{2}\phi^\top \rho\phi\right)\mathrm{d}v\right\} \\
&\mathbb{E}_t\left[\exp\left\{\int_0^h \phi_S \mathrm{d}\mathbf{W}_{t+v}^S + \int_0^h \left(\phi_r - \sigma_r D^{\kappa,h-v}\right)\mathrm{d}\mathbf{W}_{t+v}^r\right\}\right. \\
&\left.\exp\left\{\int_0^h \left(\phi_\pi - \bar{\alpha}\sigma_\pi D^{\theta,h-v}\right)\mathrm{d}\mathbf{W}_{t+v}^\pi\right\}\right] \\
&= \exp\left\{-\int_0^h \left(\mathbb{E}_t\left[\mathbf{r}_{t+v} + \bar{\alpha}\boldsymbol{\pi}_{t+v}\right] - \lambda_r\sigma_r D^{\kappa,v} - \bar{\alpha}\lambda_\pi\sigma_\pi D^{\theta,v}\right.\right. \\
&\left.\left. - \frac{1}{2}\left(\sigma_r D^{\kappa,v}\right)^2 - \frac{1}{2}\left(\bar{\alpha}\sigma_\pi D^{\theta,v}\right)^2 - \bar{\alpha}\rho_{r\pi}\sigma_r\sigma_\pi D^{\kappa,v}D^{\theta,v}\right)\mathrm{d}v\right\} \\
&= \exp\left\{-\int_0^h r_{\alpha,t}^v \mathrm{d}v\right\}.
\end{aligned} \tag{7.10.8}$$

Here $\bar{\alpha} \equiv 1 - \alpha$. The instantaneous forward interest rate $r_{\alpha,t}^v$ is defined as follows:

$$\begin{aligned}
r_{\alpha,t}^v \equiv\ & \mathbb{E}_t\left[\mathbf{r}_{t+v} + \bar{\alpha}\boldsymbol{\pi}_{t+v}\right] - \lambda_r\sigma_r D^{\kappa,v} - \bar{\alpha}\lambda_\pi\sigma_\pi D^{\theta,v} \\
& - \frac{1}{2}\left(\sigma_r D^{\kappa,v}\right)^2 - \frac{1}{2}\left(\bar{\alpha}\sigma_\pi D^{\theta,v}\right)^2 - \bar{\alpha}\rho_{r\pi}\sigma_r\sigma_\pi D^{\kappa,v}D^{\theta,v}.
\end{aligned} \tag{7.10.9}$$

The log bond price is given by (this follows from (7.10.4), (7.10.5), (7.10.8) and (7.10.9))

$$\begin{aligned}\log P^h_{\alpha,t} = &-\int_0^h \Big(r_t + \kappa\left(\bar{r} - r_t\right) D^{\kappa,v} + \bar{\alpha}\pi_t \\ &+ \bar{\alpha}\theta\left(\bar{\pi} - \pi_t\right) D^{\theta,v} - \lambda_r\sigma_r D^{\kappa,v} - \bar{\alpha}\lambda_\pi\sigma_\pi D^{\theta,v} \\ &- \frac{1}{2}\left(\sigma_r D^{\kappa,v}\right)^2 - \frac{1}{2}\left(\bar{\alpha}\sigma_\pi D^{\theta,v}\right)^2 - \bar{\alpha}\rho_{r\pi}\sigma_r\sigma_\pi D^{\kappa,v} D^{\theta,v}\Big)\mathrm{d}v.\end{aligned} \tag{7.10.10}$$

Solving the integral (7.10.10) yields[102]

$$\begin{aligned}\log P^h_{\alpha,t} = &-r_t h - \left(\bar{r} - r_t\right)\left(h - D^{\kappa,h}\right) - \bar{\alpha}\pi_t h - \bar{\alpha}\left(\bar{\pi} - \pi_t\right)\left(h - D^{\theta,h}\right) \\ &+ \frac{\lambda_r\sigma_r}{\kappa}\left(h - D^{\kappa,h}\right) + \frac{\bar{\alpha}\lambda_\pi\sigma_\pi}{\theta}\left(h - D^{\theta,h}\right) \\ &+ \frac{1}{2}\left(\frac{\sigma_r}{\kappa}\right)^2\left(h - 2D^{\kappa,h} + \frac{1}{2}D^{\kappa,2h}\right) + \frac{1}{2}\left(\frac{\bar{\alpha}\sigma_\pi}{\theta}\right)^2\left(h - 2D^{\theta,h} + \frac{1}{2}D^{\theta,2h}\right) \\ &+ \frac{\bar{\alpha}\rho_{r\pi}\sigma_r\sigma_\pi}{\kappa\theta}\left(h - D^{\kappa,h} - D^{\theta,h} + D^{\kappa+\theta,h}\right) \\ = &-r_t D^{\kappa,h} - \bar{\alpha}\pi_t D^{\theta,h} - E^h_\alpha,\end{aligned}$$

where the horizon-dependent constant E^h_α is defined as follows:

$$\begin{aligned}E^h_\alpha \equiv &\left(\bar{r} - \frac{\lambda_r\sigma_r}{\kappa} - \frac{1}{2}\left[\frac{\sigma_r}{\kappa}\right]^2\right)\left(h - D^{\kappa,h}\right) + \frac{1}{4\kappa}\left(\sigma_r D^{\kappa,h}\right)^2 \\ &+ \bar{\alpha}\left(\bar{\pi} - \frac{\lambda_\pi\sigma_\pi}{\theta} - \frac{1}{2}\bar{\alpha}\left[\frac{\sigma_\pi}{\theta}\right]^2\right)\left(h - D^{\theta,h}\right) + \frac{1}{4\theta}\left(\bar{\alpha}\sigma_\pi D^{\theta,h}\right)^2 \\ &+ \frac{\bar{\alpha}\rho_{r\pi}\sigma_r\sigma_\pi}{\kappa\theta}\left(h - D^{\kappa,h} - D^{\theta,h} + D^{\kappa+\theta,h}\right).\end{aligned} \tag{7.10.11}$$

In order to calculate how the value of the bond with a fixed maturity $t + h$ develops as time proceeds (i.e., $t + h$ is fixed but t changes), we apply Itô's Lemma to

$$P^h_{\alpha,t} = \exp\left\{-r_t D^{\kappa,h} - \bar{\alpha}\pi_t D^{\theta,h} - E^h_\alpha\right\}.$$

[102] The first equality follows from $\left(D^{\kappa,v}\right)^2 = \left(1 - 2e^{-\kappa v} + e^{-2\kappa v}\right)/\left(\kappa\right)^2$ and the second equality follows from $\left(D^{\kappa,h}\right)^2 = \left(2D^{\kappa,h} - D^{\kappa,2h}\right)/\kappa$.

We find

$$\begin{aligned}\frac{\mathrm{d}\mathbf{P}_{\alpha,t}^{h}}{P_{\alpha,t}^{h}} &= \left(r_{\alpha,t}^{h} - \kappa\left(\bar{r} - r_t\right) D^{\kappa,h} - \bar{\alpha}\theta\left(\bar{\pi} - \pi_t\right) D^{\theta,h} + \frac{1}{2}\left(\sigma_r D^{\kappa,h}\right)^2 + \frac{1}{2}\left(\bar{\alpha}\sigma_\pi D^{\theta,h}\right)^2\right. \\ &\quad \left. + \bar{\alpha}\rho_{r\pi}\sigma_r\sigma_\pi D^{\kappa,h} D^{\theta,h}\right)\mathrm{d}t - \sigma_r D^{\kappa,h}\mathrm{d}\mathbf{W}_t^r - \bar{\alpha}\sigma_\pi D^{\theta,h}\mathrm{d}\mathbf{W}_t^\pi \\ &= \left(r_t + \bar{\alpha}\pi_t - \lambda_r\sigma_r D^{\kappa,h} - \bar{\alpha}\lambda_\pi\sigma_\pi D^{\theta,h}\right)\mathrm{d}t - \sigma_r D^{\kappa,h}\mathrm{d}\mathbf{W}_t^r - \bar{\alpha}\sigma_\pi D^{\theta,h}\mathrm{d}\mathbf{W}_t^\pi.\end{aligned}$$

