

Tilburg University

Saving and investing over the life cycle and the role of collective pension funds

Bovenberg, A.L.; Koijen, R.S.J.; Nijman, T.E.; Teulings, C.N.

Published in:
De Economist

Publication date:
2007

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):

Bovenberg, A. L., Koijen, R. S. J., Nijman, T. E., & Teulings, C. N. (2007). Saving and investing over the life cycle and the role of collective pension funds. *De Economist*, 155(4), 347-415.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

SAVING AND INVESTING OVER THE LIFE CYCLE AND THE ROLE OF COLLECTIVE PENSION FUNDS***

BY

LANS BOVENBERG*, RALPH KOIJEN*, THEO NIJMAN*, AND COEN TEULINGS**

Summary

This paper surveys the academic literature on optimal saving and investment over an individual's life cycle. We start out with a simple benchmark model with separable and smooth preferences, one aggregate risk factor and riskless wage income. Within this simple setting, optimal saving and investment behavior are explored from the perspective of individuals. Subsequently, we investigate various constraints to optimal individual decision making. We discuss how collective pension schemes may help to relieve some of the market incompleteness that arises from these constraints while at the same time introducing new types of constraints. Finally, various extensions to the benchmark setting are analyzed: a more elaborate modelling of human capital, additional risk factors, and other types of preferences.

Key words: saving, investment, life cycle, pension schemes, defined contribution, defined benefit

JEL Code(s): D91, G11, G23

1 INTRODUCTION

This paper surveys the recent academic literature on optimal financial planning of individuals over the life cycle and relates this to the optimal design of collective pension schemes in facilitating this planning. Collective pension schemes can relieve borrowing constraints and enable intergenerational risk sharing but usually impose uniform rules on heterogeneous participants. We explore the costs and benefits of collective pension schemes versus individual schemes. Compared to other surveys of the literature on life-cycle consumption and saving (see e.g. Browning and Lusardi (1996) and Browning and Crossley (2001)), we explore how consumption decisions interact with risk taking and focus on the role of pension schemes in facilitating life-cycle financial planning.

* Corresponding author: Theo Nijman, Netspar and Tilburg University, P.O. Box 90153, 5000 LE, Tilburg, The Netherlands, e-mail: Nyman@uvt.nl.

** Netspar and University of Amsterdam, Roetersstraat 29, 1018 WB, Amsterdam, The Netherlands, e-mail: C.N.Teulings@uva.nl.

*** We thank Peter Kooreman for helpful comments on an earlier version and Roel Mehlkopf for research assistance.

Individuals face two main decisions in their financial planning over the life cycle. Through the saving decision, they decide how to smooth consumption over time by setting the pension premium and the pension benefits. Through the investment decision, individuals decide how to invest the premia in the various financial assets so as to smooth consumption across various future contingencies that may arise in the future. Under a set of "golden assumptions," the young should invest a larger fraction of their financial wealth in risky assets than the elderly (see e.g. Merton (1971) and Merton and Samuelson (1974)). Indeed, life-cycle funds and target-date funds, which are based on these principles, are popular investment vehicles in many countries.

Teulings and de Vries (2006) outline a stylized model to explore optimal financial planning over the life cycle. We outline in detail the arguments underlying their recommendations for individual saving and investment decisions. In addition, we investigate how robust their recommendations are with respect to various extensions of their basic model. In particular, we explore alternative models for labor income, financial risks and preferences.

In addition to analyzing optimal saving and investment decisions over the life cycle, we explore the role of financial intermediaries in general and pension funds in particular. Individuals delegate financial decision making to financial intermediaries for a variety of reasons. First, individuals lack the expertise to implement a financial plan for their lives. Second, financial intermediaries reduce the costs of long-term financial planning by benefiting from the scale economies associated with specialization in acquiring financial expertise (e.g. asset management) and accessing financial markets. Compulsory participation in collective pension schemes allows these schemes to smooth consumption across generations by shifting surpluses and deficits to future generations; the trustees of the pension scheme can decide whether to pay out the surplus as dividends to current stakeholders (in the form of additional pension benefits or lower pension premia) or to save the surplus for future participants.¹ We explore the literature on the potential benefits of this intergenerational risk sharing and investigate under which conditions these benefits dominate the costs of the homogeneous decision rules that these pension schemes impose on heterogeneous participants. Financial intermediaries like mutual funds and insurers offer individual plans and structured products that are more tailor-made to the specific preferences and circumstances of an individual but this customization typically involves additional costs. Insurers and mutual funds, however, also execute and insure collective pension plans based on compulsory participation, which involve risk sharing with future generations but are typically not tailor-made to the individual.

1 Profit participating contracts issued by insurers in many countries are in fact very closely related (see Nordahl and Dokseland (2006)). In particular, new generations may find these contracts attractive if they can benefit from surpluses that have not been distributed in the past.

Pension schemes may smooth income not only across generations but also within contemporaneous groups of participants through redistribution.² To illustrate, pension contracts usually charge uniform prices for deferred annuities irrespective of the characteristics of individuals, such as sex and age (see e.g. Brown (2002), and Aarssen and Kuipers (2006)). Especially uniform contribution rates for deferred annuities with different deferral maturities tend to imply substantial value redistribution across younger and older workers (see e.g. Boeijen et al. (2006)). Redistribution of value can be analyzed by valuing all relevant claims at market value (see e.g. Kortleve and Ponds (2006)).³ In this paper, we largely abstract from the task of redistribution. Rather than financial intermediaries, the government is often in a better position to redistribute across individuals. The reason is that governments are endowed with tax power over a larger pool of people: the nation as a whole. Collective pension schemes, in contrast, wield less effective tax power – even if all workers in a sector are forced to participate in a sectoral scheme. The reason is that labor-market mobility within a country is generally larger than labor mobility between countries.

Our analysis also abstracts from the employer as a risk-bearing agent. The implicit assumption is that the employer does not add to the risk-trading opportunities on the capital market, for example by providing guaranteed benefits at a lower price than what is available in the market. Indeed, firms increasingly mark-to-market the pension guarantees they provide to workers because international accounting rules force companies to put their pension risks on their balance sheets. Moreover, in trading risk with their employers, employees will be saddled with the credit risk of the company they work for. This makes risk trading between employer and workers within a firm less than optimal. Hence, employees trade risks on the capital markets and between one another in collective pension schemes. A defined-benefit plan thus buys guarantees either on the capital market or from young and future participants rather than from the firm that employs the workers.

The rest of this paper is structured as follows. Section 2 sets up our benchmark framework for analyzing life-cycle financial planning. Within this framework, section 3 explores the optimal saving and investment behavior from the perspective of the individual. This section abstracts from constraints on intertemporal consumption smoothing and risk taking so that the resulting allocation is first best. Section 4 turns to various constraints on optimal individual decision making due to imperfections in markets and individual

2 Risk sharing *ex ante* (i.e. before uncertainty is resolved) becomes redistribution *ex post* (i.e. after uncertainty is resolved). We thus explore redistribution as the *ex-post* outcome of an insurance or risk-sharing contract.

3 The analysis in this paper is crucially different because in these papers the assumption is made that agents unwind the contribution and investment rules imposed by the pension fund.

decision making. Collective pension schemes may help to relieve some of these constraints on individual decision making and individual behavioral biases but are also likely to introduce new types of constraints. Section 5 investigates the optimal design of collective pension schemes within our benchmark framework, including risk sharing across generations through the transfer of (possibly negative) surpluses over time. This section discusses when and how compulsory pension plans can securitize human capital of current and future participants. By thus allowing young generations and future generations to borrow against their human capital, these generations can participate in risk taking and therefore take advantage of the equity risk premium. Section 6 surveys the recent literature on various extensions to the benchmark model: a more elaborate modelling of human capital, financial and actuarial risk factors, and preferences. Mathematical equations are kept to a minimum in the main text. Technical derivations are available in an appendix to the working paper version of this paper (see Bovenberg et al. (2007)).

2 THE BENCHMARK MODEL

This section lays out our benchmark framework for exploring optimal life-time saving and investment (see also Merton (1971), Merton and Samuelson (1974); Teulings and de Vries (2006)). It describes our benchmark assumptions on financial markets, labor markets and preferences and discusses the parameter values employed in our numerical simulations.

2.1 *Financial Markets*

- A risk-free asset (a bond) is available.
- Equity-market risk is the only aggregate risk factor, which is traded through equity. Housing is abstracted from.
- The interest rate, inflation, the volatility of equity, and the equity risk premium are constant over time so that mean reversion and stochastic volatility are absent. These prices are not affected by the decisions of individuals or financial intermediaries. We thus take a partial-equilibrium perspective of a small open economy, which takes prices as given on the world market.
- Log stock returns are identically and independently distributed according to a normal distribution.
- Financial markets are dynamically complete if households and pension funds can continuously trade stocks and bonds without constraints. In some cases, households and pension funds face constraints.

- Death is predictable or perfect insurance of individual longevity risk is available. Aggregate longevity risk is thus absent.⁴
- Whenever we consider the presence of collective pension schemes we assume that participants have access to the capital markets only through their pension scheme. From the household's perspective markets are thus incomplete. Hence, individuals cannot offset the policies of their pension scheme by engaging in offsetting capital-market transactions. Consumption during the active period (i.e. when working) equals wage income minus the pension premium. Consumption in retirement is given by the pension benefit.

2.2 Labor Markets

- The after-tax wage during the working career is constant and riskless with a fixed, exogenous retirement age.⁵ Human capital is thus paid out in the form of constant wage income until it is fully depleted at the age of retirement. Moreover, labor-market risks are absent.
- Labor supply is fixed.
- Wages are exogenous: pension premia thus reduce disposable incomes one for one.

2.3 Preferences

- Individuals aim to maximize lifetime utility, which is the weighted sum over time of expected utility at each point in time. Utility at a point in time depends only on consumption at that time. The weights of future expected utilities decline exponentially at the so called rate of time preference. Hence, people are impatient: at equal levels of consumption a marginal unit of future consumption adds less to utility than current consumption
- Preferences feature positive and constant relative risk aversion. Consumption in each contingency adds to utility. At larger levels of consumption, however, additional consumption adds less to utility. The negative elasticity of marginal utility with respect to the level of consump-

4 Even with idiosyncratic longevity risk but the absence of life insurance, the results below survive if the length of individual life time is distributed exponentially (see Viceira (2001)).

5 We abstract from taxes on capital income. A constant rate of consumption tax reduces consumption in each period and each contingency proportionally and will thus not affect saving and investment decisions (except that all savings and investments are scaled back proportionally). Such a consumption tax is equivalent to an income tax with a constant marginal rate that treats pension saving on a so-called cash-flow basis (i.e. premia are tax deductible while benefits are taxed) if the tax rate against which premia can be deducted is equal to the tax rate at which the benefits are taxed.

tion is known as the coefficient of relative risk aversion. This measure for the sensitivity of marginal utility with respect to the level of consumption is constant and thus does not depend on the consumption level.

Positive risk aversion implies that individuals have a taste for moderation across time and across contingencies. They prefer a smooth consumption level (with a low variance) rather than a highly volatile consumption stream (with a high variance) over time or across contingencies. More general models (than expected utility) distinguish the taste for moderation across *contingencies* from the taste for moderation across *time*. Risk aversion measures the taste for moderation across random outcomes. Highly risk-averse individuals feature a strong preference for stable consumption levels across various contingencies. They want their consumption level to be certain ('stable') irrespective of what happens (which is reflected in the costs of maintaining the consumption level in a particular contingency). The taste for moderation across time is related to the intertemporal elasticity of substitution. An individual exhibiting a low intertemporal elasticity of substitution prefers a stable level of consumption over time.

- Bequest motives are absent.
- Individuals start consuming when they enter the labor market.

2.4 *Benchmark Parameters in Numerical Simulations*

We illustrate our results with numerical simulations. Following Teulings and de Vries (2006), these simulations are based on a constant coefficient of risk aversion of 5, a working life of 45 years, an expected retirement period of 15 years, a rate of time preference and a risk-free interest rate of 2% per year, an equity risk premium of 4% a year, and a standard deviation of stock returns of 20% per year.

3 SAVING AND INVESTING OVER THE LIFE CYCLE: THE FIRST BEST

3.1 *Introduction*

This section derives the first-best solution to the financial planning problem of the investor under the assumptions as outlined in section 2. We analyze this case in depth because it generates intuition and benchmark results for the findings in more elaborate models. Section 3.2 explores consumption smoothing (i.e. saving decisions) in the absence of risk. Section 3.3 investigates optimal asset allocation in case of risky investment opportunities. Section 3.4 returns to optimal consumption smoothing but considers the case with risky assets.

3.2 *Intertemporal Consumption Smoothing Without Risk*

This section focuses on consumption smoothing over time in the absence of risk. Together with the depreciation of human capital due to aging, the preference for a smooth consumption stream over time gives rise to a demand for pension saving. Individuals want to move part of the income from their human capital when they work to the periods in which they still would like to consume even though they do not collect any labor income anymore. Capital markets allow individuals to engage in intertemporal trade,⁶ with the interest rate measuring the reward for transferring resources to a later date. In particular, through capital markets, one can exchange resources in the active periods of life, when labor resources are relatively abundant but consumption is not so valuable at the margin, to the inactive periods of life when consumption is relatively more valuable at the margin but labor resources have already been depreciated. By investing pension premiums collected from an active individual in financial assets, financial institutions (pension funds, mutual funds, insurers) facilitate this intertemporal trade through capital markets. The financial institution in effect transforms part of human capital of the individual into a claim of that individual on that institution. These claims are secured by financial capital, which can be sold in retirement so that the individual can maintain his standard of living.

3.2.1 *Perfect Consumption Smoothing Without Impatience*

The simplest case to consider is when the real interest rate is zero,⁷ individuals are not impatient (i.e. the rate of time preference is zero), and wage income is constant during the active life. In that case, complete consumption smoothing is optimal: consumption should remain constant during the life cycle. The life-time budget constraint limits the level of consumption. Overall wealth, which consists of human and financial wealth, measures lifetime resources. The individual begins the working life without any financial wealth so that he relies on human wealth only. As the discounted value of life-time wage income, human wealth at the beginning of the active life is simply the number of active years times the annual wage income; labor income today is equally valuable as labor income tomorrow because the real interest is zero.

The life-time budget constraint implies that the discounted value of consumption (the sum of all consumption flows during the life course in the case of a zero discount rate) cannot exceed human wealth. A constant consumption flow equal to the share of the active life in the remaining life time times the labor income flow during the active life exactly exhausts human capital at

6 This trade is similar to trade in other markets except that goods at different dates are exchanged.

7 Recall that risk is absent. Hence, this is the risk-free interest rate.

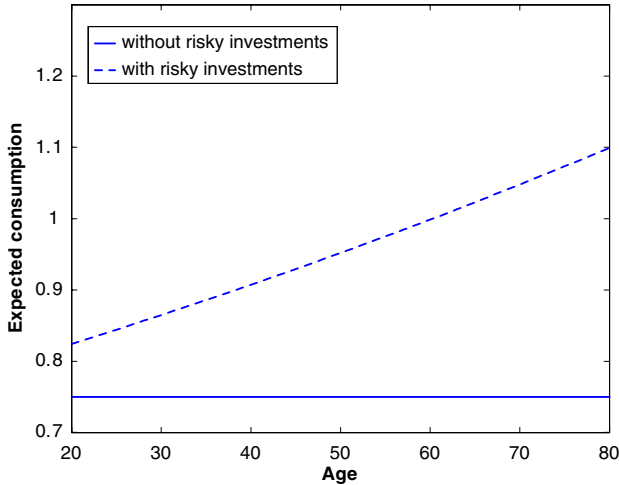


Figure 1 – Expected consumption path (C_t) over the life-cycle if the real interest rate is zero ($r=0$) and the individual is not impatient ($\rho=0$). Investment risk is absent

the end of life. The savings rate is thus equal to the share of retired life in the overall adult life. Whereas human wealth is depleted at the rate of the labor income flow, overall wealth is depleted at a lower speed because part of the labor income flow is transformed into financial wealth. At the age of retirement, human capital is exhausted so that wealth consists of financial wealth only. Financial wealth is at its maximum at the retirement age after which it is gradually depleted. However, since part of human capital is used for consumption during the working career, total wealth is less at the retirement age than at the start of the career: as a ratio of initial human wealth, financial wealth at retirement equals the share of the inactive life in the overall adult life.

Figures 1 and 2 illustrate the consumption decision and its implications for the trajectory of financial, human and overall wealth. The solid line in Figure 1 illustrates that in this simplest case without risky investments the consumption pattern is flat and that a quarter of labor income is saved for retirement. the dotted line refers to the case with risky investments and will be discussed in section 3.4. Figure 2 shows that overall wealth is depleted at three quarters of the speed at which human capital is depreciated. At its maximum, financial wealth equals a quarter of initial human wealth. The figure indicates that human wealth dominates financial wealth (i.e. pension wealth) for most part of the active working life. Labor markets thus play a key role in financial planning over the life cycle. Younger people are *ceteris paribus* wealthier than older people because they have a longer expected life time in front of them and have consumed less of their human capital.