Derivation of (7.4.2), (7.4.5), (7.4.6), (7.4.9) and (7.4.11)

The market-consistent value of $\mathbf{B}_{y,t+h}$ is given by (where the first equality follows from substituting equation (7.3.2) into (7.4.1) to eliminate $\mathbf{B}_{y,t+h}$)

$$\begin{aligned}V_{y,t}^{h} &= {}_{h}p_{t-y} B_{y,t} \exp\left\{\int_0^h \omega^{*\top}\left(\lambda - \lambda^*\right)\mathrm{d}v\right\} \\ &\quad \mathbb{E}_t\left[\exp\left\{-\int_0^h \left(\mathbf{r}_{t+v} + \boldsymbol{\pi}_{t+v} + \frac{1}{2}\phi^\top\rho\phi\right)\mathrm{d}v + \int_0^h \phi^\top \mathrm{d}\mathbf{W}_{t+v}\right\}\right. \\ &\quad \left.\exp\left\{\left(\beta + \psi\bar{\beta}\right)\int_0^h \boldsymbol{\pi}_{t+v}\mathrm{d}v + \psi\int_0^h \mathbf{r}_{t+v}\mathrm{d}v + g\cdot h + \int_0^h \omega^{*\top}\mathrm{d}\mathbf{W}_{t+v}\right\}\right] \\ &= {}_{h}p_{t-y} B_{y,t} \exp\left\{\int_0^h \omega^{*\top}\left(\lambda - \lambda^*\right)\mathrm{d}v\right\} \\ &\quad \exp\left\{-\bar{\psi}\int_0^h \left(\mathbb{E}_t\left[\mathbf{r}_{t+v}\right] + \bar{\beta}\mathbb{E}_t\left[\boldsymbol{\pi}_{t+v}\right]\right)\mathrm{d}v - \frac{1}{2}\phi^\top\rho\phi\cdot h + g\cdot h\right\} \\ &\quad \mathbb{E}_t\left[\exp\left\{\int_0^h \left(\phi_r - \bar{\psi}\sigma_r D^{\kappa,h-v} + \omega_r^*\right)\mathrm{d}\mathbf{W}_{t+v}^r\right.\right. \\ &\quad \left.\left. + \int_0^h \left(\phi_\pi - \bar{\psi}\bar{\beta}\sigma_\pi D^{\theta,h-v} + \omega_\pi^*\right)\mathrm{d}\mathbf{W}_{t+v}^\pi + \int_0^h \left(\phi_S + \omega_S^*\right)\mathrm{d}\mathbf{W}_{t+v}^S\right\}\right].\end{aligned}$$

Here $\bar{\psi} \equiv 1 - \psi$ and $\bar{\beta} \equiv 1 - \beta$.

Straightforward computations yield

$$
\begin{aligned}
V_{y,t}^{h} = {}& {}_{h}p_{t-y} B_{y,t} \exp\left\{ \int_0^h \omega^{*\top} \left(\lambda - \lambda^*\right) \mathrm{d}v \right\} \\
& \exp\left\{ -\bar{\psi} \int_0^h \left(\mathbb{E}_t\left[\mathbf{r}_{t+v}\right] + \bar{\beta}\mathbb{E}_t\left[\boldsymbol{\pi}_{t+v}\right] - \lambda_r \sigma_r D^{\kappa,v} - \frac{1}{2}\bar{\psi}\left(\sigma_r D^{\kappa,v}\right)^2 \right) \mathrm{d}v \right\} \\
& \exp\left\{ -\bar{\psi} \int_0^h \left(-\bar{\beta}\lambda_\pi \sigma_\pi D^{\theta,v} - \frac{1}{2}\bar{\psi}\left(\bar{\beta}\sigma_\pi D^{\theta,v}\right)^2 \right) \mathrm{d}v + g \cdot h \right\} \\
& \exp\left\{ -\bar{\psi} \int_0^h -\bar{\psi}\bar{\beta}\rho_{r\pi}\sigma_r\sigma_\pi D^{\kappa,v} D^{\theta,v} \mathrm{d}v \right\} \\
& \exp\left\{ -\int_0^h \left(\lambda_r + \bar{\psi}\sigma_r D^{\kappa,v} + \bar{\psi}\bar{\beta}\rho_{r\pi}\sigma_\pi D^{\theta,v} \right) \omega_r^* \mathrm{d}v \right\} \\
& \exp\left\{ -\int_0^h \left(\lambda_\pi + \bar{\psi}\bar{\beta}\sigma_\pi D^{\theta,v} + \bar{\psi}\rho_{r\pi}\sigma_r D^{\kappa,v} \right) \omega_\pi^* \mathrm{d}v \right\} \\
& \exp\left\{ -\int_0^h \left(\lambda_S + \bar{\psi}\rho_{rS}\sigma_r D^{\kappa,v} + \bar{\psi}\bar{\beta}\rho_{\pi S}\sigma_\pi D^{\theta,v} \right) \omega_S^* \mathrm{d}v \right\} \\
& \exp\left\{ \int_0^h \frac{1}{2}\omega^{*\top}\rho\omega^* \mathrm{d}v \right\} = B_{y,t} \exp\left\{ -\int_0^h \delta_{y,t}^v \mathrm{d}v \right\}.
\end{aligned}
$$

Here

$$
\begin{aligned}
\delta_{y,t}^v = {}& \mu_{t-y+v} + \bar{\psi}\left(\mathbb{E}_t\left[\mathbf{r}_{t+v}\right] + \bar{\beta}\mathbb{E}_t\left[\boldsymbol{\pi}_{t+v}\right] - \lambda_r\sigma_r D^{\kappa,v} - \bar{\beta}\lambda_\pi\sigma_\pi D^{\theta,v} \right) \\
& + \omega^{*\top}\lambda^* + \widehat{\xi}_v - g \\
= {}& \bar{\psi} r_{\beta,t}^v + \omega^{*\top}\lambda^* + \xi_v + \mu_{t-y+v} - g
\end{aligned}
$$

where[103]

$$
\begin{aligned}
\widehat{\xi}_v \equiv {}& \bar{\psi}\left[\left(\sigma_r D^{\kappa,v} + \bar{\beta}\rho_{r\pi}\sigma_\pi D^{\theta,v}\right)\omega_r^* + \left(\rho_{r\pi}\sigma_r D^{\kappa,v} + \bar{\beta}\sigma_\pi D^{\theta,v}\right)\omega_\pi^* \right. \\
& \left. + \left(\rho_{rS}\sigma_r D^{\kappa,h} + \bar{\beta}\rho_{\pi S}\sigma_\pi D^{\theta,v}\right)\omega_S^* \right] - \frac{1}{2}\omega^{*\top}\rho\omega^* \\
& - \frac{1}{2}\bar{\psi}^2 \left[\left(\sigma_r D^{\kappa,v}\right)^2 + \left(\bar{\beta}\sigma_\pi D^{\theta,v}\right)^2 + 2\bar{\beta}\rho_{r\pi}\sigma_r\sigma_\pi D^{\kappa,v} D^{\theta,v} \right], \\
\xi_v \equiv {}& \widehat{\xi}_v + \frac{1}{2}\bar{\psi}\left[\left(\sigma_r D^{\kappa,v}\right)^2 + \left(\bar{\beta}\sigma_\pi D^{\theta,v}\right)^2 + 2\bar{\beta}\rho_{r\pi}\sigma_r\sigma_\pi D^{\kappa,v} D^{\theta,v} \right].
\end{aligned}
\tag{7.10.12}
$$

The market-consistent value is given by (where the first and second equality follow from

[103] The term $\widehat{\xi}_v$ arises because we measure security price performance in terms of log (continuously compounded) returns. Indeed, with log returns, the portfolio return is *not* equal to the weighted sum of the individual returns (i.e., log returns do not aggregate across securities). However, log returns do aggregate across time.

equation (7.4.1))

$$\log V_{y,t}^{h} = \log B_{y,t} + \log A_{y,t}^{h} = \log B_{y,t} - \int_{0}^{h} \delta_{y,t}^{v} \mathrm{d}v. \tag{7.10.13}$$

Here

$$\log B_{y,t} = \log B_{y,t_0} + \beta \int_{t_0}^{t} \pi_s \mathrm{d}s + \int_{t_0}^{t} \psi \left(r_s + \bar{\beta}\pi_s \right) \mathrm{d}s + g \cdot (t - t_0) + \int_{t_0}^{t} \omega^{*\top} \mathrm{d}W_s^{*}.$$

Applying Itô's Lemma to equation (7.10.13) yields

$$\begin{aligned}
\frac{\partial \log \mathbf{A}_{y,t}^{h}}{\partial \mathbf{W}_t^{r}} \frac{1}{\sigma_r} &= -\bar{\psi} D^{\kappa,h}, \\
\frac{\partial \log \mathbf{A}_{y,t}^{h}}{\partial \mathbf{W}_t^{\pi}} \frac{1}{\sigma_\pi} &= -\bar{\psi}\bar{\beta} D^{\theta,h}.
\end{aligned}$$

Taking the partial derivative of $\log A_{y,t} = \log \int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t} \exp \left\{ \log A_{y,t}^{h} \right\} \mathrm{d}h$ with respect to $\log A_{y,t}^{h}$ yields $A_{y,t}^{h}/A_{y,t} \equiv \gamma_{y,t}^{h}$. Equations (7.4.9) and (7.4.11) now follow from Itô's Lemma.

Derivation of (7.5.1), (7.5.2) and (7.5.3)

Assets $X_{y,t}$ are subject to the following dynamic equation (this follows from equations (7.2.3) and (7.2.11)):[104]

$$\begin{aligned}
\frac{\mathrm{d}\mathbf{X}_{y,t}}{X_{y,t}} = \Big(r_t + \pi_t + \mu_{t-y} &- \left(\varpi_{y,t}^{1} D^{\kappa,n_1} + \varpi_{y,t}^{2} D^{\kappa,n_2} \right) \lambda_r \sigma_r \\
&- \left(\varpi_{y,t}^{1} D^{\theta,n_1} + \varpi_{y,t}^{2} D^{\theta,n_2} \right) \lambda_\pi \sigma_\pi + \varpi_{y,t}^{3} \lambda_S \sigma_S \Big) \mathrm{d}t \\
&- \left(\varpi_{y,t}^{1} D^{\kappa,n_1} + \varpi_{y,t}^{2} D^{\kappa,n_2} \right) \sigma_r \mathrm{d}\mathbf{W}_t^{r} \\
&- \left(\varpi_{y,t}^{1} D^{\theta,n_1} + \varpi_{y,t}^{2} D^{\theta,n_2} \right) \sigma_\pi \mathrm{d}\mathbf{W}_t^{\pi} + \varpi_{y,t}^{3} \sigma_S \mathrm{d}\mathbf{W}_t^{S}.
\end{aligned}$$

[104]Without loss of generality, we assume that the current time t is smaller than $y + x_r$.

It follows from Itô's Lemma that

$$
\begin{aligned}
\frac{\partial \log \mathbf{X}_{y,t}}{\partial \mathbf{W}_t^r}\frac{1}{\sigma_r} &= -\left(\varpi_{y,t}^1 D^{\kappa,n_1} + \varpi_{y,t}^2 D^{\kappa,n_2}\right), \\
\frac{\partial \log \mathbf{X}_{y,t}}{\partial \mathbf{W}_t^\pi}\frac{1}{\sigma_\pi} &= -\left(\varpi_{y,t}^1 D^{\theta,n_1} + \varpi_{y,t}^2 D^{\theta,n_2}\right), \\
\frac{\partial \log \mathbf{X}_{y,t}}{\partial \mathbf{W}_t^S}\frac{1}{\sigma_S} &= \varpi_{y,t}^3.
\end{aligned}
$$

Applying the principle of liability-driven investment yields equations (7.5.1), (7.5.2) and (7.5.3). The replicating portfolio strategy can be computed explicitly. We find

$$
\varpi_{y,t}^{1*} = k\left(n_1, n_2\right)\left[\left(\frac{\omega_r^*}{\sigma_r} - \frac{\omega_\pi^*}{\sigma_\pi}\frac{D^{\kappa,n_2}}{D^{\theta,n_2}}\right) - \bar{\psi}\left(\widehat{D}_{y,t}^{\kappa} - \bar{\beta}\widehat{D}_{y,t}^{\theta}\frac{D^{\kappa,n_2}}{D^{\theta,n_2}}\right)\right], \tag{7.10.14}
$$

$$
\varpi_{y,t}^{2*} = k\left(n_2, n_1\right)\left[\left(\frac{\omega_r^*}{\sigma_r} - \frac{\omega_\pi^*}{\sigma_\pi}\frac{D^{\kappa,n_1}}{D^{\theta,n_1}}\right) - \bar{\psi}\left(\widehat{D}_{y,t}^{\kappa} - \bar{\beta}\widehat{D}_{y,t}^{\theta}\frac{D^{\kappa,n_1}}{D^{\theta,n_1}}\right)\right], \tag{7.10.15}
$$

$$
\varpi_{y,t}^{3*} = \frac{\omega_S^*}{\sigma_S}. \tag{7.10.16}
$$

Here

$$
k\left(n_1, n_2\right) \equiv \frac{D^{\theta,n_2}}{D^{\kappa,n_2}D^{\theta,n_1} - D^{\kappa,n_1}D^{\theta,n_2}}.
$$

Derivation of (7.6.6) and (7.6.15)

The aggregate annuity factor does not change as a result of changes in the investment strategy. It follows that

$$
\mathrm{d}\log \mathbf{B}_{y,t} = \mathrm{d}\log \mathbf{V}_{y,t} - \mathrm{d}\log \mathbf{A}_{y,t},
$$

where

$$
\begin{aligned}
\mathrm{d}\log \mathbf{V}_{y,t} = &\left(r_t + \pi_t + \mu_{t-y} + \omega_t^{\top}\lambda - \bar{\psi}\sigma_r\lambda_r\widehat{D}_{y,t}^{\kappa} - \bar{\psi}\bar{\beta}\sigma_\pi\lambda_\pi\widehat{D}_{y,t}^{\theta} + \widetilde{\xi}_{y,t}\right)\mathrm{d}t \\
&+ \left(\omega_t^r - \bar{\psi}\sigma_r\widehat{D}_{y,t}^{\kappa}\right)\mathrm{d}\mathbf{W}_t^r + \left(\omega_t^\pi - \bar{\psi}\bar{\beta}\sigma_\pi\widehat{D}_{y,t}^{\theta}\right)\mathrm{d}\mathbf{W}_t^\pi + \omega_t^S\mathrm{d}\mathbf{W}_t^S, \\
\mathrm{d}\log \mathbf{A}_{y,t} = &\left(\bar{\psi}\left(r_t + \bar{\beta}\pi_t\right) - g + \mu_{t-y} + \omega^{*\top}\lambda^* - \bar{\psi}\sigma_r\lambda_r\widehat{D}_{y,t}^{\kappa} - \bar{\psi}\bar{\beta}\sigma_\pi\lambda_\pi\widehat{D}_{y,t}^{\theta} + \widehat{\xi}_{y,t}\right)\mathrm{d}t \\
&- \bar{\psi}\sigma_r\widehat{D}_{y,t}^{\kappa}\mathrm{d}\mathbf{W}_t^r - \bar{\psi}\bar{\beta}\sigma_\pi\widehat{D}_{y,t}^{\theta}\mathrm{d}\mathbf{W}_t^\pi.
\end{aligned}
$$