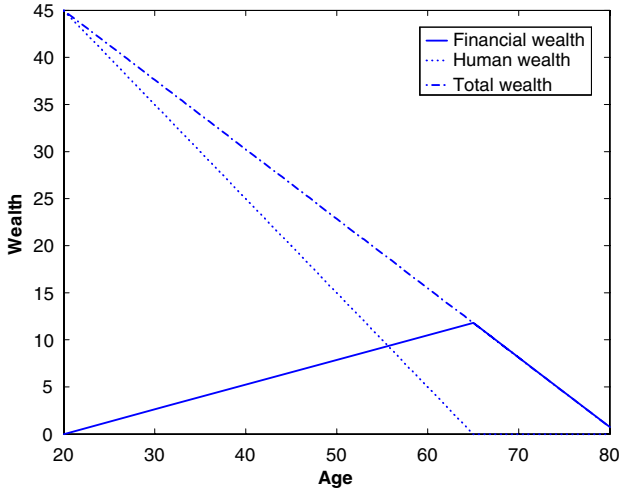


Figure 2 – Trajectories for financial wealth (F_t), human wealth (H_t) and total wealth (W_t) over the life-cycle in the case with a zero real interest rate ($r=0$) and without impatient individuals ($\rho=0$) and risk

3.2.2 Perfect Consumption Smoothing with Impatience If the Real Rate Equals the Time Preference

Individuals are typically impatient. This implies that individuals should be rewarded for saving through positive interest rates. If the real interest rate equals the rate of time preference, the individual still finds a constant consumption stream optimal because the reward of waiting (i.e. the interest rate) exactly balances the cost of waiting (i.e. the rate of time preference). Compared to the case with a zero interest rate and a zero rate of time preferences, the consumption level can be higher *ceteris paribus* because the pension saver benefits from positive net interest income on his accumulated savings. With a positive interest rate, consumption is higher even although the present value of human capital is lower at the beginning of life. The value of future labor income declines because a positive interest rate indicates that a resource earned tomorrow is worth less than a resource today. Hence, if the individual would consume all his income today, he would be able to consume less compared to the case with a zero interest rate. However, since the individual prefers a stable consumption path over time, consumption occurs on average at a later date than labor income is received. The individual thus enjoys a positive income effect if the interest rate rises above zero.

Figure 3 involves the case in which the rate of time preference ρ^i equals 2% and the interest rate is 2%. Financial wealth at retirement is a larger share of initial human wealth than in the case in which both the rate of time preference and the real interest rate are zero.

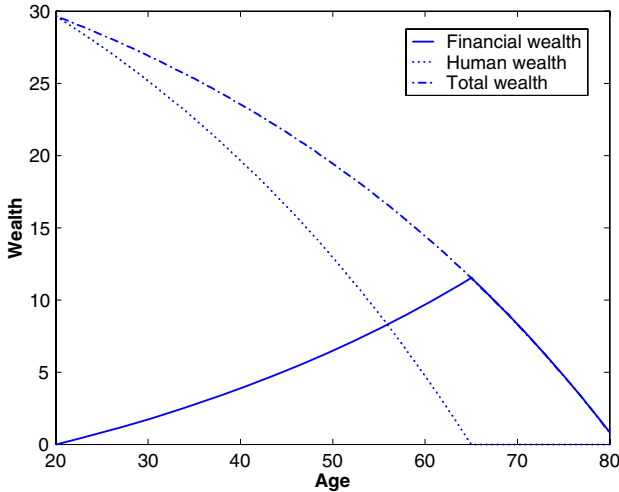


Figure 3 – Trajectories for financial wealth (F_t), human wealth (H_t) and total wealth (W_t) over the life cycle without risk and with an interest rate r and a rate of time preference ρ of 2%

3.2.3 Consumption Smoothing In The General Case, Excluding Risky Investments

In the general case, the sign of the gap between the interest rate and the rate of time preference determines whether consumption is increasing or decreasing over time. With a positive gap, the benefits from waiting offered in the capital market exceed the subjective cost of waiting. Hence, the net reward for waiting is positive. This positive reward makes a rising path for consumption c^i of individual i optimal⁸

$$\frac{dc^i/dt}{c^i} = \frac{(r - \rho^i)}{\phi^i}. \quad (1)$$

The growth rate of consumption $\frac{dc^i/dt}{c^i}$ of individual i is not very sensitive to the gap between the interest rate r and the individual's rate of time preference ρ^i if a small intertemporal substitution elasticity of that individual ($1/\phi^i$) indicates that individual i exhibits a large preference for consumption smoothing over time. Intuitively, behavior is not very sensitive to intertemporal prices (i.e. the net reward for waiting) if agents dislike large differences between consumption levels at different points in time. In that case, individuals thus need to face substantial net rewards in order to be induced to have a rising (and thus not smooth) consumption path over time.

⁸ The derivation of the equations in the text is outlined in the appendix to Bovenberg et al. (2007).

Expression (1) implies that the optimal consumption level and the optimal pension contribution depend on the parameters that describe the preferences of the individual. Collective schemes that impose identical contribution rates on participants that exhibit heterogeneous preferences thus cannot be optimal. We will return to this issue in sections 4 and 5.

3.2.4 *Shocks in Wealth and Consumption Smoothing*

The rate at which overall wealth is consumed depends on age. Since older people feature a shorter planning horizon, they consume a larger share of their overall wealth. If both a young and an old person get one additional euro, the old person will consume the euro more rapidly. However, if both agents obtain $x\%$ more wealth, both agents will increase their consumption by $x\%$ during the rest of their lives.⁹ Intuitively, the preference for moderation associated with the concave utility function implies that the agents want to spread the increase in wealth as broadly over their life time as possible. Hence, rather than spending it in a few periods, they choose to enjoy the wealth boost in equal relative consumption increases in each period during the rest of their lives

$$\frac{d \log C_s^i}{d \log W_t^i} = 1, \quad \text{where } s \geq t, \quad (2)$$

W_t^i stands for total wealth (i.e. the sum of human and financial wealth) of individual i at time t and C_s^i denotes consumption of that individual at time s . The unitary elasticity of the consumption flow (in the rest of the life time) with respect to wealth imply that both pension contributions should decline and pension benefits increase following a positive wealth shock. Hence, rather than a pension system that keeps the premium fixed (a *defined-contribution* system) or a pension system that fixes the benefits (a *defined-benefit* system), a hybrid system that adjusts both premia and benefits in response to income and wealth shocks appears to be optimal.

3.3 *Risk Taking*

Financial markets allow agents to shift consumption not only across time, but also across various future contingencies when agents face uncertainty about which contingency will actually materialize. In particular, risk-averse individuals can buy resources in bad states by giving up resources in good states. Just as in the case of optimal consumption smoothing over time, we can distinguish between the preferences for consumption smoothing (across *contingen-*

⁹ This requires that the intertemporal elasticity of substitution is fixed over the life course, which has been assumed in section 2.

cies rather than *time* so that risk aversion rather than the intertemporal elasticity of substitution measures this preference) and the prices for consumption smoothing (in this case, instead of the net interest rate $(r - \rho^i)$, the Sharpe ratio, which is defined as the expected excess return (over the risk-free return) per standard deviation of the excess return, is relevant).¹⁰

How much risk the individual optimally chooses to absorb depends on both risk aversion and the risk premium (for each unit of risk as measured by the Sharpe ratio, which can be interpreted as the reward for risk taking). The individual i chooses the optimal amount of risk in such a way that his coefficient of relative risk aversion (i.e. the negative elasticity of marginal utility with respect to consumption), θ^i , equals the marginal rate of substitution between the excess return of risk taking $f_i \bar{\mu}$ and variance $f_i^2 \sigma^2$:

$$\theta^i = \frac{f_i \bar{\mu}}{f_i^2 \sigma^2} \quad (3)$$

where $\bar{\mu}$ and σ stand for the expectation and the standard deviation of the excess return on the risk factor,¹¹ while f_i represents the individual i 's share of total wealth invested in the risky asset. This implies that

$$f_i \sigma = \frac{\lambda}{\theta^i}, \quad (4)$$

where $\lambda \equiv \bar{\mu}/\sigma$ denotes the Sharpe ratio so that the optimal expected excess return due to risk taking amounts to $f_i \bar{\mu} = \lambda^2/\theta^i$. Risk taking (i.e. volatility of consumption across various contingencies) as measured by $f_i \sigma$ increases with the reward to risk taking λ and decreases with the preference for consumption smoothing (i.e. a certain consumption level independent of contingencies) as measured by relative risk aversion θ^i . Note that f_i can be larger than unity. In that case, the worker should go short in bonds to buy the risk factor. With our benchmark parameters, the Sharpe ratio λ and the standard deviation σ are both 0.2 while relative risk aversion amounts to five. The optimal investment share f_i is thus 20%.

The expression for optimal consumption smoothing across contingencies (4) is similar to that for optimal consumption smoothing across time (1). In

10 In complete capital markets, resources in each contingency have a single, unique price. Capital markets are complete if the number of not perfectly correlated assets is equal to the number of risk factors and these assets can be traded continuously. If agents can freely trade in complete capital markets (as we assume in section 3), we can measure the utility value of the property rights of individuals on resources by a single metric: wealth. In the presence of constraints (as in Section 4), in contrast, individuals do not equate their marginal rates of substitution (the ratios of their marginal utilities) to market prices. Hence, the market value of assets do not fully describe the utility value of these resources.

11 Assume that the relative change in the risk factor (in excess of the risk-free return) is distributed normally: $\ln(\Delta \text{risk factor}) \sim N(\mu, \sigma^2)$. Then: $\bar{\mu} \equiv \ln(E\Delta \text{risk factor}) = \mu + \frac{1}{2}\sigma^2$.

both cases, the right-hand side involves the net price for smooth consumption (i.e. the Sharpe ratio λ in the case of (4) and the net interest rate $(r - \rho^i)$ in the case of (1)) and the preference for consumption smoothing (i.e. relative risk aversion θ^i in the case of (4) and the reciprocal inverse intertemporal substitution elasticity $(1/\phi_i)$ in the case of (1)). Prices and preferences together determine the optimal inequality in consumption (across *contingencies* as measured by $f_i\sigma$ or *time* as measured by the growth rate of consumption) at the left-hand side of these expressions.

3.3.1 *Optimal Investment Share Age Invariant*

The share f_i of total wealth invested in the risk factor does not depend on age, or more generally, the investment horizon. Suppose that we invest for t years rather than one year. In that case, both the excess return $f_i\bar{\mu}t$ and the variance of the excess return $f_i^2\sigma^2t$ in (3) vary proportionally with the length of the investment period so that the optimal investment share is not affected by the time horizon. In other words, both the marginal benefits and the marginal costs of investing more in equity rise linearly with time.¹² This reasoning shows that the familiar argument that time diversification allows young people to take more risk is fallacious and relies on a wrong interpretation of the law of large numbers. The sum of n independent and identically distributed random variables has a variance that is n times larger than the variance of each of the separate risks. The law of large numbers, in contrast, states that the variance of the *average* (rather than the *sum*) of n independent and identically distributed random variables goes to zero if n becomes very large.¹³

Mossin (1968), Merton (1969) and Samuelson (1969) first independently derived the result that the investment share in the risky asset is independent of age, in contrast to the speed with which wealth is consumed. As we noted in the subsection on consumption smoothing over time above, the elasticity of consumption with respect to wealth is unity and thus independent of age. Hence, a young person is equally vulnerable (in terms of the *relative* change in the consumption flow, which is relevant in case *relative* risk aversion is constant) to the same *relative* change in wealth and should thus *ceteris paribus* hold the same wealth *share* (as opposed to the absolute amount of wealth) in

12 This is the case with smooth preferences. Subsection 6.3.4 shows that if preferences are characterized by loss aversion the costs of risky investment rise less rapidly with the horizon than the benefits do.

13 Another reasoning maintains that the benefit of investing in equity is the excess return, which increases linearly with the length of time, while the cost is the standard deviation of that return, which rises only with the square root of that length because the drawings are independent across time. With a smooth twice differentiable utility function, however, the costs of risk are captured by the variance rather than the standard deviation. See also sub-section 6.3.4 on loss aversion in which case the costs of risk are in fact closely related to the standard deviation and thus rise less rapidly with the time horizon than the benefits do.

risk-bearing assets.¹⁴ If the investment share f_i would depend on age, people would bear more consumption risk in some parts of their lives than in others. With constant relative risk aversion, this would generate opportunities for more efficient interpersonal allocation of risk.

3.3.2 *Optimal Risk Sharing Across Individuals And Time*

Optimal risk sharing when a shock hits implies that marginal utilities of consumption in the remaining lives of all agents changes with the same percentage. The combination of optimal consumption smoothing (i.e. optimal saving, see (2)) and optimal investment (see (4)) accomplishes exactly that. In particular, a standard deviation in the risk factor changes wealth by $100\lambda/\theta^i\%$ (according to (4)) and thus (from (2)) also consumption during the rest of the life of all living generations by that percentage. This yields a relative change in marginal utility of consumption of $100\lambda\%$, which is the same for all households. We thus see that (2) and (4) ensure that a shock is distributed as broadly as possible over the currently living individuals. *All* these individuals are affected and they are affected during their *entire* remaining lives. Consumption of these individuals behaves in the same fashion as the risk factor, namely as a random walk.

3.3.3 *Non-Tradable Human Wealth: The Investment Share Of Tradable Financial Wealth*

Human wealth is non tradable.¹⁵ Hence, financial rather than human wealth should be adjusted to achieve the right exposure to risk factors. If human capital is riskless, it acts like a risk-free asset and all the exposure to risk should come from financial wealth. As the wealth share of financial wealth increases from zero to one during the working life, the share of *financial* wealth invested by individual i in risk-bearing assets, f_i^* , falls from infinity at the beginning to f_i at retirement:

$$f_{it}^* = \frac{S_t^i}{F_t^i} = f_i \frac{W_t^i}{F_t^i} = f_i \left(1 + \frac{H_t^i}{F_t^i}\right) = \frac{1}{\theta^i} \frac{\lambda}{\sigma} \left(1 + \frac{H_t^i}{F_t^i}\right), \quad (5)$$

where S_t^i , F_t^i and H_t^i denote, respectively, wealth invested in the risk factor, financial wealth, and human wealth at time t by individual i so that $W_t^i =$

14 This assumes that the young and the old person share the same *relative* risk aversion. Gollier and Zeckhauser (2002) show that this requires that absolute risk tolerance is proportional to wealth. If this is not the case, the wealthier young person does not feature the same relative risk aversion as an older person who shares the same utility function and labor income path over the life cycle. Fortunately, the possible non-proportionality of absolute risk tolerance yields only a marginal effect on the optimal portfolio of young investors (see Gollier (2005)).

15 Other illiquid assets are owner-occupied housing and privately owned businesses. These assets may have similar risk characteristics as equity.

$H_t^i + F_t^i$. This financial wealth share tends¹⁶ to fall for two reasons. First, the absolute amount of wealth invested in equity (or risk-bearing assets) tends to fall with time as an individual consumes part of human wealth during the working life (so that W_t tends to decline with time). Second, the stock of financial wealth F_t tends to increase as the individual saves part of their human capital. The economic intuition why the young should hold a larger component of their financial wealth in stocks is that the young are less dependent on financial wealth for their consumption because they have an alternative income source in the form of labor income. They thus can afford to take more risk with financial wealth than elderly agents who depend almost entirely on this type of wealth for their livelihood.

3.3.4 *Young Go Short to Acquire Optimal Risk Exposure*

Equation (5) indicates that the optimal share of financial wealth that is invested in the risk factor can well be above one if financial capital is small and human capital substantial. This is typically the case early in the life cycle. At the beginning of one's career, one should thus borrow to acquire the risk-bearing assets. On the basis of benchmark parameters, Teulings and de Vries (2006) and Bodie et al. (1992) find that a young worker may want to borrow as much as six times his annual salary and invest this in the equity market. Investment in housing financed by mortgages can contribute to efficient risk bearing at younger ages if housing risk is correlated with equity risk.

Figure 4 illustrates the horizon dependence of the asset allocation of financial wealth. The solid line specifies the expected fraction of financial wealth invested in equity over the life cycle. The dotted lines provide 10%- and 90%-quantiles of the optimal fraction of financial wealth that is invested in equity. These quantiles indicate that, depending on the actual investment returns experienced in the past, the optimal asset allocation can deviate from the a priori expected allocation. Figure 4 shows that financial wealth can get negative in which case the optimal exposure can be negative. It also presents the optimal fraction of wealth that is invested in equity for two randomly selected

16 We use the word 'tends' here because unexpectedly positive shocks may temporarily raise financial wealth so much that it offsets the depreciation of human capital. Conversely, adverse shocks may cause financial wealth to fall even though the individual saves part of his labor income.

Another reason why the ratio H_t/F_t may not decline that steeply is that retired agents may benefit from a first-pillar pension indexed to wages. In that case, a part of pension wealth is like human capital so that also retirees rely to some extent on human capital.

Finally, especially high-skilled workers can experience rapid labor-income growth in the beginning of their career. Hence, human capital may increase in the beginning of the career. At the same time, financial wealth remains small (and even become negative) as intertemporal consumption smoothing gives rise to low saving rates. As human capital thus increases compared to financial wealth in the beginning of life, also the equity share in financial wealth f^* increases (see Cocco et al. (2005)).

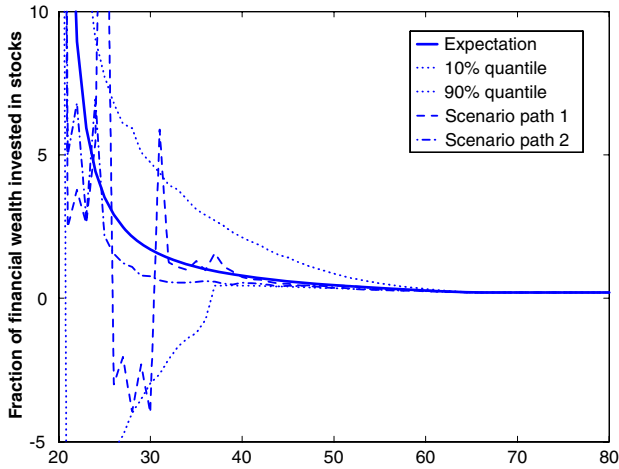


Figure 4 – Share of financial wealth invested in risky assets (f_t^*) over the life-cycle with the benchmark parameters ($r=2\%$, $\rho=2\%$, $\lambda=0.2$)

scenarios. The same scenarios will be used in subsequent graphs to illustrate the evolution of consumption and wealth. The scenario's in figure 4 show that the optimal equity share in financial wealth is very volatile because financial wealth is very small early in the life cycle.