Here,

$$
\begin{aligned}
\omega_t^r &\equiv \varphi_t^r + \sigma_r \left(\bar{\psi}\widehat{D}_t^\kappa - H_t^r\right), \\
\omega_t^\pi &\equiv \varphi_t^\pi + \sigma_\pi \left(\bar{\psi}\bar{\beta}\widehat{D}_t^\theta - H_t^\pi\right), \\
\omega_t^S &\equiv \varphi_t^S, \\
\widetilde{\xi}_{y,t} &\equiv \bar{\psi}\left[\left(\sigma_r \widehat{D}_{y,t}^\kappa + \bar{\beta}\rho_{r\pi}\sigma_\pi \widehat{D}_{y,t}^\theta\right)\omega_t^r + \left(\rho_{r\pi}\sigma_r \widehat{D}_{y,t}^\kappa + \bar{\beta}\sigma_\pi \widehat{D}_{y,t}^\theta\right)\omega_t^\pi \right. \\
&\quad \left. + \left(\rho_{rS}\sigma_r \widehat{D}_{y,t}^\kappa + \bar{\beta}\rho_{\pi S}\sigma_\pi \widehat{D}_{y,t}^\theta\right)\omega_t^S\right] - \frac{1}{2}\omega_t^\top \rho \omega_t \\
&\quad - \frac{1}{2}\bar{\psi}^2 \left[\left(\sigma_r \widehat{D}_{y,t}^\kappa\right)^2 + \left(\bar{\beta}\sigma_\pi \widehat{D}_{y,t}^\theta\right)^2 + 2\bar{\beta}\rho_{r\pi}\sigma_r\sigma_\pi \widehat{D}_{y,t}^\kappa \widehat{D}_{y,t}^\theta\right], \\
\widehat{\xi}_{y,t} &\equiv \bar{\psi}\left[\left(\sigma_r \widehat{D}_{y,t}^\kappa + \bar{\beta}\rho_{r\pi}\sigma_\pi \widehat{D}_{y,t}^\theta\right)\omega_r^* + \left(\rho_{r\pi}\sigma_r \widehat{D}_{y,t}^\kappa + \bar{\beta}\sigma_\pi \widehat{D}_{y,t}^\theta\right)\omega_\pi^* \right. \\
&\quad \left. + \left(\rho_{rS}\sigma_r \widehat{D}_{y,t}^\kappa + \bar{\beta}\rho_{\pi S}\sigma_\pi \widehat{D}_{y,t}^\theta\right)\omega_S^*\right] - \frac{1}{2}\omega^{*\top} \rho \omega^* \\
&\quad - \frac{1}{2}\bar{\psi}^2 \left[\left(\sigma_r \widehat{D}_{y,t}^\kappa\right)^2 + \left(\bar{\beta}\sigma_\pi \widehat{D}_{y,t}^\theta\right)^2 + 2\bar{\beta}\rho_{r\pi}\sigma_r\sigma_\pi \widehat{D}_{y,t}^\kappa \widehat{D}_{y,t}^\theta\right],
\end{aligned}
$$

with $\omega_t \equiv \left(\omega_t^r, \omega_t^\pi, \omega_t^S\right)$. Hence,

$$
\begin{aligned}
\mathrm{d}\log \mathbf{B}_{y,t} &= \left(\beta\pi_t + \psi\left(r_t + \bar{\beta}\pi_t\right) + g + \omega_t^\top \lambda - \omega^{*\top}\lambda^* + \widetilde{\xi}_{y,t} - \widehat{\xi}_{y,t}\right)\mathrm{d}t \\
&\quad + \omega_t^S \mathrm{d}\mathbf{W}_t^S + \omega_t^r \mathrm{d}\mathbf{W}_t^r + \omega_t^\pi \mathrm{d}\mathbf{W}_t^\pi \\
&= \left(\beta\pi_t + \psi\left(r_t + \bar{\beta}\pi_t\right) + g + \widetilde{\xi}_{y,t} - \widehat{\xi}_{y,t}\right)\mathrm{d}t + \mathrm{d}\log \mathbf{M}_t.
\end{aligned}
$$

Equation (7.6.15) can be derived in a similar fashion as above. We now have

$$
\begin{aligned}
\omega_{y,t}^r &\equiv \varphi_{y,t}^r + \sigma_r \left(\bar{\psi}\widehat{D}_{y,t}^\kappa - H_{y,t}^r\right), \\
\omega_{y,t}^\pi &\equiv \varphi_{y,t}^\pi + \sigma_\pi \left(\bar{\psi}\bar{\beta}\widehat{D}_{y,t}^\theta - H_{y,t}^\pi\right), \\
\omega_{y,t}^S &\equiv \varphi_{y,t}^S, \\
\widetilde{\xi}_{y,t} &\equiv \bar{\psi}\left[\left(\sigma_r \widehat{D}_{y,t}^\kappa + \bar{\beta}\rho_{r\pi}\sigma_\pi \widehat{D}_{y,t}^\theta\right)\omega_{y,t}^r + \left(\rho_{r\pi}\sigma_r \widehat{D}_{y,t}^\kappa + \bar{\beta}\sigma_\pi \widehat{D}_{y,t}^\theta\right)\omega_{y,t}^\pi \right. \\
&\quad \left. + \left(\rho_{rS}\sigma_r \widehat{D}_{y,t}^\kappa + \bar{\beta}\rho_{\pi S}\sigma_\pi \widehat{D}_{y,t}^\theta\right)\omega_{y,t}^S\right] - \frac{1}{2}\omega_{y,t}^\top \rho \omega_{y,t} \\
&\quad - \frac{1}{2}\bar{\psi}^2 \left[\left(\sigma_r \widehat{D}_{y,t}^\kappa\right)^2 + \left(\bar{\beta}\sigma_\pi \widehat{D}_{y,t}^\theta\right)^2 + 2\bar{\beta}\rho_{r\pi}\sigma_r\sigma_\pi \widehat{D}_{y,t}^\kappa \widehat{D}_{y,t}^\theta\right],
\end{aligned}
$$

with $\omega_{y,t} \equiv \left(\omega_{y,t}^r, \omega_{y,t}^\pi, \omega_{y,t}^S\right)$.

Derivation of (7.7.7), (7.7.10), (7.7.11), (7.7.12), (7.7.13), (7.7.14) **and** (7.7.16)

The market-consistent value at time $t+h$ is given by

$$\begin{aligned}\mathbf{V}_{y,t+h} &= \mathbf{B}_{y,t+h}\\ &= B_{y,t}\left(\frac{\mathbf{\Pi}_{t+h}}{\Pi_t}\right)^{\beta}\exp\left\{\psi\int_0^h\left(\mathbf{r}_{t+v}+\bar{\beta}\boldsymbol{\pi}_{t+v}\right)\mathrm{d}v+g\cdot h\right\}\\ &\exp\left\{\int_0^h\omega^{*\top}\left(\lambda_t-\lambda_t^*\right)\mathrm{d}v\right\}\exp\left\{\int_0^h\omega^{*\top}\mathrm{d}\mathbf{W}_{t+h-v}\right\}.\end{aligned} \tag{7.10.17}$$

Denote by $X_{y,t}^h$ wealth that finances $\mathbf{B}_{y,t+h}$. We have

$$\begin{aligned}\frac{\mathrm{d}\mathbf{X}_{y,t}^h}{X_{y,t}^h} &= \Big(r_t+\pi_t+\mu_{t-y}-\left(\varpi_{h,1}D^{\kappa,n_1}+\varpi_{h,2}D^{\kappa,n_2}\right)\lambda_r\sigma_r\\ &-\left(\varpi_{h,1}D^{\theta,n_1}+\varpi_{h,2}D^{\theta,n_2}\right)\lambda_\pi\sigma_\pi+\varpi_{h,3}e_t\Big)\,\mathrm{d}t\\ &-\left(\varpi_{h,1}D^{\kappa,n_1}+\varpi_{h,2}D^{\kappa,n_2}\right)\sigma_r\mathrm{d}\mathbf{W}_t^r\\ &-\left(\varpi_{h,1}D^{\theta,n_1}+\varpi_{h,2}D^{\theta,n_2}\right)\sigma_\pi\mathrm{d}\mathbf{W}_t^\pi+\varpi_{h,3}\sigma_S\mathrm{d}\mathbf{W}_t^S.\end{aligned}$$

Here $\varpi_{h,i}$ denotes the fraction of $X_{y,t}^h$ invested in a nominal bond with time to maturity n_i $(i=1,2)$, and $\varpi_{h,3}$ represents the fraction of $X_{y,t}^h$ invested in the risky stock.

Straightforward application of Itô's Lemma yields

$$\begin{aligned}\frac{\mathbf{X}_{y,t+h}}{X_{y,t}^h} &= \left(\frac{\mathbf{\Pi}_{t+h}}{\Pi_t}\right)\exp\left\{\int_0^h\left(\mathbf{r}_{t+v}+\mu_{t-y+v}-\left(\varpi_{v,1}D^{\kappa,n_1}+\varpi_{v,2}D^{\kappa,n_2}\right)\lambda_r\sigma_r\right)\mathrm{d}v\right\}\\ &\exp\left\{\int_0^h\left(-\left(\varpi_{v,1}D^{\theta,n_1}+\varpi_{v,2}D^{\theta,n_2}\right)\lambda_\pi\sigma_\pi+\varpi_{v,3}\mathbf{e}_{t+h-v}+\widehat{\xi}_v\right)\mathrm{d}v\right\}\\ &\exp\left\{-\int_0^h\left(\varpi_{v,1}D^{\kappa,n_1}+\varpi_{v,2}D^{\kappa,n_2}\right)\sigma_r\mathrm{d}\mathbf{W}_{t+h-v}^r\right\}\\ &\exp\left\{-\int_0^h\left(\varpi_{v,1}D^{\theta,n_1}+\varpi_{v,2}D^{\theta,n_2}\right)\sigma_\pi\mathrm{d}\mathbf{W}_{t+h-v}^\pi+\int_0^h\varpi_{v,3}\sigma_S\mathrm{d}\mathbf{W}_{t+h-v}^S\right\},\end{aligned} \tag{7.10.18}$$

where

$$
\begin{aligned}
\widehat{\xi}_v \equiv & -\frac{1}{2}\left(\varpi_{v,1}D^{\kappa,n_1}+\varpi_{v,2}D^{\kappa,n_2}\right)^2\sigma_r^2-\frac{1}{2}\left(\varpi_{v,1}D^{\theta,n_1}+\varpi_{v,2}D^{\theta,n_2}\right)^2\sigma_\pi^2 \\
& -\frac{1}{2}\left(\varpi_{v,3}\sigma_S\right)^2+\rho_{rS}\sigma_r\sigma_S\left(\varpi_{v,1}D^{\kappa,n_1}+\varpi_{v,2}D^{\kappa,n_2}\right)\varpi_{v,3} \\
& -\rho_{r\pi}\sigma_r\sigma_\pi\left(\varpi_{v,1}D^{\kappa,n_1}+\varpi_{v,2}D^{\kappa,n_2}\right)\left(\varpi_{v,1}D^{\theta,n_1}+\varpi_{v,2}D^{\theta,n_2}\right) \\
& +\rho_{\pi S}\sigma_\pi\sigma_S\left(\varpi_{v,1}D^{\theta,n_1}+\varpi_{v,2}D^{\theta,n_2}\right)\varpi_{v,3}.
\end{aligned}
$$

We note that

$$
\begin{aligned}
\mathbf{e}_{t+v} = & \, e_t - a\kappa\left(\bar{r}-r_t\right)D^{\kappa,v}-b\theta\left(\bar{\pi}-\pi_t\right)D^{\theta,v} \\
& - a\sigma_r\int_0^v e^{-\kappa(v-u)}\mathrm{d}\mathbf{W}^r_{t+u}-b\sigma_\pi\int_0^v e^{-\theta(v-u)}\mathrm{d}\mathbf{W}^\pi_{t+u}.
\end{aligned}
$$

Simple algebra yields

$$
\begin{aligned}
\frac{1}{h}\int_0^h \mathbf{e}_{t+v}\mathrm{d}v = & \frac{1}{h}\int_0^h \mathbb{E}_t\left[\mathbf{e}_{t+v}\right]\mathrm{d}v \\
& -\frac{a\sigma_r}{h}\int_0^h D^{\kappa,h-v}\mathrm{d}\mathbf{W}^r_{t+v}-\frac{b\sigma_\pi}{h}\int_0^h D^{\theta,h-v}\mathrm{d}\mathbf{W}^\pi_{t+v},
\end{aligned}
\tag{7.10.19}
$$

where $\mathbb{E}_t\left[\mathbf{e}_{t+v}\right]=e_t-a\kappa\left(\bar{r}-r_t\right)D^{\kappa,v}-b\theta\left(\bar{\pi}-\pi_t\right)D^{\theta,v}$.

Since ω^*_S is constant by assumption, it follows that $\varpi^*_{v,3}=\omega^*_S/\sigma_S$ (this follows from comparing equation (7.10.17) with equation (7.10.18)). Substituting equations (7.10.6), (7.10.7), (7.10.19) and $\varpi_{v,3}^{*}=\omega_S^{*}/\sigma_S$ into equation (7.10.18) yields

$$\begin{aligned}\frac{\mathbf{X}_{y,t+h}}{X_{y,t}^{h}}=&\left(\frac{\mathbf{\Pi}_{t+h}}{\Pi_t}\right)^{\beta}\exp\left\{\psi\int_0^h\left(\mathbf{r}_{t+v}+\bar{\beta}\boldsymbol{\pi}_{t+v}\right)\mathrm{d}v+g\cdot h\right\}\\&\exp\left\{-\int_0^h\left(\frac{\omega_S^{*}a^{*}}{\sigma_S}D^{\kappa,v}-\bar{\psi}D^{\kappa,v}+\varpi_{v,1}D^{\kappa,n_1}+\varpi_{v,2}D^{\kappa,n_2}\right)\sigma_r\mathrm{d}\mathbf{W}_{t+h-v}^{r}\right\}\\&\exp\left\{-\int_0^h\left(\frac{\omega_S^{*}b^{*}}{\sigma_S}D^{\theta,v}-\bar{\psi}\bar{\beta}D^{\theta,v}+\varpi_{v,1}D^{\theta,n_1}+\varpi_{v,2}D^{\theta,n_2}\right)\sigma_\pi\mathrm{d}\mathbf{W}_{t+h-v}^{\pi}\right\}\\&\exp\left\{\int_0^h\omega_S^{*}\mathrm{d}\mathbf{W}_{t+h-v}^{S}\right\}\exp\left\{\bar{\psi}\int_0^h\left(\mathbb{E}_t\left[\mathbf{r}_{t+v}\right]+\bar{\beta}\mathbb{E}_t\left[\boldsymbol{\pi}_{t+v}\right]\right)\mathrm{d}v-g\cdot h\right\}\\&\exp\left\{-\int_0^h\left(\varpi_{v,1}D^{\kappa,n_1}+\varpi_{v,2}D^{\kappa,n_2}\right)\lambda_r\sigma_r\mathrm{d}v\right\}\\&\exp\left\{-\int_0^h\left(\varpi_{v,1}D^{\theta,n_1}+\varpi_{v,2}D^{\theta,n_2}\right)\lambda_\pi\sigma_\pi\mathrm{d}v\right\}\exp\left\{\int_0^h\frac{\omega_S^{*}}{\sigma_S}\mathbb{E}_t\left[e_{t+h-v}^{*}\right]\mathrm{d}v\right\}\\&\exp\left\{\int_0^h\left(\frac{\omega_S^{*}}{\sigma_S}\left(\mathbf{e}_{t+h-v}-\mathbf{e}_{t+h-v}^{*}\right)+\widehat{\xi}_v+\mu_{t-y+v}\right)\mathrm{d}v\right\}.\end{aligned}\tag{7.10.20}$$

It follows from comparing equation (7.10.17) with equation (7.10.20) that

$$\begin{aligned}\omega_r^{*}&=-\left(\frac{\omega_S^{*}a^{*}}{\sigma_S}D^{\kappa,h}-\bar{\psi}D^{\kappa,h}+\varpi_{h,1}^{*}D^{\kappa,n_1}+\varpi_{h,2}^{*}D^{\kappa,n_2}\right)\sigma_r,\\\omega_\pi^{*}&=-\left(\frac{\omega_S^{*}b^{*}}{\sigma_S}D^{\theta,h}-\bar{\psi}\bar{\beta}D^{\theta,h}+\varpi_{h,1}^{*}D^{\theta,n_1}+\varpi_{h,2}^{*}D^{\theta,n_2}\right)\sigma_\pi.\end{aligned}$$