3.4 Intertemporal Consumption Smoothing with Risk

Section 3.2 explored optimal saving without risk. This subsection reconsiders optimal saving behavior in the presence of risky investment opportunities (see also figure 1, which compares consumption behavior in both the absence and the presence of risk if agents are not impatient (i.e. $\rho^i=0$) and the real interest rate is zero (i.e. $r=0$)). The introduction of risk affects optimal saving through two channels. First, risk taking enhances welfare because investors can now capture the equity premium. This positive income effect depends on the risk premium $f_i \bar{\mu}$ and raises optimal consumption and reduces the optimal pension premium (i.e. saving out of labor income). Intuitively, part of retirement saving is financed out of the risk premium. Second, risk taking introduces a precautionary saving motive. The sign of this motive depends on whether marginal utility is convex (i.e. on the sign of the third derivative of the utility function). In particular, risk implies an additional precautionary saving motive if marginal utility is convex so that the expected marginal utility of consumption (which determines the savings motive if the future is uncertain) exceeds the marginal utility of expected consumption (which determines the savings motive if the future is certain). Intuitively, transferring

resources to the future adds more to utility if the future becomes more uncertain. If the third derivative of the utility function is positive, the consumer is prudent (see Leland (1968)).

With a constant relative risk aversion, marginal utility is indeed convex so that the investor is prudent. As a direct consequence, the introduction of risk unambiguously increases the growth rate of expected consumption. With a variance of $\psi_i^2 = f_i^2 \sigma^2$ on aggregate wealth, we derive for a CRRA utility function (with $\phi^i = \theta^i$, which we impose for the rest of this section and sections 4 and 5) and the optimal investment share $f_i = \frac{\lambda}{\sigma \theta^i}$ that the expected growth rate of consumption can be written as¹⁷

$$E \frac{dc^i / dt}{c^i} = \frac{(r - \rho^i + \frac{(1+\theta^i)}{2} \theta^i \psi_i^2)}{\theta^i} = \frac{(r - \rho^i + \frac{(1+\theta^i)}{2} \frac{\lambda^2}{\theta^i})}{\theta^i} \tag{6}$$

Hence, more risk ψ_i *ceteris paribus* raises the rate of time preference (or the growth rate of labor income) that is required to ensure that expected consumption is constant over the life cycle so that the expected pension premium is constant during the working life.¹⁸ With our benchmark parameters the expected growth rate of consumption amounts to 0.48% per year. In the present setting, the actual consumption and wealth trajectories can very well deviate substantially from their expectation. Figures 5 and 6 present the expectation, some quantiles and two scenario's of the consumption path and the corresponding wealth trajectories over the life-cycle.

3.4.1 Cohort Effects in Premium Rates

For individuals with the same preferences but different birth dates, the optimal premium rate starts at the same level ($F_t = 0$). Subsequently, the optimal premium rates move parallel in response to shocks. This implies that individuals of different cohorts pay different premium rates at the same age because they have experienced different shocks. Figures 7 and 8 provide examples of how consumption and financial wealth are affected by shocks. At any point

17 The overall effect on consumption at the beginning of the working life depends on the sign of $\frac{1}{2} f_i \bar{\mu} - \frac{(1+\theta^i)}{4} f_i^2 \sigma^2$. At the optimal investment share (4) this is given by $\frac{\lambda^2(\theta^i-1)}{2\theta^{i2}}$.

The first term $\frac{1}{2} f_i \bar{\mu}$ represents the welfare gain (i.e. the Harberger triangle) from being able to invest in risky assets. The second term $\frac{(1+\theta^i)}{4} f_i^2 \sigma^2$ stands for the impact of more precautionary saving. This term is half of the corresponding term in the growth rate of consumption (6) because the other half implies higher consumption at the end of the working life.

The equity risk premium does not directly raise the growth rate of consumption. This is because at the margin the benefit of the risk premium is exactly offset by the cost of risk if the investment share is optimal.

18 With the optimal investment share f_i from (4), we have $\psi = f_i \bar{\mu} = \lambda^2 / \theta^i$ so that $\rho^i = r + \frac{(1+\theta^i)\lambda^2}{2\theta^i} - \theta^i g^i$ (where g^i is the growth rate of wage income) implies that the premium rate does not depend on age. Note that we have assumed in section 2 that $g^i = 0$.

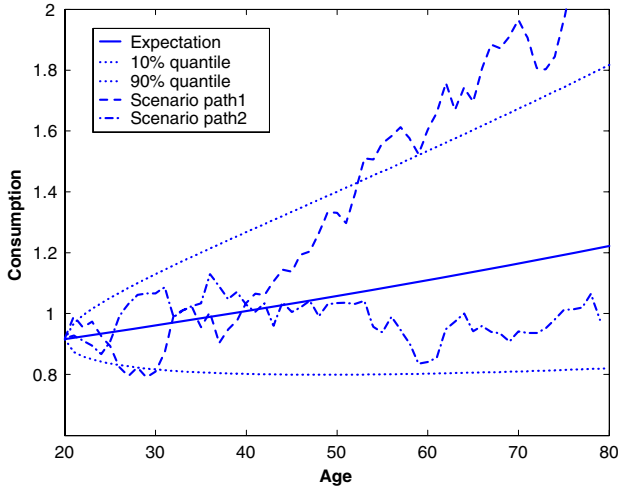


Figure 5 – Expectation, 10% and 90% quantiles and two simulated scenario paths of the consumption path (C_t) over the life cycle for the benchmark parameters ($r = 2\%$, $\rho = 2\%$, $\lambda = 0.2$)

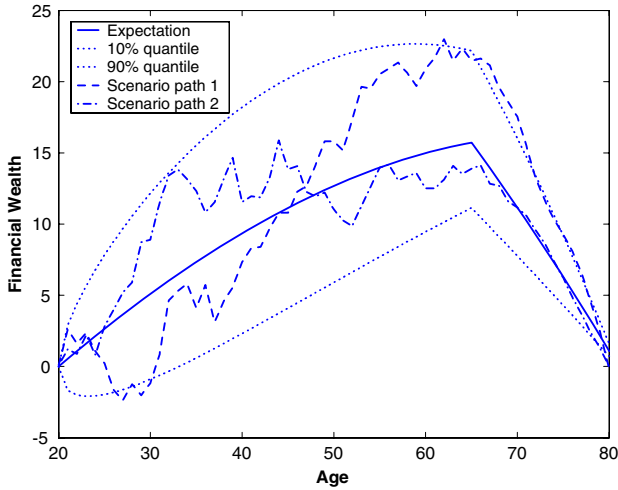


Figure 6 – Expectation, 10% and 90% quantiles and two simulated scenario paths of the financial wealth path (F_t) over the lifecycle for the benchmark parameters ($r = 2\%$, $\rho = 2\%$, $\lambda = 0.2$)

in time, generations of different ages pay different premium rates even if the expected premium rates are constant over the working life (i.e. the right-hand side of (6) is zero). To illustrate, a generation that started to work at $t = 1$ participates in the risks that materialized between $t = 1$ and $t = 31$. A young-

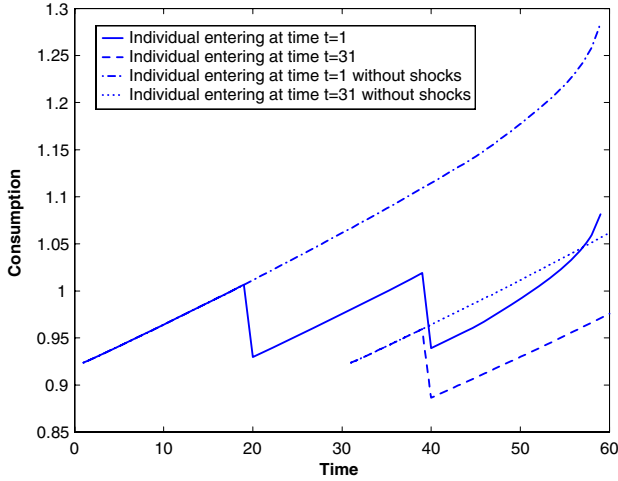


Figure 7 – Trajectories of the consumption paths (C_t) of individuals entering at respectively time $t = 1$ and $t = 31$. Shocks in the asset price are absent except at time $t = 20$ and $t = 40$, when a negative shock is imposed. All trajectories correspond to the benchmark parameters ($r = 2\%$, $\rho = 2\%$, $\lambda = 0.2$)

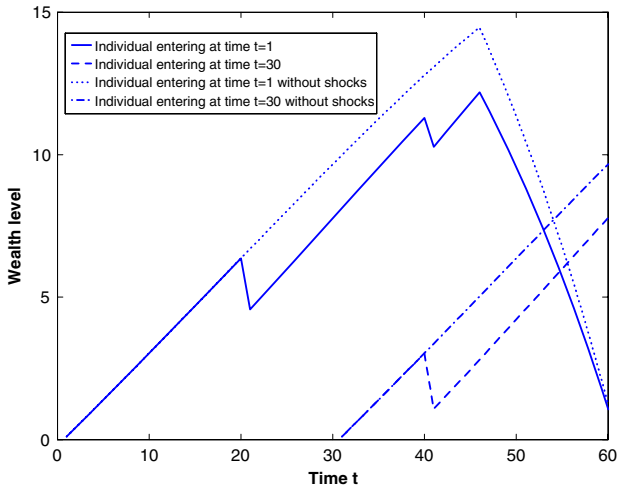


Figure 8 – Trajectories of the financial wealth paths (F_t) of individuals entering at respectively time $t = 1$ and $t = 31$. Shocks in the asset price are absent except at time $t = 20$ and $t = 40$, when a negative shock is imposed. All trajectories correspond to the benchmark parameters ($r = 2\%$, $\rho = 2\%$, $\lambda = 0.2$)

ger generation who enters the labor market at $t = 31$, in contrast, does not. The negative shock in the financial market exerts a proportional effect on consumption while the financial wealth of the younger generation is reduced by more (in absolute value) than that of the older generation.

4 SAVING AND INVESTING OVER THE LIFE CYCLE WITH CONSTRAINTS

4.1 *Introduction*

The first-best contribution and asset allocation strategies derived in the previous section require that individuals can borrow against their human capital. Moreover, individuals must be able to implement rather complicated saving and investment strategies that depend on age and the level of financial wealth. This section investigates how constraints faced by individual investors affect their strategies and welfare levels. In particular, after Subsection 4.2 considers the implications of annuities, Subsection 4.3 analyzes what happens if agents can not borrow against their human capital. Subsection 4.4 explores the implications of constant contribution rates and asset allocations. Subsection 4.5 analyzes the impact of other behavioral limitations to individual decision making, including possible underdiversification and infrequent rebalancing.

4.2 *Annuities*

To protect themselves against individual longevity risk, individual agents can buy annuities at retirement. Standard annuities typically do not allow taking investment risk during retirement. Individuals thus fail to exploit their risk-bearing capacity for shocks that occur after retirement. The risk exposure of retired individuals is too small: they use an implicit Sharpe ratio of zero (or an infinite implicit relative risk aversion). Shocks are thus absorbed only during working life. All capital is risk bearing as regards the shocks that occur during the working life. We find that the wrong risk exposure of the retired generation yields a welfare loss of 0.5% of ex-ante life-time utility in terms of certainty equivalent consumption.¹⁹ The utility loss is small²⁰ because of two reasons. First, shocks that occur during the retirement period are discounted heavily (since they occur late in life). Second, these shocks affect only a small part of overall wealth: most wealth has already been depleted at the time of retirement. In other words, the shocks during retirement can be smoothed

19 We measure welfare levels by the constant, certain consumption level that achieves the same utility level as the various stochastic consumption streams.

20 The losses are larger if we measure them compared to the certainty equivalent consumption that remains at the time of retirement (see Koijen et al. (2006a)).

only over a relatively short time period as the remaining life is relatively short. The rest of this section assumes annuities so that retired agents do not take investment risk.

4.3 *Borrowing Constraints*

As discussed in section 3.3, the first-best asset allocation implies that one should borrow at the beginning of one's career and invest the proceeds in the stock market to acquire sufficient exposure to the equity market. Adverse selection and moral hazard, however, typically preclude borrowing against future labor income. Financial institutions cannot use human capital as a collateral to ensure that the loan is paid back.²¹ This subsection therefore considers the more realistic case with an exogenous²² borrowing constraint so that equity exposure cannot exceed unity (i.e. $f_{it}^* \leq 1$). With these constraints, agents must get all their risk exposure from positive financial capital. Hence, in contrast to older workers who can accumulate substantial financial capital through saving, young workers cannot get enough exposure to risk.²³ The restricted access of younger workers to capital markets that is implied by the borrowing constraint harms their welfare and thus functions like an implicit tax on wealth. This negative income effects reduces consumption and thus boosts saving initially. Another reason why borrowing constraints raise saving is that additional saving allows the young worker to acquire more access to equity markets so that they can take more advantage of the equity premium.

In case of borrowing constraints, the optimal contribution rates and asset allocations can no longer be determined analytically so that we must rely on numerical simulations (see e.g. Kojen et al. (2007)). These numerical simulations assume that agents can trade only at an annual frequency. The transition from continuous to discrete trading (at annual frequency) yields a welfare loss of 0.8% of certainty-equivalent consumption.

The presence of the borrowing constraints implies an additional welfare loss of approximately 2.8% of certainty equivalent consumption. This welfare loss associated with borrowing constraints is thus substantially larger than the welfare loss due to annuities. At the beginning of life, limits on risk expo-

21 In principle investments in derivatives with extreme exposures (such as far out of the money call options) can be used to obtain adequate equity market exposure with small monetary investments only. However, very far out of the money derivatives are not traded in liquid markets. If the real estate market and the equity market are correlated, strategies based on investments in real estate financed by mortgages can be used as an alternative strategy to build substantial equity exposure even though investors do not own much financial wealth.

22 Cocco et al. (2005) show that exogenous borrowing constraints can be viewed as approximations to endogenous borrowing constraints derived from recent literature on endogenous credit-market imperfections.

23 See also Constantinides et al. (2002).

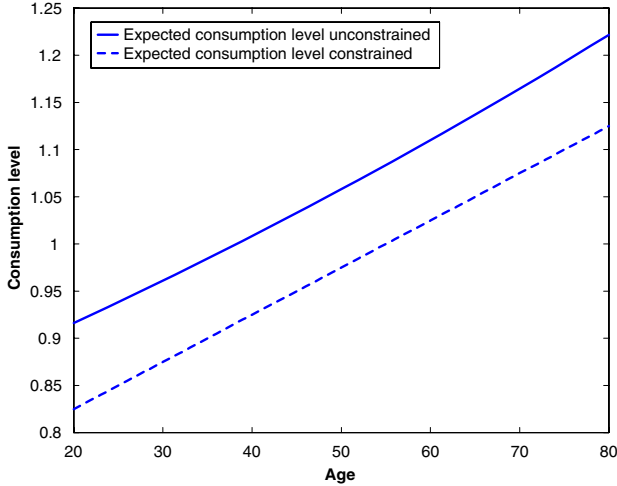


Figure 9 – Expected consumption level C_t under the benchmark parameters ($r = \rho = 2\%$, $\lambda = 0.2$) in the absence and in the presence of borrowing constraints

sure are more costly than at the end of life because at the beginning of life shocks can be smoothed out over a relatively long time period. The risk-bearing capacity of the young is particularly valuable also because shocks that are absorbed when young are not discounted much.

For our benchmark case, the quantiles of the optimal solutions with borrowing constraints are presented in Figures 9 and 10. Figure 9 presents the optimal consumption path. A comparison with the solid line in Figure 5 (which is added to Figure 9 for ease of reference) illustrates that the restricted access of younger workers reduces consumption levels and raises saving.

To put the welfare loss of borrowing constraints and other constraints into perspective, we compute the welfare costs of taking no risks at all during the life cycle. This is the most extreme form of limits on risk sharing. Individuals in effect do not have any risk-bearing capital at all. Teulings and de Vries (2006) compute the following approximation for the welfare loss (in terms of the relative decline in certainty equivalent consumption) associated with this extreme limitation on risk taking:

$$\frac{1}{4} \frac{\lambda^2}{\theta^i} D. \quad (7)$$

The intuition behind this expression is that the average duration of risky investments is half the lifetime D and that the welfare loss for a year riskless

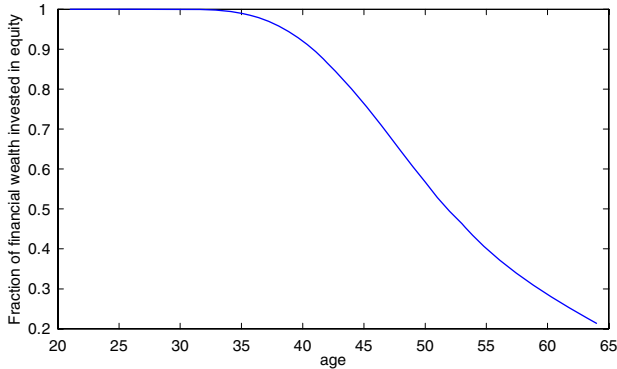


Figure 10 – Expectation of the optimal equity share in financial wealth f_i^* over the life cycle in the presence of a borrowing constraint case under the benchmark parameters ($r = \rho = 2\%$, $\lambda = 0.2$)

investment is given by half of the potential reward for risk²⁴ $\frac{1}{2} f_i \bar{\mu} = \frac{1}{2} \frac{\lambda^2}{\theta^i}$. The welfare loss thus rises with the reward to risk taking (i.e. the Sharpe ratio λ) and declines with the willingness to do so (i.e. relative risk aversion θ^i). With our benchmark parameters, this approximation yields a welfare loss of 12% (in terms of certainty equivalent consumption).

4.4 Time-Invariant Pension Premium and Asset Allocation

The savings and investment decisions require substantial investor sophistication. This section considers the welfare losses imposed by a number of simpler strategies in the case with borrowing constraints.

4.4.1 Defined Contribution

The first strategy we consider is a standard defined contribution (DC) solution where the premium rate is fixed independently of age or wealth. The level of this premium rate is determined optimally and the asset allocation optimally varies with age and shocks to financial wealth. For the benchmark parameters, the optimal consumption rate is 87% of labor income.

If the contribution rate is fixed *a priori* and does not respond to shocks, the part of wealth that is dedicated to pre-retirement consumption does not contribute to risk sharing. Only the part of wealth that is dedicated

24 Only half of the reward to risk is actually a welfare gain. The reason is that the individual also bears additional risk, which harms welfare. In fact, one can interpret the welfare loss as a Harberger triangle due to a tax on risk taking. The tax on risk taking is given by $\bar{\mu}$ (since the reward for taking risk is taken away), while the behavioral response (as a share of overall wealth) is given by f_i . The Harberger triangle is thus given by $\frac{1}{2} f_i \bar{\mu}$.

to post-retirement consumption is exposed to shocks, which implies that risk exposure is not optimal and that the individual does not fully take advantage of the risk premium. The elasticities of consumption with respect to wealth shocks indicate inefficient consumption smoothing. Indeed, consumption during the working life does not react to wealth shocks at all. The welfare loss of this individual DC plan is 3% (in certainty equivalent consumption) relative to the optimal individual plan (with borrowing constraints, annuities and discrete trading) in Subsection 4.3. Compared to the first best in section 3, the welfare loss is thus a little more than 7% (in certainty equivalent consumption) because the optimal individual scheme yields a welfare loss of 4.1% relative to the first best (0.5% due to annuities, 0.8% due to discrete trading, and 2.8% due to borrowing constraints). This welfare loss of 7% can be compared to the case when the individual is entirely precluded from investing in equity (and not just as regards the part of wealth that is dedicated to post-retirement consumption) but sets an optimal age dependent contribution rate. This welfare loss is 8.5%. The welfare losses of the constraint that the DC plan imposes on risk taking before retirement is particularly large if a small coefficient of risk aversion (and a large Sharpe ratio) indicates that young workers would like to have taken risk on their human wealth dedicated to before-retirement consumption.

Just as in the optimal individual system (see Figure 10), in the DC system the share of financial wealth invested in equity declines during the working life. The reason is that pension wealth (i.e. the discounted value of contributions) is gradually transformed into financial wealth so that financial wealth tends to increase during the working life. With suboptimal consumption smoothing under the DC system, however, the absolute amount of equity at the beginning of life is substantially smaller as pension wealth (i.e. the discounted value of the fixed pension contributions, which are equal to the discounted value of retirement consumption) is significantly smaller than human capital (i.e. the discounted value of wages, which is equal to discounted consumption during not only retirement but also the working life). Indeed, a pension system that excludes part of life-time consumption from absorbing risks limits the risk-bearing capacity of individuals.²⁵

25 For that same reason, defined-benefit (DB) systems, which fix the retirement benefit, limit the the risk-bearing capacity of individuals. Compared to DC systems, they damage the risk-bearing capacities of old agents relatively more. Indeed, compared to DC systems, DB schemes start to constrain the risk-bearing capacity of individuals more at the age at which the value of pension wealth (the discounted value of retirement consumption) starts to exceed the discounted value of consumption during the remaining working life). The larger is the working life compared to the inactive life, the larger are the welfare losses from a DC scheme compared to those of a DB scheme.

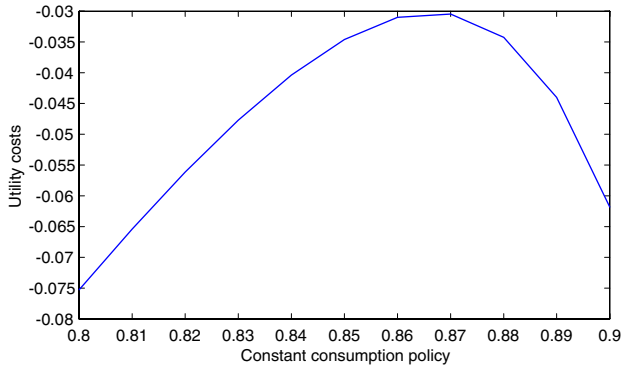


Figure 11 – Welfare loss due to suboptimal constant saving rates. The asset allocation is optimally adjusted given the constant saving rates

Figure 11 indicates how the welfare losses on account of defined-contribution systems depend on the selection of the optimal contribution rate. The optimal pension contribution is 13% (corresponding to a savings rate of 87%) If the savings rate during the working life is set 7%-points too high (at a consumption rate of 80%) an additional welfare loss approximately of 4.5% (in terms of certainty equivalent consumption) is incurred.

4.4.2 Fixed Equity Share of Financial Wealth

The financial planning of the investor can be simplified also by imposing a constant portfolio allocation in financial wealth. Figure 12 shows how the welfare level varies with the constant portfolio allocation. An equity share of 53% of financial wealth minimizes the welfare loss from a fixed portfolio allocation. The welfare loss is 1.7% compared to the optimal strategy of an individual facing borrowing constraints. This loss is smaller than the costs of borrowing constraints. This again reveals that constraints on risk taking are especially large at the beginning of life when the individual is especially wealthy.

If we set not only a constant asset allocation but also a constant contribution rate, the optimal share of financial wealth invested in equity declines to 30%. The reason is that individuals no longer employ consumption when working as a risk absorber so that they can afford to take on only a limited amount of risk. The welfare loss in this case relative to the DC case where the asset allocation was optimized dependent on age and wealth is only 0.4%. The welfare loss is relatively small because large risk taking in the beginning of life is no longer optimal if young consumption cannot adjust to shocks. A constant asset allocation through life is thus less costly.

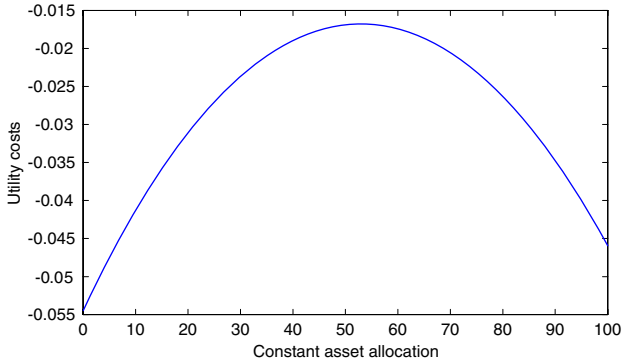


Figure 12 – Welfare loss due to suboptimal constant portfolio choice (i.e. equity as share of financial wealth). The savings rate is adjusted optimally given the constraints on the portfolio choice

4.5 *The Ability of Individuals to Choose their Optimal Pension Contract*

Section 3 considered the first-best life-time financial plan and Subsection 4.3 explored the optimal individual contract in case of borrowing constraints. In both cases, we assumed that the contract was optimally tuned to the preferences of the individual. Subsection 4.4 analyzed how robust the optimal solution is to the use of static rather than dynamic contribution rules and asset allocation strategies. This analysis suggests that the individuals suffer substantial welfare losses if they lack the expertise to select and implement optimal savings and investment strategies during their life time. As a direct consequence, adequate financial advice and access to the structured products offered by financial intermediaries are quite valuable in order to help individuals with their life-cycle financial planning. A large literature on how individuals actually decide on saving and investment over their life time has developed recently. This subsection provides a short introduction to this literature.

Individuals spend little time and effort on planning for retirement. Van Rooij et al. (2007) find that individuals in the Netherlands are not interested in pension related issues and think they are not able to take pension related decisions. Lusardi (1999) reports that even in the US, where individuals have many more pension related decisions to make, one-third of the workers had “hardly thought about retirement” only ten years before retirement. Ameriks et al. (2003) reports, also for the US, that more than 50% of a sample of highly educated wealthy individuals had “not spent a great deal of time developing a financial plan”. Lack of planning has serious consequences: those who do not plan have lower wealth holdings and are less likely to report that they experience a satisfying retirement. There are many recent papers that

analyze the impact of improved information (e.g. through benefit information fairs) and financial education (see e.g. Venti (2006); Clark et al. (2006)). At best, the results indicate small improvements in the effort spent on pension related decisions and on the quality of these decisions. Choi et al. (2005) report that tax-incentives for retirement saving were not utilized, even after they had been explicitly pointed out in financial training courses.

Individuals selecting their pension contract face two decisions: how much to save and how to invest the financial wealth that is accumulated. The existing evidence, which is primarily on the US, indicates that individuals undersave and that they experience unanticipated drops in consumption at retirement. Moreover a number of surveys have found that the vast majority of individuals thinks that they should be saving more for retirement (Laibson et al. (1998); Choi et al. (2002)). One explanation that is well documented in psychology is that people lack the self control that is required to implement a savings plan. People want to save for the future but they lack the capacity to carry out their intention. A convenient way to model actual behavior is hyperbolic discounting. This model assumes that nearby discount rates are much larger than long-term discount rates. As a consequence, consumers exhibit time-inconsistent behavior while actual behavior diverges from planned behavior.

With respect to the asset allocation, an even more extensive literature indicates that individuals seem to make significant mistakes. Many households take decisions that are difficult to reconcile with the advice given to them by financial planners or other experts. Of course it could be that investors behave optimally but should be characterized by non-standard preferences such as loss aversion or habit formation. These extensions of the basic model will be discussed in section 6. This section takes the standard preference specification introduced in section 2 as starting point and interprets deviations in observed behaviour as investment mistakes.

A well-established stylized fact in the literature is that only a small part of the population in countries like Italy (Guiso and Japelli (2005)) or the US (Haliassos and Bertaut (1995)) hold stocks, directly or indirectly. Participants in retirement saving plans rarely rebalance their portfolio or alter the allocation of their contributions over the life cycle (Ameriks and Zeldes (2004)). Many households do not diversify and hold only a few stocks, often with a local bias and in 401(k) schemes often even with a bias to the stock of their own employer (Huberman (2001); Mitchell and Utkus (2003); Munnell and Sunden (2004)). Participants in 401(k) plans display a tendency to split their contributions evenly among investment options, irrespective of the type of options that is offered (Benartzi and Thaler (2001)). Such a "1/n"-rule of thumb is clearly at odds with optimal diversification and adequate risk taking. Financial education could play a role. Campbell (2006) reports that more

educated households in Sweden diversify their portfolio more efficiently than less educated households do.

The overwhelming evidence that individuals have difficulties in taking adequate savings and investments decisions raises the question how to address this problem (see also Kooreman and Prast (2007)). One option is to force mandatory participation in a collective scheme where decisions are taken by professionals. One of the main drawbacks (see also section 5) is that collective schemes typically do not take into account individual characteristics and preferences. Moreover, important decisions in these schemes are usually taken by pension fund trustees and the evidence on the competence of trustees suggests at least that it is heterogeneous (Clark et al. (2006)). Another option is to provide adequate financial advice and to develop innovative products that stimulate individuals to take adequate pension related decisions. A well known example is the “Save More Tomorrow” plan proposed by Thaler and Benartzi (2004). The commitment mechanism allows individuals to commit to increasing their savings rate at some later date. Other examples include the use of automatic enrollment in savings plans, automatic payroll deduction in employer-based plans and the use of adequate defaults with respect to savings rates and asset allocation. Yet another example is the use of peer effects by setting up benefits information fairs for some workers, which has been shown to have an impact also on the saving behavior of colleagues that had not been invited to attend (Duflo and Saez (2003)). No doubt in the years to come many other structured plans and products will be developed that aim to facilitate individual decision making.

Figure 13 illustrate the potential costs of sub-optimal individual behavior due to behavioral biases. In particular, this Figure contains the welfare losses if an individual adopts wrong risk aversion and time-preference parameters when selecting structured products to implement the optimal investment and saving decisions in individual plans (i.e. with borrowing constraints and annuities) derived in Subsection 4.3. Figure 13 is based on the assumption that product that is offered is based on a risk aversion of 5% and a time preference parameter of 2%. The costs of inadequate financial advise, i.e. of adopting strategies that are optimal for individuals with other preference parameters can be substantial. The costs of the inadequate assumptions on risk preferences²⁶ are particularly large if the individual turns out to have a time preference that substantially exceeds the time preference on which the financial plan is based. The welfare loss can be as large 4.0% (in terms of certainty equivalent consumption) for a risk preference of 4% rather 2% and even sub-

26 With a CRRA utility function, a wrong risk aversion parameter implies that also the intertemporal substitution elasticity is not adequate. Hence, it distorts both risk taking and intertemporal consumption smoothing.

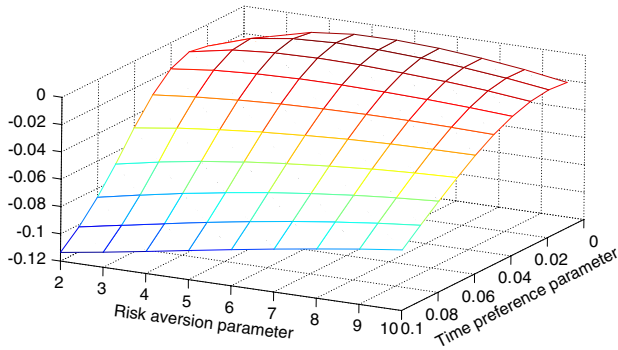


Figure 13 – Welfare costs of suboptimal individual behavior due to incorrect risk aversion and/or time preference assumptions

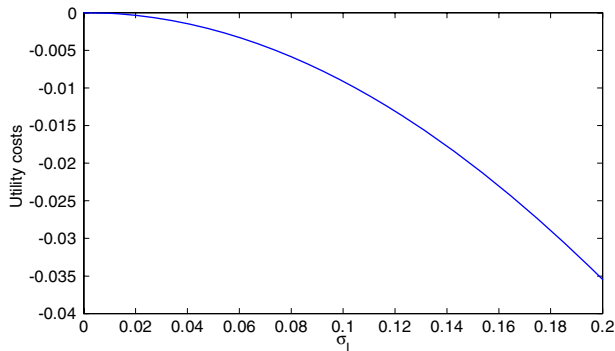


Figure 14 – Welfare costs of idiosyncratic risk taking

stantially larger for more extreme deviations between the assumed and actual preference parameters.

In the result so far we assumed that agents hold optimally diversified portfolios when taking equity exposure. In reality individuals will often hold a few stocks only which implies that idiosyncratic risk is taken that does not generate an equity premium. Figure 14 illustrates the welfare costs of carrying idiosyncratic risk. If both the level of systematic and idiosyncratic risk equals 20% the over-all volatility of the equity investment is 28.2% with an associated 3.5% welfare loss relative to an optimally diversified equity portfolio with a 20% volatility

5 COLLECTIVE PENSION SCHEMES

5.1 *Introduction*

Sections 3 and 4 considered the optimal individual pension contract in the first-best case as well as under a number of more realistic constraints.

This section turns to the cost and benefits of mandatory, collective pension schemes.

A number of arguments have been put forward as advantages collective pension schemes:

- (i) Collective pension schemes enable young participants to alleviate borrowing constraints
- (ii) Collective pension schemes are cost efficient
- (iii) Collective pension schemes enable risk sharing between non-overlapping generations
- (iv) Collective pension schemes complete markets by arranging trades of assets that are not traded on financial markets
- (v) Collective pension funds take over decisions of boundedly rational participants

Other arguments can be put forward as disadvantages of collective pension schemes:

- (i) Collective pension schemes are not tailor made to the characteristics of the individual
- (ii) Collective pension schemes impose uniform contribution rates and asset allocations on heterogeneous participants
- (iii) Collective pension schemes involve suboptimal consumption smoothing as shocks are not spread out over the remaining life time
- (iv) Collective pension contracts are incomplete and not transparent
- (v) Collective pension funds reduce competition between funds

The first argument implies that collective pension schemes allow individuals to implement their optimal saving and investment plans in a more optimal way than individual pension contracts. The reason is that compulsory participation in collective pension schemes combat adverse selection and moral hazard if young workers borrow against the unsecured collateral of their human capital. The second argument in favor of collective schemes is that their costs are usually lower than the costs associated with individual pension schemes. In particular, collective schemes avoid time consuming search incurred by individuals and marketing cost by financial intermediaries. Subsection 5.2 elaborates on the first two arguments in favor of collective schemes.

Mandatory collective pension schemes also allow risks to be shared with future generations, i.e. generations that do not yet participate in the labor market at the time a shock materializes. Subsection 5.3 shows that the benefits of additional intergenerational risk sharing are potentially substantial if all generations are committed to the contract; in this case, we can abstract from the so-called discontinuity risk that generations entering the labor market decide not to enter the contract if they have to make up a deficit. This

subsection considers also the welfare gains of intergenerational risk sharing in case the potential losses for newly entering generations in case of adverse shocks are limited through dynamic asset allocation strategies.

An important drawback of collective pension schemes is that they are not tailor made to the preferences or specific circumstances of each individual. As discussed in Subsection 4.5, the use of the wrong preference parameters generates a substantial welfare loss compared to the case in which the individual could have optimally implemented a pension contract with the adequate specifications. Indeed, individual characteristics such as the presence of partner pension or non-pension wealth are typically not taken into account in collective pension schemes.²⁷ Subsection 5.4 explores the welfare losses due to uniform pension contracts that are not tailor-made to individual features.

Subsection 5.5 explores the implications of the collective pension contracts that are implemented in practice for welfare, consumption smoothing and individual risk exposures. The advantages of these schemes are that they allow for risk sharing with future generations and that they allow young agents facing borrowing constraints to acquire substantial exposure to equity risk. The disadvantage is that these schemes involve suboptimal intertemporal consumption smoothing as shocks are not spread out over the remaining life time. Moreover, the risk exposures of participants are not optimal as they depend the funding rate of the scheme as a whole rather than on the wealth position of the individual.

Collective pension schemes have other advantages and disadvantages that we do not explore further in this section. In the benchmark model, the single financial risk factor is the only risk factor and this risk factor is traded in frictionless markets. In more realistic economies, also other risk factors that can not easily be traded in capital markets (such as price inflation, wage inflation and longevity) are important. A potentially important advantage of collective pension schemes is that risks that can not be traded in capital markets can be shared through the pension contract. In incomplete financial markets, pension contracts may create new, non-tradable financial assets that can enhance efficient risk sharing.

This section assumes that collective pension contracts are complete, i.e. all pay-offs in future contingencies are set in advance. In reality, however, many collective pension contracts are incomplete and the property rights on the surplus of the pension fund are often unclear. In particular, the board of the pension scheme has the discretion to pay out the surplus as dividends to current stakeholders (in the form of additional pension benefits or additional future benefit claims that are not fully covered by additional premia) or to

27 Also individual differences in the ability to postpone retirement or in life expectancy (e.g due to socioeconomic characteristics) are important determinants of optimal individual pension contracts.

save the surplus for future participants. This discretionary power of the trustees has the advantage that all the possible contingencies do not have to be thought through *ex ante*, thereby reducing transaction costs. A disadvantage is that the collective decision making within the board of the scheme may give rise to political risks and less secure individual property rights, thereby harming the welfare of risk-averse participants. Indeed, budget constraints apply to the pension scheme as a whole rather than to individual participants. Hence, a collective pension scheme can redistribute market value across generations as participants do not hold tradable individual property rights on the pension fund. Rather, they have non-tradable claims on the pension scheme when they work in the sector or have accumulated pension rights in that sector. These claims are not necessarily matched with tradable financial assets owned by the pension scheme. Moreover, the market-value of these claims may be affected by the discretionary decisions by the board of the pension scheme (see Hoevenaars and Ponds (2006)). Indeed, the pension fund is an independent legal entity, which has the discretionary power to shift resources between generations.