Hence,

$$\begin{aligned}\varpi_{h,1}^{*}=&\,k\left(n_1,n_2\right)\left(\frac{\omega_r^{*}}{\sigma_r}-\frac{\omega_\pi^{*}}{\sigma_\pi}\frac{D^{\kappa,n_2}}{D^{\theta,n_2}}\right)+\\&k\left(n_1,n_2\right)\left[-\bar{\psi}\left(D^{\kappa,h}-\bar{\beta}D^{\theta,h}\frac{D^{\kappa,n_2}}{D^{\theta,n_2}}\right)+\frac{\omega_S^{*}a^{*}}{\sigma_S}\left(D^{\kappa,h}-\frac{b^{*}}{a^{*}}D^{\theta,h}\frac{D^{\kappa,n_2}}{D^{\theta,n_2}}\right)\right],\\\varpi_{h,2}^{*}=&\,k\left(n_2,n_1\right)\left(\frac{\omega_r^{*}}{\sigma_r}-\frac{\omega_\pi^{*}}{\sigma_\pi}\frac{D^{\kappa,n_1}}{D^{\theta,n_1}}\right)+\\&k\left(n_2,n_1\right)\left[-\bar{\psi}\left(D^{\kappa,h}-\bar{\beta}D^{\theta,h}\frac{D^{\kappa,n_1}}{D^{\theta,n_1}}\right)+\frac{\omega_S^{*}a^{*}}{\sigma_S}\left(D^{\kappa,h}-\frac{b^{*}}{a^{*}}D^{\theta,h}\frac{D^{\kappa,n_1}}{D^{\theta,n_1}}\right)\right].\end{aligned}$$

The aggregate portfolio weight $\varpi_{y,t}^{i*}$ is given by $\varpi_{y,t}^{i*}=\int_{\max\{x_r+y-t,0\}}^{x_{\max}+y-t}\gamma_{y,t}^{h}\varpi_{h,i}^{*}\mathrm{d}h$ $(i=1,2,3)$. Equations (7.7.12), (7.7.13) and (7.7.14) now follow. The market value $V_{y,t}^{h}=$

$X_{y,t}^h$ is given by (this follows from comparing equation (7.10.17) with equation (7.10.20))

$$
\begin{aligned}
V_{y,t}^h = B_{y,t}^h \exp&\left\{\bar{\psi}\int_0^h \left(\mathbb{E}_t\left[\mathbf{r}_{t+v}\right] + \bar{\beta}\mathbb{E}_t\left[\boldsymbol{\pi}_{t+v}\right]\right)\mathrm{d}v - g\cdot h\right\} \\
&\exp\left\{-\int_0^h \left(\mu_{t-y+v} - \left(\varpi_{v,1}^* D^{\kappa,n_1} + \varpi_{v,2}^* D^{\kappa,n_2}\right)\lambda_r\sigma_r\right)\mathrm{d}v\right\} \\
&\exp\left\{-\int_0^h \left(\omega_r^*\left(\lambda_r^* - \lambda_r\right) - \left(\varpi_{v,1}^* D^{\theta,n_1} + \varpi_{v,2}^* D^{\theta,n_2}\right)\lambda_\pi\sigma_\pi\right)\mathrm{d}v\right\} \\
&\exp\left\{-\int_0^h \left(\omega_\pi^*\left(\lambda_\pi^* - \lambda_\pi\right) + \omega_S^*\mathbb{E}_t\left[\boldsymbol{\lambda}_{t+h-v}^{S*}\right] + \widehat{\xi}_v\right)\mathrm{d}v\right\} \\
&= B_{y,t}^h A_{y,t}^h = B_{y,t}^h \exp\left\{-\int_0^h \delta_{y,t}^{v,h}\mathrm{d}v\right\}.
\end{aligned}
$$

Here,

$$
\delta_{y,t}^{v,h} = \bar{\psi} r_{\beta,t}^v + \lambda_r\sigma_r\frac{\omega_S^* a^*}{\sigma_S}D^{\kappa,v} + \lambda_\pi\sigma_\pi\frac{\omega_S^* b^*}{\sigma_S}D^{\theta,v} + \mu_{t-y+v} - g + {\omega^*}^\top\mathbb{E}_t\left[\boldsymbol{\lambda}_{t+h-v}^*\right] + \xi_v
$$

with

$$
\xi_v = \widehat{\xi}_v + \frac{1}{2}\bar{\psi}\left[\left(D^{\kappa,v}\sigma_r\right)^2 + \left(\bar{\beta}D^{\theta,v}\sigma_\pi\right)^2 + 2\bar{\beta}\rho_{r\pi}\sigma_r\sigma_\pi D^{\kappa,v}D^{\theta,v}\right].
$$

Equations (7.7.10) and (7.7.11) follow from Itô's Lemma. The bonus rule (7.7.16) is derived in a similar fashion as (7.6.6).

Bibliography

Abdellaoui, M. 2000. Parameter-Free Elicitation of Utility and Probability Weighting Functions. *Management Science* 46:1497–1512.

Abdellaoui, M., H. Bleichrodt, and O. L'Haridon. 2008. A Tractable Method to Measure Utility and Loss Aversion under Prospect Theory. *Journal of Risk and Uncertainty* 36:245–266.

Abdellaoui, M., H. Bleichrodt, and C. Paraschiv. 2007. Loss Aversion under Prospect Theory: A Parameter-Free Measurement. *Management Science* 53:1659–1674.

Abdellaoui, M., F. Vossmann, and M. Weber. 2005. Choice-Based Elicitation and Decomposition of Decision Weights for Gains and Losses under Uncertainty. *Management Science* 51:1384–1399.

Abel, A. B. 1990. Asset Prices under Habit Formation and Catching up with the Joneses. *The American Economic Review* 80:38–42.

Ang, A., G. Bekaert, and M. Wei. 2008. The Term Structure of Real Rates and Expected Inflation. *The Journal of Finance* 63:797–849.

Antolín, P., S. Payet, E. R. Whitehouse, and J. Yermo. 2011. The Role of Guarantees in Defined Contribution Pensions. OECD Working Papers on Finance, Insurance and Private Pensions.

Back, K. E. 2010. *Asset Pricing and Portfolio Choice Theory.* Oxford University Press.

Barberis, N., and M. Huang. 2008. Stocks as Lotteries: The Implications of Probability Weighting for Security Prices. *American Economic Review* 98:2066–2100.

Barberis, N., M. Huang, and T. Santos. 2001. Prospect Theory and Asset Prices. *The Quarterly Journal of Economics* 116:1–53.

Barberis, N., and W. Xiong. 2009. What Drives the Disposition Effect? An Analysis of a Long-Standing Preference-Based Explanation. *The Journal of Finance* 64:751–784.

Basak, S., and A. Shapiro. 2001. Value-at-Risk Based Risk Management: Optimal Policies and Asset Prices. *The Review of Financial Studies* 14:371–405.

Bell, D. E. 1982. Regret in Decision Making under Uncertainty. *Operations Research* 30:961–981.

Benartzi, S., and R. Thaler. 1995. Myopic Loss Aversion and the Equity Premium Puzzle. *The Quarterly Journal of Economics* 110:73–92.

Benzoni, L., P. Collin-Dufresne, and R. S. Goldstein. 2007. Portfolio Choice over the Life-Cycle when the Stock and Labor Markets Are Cointegrated. *The Journal of Finance* 62:2123–2167.

Berkelaar, A. B., R. Kouwenberg, and T. Post. 2004. Optimal Portfolio Choice under Loss Aversion. *The Review of Economics and Statistics* 86:973–987.

Bleichrodt, H., and J. L. Pinto. 2000. A Parameter-Free Elicitation of the Probability Weighting Function in Medical Decision Analysis. *Management Science* 46:1485–1496.

Bodie, Z., J. B. Detemple, S. Otruba, and S. Walter. 2004. Optimal Consumption-Portfolio Choices and Retirement Planning. *Journal of Economic Dynamics and Control* 28:1115–1148.

Bodie, Z., R. C. Merton, and W. F. Samuelson. 1992. Labor Supply Flexibility and Portfolio Choice in a Life Cycle Model. *Journal of Economic Dynamics and Control* 16:427–449.