5.2 Borrowing Constraints and Cost Efficiency

As discussed in section 3.3, the first-best asset allocation implies that one should borrow at the beginning of one's career and invest the proceeds in the stock market to acquire sufficient exposure to the equity market. As discussed in section 4.3, when trying to implement the first-best solution derived in section 3.3, young workers typically face borrowing constraints. The reason is that the lender faces substantial credit risk on an unsecured loan due to adverse selection and moral hazard. Indeed, if the stock market crashes, the lender does not have collateral. Compulsory participation in collective pension schemes can alleviate adverse selection and moral hazard when young workers borrow against their human capital to invest in equity. In effect, mandatory participation helps to secure the human capital of younger generation as a collateral, thereby limiting bankruptcy risk and thus relaxing the credit constraints faced by young generations. A collective defined-benefit pension fund that carries mismatch risk by investing in equity can be viewed as younger participants issuing risk-free debt to the older generations and using the borrowed funds to invest in the equity market. Indeed, the mismatch risk of the pension scheme is born by the younger generations in the form of volatile recovery premia. In fact, the risky pension contribution allows the young to transform their riskless human capital in a risk-bearing asset. In this way, they acquire the optimal exposure to equity risk (see Beetsma and Bovenberg (2007)). Section 4.3 reported that for the benchmark parameters the welfare gains of being able to completely avoid the borrowing constraints are 2.8% of certainty equivalent consumption.

TABLE 1 – WELFARE LOSSES DUE TO ANNUAL IMPLEMENTATION COSTS RELATIVE TO FRICTIONLESS FIRST-BEST SOLUTION (from section 3)

	Annual implementation costs (%)			
	0%	0.3%	0.5%	1.0%
First best	0%	-1.2%	-2.0%	-4.0%
rule of thumb	0%	-1.6%	-2.8%	-5.6%
constrained	-3.6%	-4.6%	-5.3%	-7.0%

Relieving the borrowing constraints through compulsory participation is not costless if, in contrast to our benchmark model, labor supply in the sector to which the compulsory participation applies is endogenous. In that case, mandatory participation gives rise to labor-market distortions. In particular, agents reduce their labor supply to the sector if adverse developments in equity markets saddle them with negative pension wealth (i.e. pension debt) that they have to finance with recovery premia on basis of their labor supply. Sectoral labor supply can be reduced by moving to other sectors or countries, by becoming self-employed, by working part-time or by retiring early.

Mandatory collective pension schemes allow cost-effective pension solutions because of economies of scale as well as lower marketing and search costs. Bikker and De Dreu (2006) report a cost average of 0.30% per annum for medium sized and large collective pension schemes, whereas mutual funds and insurers typically charge 0.80–1.00% per annum or more for individual plans (see e.g. Poterba et al. (2005)). So far, we have ignored implementation costs. Table 1 computes the welfare losses (in terms of certainty equivalent consumption) of the various costly pension deals relative to the case in which implementation costs are negligible. In order to derive the welfare loss of implementation costs, the annual return on both riskless and risky assets is decreased by the annual implementation costs. The welfare costs (in percentage changes of certainty equivalent consumption) associated with a fixed lower annual return of x percentage points on financial wealth can be approximated by

$$\frac{x}{2}(D - T), \quad (8)$$

where T stands for the length of the working period so that $D - T$ equals the length of the retirement period. Note that in the constrained case both borrowing constraints and inadequate risk taking after retirement due to an insufficiently rich annuity menu are taken into account.

5.3 *Risk Taking Before Working*

Section 3 considered the optimal pension contract for an individual who enters adult life, i.e. enters the labor market and starts to consume out of his own labor income. As shown by Teulings and de Vries (2006), this contract can be improved further if we allow for the possibility that agents assume financial risks already before they are adults. Indeed, an individual is already endowed with human capital at birth and can in theory then already participate in stock-market risks by borrowing against his human capital to invest in the stock market. Teulings and de Vries (2006) demonstrate that the potential welfare benefits of participation in the stock market before entering the labor force are substantial.

Teenagers and babies do not usually directly invest in equity markets. However, compulsory pension schemes allow them to participate in the stock market in an indirect way. Collective pension schemes share risks with generations that do not yet participate in the scheme by transferring buffers (which may be negative) from current participants to future generations. A buffer can be defined as the assets in the scheme minus the assets of the generations that have already started to work and pay pension premiums. A buffer thus corresponds to the aggregate financial wealth of future workers. These workers have in fact borrowed to buy stocks, thereby participating in stock market risks even before they start to work (see also Beetsma and Bovenberg (2007)). In particular, collective pension schemes typically levy equal contributions on all active generations at a particular point in time with the level of the uniform contribution depending on the funding rate of the collective scheme. In this way, the premium cohorts pay when they enter the scheme depends on the investment history of the scheme before they entered and thus on the shocks that materialized before they started to work.

Through the transfer of buffers between generations a collective pension scheme in fact helps to complete capital markets by allowing younger generations (and even unborn generations) to buy risk-bearing assets. In decentralized capital markets, individuals can obviously buy risk-bearing assets only after they have been born. They are thus not able to trade risks that occur before they are born. In other words, decentralized capital markets allow only overlapping generations (i.e. generations that are in the market at the same time) to trade risks that emerge during the rest of their life times. Risk sharing among non-overlapping generations thus has to occur through other institutions, such as compulsory pension schemes but also tax and public debt policy.

It is theoretically optimal to diversify risk over as many generations and as many time periods as possible. Only generations who consume after the shocks occur can possibly share in the risk: generations who are deceased when the shock occurs have depleted their wealth already and thus do not

have any risk-bearing capacity any more. Future generations who have not yet started consuming, in contrast, can share in current risks. Optimal risk sharing implies that a current shock affects consumption during the rest of the life of all living *and future* generations by that percentage. The shock is thus smoothed out as widely as possible across both individuals and periods so that each separate individual in each period is affected as little as possible. Collective pension schemes can thus be interpreted as the pension scheme having invested on behalf of all its current contributors before these contributors actually started to work and pay a pension premium.

In our partial equilibrium setting, the benefit of additional risk sharing is not to reduce risk but rather to increase the expected payoff from risky investments by generations who have not entered the labor market yet.²⁸ In a general equilibrium context in which the supply of risk-bearing assets is less than infinitely elastic, the additional risk-bearing capacity of younger generations reduces the price of risk. This, in turn, reduces the risk born by the current generations.

With intergenerational risk sharing across non-overlapping generations, the expectation of the buffer that is transferred to future workers is positive but can be negative for some bad draws. This risk sharing thus works only if future generations are forced to pay back their unsecured loans in these cases. Mandatory participation in sectoral pension schemes helps to enforce this. In this way, the government allows pension schemes to lay a claim on the collateral of workers: their human capital.

Just as in the case of relieving the borrowing constraints of young workers, mandatory participation is a necessary but not a sufficient condition for making risk sharing among non-overlapping generations feasible by securing the collateral of future workers. In particular, employers in a sector face competition from other sectors on the labor market. With mobile labor across sectors, these employers may have to pay compensating wage differentials to attract workers to their sector if these workers are forced to pay back implicit debt extended by the sectoral pension scheme. In that case, workers can in fact evade the compulsory participation and shift the implicit burden on employment in the sector by seeking employment in other sectors. As a direct consequence of the mobility of young workers, the buffers are in fact benefiting the shareholders of the firms in the sector (in case of tradable goods), the consumers of the commodities produced by the sector (in case of non-tradable goods such as health care), or the current retirees in the pension scheme (in case of mobile workers, capital and consumers). Indeed, depending

28 Gollier (2006) stresses this result and formulates a simple model of risk sharing between two agents facing independent risks. In an intergenerational context, the young can share in the risks of the old. The old, however, do not share in the risks of the young. The possibilities for sharing risks among non-overlapping generations are thus asymmetric.

on the various elasticities, competition limits the possibilities for risk sharing among non-overlapping generations, especially if behavior is sensitive to costs (as indicated, for example, by the elasticity of sectoral labor supply).

If young workers face substantial negative buffers on entry into a pension scheme, they may vote with not only their feet (by seeking employment elsewhere or reducing labor supply) but also their voice. In particular, they may encourage the pension scheme to default on the pension obligations to older workers and retirees, for example by no longer indexing their pensions to inflation. In addition to the economic considerations of competition on labor, capital and commodity markets, political economy considerations thus make current participants vulnerable to large negative buffers. This is why regulators adopt the so-called discontinuity principle: a pension fund should be able to comply to its obligations also if, starting today, no new generations would be willing to enter the scheme. Recognizing that pension schemes cannot easily secure the human capital of future generations, this discontinuity principle limits the scope for risk sharing among non-overlapping generations.²⁹

Risk sharing among non-overlapping generations can probably best be implemented by the largest pool we have to enforce this risk sharing: the national state. In particular, in combination with public debt, the government can employ its tax power to commit future generations to share risks with current generations. To illustrate, by issuing indexed bonds and longevity bonds, the government protects current generations against inflation and longevity risks and shifts these risks in part to future generations of tax payers. Even the national state, however, is restricted by international mobility and other behavioral responses to taxes (e.g. the wage elasticity of labor supply). Moreover, in a democracy, the state can commit only imperfectly. Young generations can always exert political pressure to change laws that force them to pick up the tab of risks that the older generations have shifted unto them. Similarly, current generations, who control the votes, can always consume the buffers that were initially assigned to the future generations. This illustrates how difficult it is to actually implement and enforce credible contracts between non-overlapping generations.

Table 2 illustrates the potential welfare gains of intergenerational risk sharing as well as the distribution of the buffer that new generations face upon entry to the labor market. The first column of Table 2 reports the welfare gain (in terms of certainty equivalent consumption) when a cohort starts to optimally participate in equity-market risk B years before entering the labor mar-

29 Also positive buffers may in fact benefit current rather than future generations because of political or economic reasons. In particular, if buffers are large, current generations (and especially the older generations) may put pressure on the fund to use these buffers to improve their pensions. Moreover, young workers are willing to work for the sector at low wages, thereby benefitting shareholders and consumers.

TABLE 2 – WELFARE GAINS OF INTERGENERATIONAL RISK SHARING WITH FULL AND REDUCED DISCONTINUITY RISK

Years	Optimal (%)	Buy-and-hold with-out options (%)	with- Maximum loss 2 years (%)	Maximum loss 1 year (%)	Maximum loss 6 months (%)
1	0.5	0.4	0.4	0.4	0.3
5	2.0	1.9	1.8	1.5	1.0
10	4.1	3.7	3.2	2.6	1.9
15	6.2	5.3	4.6	3.6	2.6
20	8.3	6.9	5.9	4.5	3.3

ket compared to the first-best contract explored in section 3 (in which case B is in fact 0 because agents do not participate in risks before they enter the labor market). These calculations assume that the optimal exposure is acquired immediately at that time and that the pension scheme can costlessly force a generation to pay back possibly negative buffers (i.e. pension debt) when entering the labor force. The welfare gain from optimally participating in stock-market risk B years before entering the work force can be approximated by³⁰ (see Teulings and de Vries (2006))

$$\frac{1}{2} \frac{\lambda^2}{\theta} B. \quad (9)$$

The intuition behind this expression is that the duration of the risky investments is B and that the welfare gain for a year of optimal participation in the stock market is given by half of the potential reward for risk³¹ $\frac{1}{2} f \bar{\mu} = \frac{1}{2} \frac{\lambda^2}{\theta}$. The welfare loss thus rises with the reward to risk taking (i.e. the Sharpe ratio λ) and declines with the willingness to do so (i.e. relative risk aversion θ).

Table 2 shows that participation before entry to the labor market can raise welfare substantially. To illustrate, participation 15 years before entry to the labor market increases welfare by 6.2%. As discussed before, the main drawback of intergenerational risk sharing is discontinuity risk: new generations facing substantial deficits when they start working may decide not to participate in the contract. The probability that the individual will start adult life with a loss is 35%, the probability that the loss exceeds two working years is 28%.

The credibility of intergenerational risk sharing can be improved if the discontinuity risk is reduced through dynamic investment strategies that reduce

30 This equation is closely related to Equation (7). Note that the average length of the investment period is different.

31 Only half of the reward to risk is actually a welfare gain because the individual also bears additional risk, which harms welfare.

TABLE 3 – PROBABILITY THAT THE MAXIMUM LOSS OF DYNAMIC STRATEGY TO SHARE RISK WITH NON-OVERLAPPING GENERATION IS REACHED

Y	Maximum loss 2 years (%)	Maximum loss 1 year (%)	Maximum loss 6 months (%)
1	3	25	43
5	15	29	57
10	17	49	68
15	14	42	58
20	11	29	45

the probability of substantial funding deficits when young cohorts start their career on the labor market. We model these investment strategies as buy and hold investments in equity as well as put options that limit the maximum loss to either 24, 12 or 6 months of salary. The fraction of annual income that is borrowed B years before entering the labor market is optimized. Tables 2 and 3 report the welfare gains of such strategies compared to the first-best case of section 3, as well as the probability that the maximum loss is incurred.

Table 2 also shows that in case of participation 15 years before entry to the labor market a welfare gain of 2.6% is left if new generations are only willing to participate in the contract if their initial funding deficit is limited to six months income. Table 3 shows that the probability that this maximum loss will materialize is 58%. If new participants set two years of salary at risk, the welfare gain amounts to 4.6%, while the probability that as much as two years of salary will be lost is 14%. The investment strategy associated with these guarantees closely corresponds to solvency rules imposed on collective pension schemes by supervisors in order to ensure that the discontinuity risk does not become large.

Also Gollier (2006) has considered the benefits of intergenerational risk sharing. He puts a constraint on how large the negative buffer can become by requiring that at any point in time, the fund must be large enough to repay the contributions made in the past. This in fact puts a constraint on how large the negative buffer can become in absolute value. The solvency constraint thus reduces the ability of the fund to time diversify the portfolio risk. This has consequences for optimal saving and investment. As regards saving, the solvency constraint strengthens precautionary saving, especially if the fund is close to the solvency constraint. Indeed, the fund then puts aside additional funds to escape the limit on intergenerational risk sharing; the financial reserves of the fund in fact act like a buffer stock. As regards investment, the fund invests mainly in bonds if it has a small amount of assets and is thus close to the solvency constraint. Intuitively, the pension fund constructs a put option by trading dynamically and exchanging its equity for

bonds if equity declines in value in line with a dynamic strategy that replicates the put option. The fund thus behaves as if it exhibits decreasing absolute risk aversion. This behavior of selling equity when the stock market goes down contrasts with an individual who invests optimally in the absence of liquidity constraints. Such an investor buys additional equity to keep constant the share of equity in overall wealth.³² Gollier (2006) shows that the welfare gains from intergenerational risk sharing are substantially reduced when solvency constraints are imposed.

5.4 *Welfare Losses Due to Uniform Premium Rates and Asset Allocations*

So far we have assumed that the collective pension scheme offers every participant the pension contract that fits personal references and circumstances. In reality, pension schemes usually impose identical contributions rates and asset allocations for all participants even though agents typically differ substantially in their rate of time preference ρ^i and risk aversion θ^i (see e.g. Cocco et al. (2005)). Uniform and age independent contribution rates and asset allocations are optimal only under very stringent assumptions. To illustrate, in the simple model without risk discussed in section 3.2, the optimal contribution rate is constant if the real interest rate and the rate of time preference coincide. Equation (1) shows that consumption is constant over time in that case so that, with constant labor income, the optimal saving rate and thus the optimal pension premium is also constant.³³ The optimal share of financial wealth invested in equity is age-independent only if human wealth has the optimal exposure to the risk factor given by (4). This is not the case in our benchmark model with riskless human capital. Uniformity of the asset allocation over agents with different preferences requires even stronger restrictions on the parameters.

With a stationary population, one can aggregate over the various age cohorts to arrive at a first-order approximation for the ratio between aggregated wealth of all participants in the pension scheme to aggregate saving (see Teulings and de Vries (2006)):

$$\frac{\bar{W}_t}{\bar{F}_t} \simeq \frac{D}{D - T}, \quad (10)$$

32 Also this investor decreases the absolute size of its equity holdings as overall wealth declines.

33 In models that include wage growth, the optimal saving rate out of labor income (i.e. the optimal premium rate) remains constant over the life cycle and thus does not depend on age if the optimal growth rate of consumption equals the growth rate of labor income. This implies a particular relationship between the growth rate of wage income, the interest rate and preference parameters (in particular, the intertemporal elasticity of substitution and the rate of time preference).

TABLE 4 – WELFARE LOSSES WITH RESPECT TO THE FIRST-BEST SOLUTION IN THE CASE IN WHICH CONSUMPTION AND INVESTMENT DECISIONS ARE BASED UPON THE WRONG RATE OF TIME PREFERENCE ρ OR THE WRONG PARAMETER OF RISK AVERSION θ . THE ACTUAL PREFERENCE PARAMETERS ARE GIVEN BY $P=0.02$ AND $\theta=5$

	Rate of time preference ρ			
	1%	2%	3%	5%
Risk aversion θ				
2	-26.7%	-31.2%	-40.3%	-74.8%
3	-4.6%	-4.9%	-6.8%	-16.3%
5	-0.3%	0%	-0.2%	-2.5%
10	-2.5%	-2.3%	-2.3%	-2.7%

The actual preference parameters are given by $\rho=0.02$ and $\theta=5$

where a bar represents the value for a variable aggregated over all participants of a pension scheme. The ratio of aggregate overall wealth to aggregate financial wealth thus approximately corresponds to total life time divided by the length of the retirement period. Using (9) and (4) in (5), one can approximate the investment share in equity for a pension scheme of participants with homogeneous preferences. On the basis of our benchmark values, we find that a pension scheme should invest 80% of its financial assets in equity.

We do not address the optimal asset allocation of a collective pension scheme with heterogeneous agents, which in addition to the age composition of the scheme depends on the distribution of preference parameters as well as relative weights in the collective welfare function. Rather, in line with the analysis in Subsection 4.5, we compute the welfare losses if consumption smoothing and risk exposures are based on the wrong preference parameters. Subsection 4.5 reported results for borrowing constrained investors. We interpreted these results as misspecified individual contracts because of misperceived preferences and poor investment advice.

Table 4 presents similar results for the case in which collective pension schemes succeed in alleviating borrowing constraints for young, working generations (but do not allow future generations to participate in equity-market risk). The table shows the impact of using the wrong risk aversion parameter or rate of time preference if constraints (including borrowing constraints) are absent so that premium and portfolio shares can adjust flexibly to shocks. These simulations indicate what happens if a collective scheme cannot tailor its behavior to the actual preferences of each individual participant, for example because it does not have access to information about the preferences of the individual (since the individual does not implement its own decisions but

has delegated these to the pension scheme). These losses assume that individuals do not have any access to capital markets outside the pension scheme.³⁴ Furthermore, the assumed CRRA utility functions imply that the coefficient of risk aversion is equal to the inverse of elasticity of intertemporal substitution. The wrong coefficient of relative risk aversion thus distorts not only optimal investment behavior (or risk sharing) but also optimal saving behavior (or intertemporal consumption smoothing).

5.5 Risk Exposure and Welfare Effects in Some Stylized Collective Models

Subsections 5.2 and 5.3 considered the idealized collective model that removes the borrowing constraints early in the life cycle and allows intergenerational risk sharing while implementing the individually optimal contract. In practice, collective pension schemes engage in risk sharing among non-overlapping generations by levying equal contributions on all active generations at a particular point in time. Furthermore, the recovery premia and indexation cuts imposed in case of funding deficits typically decline over time as the funding deficits are gradually closed. Rather than behaving as a random walk as in the optimal contract, consumption thus features mean reversion. Hence, risks are not optimally smoothed over the life cycle. Moreover, from the point of view of overall remaining wealth, the generations alive at the time the shock hits bear a larger part of the overall burden than the generations who are not yet participating in the labor force even though all cohorts face the same consumption level at each point in time. The uniform, age-independent premia and benefits, which depend only on the aggregate funding status of the fund, thus imply imperfect consumption smoothing and imperfect risk sharing. This subsection considers implications of actual pension contracts for risk exposures and welfare.

Boender et al. (2006) compares the performance of a number of different pension schemes using a semi-variance criterion to evaluate consumption during retirement.³⁵ The study implicitly assumes that agents are risk neutral with respect to consumption during the working life because only the expected premium level is taken into account when evaluating policies.³⁶ Working agents are therefore willing to bear the risk of fluctuating premia with small rewards in the form of risk premia. The retired generations can thus take advantage of the equity premium even though they do not bear much risk. One of the schemes that is considered is an individual scheme

34 In practice, it is easier for an individual to correct undersaving of the pension fund by saving more than to correct oversaving of the fund by borrowing against the pension wealth.

35 This criterion is closely associated with preferences that are associated with loss aversion (see Subsection 6.3.4).

36 The paper assumes mean reversion in stock prices (see section 6.2.5).

with an age dependent asset mix.³⁷ The paper explores also a hybrid collective scheme where both the contribution rate and the (real) benefit level of all participants adjust in identical ways to fluctuations in the collective funding ratio. The paper argues that the hybrid collective scheme is welfare improving compared to the individual scheme due to the value of the buffers in intergenerational risk sharing. The study abstracts from heterogeneity between agents. For an extensive evaluation of this study, see Boender et al. (2006).

Cui et al. (2006) consider the welfare gains due to intergenerational risk sharing in four pension deals. These deals are approximately fair in value terms and assume time-invariant asset allocations for the collective contracts. Their main finding is that these collective schemes can add value from the perspective of new entrants even relative to the optimal unrestricted individual scheme without intergenerational risk sharing and even if the scheme is initially underfunded. The welfare gain is estimated to be 1–4% in terms of certainty equivalent consumption. The paper also analyzes the impact of solvency constraints, which ensure that the probability of underfunding should remain small. The paper shows that the reduction in risk taking implied by this constraint does not substantially reduce the welfare gains of the collective scheme. The results of Cui et al. (2006) are based on homogeneous populations and imply significant risk taking by the fund (up to 100% in equity) in the unconstrained case. The paper does not report how long it takes before situations of underfunding are recovered and how deep underfunding can be. This makes it difficult to assess how serious the discontinuity risk actually is.

Koijen et al. (2007) consider stylized collective pension schemes using asset allocation rules that vary with the funding ratio of the scheme. The funding ratio, f_t is defined as the ratio of the market value of the assets over the market value of time invariant basis benefits b^b . Benefit improvements or reductions b_t^a occur during the entire retirement period whenever the funding ratio is more or less than one at the time of retirement.³⁸ The basis premium π^b is actuarially fair for the basis benefits.³⁹ If the funding ratio is below one, additional pension contributions π_t^a will be imposed, and likewise reductions are given if the funding ratio is high. The total pension contribution and the

37 In contrast to the model in section 2, the asset mix does not depend on past shocks in the financial markets. Hence, individuals can not construct options by dynamically trading in bonds and stocks. Since options are not available in the model, financial markets are in effect dynamically incomplete. In these circumstances, collective pension funds can help to complete markets by in effect creating new assets through pension contracts.

38 Agents thus obtain fixed annuities at the time of retirement. After they have obtained these annuities, they are no longer exposed to risk. Agents are therefore exposed to the risk of underfunding at the time of retirement. In fact, if the pension fund is underfunded at the time they retire, the value of their annuity is reduced so that the consumption level is scaled down during the rest of their life.

39 Since both the basis premium and the basis benefits are risk free, the risk-free interest rate is used to compute the cost price of the basis benefits.

total benefit are defined as

$$\pi_t = \pi^b + \pi_t^a \quad (11)$$

$$b_t = b^b + b_t^a \quad (12)$$

The asset mix x_t as well as the benefits and contributions adjustments depend linearly on the funding ratio:

$$x_t = \alpha_0 + \alpha_1(1 - f_t) \quad (13)$$

$$\pi_t^a = \beta_1(1 - f_t) \quad (14)$$

$$b_t^a = \zeta_1(1 - f_t) \quad (15)$$

The paper considers three stylized schemes: a collective defined contribution scheme with a fixed contribution rate ($\beta_1 = 0$), a collective defined benefit scheme with a fixed benefit level ($\zeta_1 = 0$) and a hybrid collective scheme in which both the contribution and benefit levels can fluctuate with the funding rate. The parameters in (13)–(15) are optimized conditional on a condition on the half-time of any deficits in the fund.

Table 5 summarizes the main results. The DC schemes perform worse than DB schemes. The intuition is that consumption during the working life contributes more to the risk-sharing capacity of an individual than consumption during retirement because the working life lasts longer than the retirement period does and is also discounted less heavily.⁴⁰ The hybrid systems perform best because these schemes utilize the risk-bearing capacities of individuals during both their working and retirement periods.

The results in Table 5 illustrate the importance of the recovery period. With a relatively short half-life of three years (i.e. the period during which half of the funding imbalance is made up), collective schemes do worse in terms of ex-ante welfare than the optimal individual scheme from Subsection 4.3 (i.e. including borrowing constraints, annuities and discrete trading but with fully flexible premia, annuity levels, and portfolio choices). For longer recovery periods (a half-life of 10 years), in contrast, the collective hybrid scheme can realize a certainty equivalent consumption level that is 4.7% above that for the optimal individual scheme.

40 An argument that goes the other way is that consumption when young cannot respond to shocks that occur after consumption has taken place. Hence, retirement consumption can respond to more shocks. Flexibility in retirement consumption is thus more valuable than flexibility in working-life consumption. If the length of the working period becomes smaller compared to the expected retired period, the DC scheme becomes more efficient compared to a DB scheme.

TABLE 5 – CHARACTERISTICS OF COLLECTIVE SCHEMES AND WELFARE GAINS RELATIVE TO OPTIMAL INDIVIDUAL CONTRACT DISCUSSED IN SUBSECTION 4.3 (I.E. INCLUDING BORROWING CONSTRAINTS, ANNUITIES AND DISCRETE TRADING BUT WITH FULLY FLEXIBLE PREMIA, ANNUITY LEVELS, AND PORTFOLIO CHOICES)

	Half-life 3 years			Half life $6\frac{1}{2}$ years			Half-life 10 years		
	CDC	CDB	CH	CDC	CDB	CH	CDC	CDB	CH
α_0	9.2%	28.8%	39.1%	17.9%	54%	76.6%	25.4%	73.5%	94.9%
α_1	-0.35	-0.08	-0.01	-0.23	0.14	0.24	-0.08	0.26	0.24
π^b	14.6%	15.5%	15.0%	14.1%	15.9%	14.9%	13.6%	16.2%	15.0%
b^b	82.1%	87.3%	84.7%	79.3%	89.4%	84.0%	76.8%	91.3%	84.4%
β_1	0	0.99	0.73	0	0.53	0.38	0	0.38	0.27
ς_1	-3.25	0	-0.79	-1.65	0	-0.42	-1.12	0	-0.29
Welfare gain	-4.7%	-2.4%	-1.8%	-4.1%	0.4%	1.8%	-3.4%	2.9%	4.7%

Table 5 contains also relevant information on the risk-sharing properties of the collective schemes. In the optimized hybrid scheme with a half-life of three years a 10% funding deficit implies a reduction in benefit level of 7.9 percentage points to 76.8%. Likewise the contribution rate would be 7.3 percentage points above the base level of 15.0%. In this particular case, the asset mix would be almost fixed at an equity share of 39%. In case of longer recovery periods, more investment risk is taken and the contribution and benefit levels will be substantially more stable.

6 EXTENSIONS

6.1 *Human Capital*

Sections 3, 4 and 5 restricted the analysis to two financial assets, namely stocks and risk-free bonds. We assumed a constant and riskless wage so that in the unconstrained case human capital was in fact equivalent to a risk-free bond. Sections 4 and 5 allowed for the non-tradability of human capital: moral hazard and adverse selection prevent individuals from trading claims against their future labor income. The non-tradability of human capital affects both intertemporal consumption smoothing and risk sharing.

This section considers human capital more closely. Rather than assuming a constant and riskless wage over the life cycle (until retirement), we introduce realistic career patterns (i.e. wage income depends on age) and allow wage income to be stochastic. Human capital risk can be decomposed into two components: an individual (idiosyncratic) component that is specific to

the individual and an aggregate component that applies to all individuals alike. Aggregate risk may be more important for some individuals than others. Entrepreneurs, for example, tend to bear more aggregate risk than civil servants.

Subsection 6.1.1 first discusses more realistic career patterns and the constraints on intertemporal consumption smoothing that result in the presence of borrowing constraints, then Subsection 6.1.2 introduces non-insurable idiosyncratic human capital risk, and Subsection 6.1.3 combines this risk with borrowing constraints. The implications of aggregate human capital are explored in Subsection 6.1.4, while Subsection 6.1.5 analyzes what happens if labor supply becomes endogenous.

6.1.1 *Intertemporal Consumption Smoothing and Borrowing Constraints*

Until now, we have stressed the consequences of borrowing constraints for risk taking. Especially young agents often cannot borrow against the collateral of their human capital to invest in equity. As a result of these short-sale constraints, the shadow value of equity for young agents exceeds the market value. Borrowing constraints, however, constrain not only risk taking but also intertemporal consumption smoothing for young agents. This is especially relevant for young high-skilled agents facing a steep career pattern: the youngest years feature low earnings but rapid growth of these earnings. Hence, these agents would like to borrow against the future value of their human capital for the purpose of consumption smoothing. The desire to borrow is particularly large if precautionary saving motives are weak on account of a low level of risk. In the presence of borrowing constraints, young agents can not capitalize the value of their human capital to raise their consumption in youth. They in effect feature shadow interest rates r_i that exceed the market rate.⁴¹

6.1.2 *Idiosyncratic Risks*

The idiosyncratic component of human capital risk is uncorrelated to any other risk in the economy. This component can thus in principle be diversified in financial markets. However, this would give rise to serious moral hazard problems; individuals would reduce their effort and claim that the resulting low wage income has been due to misfortune instead of shirking. Indeed, human capital is special in that it is not really tradable in financial markets and introduces market incompleteness. Nevertheless, social insurances (such as disability and unemployment insurances) do provide partial insurance against idiosyncratic human capital risk but moral hazard reduces the

41 This shadow interest rate is determined such that (2) is met with the constrained consumption growth rate. Hence, rather than that (2) determines the growth rate of consumption at a fixed interest rate, the expression determines the implicit individual-specific interest rate r_i at a given growth rate of consumption.

optimal coverage, thereby leaving a substantial amount of risk with the individual.

An important topic for future research is to estimate idiosyncratic human capital risk in the Netherlands and how it has developed over time. On the one hand, reduced coverage of social insurance and more competition and the associated creative destruction may have increased this risk. On the other hand, a more flexible labor market may reduce the length of unemployment spells and increase the opportunities for outsiders to find work, thereby reducing persistent labor-market risks. Within idiosyncratic risks, one can distinguish between temporary and permanent wage shocks. Abowd and Card (1989) and Topel and Ward (1992) have shown that log wages in the United States are close to a random walk, where the standard deviation of innovations is large, namely about 10% per year.

The uninsurable idiosyncratic uncertainty about human capital affects both optimal saving (i.e. intertemporal risk sharing over time) and optimal investment (i.e. risk sharing). Persistent shocks that substantially affect the value of human capital can have substantial effects, especially on precautionary savings (see Koo (1995) and Viceira (2001)).⁴² If we extend the basic model laid out in section 2 with uninsurable uncertainty about human capital, prudent investors increase precautionary saving.

As regards investment, the introduction of idiosyncratic risk reduces the value of human capital.⁴³ In the beginning of the working career, overall wealth consists mainly of human capital. With fixed relative risk aversion (CRRA) and thus a fixed optimal share of overall wealth invested in risk-bearing assets (see (4)),⁴⁴ a reduction in human wealth as a result of idiosyncratic risk decreases the initial absolute holdings of equity. Later on during the life cycle, the stock of financial wealth increases as a result of larger precautionary savings. This tends to raise the absolute size of stock holdings. As agents age, human capital becomes less important compared to financial capital and eventually the effect on absolute equity holdings of larger financial wealth dominates the effect of smaller human capital. The young holding less equity and the old holding more equity implies that the age pattern of equity holdings becomes less skewed towards the young as a result of uninsurable

42 Hence, more idiosyncratic risk as a result of increased labor-market flexibility boosts life-course saving for precautionary motives.

43 In the presence of undiversifiable wage risk, we can no longer rely on market prices to measure human capital but must employ individual-specific risk premia (or pricing kernels) to discount future wage income. If wages behave like a random walk, the reduction in human wealth is particularly large at the beginning of the career when households face uncertainty that accumulates over a long horizon.

44 Gollier and Prat (1996) derive the conditions on more general utility functions for equity holdings to decline if investors bear another independent background risk. In particular, investors should be risk vulnerable.

human-capital risk: whereas human capital increases the demand for equity, human-capital risk reduces it. As a fraction of financial assets, stock holdings unambiguously decline during the entire working life as a result of undiversifiable background risk in human capital. The reason is that human capital becomes less important compared to financial capital. Hence, a smaller share of financial wealth needs to be exposed to equity risk to obtain the correct exposure of overall wealth to that risk (see the impact of a smaller ratio H_t/F_t on f_t^* in (5)).

6.1.3 *Idiosyncratic Risks and Limited Intertemporal Consumption Smoothing*

Combined with borrowing constraints (and steep career patterns) limiting intertemporal consumption smoothing, idiosyncratic human-capital risk introduces other complications. With borrowing constraints, the individual is not able to optimally diversify labor income risks over time. As a direct consequence of the reduced ability to smooth risks intertemporally, the individual will take less risk on its investment portfolio if the household owns positive financial wealth.⁴⁵ The effects of background labor-income risks on investment behavior seems relatively minor but become more important if borrowing constraints limit intertemporal consumption smoothing of large transitory shocks. Indeed, the flexibility of intertemporal consumption smoothing is a key determinant of the risk that agents can afford to take on.

On the basis of numerical simulations, Gollier (2005) shows that this may cause young households with small financial reserves to invest less in equity than older households who are less constrained by borrowing constraints in diversifying risks over a longer period. In a similar vein, Cohen and Einav (2005) find that young investors choose portfolios that are less tilted towards equity than the portfolios of middle-aged investors. Thus, the combination of labor-market risks and a steep career pattern preventing optimal intertemporal consumption smoothing (or the potential of large drops in labor income) thus turn around the result from unconstrained models that young households should hold substantial amounts of equity that exceed the amounts (both in absolute values and as a share of financial wealth) of older households. Indeed, the desire to hold a safe, liquid stock of precautionary savings has become one of the explanations for the equity premium puzzle and for why young households do not participate much in the stock market (see Constantinides and Duffie (1996) and Brav et al. (2002)). Social insurance that provide insurance of human capital risk can thus boost equity investments and thus increase the risk-bearing capacity of the economy. Hence, pension schemes can raise the demand for equity not only by relieving short-sale constraints but also by providing insurance against human-capital risks.

45 This is similar to the impact of a defined-contribution pension system that forces individuals to absorb risk in the limited period of their retirement.

Liquidity constraints not only make investment behavior more conservative but also strengthen precautionary saving motives – even if financial wealth is positive. Indeed, by investing conservatively and setting aside resources through saving, individuals ensure the presence of a financial buffer that helps them to optimally time diversify temporary risks.⁴⁶ Precautionary motives rather than saving for retirement tend to be the main reason why young households save (see Cocco et al. (2005)). As a direct consequence of the additional saving, financial wealth will be higher later on in life, which will tend to increase equity exposure when old. Hence, precautionary saving further weakens the result that the young should have much more equity exposure than the old if idiosyncratic risk is combined with borrowing constraints and a steep career pattern preventing intertemporal consumption smoothing at young ages.