Bodie, Z., and R. Taqqu. 2011. *Risk Less and Prosper: Your Guide to Safer Investing.* Wiley.

Booij, A. S., and G. van de Kuilen. 2009. A Parameter-Free Analysis of the Utility of Money for the General Population under Prospect Theory. *Journal of Economic Psychology* 30:651–666.

Bovenberg, A. L., and Th. E. Nijman. 2015. Personal Pensions with Risk Sharing: Affordable, Adequate and Stable Private Pensions in Europe. Netspar Discussion Paper.

Brennan, M. J., and Y. Xia. 2002. Dynamic Asset Allocation under Inflation. *Journal of Finance* 57:1201–1238.

Brown, J. R., J. R. Kling, S. Mullainathan, and M. V. Wrobel. 2008. Why Don't People Insure Late-Life Consumption? A Framing Explanation of the Under-Annuitization Puzzle. *American Economic Review* 98:304–309.

Brown, J. R., J. R. Kling, S. Mullainathan, and M. V. Wrobel. 2013. Framing Lifetime Income. *The Journal of Retirement* 1:27–37.

Buraschi, A., P. Porchia, and F. Trojani. 2010. Correlation Risk and Optimal Portfolio Choice. *The Journal of Finance* 65:393–420.

Campbell, J. Y., and J. Cochrane. 1999. By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior. *Journal of Political Economy* 107:205–251.

Campbell, J. Y., and N. G. Mankiw. 1991. The Response of Consumption to Income: A Cross-Country Investigation. *European Economic Review* 35:723–767.

Chacko, G., and L. M. Viceira. 2005. Dynamic Consumption and Portfolio Choice with Stochastic Volatility in Incomplete Markets. *The Review of Financial Studies* 18:1369–1402.

Chai, J., W. J. Horneff, R. H. Maurer, and O. S. Mitchell. 2011. Optimal Portfolio Choice over the Life Cycle with Flexible Work, Endogenous Retirement, and Lifetime Payouts. *Review of Finance* 15:875–907.

Chateauneuf, A., and M. Cohen. 1994. Risk Seeking with Diminishing Marginal Utility in a Non-Expected Utility Model. *Journal of Risk and Uncertainty* 9:77–91.

Cocco, J. F., F. J. Gomes, and P. J. Maenhout. 2005. Consumption and Portfolio Choice over the Life Cycle. *The Review of Financial Studies* 18:491–533.

Cochrane, J. H. 2001. *Asset Pricing.* Princeton University Press.

Constantinides, G. M. 1990. Habit Formation: A Resolution of the Equity Premium Puzzle. *Journal of Political Economy* 98:519–543.

Corrado, L., and S. Holly. 2011. Multiplicative Habit Formation and Consumption: A Note. *Economics Letters* 113:116–119.

Cox, J. C., and C. Huang. 1989. Optimal Consumption and Portfolio Policies when Asset Prices Follow a Diffusion Process. *Journal of Economic Theory* 49:33–83.

Cox, J. C., and C. Huang. 1991. A Variational Problem Arising in Financial Economics. *Journal of Mathematical Economics* 20:465–487.

Deaton, A. 1992. *Understanding Consumption.* Oxford University Press.

Deelstra, G., M. Grasselli, and P. Koehl. 2003. Optimal Investment Strategies in the Presence of a Minimum Guarantee. *Insurance: Mathematics and Economics* 33:189–207.

Detemple, J. B., and I. Karatzas. 2003. Non-Addictive Habits: Optimal Consumption-Portfolio Policies. *Journal of Economic Theory* 113:265–285.

Detemple, J. B., and F. Zapatero. 1991. Asset Prices in an Exchange Economy with Habit Formation. *Econometrica* 59:1633–1657.

Detemple, J. B., and F. Zapatero. 1992. Optimal Consumption-Portfolio Policies With Habit Formation. *Mathematical Finance* 2:251–274.

Duffie, D., and L. G. Epstein. 1992. Stochastic Differential Utility. *Econometrica* 60:353–394.

Dus, I., R. Maurer, and O. S. Mitchell. 2005. Betting on Death and Capital Markets in Retirement: A Shortfall Risk Analysis of Life Annuities. *Financial Services Review* 14:169–196.

Epstein, L. G., and S. E. Zin. 1989. Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica* 57:937–969.

Etchart-Vincent, N. 2004. Is Probability Weighting Sensitive to the Magnitude of Consequences? An Empirical Investigation on Losses. *Journal of Risk and Uncertainty* 28:217–235.

Fama, E. F., and K. R. French. 1989. Business Conditions and Expected Returns on Stocks and Bonds. *Journal of Financial Economics* 25:23–49.

Ferson, W. E., and C. R. Harvey. 1991. The Variation of Economic Risk Premiums. *Journal of Political Economy* 99:385–415.

Fuhrer, J. C. 2000. Habit Formation in Consumption and Its Implications for Monetary-Policy Models. *American Economic Review* 90:367–390.

Gomes, F., and A. Michaelides. 2003. Portfolio Choice with Internal Habit Formation: A Life-Cycle Model with Uninsurable Labor Income Risk. *Review of Economic Dynamics* 6:729–766.

Gomes, F., and A. Michaelides. 2008. Asset Pricing with Limited Risk Sharing and Heterogeneous Agents. *The Review of Financial Studies* 21:415–448.

Gomes, F. J. 2005. Portfolio Choice and Trading Volume with Loss Averse Investors. *The Journal of Business* 78:675–706.

Gomes, F. J., L. J. Kotlikoff, and L. M. Viceira. 2008. Optimal Life-Cycle Investing with Flexible Labor Supply: A Welfare Analysis of Life-Cycle Funds. *American Economic Review* 98:297–303.

Guasoni, P., G. Huberman, and D. Ren. 2014. Shortfall Aversion. Working Paper.

Guillén, M., P. L. Jørgensen, and J. P. Nielsen. 2006. Return Smoothing Mechanisms in Life and Pension Insurance: Path-Dependent Contingent Claims. *Insurance: Mathematics and Economics* 38:229–252.

Harvey, C. R. 1989. Time-Varying Conditional Covariances in Tests of Asset Pricing Models. *Journal of Financial Economics* 24:289–317.

He, X. D., and X. Y. Zhou. 2011. Portfolio Choice via Quantiles. *Mathematical Finance* 21:203–231.

He, X. D., and X. Y. Zhou. 2014. Hope, Fear and Aspirations. To Appear in Mathematical Finance.

Holzmann, R., E. Palmer, and D. A. Robalino. 2012. *Nonfinancial Defined Contribution Pension Schemes in a Changing Pension World*. World Bank Publications.

Horneff, W. J., R. H. Maurer, O. S. Mitchell, and I. Dus. 2008. Following the Rules: Integrating Asset Allocation and Annuitization in Retirement Portfolios. *Insurance: Mathematics and Economics* 42:396–408.

Horneff, W. J., R. H. Maurer, O. S. Mitchell, and M. Z. Stamos. 2009. Asset Allocation and Location over the Life Cycle with Investment-Linked Survival-Contingent Payouts. *Journal of Banking and Finance* 33:1688–1699.

Horneff, W. J., R. H. Maurer, O. S. Mitchell, and M. Z. Stamos. 2010. Variable Payout Annuities and Dynamic Portfolio Choice in Retirement. *Journal of Pension Economics and Finance* 9:163–183.

Investment Company Institute. 2014. Investment Company Fact Book. A Review of Trends and Activities in the U.S. Investment Company Industry.

Jin, H., and X. Y. Zhou. 2008. Behavioral Portfolio Selection in Continuous Time. *Mathematical Finance* 18:385–426.

Kahneman, D., and A. Tversky. 1979. Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 47:263–292.

Karatzas, I., J. P. Lehoczky, and S. E. Shreve. 1987. Optimal Consumption and Portfolio Decisions for a “Small Investor” on a Finite Horizon. *SIAM Journal of Control and Optimization* 25:1557–1586.

Karatzas, I., and S. E. Shreve. 1998. *Methods of Mathematical Finance*, vol. 39. Springer.

Köbberling, V., and P. P. Wakker. 2005. An Index of Loss Aversion. *Journal of Economic Theory* 122:119–131.

Kőszegi, B., and M. Rabin. 2006. A Model of Reference-Dependent Preferences. *The Quarterly Journal of Economics* 121:1133–1165.

Kőszegi, B., and M. Rabin. 2007. Reference-Dependent Risk Attitudes. *The American Economic Review* 97:1047–1073.

Kőszegi, B., and M. Rabin. 2009. Reference-Dependent Consumption Plans. *American Economic Review* 99:909–936.

Kozicki, and Tinsley. 2002. Dynamic Specifications in Optimizing Trend-Deviation Macro Models. *Journal of Economic Dynamics and Control* 26:1585–1611.

Kreps, D. M., and E. L. Porteus. 1978. Temporal Resolution of Uncertainty and Dynamic Choice Theory. *Econometrica* 46:185–200.

Laeven, R. J. A., and M. A. Stadje. 2014. Robust Portfolio Choice and Indifference Valuation. *Mathematics of Operations Research* 39:1109–1141.

Laughhunn, D. J., J. W. Payne, and R. Crum. 1980. Managerial Risk Preferences for Below-Target Returns. *Management Science* 26:1238–1249.

Lee, R., and L. Carter. 1992. Modeling and Forecasting U.S. Mortality. *Journal of the American Statistical Association* 87:659–671.

Lettau, M., and S. C. Ludvigson. 2009. Measuring and Modeling Variation in the Risk-Return Trade-off. In Y. Aït-Sahalia and L. P. Hansen (eds.), *Handbook of Financial Econometrics: Tools and Techniques*, vol. 1, chap. 11, pp. 617–690. North Holland.

Li, Y. 2001. Expected Returns and Habit Persistence. *The Review of Financial Studies* 14:861–899.

Liu, J. 2007. Portfolio Selection in Stochastic Environments. *The Review of Financial Studies* 20:1–39.

Loomes, G., and R. Sugden. 1982. Regret Theory: An Alternative Theory of Rational Choice under Uncertainty. *The Economic Journal* 92:805–824.

Lupton, J. P. 2003. Household Portfolio Choice and the Habit Liability: Evidence from Panel Data. Working Paper.

Lynch, A. W., and S. Tan. 2011. Labor Income Dynamics at Business-Cycle Frequencies: Implications for Portfolio Choice. *Journal of Financial Economics* 101:333–359.

MacDonald, B.-J., B. Jones, R. J. Morrison, R. L. Brown, and M. Hardy. 2013. Research and Reality: A Literature Review on Drawing Down Retirement Financial Savings. *North American Actuarial Journal* 17:181–215.

Maurer, R., O. S. Mitchell, R. Rogalla, and I. Siegelin. 2014. Accounting and Actuarial Smoothing of Retirement Payouts in Participating Life Annuities. NBER Working Paper.

Maurer, R., R. Rogalla, and I. Siegelin. 2013a. Participating Payout Annuities: Lessons from Germany. *ASTIN Bulletin* 43:159–187.

Maurer, R. H., O. S. Mitchell, and R. Rogalla. 2010. The Effect of Uncertain Labor Income and Social security on Lifecycle Portfolios. In R. Clark and O. S. Mitchell (eds.), *Reorienting Retirement Risk Management*. Oxford University Press.

Maurer, R. H., O. S. Mitchell, R. Rogalla, and V. Kartashov. 2013b. Life Cycle Portfolio Choice with Systematic Longevity Risk and Variable Investment-Linked Deferred Annuities. *Journal of Risk and Uncertainty* 80:649–676.

Mehlkopf, R. J., I. Boelaars, A. L. Bovenberg, and S. van Bilsen. 2015. The Welfare Costs of Suboptimal Design in Occupational Pension Schemes. Working Paper.

Merton, R. C. 1969. Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case. *The Review of Economics and Statistics* 51:247–257.

Merton, R. C. 1971. Optimum Consumption and Portfolio Rules in a Continuous-Time Model. *Journal of Economic Theory* 3:373–413.

Merton, R. C. 1980. On Estimating the Expected Return on the Market: An Exploratory Investigation. *Journal of Financial Economics* 8:323–361.

Merton, R. C., and Z. Bodie. 2005. Design of Financial Systems: Towards a Synthesis of Function and Structure. *Journal of Investment Management* 3:1–23.

Muermann, A., O. S. Mitchell, and J. M. Volkman. 2006. Regret, Portfolio Choice, and Guarantees in Defined Contribution Schemes. *Insurance: Mathematics and Economics* 39:219–229.

Munk, C. 2008. Portfolio and Consumption Choice with Stochastic Investment Opportunities and Habit Formation in Preferences. *Journal of Economic Dynamics and Control* 32:3560–3589.

Novy-Marx, R., and J. D. Rauh. 2014. Linking Benefits to Investment Performance in US Public Pension Systems. *Journal of Public Economics* 116:47–61.

Pagel, M. 2012. Expectations-Based Reference-Dependent Life-Cycle Consumption. Working Paper.

Pliska, S. R. 1986. A Stochastic Calculus Model of Continuous Trading: Optimal Portfolios. *Mathematics of Operations Research* 11:371–382.

Polyak, I. 2005. New Advice to Retirees: Spend More at First, Cut Back Later. The New York Times.

Poterba, J. M., S. F. Venti, and D. A. Wise. 2009. The Changing Landscape of Pensions in the United States. In A. Lusardi (ed.), *Overcoming the Saving Slump: How to Increase the Effectiveness of Financial Education and Saving Programs*, chap. 1. University of Chicago Press.

Rauh, J. D. 2008. Risk Shifting versus Risk Management: Investment Policy in Corporate Pension Plans. *The Review of Financial Studies* 22:2687–2733.

Van Rooij, M. C. J., C. J. M. Kool, and H. M. Prast. 2007. Risk-Return Preferences in the Pension Domain: Are People Able to Choose? *Journal of Public Economics* 91:701–722.

Samuelson, P. A. 1969. Lifetime Portfolio Selection by Dynamic Stochastic Programming. *The Review of Economics and Statistics* 51:239–246.

Schroder, M., and C. Skiadas. 1999. Optimal Consumption and Portfolio Selection with Stochastic Differential Utility. *Journal of Economic Theory* 89:68–126.

Schroder, M., and C. Skiadas. 2002. An Isomorphism Between Asset Pricing Models With and Without Linear Habit Formation. *The Review of Financial Studies* 15:1189–1221.

Sundaresan, S. M. 1989. Intertemporally Dependent Preferences and the Volatility of Consumption and Wealth. *The Review of Financial Studies* 2:73–89.

The Melbourne Mercer Global Pension Index. 2013. Australian Centre for Financial Studies.

Tversky, A., and D. Kahneman. 1992. Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty* 5:297–323.

Vasicek, O. 1977. An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics* 5:177–188.

Wachter, J. A. 2002. Portfolio and Consumption Decisions under Mean-Reverting Returns: An Exact Solution for Complete Markets. *Journal of Financial and Quantitative Analysis* 37:63–91.

Wakker, P. P. 2010. *Prospect Theory: For Risk and Ambiguity.* Cambridge, UK: Cambridge University Press.

Wakker, P. P., and D. Deneffe. 1996. Eliciting von Neumann-Morgenstern Utilities when Probabilities Are Distorted or Unknown. *Management Science* 42:1131–1150.

Whitaker, B. 2005. Managing Retirement, After You Really Retire. The New York Times.

Wu, G., and R. Gonzalez. 1996. Curvature of the Probability Weighting Function. *Management Science* 42:1676–1690.

Xia, J. M., and X. Y. Zhou. 2014. Arrow-Debrew Equilibria for Rank-Dependent Utilities. To Appear in Mathematical Finance.

Xu, Z. Q. 2014. A Note on the Quantile Formulation. To Appear in Mathematical Finance.

Yaari, M. E. 1965. Uncertain Lifetime, Life Insurance, and the Theory of the Consumer. *The Review of Economic Studies* 32:137–150.

Yogo, M. 2008. Asset Prices Under Habit Formation and Reference-Dependent Preferences. *Journal of Business and Economics Statistics* 26:539–558.