Davis et al. (2005) formulate a model in which individuals can borrow but at a cost that exceeds the risk-free rate of return. As far as investment behavior is concerned, high borrowing rates discourage young households from acquiring a large equity exposure by leveraging their portfolio. As regards consumption smoothing, high borrowing rates discourage dissaving early in life so that agents can build up a stock of financial wealth earlier, which helps to boost equity exposure later on in life. Their model is better able to fit the empirical evidence on borrowing and individual equity holdings over the life cycle. They show that a borrowing rate equal to the expected return on equity minimizes the overall demand for equity over the life cycle. The reason is that this crowds out all borrowing aimed at buying equity but does not unnecessarily discourage borrowing for consumption-smoothing purposes, which boosts wealth and equity exposure later in life. High borrowing rates do explain why the equity exposure of older workers is larger than that of younger workers.

The behavior of individuals facing borrowing constraints resembles that of a pension fund facing solvency constraints. Both borrowing constraints and solvency constraints harm the ability to time diversify risk. They thus reduce risk-bearing capacities and encourage precautionary saving to escape these constraints. Saving occurs not only to smooth consumption in the face of the depreciation of human capital during retirement but also to help smooth consumption in the face of labor-income shocks during the active life. Especially at low levels of financial wealth, the latter precautionary motive may dominate the retirement motive.

Non-insured medical expenses can be an important source of idiosyncratic background risk for older households. Hence, also older households may face

46 In the case of permanent shocks, optimal consumption smoothing dictates that consumption is reduced in line with income. Habit formation, however, can then still give rise to a desire to reduce consumption gradually and thus a need for a financial buffer.

idiosyncratic risk with implications for saving behavior (i.e. precautionary saving) and investment behavior (and also the willingness to take out annuities). In particular, these households then also save for precautionary reasons and when they start to be subject to these risks invest more conservatively and are discouraged from taking out illiquid annuities. An important topic for future research is to distinguish between liquid precautionary saving and illiquid retirement saving. The papers discussed do not distinguish between these two categories and assume that all saving is liquid.

6.1.4 *Aggregate Risks*

An aggregate component to wage risk introduces other issues. First of all, to the extent that aggregate human-capital risk is correlated with equity risk, human capital becomes more like equity. Equity and human capital are thus substitutes. As a direct consequence, the wealth share invested in equity declines compared the case in which human capital does not carry any macro-economic risk. This is especially so at the beginning of the working life when human capital constitutes the bulk of overall wealth. The share of financial wealth invested in equity thus declines less sharply over the life cycle. As noted above, if human wealth has the optimal exposure to equity risk given by the optimal exposure (4), the part of financial wealth invested in equity is constant during the life cycle. This may be relevant for entrepreneurs who face substantial aggregate risk.

Another issue arises if wage risk is not perfectly correlated with equity risk. In that case, a second macro-economic risk factor emerges. This risk factor can in principle be traded through wage-indexed bonds. In particular, young agents who are long on this risk factor can trade this risk with older agents who do not bear (much of) this risk. In practice, however, we do not observe wage-indexed bonds. A wage-indexed defined-benefit pension may, however, act as a substitute for tradable wage-indexed bonds. Through such a pension system, the younger generations share wage risk with the retired agents (see Beetsma and Bovenberg (2007)). Hence, pension schemes allow not only young workers with ample human capital to capture the equity risk premium but also older agents to capture part of the human capital risk premium. Compulsory pension schemes may allow human capital to become in fact tradable.

Older agents are willing to absorb wage risk even without compensation if retirees care about the value of their pension relative to the average standard of living in the economy (see also Subsection 6.3.2). Moreover, with these so-called standard-of-living preferences, positive correlation between equity and wage risk makes equity a more attractive investment for old-age pension insurance, especially for older agents who rely more on financial rather

than human capital to finance their consumption.⁴⁷ Intuitively, stocks match to some extent the liability risks that are associated with changes in desired consumption based on the aggregate standard of living. Hence, whereas positive correlation between equity and wages reduces the equity holdings of young agents with ample human capital, this correlation may actually raise the equity holdings of older agents with standard-of-living utility. The conclusion is again that the age profile of equity holding is thus less steeply declining than in the benchmark model.

This conclusion is strengthened in a recent paper by Benzoni et al. (2007). This study indicates that labor income is cointegrated with equity prices. This makes human capital a closer equity substitute for younger workers with a longer investment horizon than for older workers and retirees with a shorter investment horizon. Consequently, young workers invest less in stocks than older agents do. Benzoni et al. (2007) finds that the cointegration between wages and stocks can explain why young workers do not hold positive equity. In fact, they find that young investors should want to short stocks because their human capital is effectively a leveraged security with large exposure to stock markets.

6.1.5 *Endogenous Labor Supply*

Sections 4 and 5 showed that more flexibility to adjust premia and retirement benefits can substantially enhance the capacity to bear risks. Bodie et al. (1992) show that the same holds for the flexibility to adjust labor supply and the consumption of leisure. With endogenous labor supply, human capital amounts to the discounted value of *potential* wage income (i.e. wage income if no leisure would be consumed) and is thus larger than the discounted value of *actual* wage income. Indeed, the stock of wealth is larger now because it finances consumption of not only produced goods but also leisure. As the share f in (4) applies now to a larger stock of wealth, the overall demand for risk-bearing capital rises. Indeed, shocks can be absorbed by adjusting consumption of both produced commodities and leisure. This enhances the capacity of the economy to absorb risks and thus increases the aggregate demand for risk taking.

If labor supply becomes more flexible over the entire life cycle, the share of financial wealth invested in risk-bearing assets, f^* , now varies even more with age. This changes if labor supply is especially flexible around retirement because agents can adjust the time and speed with which they retire. In that

47 A similar argument applies to pension funds providing wage-indexed pension benefits. See de Jong (2005) for the market valuation of wage-indexed liabilities, taking into account the potential correlation between equity and wages. In particular, this correlation *ceteris paribus* reduces the market value of wage-indexed liabilities by raising the implicit risk premium when discounting.

case, also older workers can afford to invest in risk-bearing assets. This points to the importance of a flexible labor market for older workers to sustain the supply of risk-bearing capital in an aging economy. A formal analysis of the impact of flexibility of the retirement date on the optimal risk taking is provided by Farhi and Panageas (2007).

6.2 *Financial Risk Factors and Longevity Risk*

6.2.1 *Introduction*

Apart from human capital risk, the main risk factors that individuals face are interest rate risk, inflation risk, equity risk and longevity risk. Until now, we have explored only equity risk and imposed that stock returns were not predictable. Subsections 6.2.2 and 6.2.3 introduce interest-rate risk as well as inflation risk. Subsection 6.2.3 briefly discusses the evidence on stock return predictability and its implications for optimal consumption smoothing and risk sharing over the life cycle. Subsection 6.2.4 considers micro as well as macro longevity risk and the value of annuity markets and longevity bonds.

6.2.2 *Interest Rate Risk*

So far we ignored inflation risk and assumed a constant interest rate. If (real) interest rates fluctuate over time, the investor is faced with an additional risk factor. Fluctuations in interest rates affect the value of the human capital as well as the value of the bond portfolio. Early in the life cycle, portfolios are to be held with long durations to hedge against the required consumption during retirement. The duration of human capital is typically shorter than required. If bonds or interest rate derivatives would be available with very long maturities, the part of financial wealth that is invested in fixed income can be invested in these assets to obtain the optimal duration for total wealth. In the more realistic case in which the maximum maturity of traded bonds and derivatives is limited, the fraction of financial wealth invested in bonds is increased at the expense of equity. In this way, the duration of the assets is more closely aligned with the duration of the consumption pattern while the remaining interest rate risk is balanced with equity risk. Campbell and Viceira (2001), Brennan and Xia (2000), Sangvinatsos and Wachter (2005) and Munk and Sorensen (2005) have considered the implications of time varying interest rates on the optimal demand for long-term bonds and analyze the welfare gains of hedging variations in real rates.

A number of recent studies report that the expected bond returns are clearly time-varying, i.e. bonds are more attractive in some periods than in others, see e.g. Cochrane and Piazzesi (2005). In particular, the real (as well as the nominal) bond risk premium is increasing in the real rate. Sangvinatsos and Wachter (2005) show that unconstrained long-term investors (as considered in section 3) can realize large gains by exploiting these time varia-

tion s in bond premia. Kojien et al. (2006b) consider the case of borrowing constrained life-cycle investors explored in section 4. They find that short-sell constraints reduce the potential gains associated with bond timing strategies because these gains can be obtained only at the expense of reduced equity exposure unless very long maturity bonds would be traded. The economic gains realized by bond timing strategies peak around the age of 50 (when agents have build up substantial financial wealth and are less constrained by borrowing constraints) and are hump-shaped over the life-cycle. If investors can avoid borrowing constraints through participation in collective pension schemes (see section 5), they may be able to realize the larger benefits reported in Sangvinatsos and Wachter (2005).

6.2.3 *Inflation Risk*

We have assumed so far that inflation risk is negligible or, equivalently, that all assets that have been considered are inflation indexed. In reality, inflation linked assets are rarely traded and inflation is an important risk factor for long term financial planning. In the absence of inflation risk, long-term nominal bonds e.g. are riskless for long-term investors if the maturity equals the horizon of the investor. If inflation risk is taken into account, however, Campbell and Viceira (2005) report that the annualized standard deviation of real returns on these bonds is as large as 8% and exceeds that of stocks for maturities of over 35 years. Risk-averse long-term investors therefore exhibit a strong preference for index-linked bonds if available. The difference between the expected return on nominal bonds and that on real bonds with the same maturity is known as the inflation risk premium. The inflation risk premium is substantial (0.7% per annum on average) and decreases in the real rate and increases with expected inflation (Burashi and Jiltsov (2005)). Less risk-averse investors should take these fluctuations of the risk premium over the business cycle (from 0.2% to 1.4%) into account in their decision whether or not to hedge inflation risk.

Campbell and Viceira (2001) report that the certainty equivalent wealth effect of access to inflation indexed bonds for short-sell constrained investors⁴⁸ can be as large as 10–30%.⁴⁹ In reality, inflation linked assets are often not available to individual investors. To a lesser extent, this is also true for institutional investors.⁵⁰ These investors therefore face the question

48 Note though that Campbell and Viceira (2001) abstract from labor income.

49 The welfare gains of access to real rather than nominal annuities have been considered e.g. by Brown et al. (2001). Kojien et al. (2006a) extend this analysis to access to real as well as equity linked annuities.

50 In the UK inflation linked bonds are traded and a number of other countries introduced them recently. Still there could be a gap between the inflation index that is traded (e.g. national price inflation) and the inflation rate that is relevant to the investor (e.g. regional wage inflation).

to what extent indexed bonds can be replicated by nominal bonds. The time variation in nominal and real interest rates can be adequately captured by fluctuations in the real interest rate and in expected inflation. Expected inflation is much more persistent than the real rate, whereas the risk premium for real interest rate risk exceeds that for fluctuations in expected inflation. Unconstrained investors can exploit the differences between bonds with different maturities with respect to the sensitivities to the real rate and expected inflation factor in order to hedge away substantial parts of the inflation risk using sophisticated dynamic trading strategies. Unexpected inflation however can not be hedged without access to inflation-linked assets. Moreover, many investors will be constrained to hedge against the real interest and expected inflation factors or will not be equipped to implement dynamic trading strategies. Institutional investors like pension funds or insurers can create value here.

If inflation sensitive assets are not available for the investor, the minimum variance portfolio for a long term investor is a horizon dependent mix of the nominal bond with the corresponding maturity, stocks and cash. This implies that when investors are forced to bear inflation risk they will shorten the maturity of the bond portfolio (see Campbell and Viceira (2001, 2005)). This implies that the optimal duration of the fixed income portfolio is smaller than that of the consumption profile. In the Netherlands most pension schemes issue nominal guarantees to their participants. The solvency requirements for these guarantees imposed by the regulator lead them to extend the maturity of their assets to that of their nominal liabilities. If individuals care about the purchasing power rather than the nominal value of their pension, even the most risk averse individuals would prefer a less than full reduction of the duration mismatch between the assets and the liabilities.

6.2.4 *Equity Risk Reconsidered*

In the analysis so far we assumed that stock prices are generated by a random walk, and that the (log) stock returns are independently normally distributed with constant mean and volatility. The assumptions that stock returns feature constant expected returns (and hence are uncorrelated) and that their risk level is constant have been challenged in the literature. In the sequel, we address each of these assumptions as well as the implications of alternative assumptions for optimal financial planning.

There is a lively debate in the literature as to whether the expected stock returns are time-varying.⁵¹ Many authors claim that the expected return on equity depends on variables such as the dividend yield or the term spread so that equity returns are to some extent predictable. The typical finding in the

51 See e.g. Ang and Bekaert (2006), Campbell and Yogo (2006), Cochrane (2006), Goyal and Welch (2006), and Lettau and van Nieuwerburgh (2006).

literature is that, at the low frequencies that are relevant for financial planning over the life cycle, stock returns are negatively correlated, implying that the stock market is mean reverting. As discussed in Subsection 3.2, both the mean and the variance of the excess return vary in the same way with the investment horizon (namely proportionally). If stock returns are mean reverting, in contrast, the variance of the excess returns rises less rapidly with the time horizon than the mean of the excess return. As a direct consequence, stocks are less risky for long-term investors than for short-term investors so that young investors should invest an even larger over their financial wealth in equities. Campbell, Chan and Viceira (2003) and Campbell and Viceira (2005) show that the annualized standard deviation of real returns of about 17% of a diversified stock portfolio is reduced to an annualized standard deviation of only 8% on a 20-year horizon, which is in fact comparable to the annualized variance of the real returns of nominal bonds with the same maturity that are held to maturity. Bond portfolios with constant maturity are also somewhat mean reverting.

A second assumption on the stock market process that we adopted so far is that the volatility of stock returns is time-invariant. The literature contains abundant evidence that this is not the case at shorter horizons and therefore can be an approximation at lower frequencies at best. Chacko and Viceira (2005) and Liu (2007) consider an incomplete markets setting with time varying volatility of equity returns. This volatility factor can not be traded. These papers show that the impact of time variation in volatility on the optimal asset allocation is small, in particular in the empirically relevant case of volatility being negatively correlated with stock returns (so that stocks become less risky if returns are high).

6.2.5 *Longevity Risk*

Longevity risk is important both for the financial planning of individual investors and for the optimal asset allocation of collective pension schemes. Two types of longevity risk have been distinguished in the literature. Macro longevity risk refers to uncertain development in future mortality rates due to improved health care, changing in nutrition and habits, etc. Macro longevity affects the average mortality of a pool of individuals, e.g. in a pension fund or insurance pool. Micro longevity risk refers to the uncertainty for an individual on whether or not he will survive, irrespective of changes in mortality rates.

Micro longevity risk can not easily be traded on financial markets because of adverse selection. Older individuals typically have superior information on their survival probabilities than the issuer of the annuity. As the asymmetry in information becomes more important for older agents with poor health conditions, annuities are irreversible. This gives rise to a fundamental trade-off between liquidity and longevity insurance. Adverse selection forces the issuer

of voluntary annuities to offer prices that are not actuarially fair for the population at large.⁵² Mandatory participation in a collective pension scheme alleviates these adverse-selection problems.

In the absence of bequest motives and liquidity problems and in the presence of a rich menu of annuities is available at not too actuarially unfair prices, annuities are welfare improving (Davidoff et al. (2005); Poterba (2006)). If real annuities or variable annuities that offer exposure to stock markets are not available, annuities become less attractive to the individual investor. In view of the absence of adverse selection, collective pension schemes offer annuities that are not available to individual investors. In particular, many collective schemes aim to offer real annuities through their indexation ambition and imply some equity exposure because indexation is reduced if investment returns are insufficient. Another important aspect of collective pension arrangements is that they usually impose deferred annuities at an early age. Deferred annuities, i.e. annuities that only generate cash flows after a deferral period, are often not available in voluntary annuity markets. Even if the option would be available, one may wonder whether agents would buy deferred annuities early in their life cycle. Deferred annuities can be attractive because they offer a hedge against the conversion risk generated by future fluctuations in interest rates and, to a lesser extent, changes in mortality rates.

Collective pension schemes usually impose uniform contribution rates on their participants. Section 5 explored the impact of heterogeneity in risk aversion and time preference, and referred to heterogeneity in non-pension wealth. Longevity risk is another dimension of heterogeneity: the life expectancy of women exceeds that of men, that of whites that of blacks and that of highly educated persons that of less educated persons. Uniform contribution rates impose a welfare loss on individuals with low life expectancy due to implicit value transfers from individuals with lower life expectancy to those with longer life expectancy (see e.g. Brown (2002) and Finkelstein, Poterba and Rothschild (2005)).

For pension funds and other issuers of annuity contracts macro longevity risk is a relevant concern. Unexpected increases in life expectancy increase the market value of their liabilities and can harm their solvency position. From a macro-economic point of view, the natural issuer of longevity protection are young and future generations. By issuing longevity bonds, the government can offer instruments to insure the longevity risk of older people by shifting it to younger cohorts, for example by conditioning the retirement age of the public pension system on longevity. By this conditioning of the retirement age, the government reduces the impact of longevity risk on its balance sheet.

52 Annuities can be attractive for individuals even if they are not actuarially fairly priced (see e.g. Mitchell et al. (1999)).

Macro longevity risk can in principle be traded on financial markets (see Blake et al. (2006)) although attempts to introduce longevity bonds have failed so far. Through financial markets, macro longevity risks can be diversified between pools, sectors and countries. Note that the scope for diversification is limited if changes in health care and life style affect all populations in approximately the same way at the same point in time.

6.3 Preferences

This subsection explores the implications of a coefficient of relative risk aversion that is not constant. Subsection 6.3.1 first of all considers the implications of minimum consumption needs. Subsections 6.3.2 and 6.3.3 subsequently consider external and internal habits, respectively. Finally, Subsections 6.3.4 and 6.3.5 investigate loss aversion and endogenous reference points.

6.3.1 Required Minimum Consumption

Preferences may imply that people require a minimum amount of consumption in order to derive any utility at all:

$$U = E \left[\int_0^L \exp(-\rho t) \frac{1}{1-\theta} (C_t - \bar{C}_t)^{1-\theta} dt \right], \quad (16)$$

where $\bar{C}_t \geq 0$ represents the *minimum* (or *necessary* or *required*) consumption level at time t .

As regards investment behavior, the optimal wealth share invested in equity is given by

$$f_t^* = f \frac{(W_t - X_t)}{W_t} = f \left(1 - \frac{X_t}{W_t} \right), \quad (17)$$

where $X_t = \int_0^L -t \exp(-rx) \bar{C}_{x+t} dx$ stands for the discounted value (using the risk-free rate as a discount rate) of the minimum consumption levels over the rest of the life cycle and f is the solution for f_i in (4) with $\theta^i = \theta$. Intuitively, the investor first invests in risk-free assets X_t to guarantee the minimum consumption level and then invests a fixed fraction f (given by (4)) of remaining wealth $(W_t - X_t)$ in risk-bearing assets. In the terminology of asset-liability management (ALM), one thus matches the fixed liability X_t with risk-free assets. Indeed, comparing the second equality in (17) (with $W_t = F_t$ since $H_t = 0$) with the corresponding equality in (5), we see that human wealth H_t acts like a risk-free asset, while the value of the minimum consumption levels X_t functions like a risk-free liability. Defined-benefit pensions, which guarantee a particular consumption level in retirement, can thus be viewed as the optimal consumption level for somebody who requires a certain minimum consumption level in retirement (i.e. $\bar{C}_t > 0$ for $t > T$). The need for a guaranteed

level of consumption in retirement forces the pension fund to take less risk and invest a larger share of retirement wealth in risk-free assets compared to the case without the need for consumption guarantees.

An alternative interpretation of optimal risk exposure is that preferences exhibit decreasing instead of constant relative risk aversion. In particular, one can compute the optimal wealth share $f_t^* = \frac{\lambda(1 - \frac{x_t}{W_t})}{\sigma\theta}$ as the solution for f_t from (4) with $\theta^i = \theta / (1 - \frac{x_t}{W_t})$. Accordingly, relative risk aversion θ^i decreases with wealth W_t . The local curvature of the utility function depends on how far overall wealth is above wealth required for minimum consumption. People thus become less willing to tolerate risk if the wealth level drops and become closer to this minimum wealth level. Accordingly, a low power coefficient θ can still imply high levels of relative risk aversion.

The implications of preferences (16) for saving behavior depend on how the minimum consumption flows \bar{C}_t develop over time and whether people correctly anticipate this. If agents correctly anticipate rising minimum consumption needs $d\bar{C}_t/dt > 0$, they may save more, especially because larger coefficients of risk aversion $\theta^i = \theta / (1 - \frac{x_t}{W_t})$ boost saving incentives.

6.3.2 External Habits

Campbell and Cochrane (1999) use the preferences (16) in which the consumption requirements \bar{C}_t vary with the history of aggregate consumption. They employ these preferences to explain a number of asset pricing phenomena, including a large equity premium that varies countercyclically. Investors demand a high return on stocks because stocks perform poorly in recessions when consumption levels C_t fall close to the minimum levels \bar{C}_t . Agents thus measure their consumption level compared to that of the rest of the population. This desire to "catch up with the Jones" implies a form of reference drift. Empirical estimates typically find that reference consumption levels are in fact close to actual consumption levels, implying high levels of risk aversion $\theta^i = \theta / (1 - \frac{x_t}{W_t})$.

6.3.3 Internal Habits

This external habit formation contrasts with internal habit formation where consumption requirements \bar{C}_t depend on the history of *individual* consumption rather than *aggregate* consumption of the economy as a whole. Constantinides (1990), for example, assumes that habits of an individual \bar{C}_t depend on actual consumption levels of that same individual as follows

$$d\bar{C}_t/dt = \beta C_t - \alpha \bar{C}_t$$

Internal habit formation implies more complicated behavior if agents anticipate how their current consumption levels C_t affect future consumption needs \bar{C}_{t+s} . For example, consumption at two neighbouring points in time become

complements: a high consumption level at time t raises, by its effect on habits \bar{C}_{t+s} , the marginal utility of consumption in the immediate future $t + s$. Hence, *ceteris paribus* θ in (15), the elasticity of intertemporal substitution declines with β as agents dislike rapid changes in consumption levels more. As time elapses and s increases, the impact of higher consumption level at time t on habits \bar{C}_{t+s} declines on account of the positive depreciation coefficient $\alpha > 0$.⁵³

This explains why, following an unexpected shock in wealth at the long-run equilibrium level $C_t = \frac{\alpha}{\beta} \bar{C}_t$, consumption adjusts does not adjust immediately to its long-run level but adjusts gradually over time. Compared to the case with preferences without habit formation (i.e. $\alpha = \beta = 0$), the immediate adjustment of consumption to a wealth shock is less, but adjustment later is larger as habits have adjusted to the changed consumption level. Innovations in consumption will now be correlated and the optimal consumption no longer follows a random walk. Internal habit formation thus explains why pension schemes adjust their premia levels only gradually following an unexpected shock. Hence, consumption is less volatile with internal habits for two reasons. First, shocks are absorbed more in the long run than in the short run. Second, agents invest less aggressively so that they have to adjust their consumption less in the short run.

Habit formation (i.e. $\alpha, \beta > 0$) implies also that the ability to absorb an unexpected shock increases if time increases. This implies that younger agents with longer horizons can adjust more easily to shocks and thus should invest a larger share of their wealth in risk-bearing assets. Intuitively, they have more time to adjust their habits to unexpected shocks. In addition to human wealth that is not exposed to equity risk and mean-reversion of returns, internal habit formation is thus a third reason why young agents should optimally invest a larger share of their portfolio in equity than older agents do and why pension contracts may optimally get more of a defined-benefit nature when people grow older. Habit formation also explains why retirement benefits may be linked to earnings (and consumption levels) close to retirement.

The implications of internal habits for saving behavior (and intertemporal consumption smoothing over time) depend on the initial habit level \bar{C}_0 . If young agents start with low habit levels, they optimally should save more since high initial consumption levels raise consumption needs later in life. As people get older, they raise consumption levels as the horizons become shorter over which the negative future external effects of current consumption

53 Internal habit formation can thus explain why risk aversion declines with time (and the intertemporal elasticity of consumption increases with time). Shocks can be easier absorbed and consumption levels can change more over time if agents have more time to adjust their consumption needs. Just as Epstein-Zin preferences, internal habit formation can imply that risk aversion may deviate for intertemporal substitution.

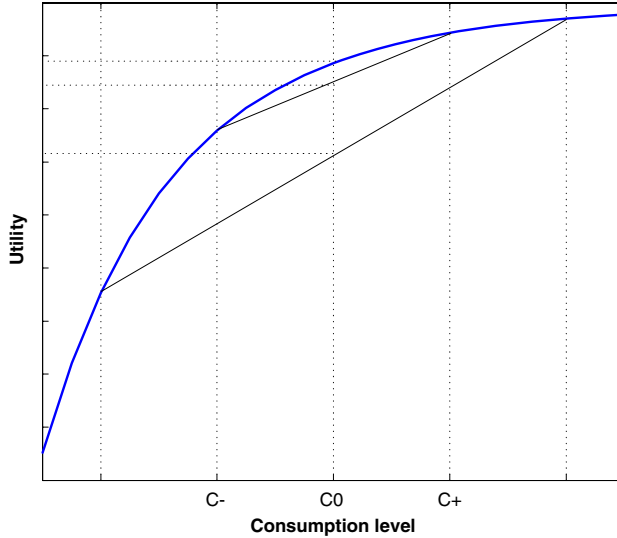


Figure 15 – A utility function with constant relative risk aversion

levels materialize. Overall then, habit formation can explain why consumption levels typically rise over the life cycle in line with income levels.

6.3.4 *Loss Aversion*

Standard utility functions, which are twice differentiable, imply that the costs of risk (see the denominator of the fraction at the right-hand side of (3)) are proportional to the variance of a portfolio, $f^2\sigma^2$. Figure 15 illustrates why this is the case. The utility function $u(C)$ is given by the concave curve. Consider the effect on utility of both a positive deviation C^+ and a negative, C^- from the mean value, C^o , keeping the expected value constant. This utility effect of risk (the required risk premium) can be shown graphically by the difference between $u(C^o)$ and the value of the line connecting $u(C^-)$ and $u(C^+)$ at C^o . A twice as large a deviation of C^+ and C^- from the mean C^o causes approximately a four time as big effect on utility.⁵⁴

Recent experimental work, however, suggests that human behavior is not always well described by twice differentiable concave utility functions. Tversky and Kahnemann (1992) therefore developed prospect theory to better describe human behavior with respect to risk. In particular, behavior seems better captured by the type of kinked utility function depicted in Figure 16. This util-

⁵⁴ Mathematically, this result relies on a second-order Taylor expansion. The cost of risk taking rises exactly quadratically with the magnitude of the deviations if the coefficient of relative risk aversion is constant and the Taylor approximation is exact.

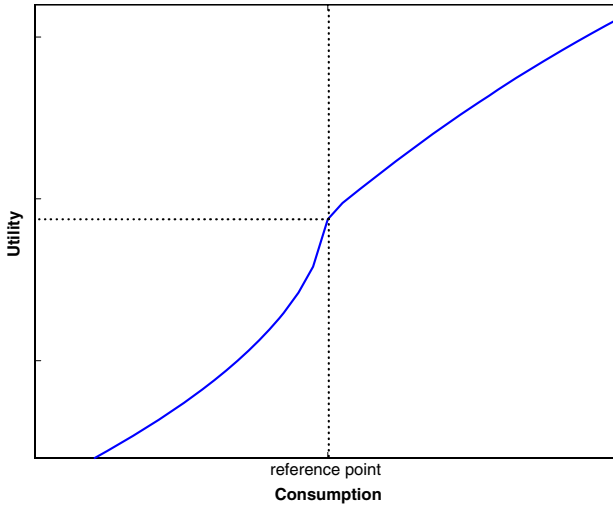


Figure 16 – A kinked utility function as proposed by Kahneman and Tversky

ity features a kink at the so-called reference point: a gain from this reference point yields only a small increase in utility, while a loss yields a large decrease.⁵⁵ This kink implies first-order risk aversion at the reference point: an individual is sensitive to small deviations from this reference point. This kink in fact implies that risk aversion is particularly large at consumption levels close to the reference point.⁵⁶ In other words, relative risk aversion is not constant but depends on the actual consumption level and the particular losses and gains considered. A second⁵⁷ element of prospect theory is that agents are risk averse in gains (twice as large gains yield less than twice as much increase in utility) and risk loving in losses (twice as big losses yield less than twice as much decrease in utility) (see Figure 16). This implies that agents in effect exhibit a large preference for consumption levels close to the target level of the reference point. If their current wealth is insufficient to finance the consumption level at the reference point, they will take risks to get close to it. If their wealth level exceeds the target level, in contrast, they will shy away from risks in order to prevent the risk of falling below the target level. This is especially so if their wealth level is just sufficient to finance con-

55 On the basis of experimental work, Tversky and Kahnemann (1992) find that losses tend to weigh about 2 1/4 times more heavily than gains.

56 This explains why it is often difficult to empirically disentangle risk aversion from loss aversion.

57 A third element is that decision-makers employ subjective decision weights that differ from the true probabilities.

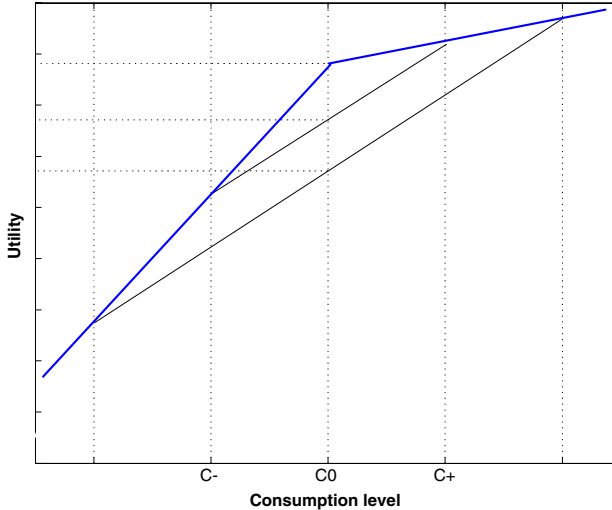


Figure 17 – Utility function with a kink at consumption level C_0

sumption at the reference point: the kink implies a relatively high local level of risk aversion.

Prospect theory, and the kink at the reference point in particular, has important implications for risk taking over time. We can best explain this by the piece-wise linear utility function with a kink at the average consumption level in Figure 17. We draw a line between $u(C^+)$ and $u(C^-)$ and repeat that exercise for a twice as large deviation from the mean C^o (*ceteris paribus* the expected value). We then observe that the cost of risk taking increases proportionally in the deviation rather than quadratically as with a smooth concave utility function. This implies that the cost of risk taking (i.e. the risk premium) increases with the square root of the time horizon because the expected standard deviations rise with the time horizon in this way as the draws are independent of time.⁵⁸ This is in contrast to the case of a smooth concave utility function when the cost of risk taking increases quadratically with the deviations and thus linearly with the time horizon, see Figure 15. With kinked utility, the marginal benefit of risk taking (i.e. the excess return $\bar{\mu}T$) increases linearly in the investment horizon while the associated marginal cost increases only with the square root of the time horizon $\sigma\sqrt{T}$. As a direct consequence, optimal risk taking increases with the time horizon. This yields another reason why young agents should bear more risk than old agents: loss

58 If the draws were perfectly positively correlated, the expected deviations would increase linearly with the time horizon.

aversion as described by prospect theory implies that young agents with long horizons should invest more in equity than old agents.

Bernatzi and Thaler (1995) argue that people are reluctant to invest in stocks because of a combination of loss aversion and a short planning horizon. People evaluate stocks over periods that are shorter than their true long-run horizon. This so-called *myopic loss aversion* explains why agents require a high risk premium on equity. Pension funds can correct this distortion by investing more in equity than individuals would do. Myopic loss aversion also suggests that pension schemes should not report their wealth levels too frequently to their members because this would stimulate the tendency of the members to evaluate their wealth on an excessively short time horizon.

Berkelaar et al. (2004) revisit the relationship between length of the investment horizon and risk taking if utility not only exhibits a kink at the reference point (or target level) for retirement wealth but also implies risk aversion in gains and risk seeking behavior in losses (as is suggested by prospect theory). As long as agents are above and close to the reference point, they invest less in risk bearing assets if the retirement date approaches. Intuitively, they have less time to make up losses if a bad shock would reduce their wealth below the target level. Hence, they converge to a portfolio insurance strategy as retirement gets close and their wealth level is just sufficient to finance the target level of retirement consumption. If wealth drops inadvertently below that needed to finance the target level, however, the investor attempts to gamble his way out by investing more in equity – especially if the planning horizon is short.

Ang et al. (2005) formulate a model of disappointment aversion based on the axiomatic disappointment-aversion framework of Gul (1991). Compared to a familiar CRRA utility, utility features a kink at the certainty equivalent outcome: losses are overweighted compared to gains. As a consequence of the kink at the reference point, risk aversion is locally higher closer to the reference point than further away from it.

A kinked, piece-wise linear utility function can be approximated as a linear combination of the excess return on a portfolio and a term that is closely related to the *semi standard deviation* of returns. With a smooth, twice differentiable, concave utility function, in contrast, utility can be approximated by a linear combination of the excess return and the *variance* of returns. Boender et al. (2006) employ a semi-variance criterion to evaluate consumption during retirement. One can view this criterion function as a special case of the loss-aversion specification in which gains above the reference point do not yield utility at all, agents are risk averse in losses, while the reference point equals the ambition level of the pension system (70% of the final wage in the case of Boender et al. (2006)). This study implicitly assumes that agents are risk neutral with respect to consumption during the working life because only the expected premium level is taken into account when evaluating policies. For an

extensive evaluation of this study and an interpretation of it in terms of utility functions, see Boender et al. (2006).

6.3.5 *Endogenous Reference Point*

The reference point in prospect literature plays a similar role as the habit level in the habit-formation literature.⁵⁹ Hence, the literature has explored what happens if the reference point adjusts over time in response to current consumption levels and rational investors anticipate this (see, e.g. Barberis et al. (2001), and Berkelaar et al. (2004)). Berkelaar et al. (2004) find that updating the reference point implies more risk taking. Intuitively, preferences are more flexible: with adverse shocks the point from which gains and losses are being evaluated and to which you want to be close is adjusted downwards.

Barberis et al. (2001) explore a setting in which investors derive direct utility from not only consumption but also from fluctuations in their financial wealth. After a run-up in stock prices, the agent gets further away from this reference point and thus becomes less risk averse. History thus determines risk aversion. Just as in Campbell and Cochrane (1999), this variation in risk aversion allows the equilibrium returns to be much more volatile than the underlying fundamentals. In contrast to the consumption-based approach of Campbell and Cochrane (1999), however, investors care about fluctuations in financial wealth even if they are unrelated to consumption.

Dybvig (1995) considers a model with extreme loss aversion and immediate adjustment of the reference point to changes in consumption: consumption is not allowed to decline.⁶⁰ Consumption develops smoothly and increases each time wealth reaches a new maximum. At each point in time, a part of wealth is set apart to match the risk-free liability of the current consumption level. The wealth in excess of this level is used to take risks so as to finance increases in consumption. The behavior of the consumer thus resembles that of a click fund or a so-called escalating annuity. This model obviously explains why consumption increases over time and tends to be low at the beginning of the life cycle. In fact, individuals want to save even if their income is expected to rise over time. Borrowing constraints are thus less relevant.

59 A difference is that consumption can fall below the reference point with loss aversion but that consumption cannot fall below the necessary consumption level. Loss aversion thus gives rise to less extreme portfolio insurance as with habits. In particular, a loss-averse investor wants to stay above the target level but this is not absolutely necessary. In bad states that are very expensive, the individual accepts losses and may actually become risk seeking to make up the losses in order to return to a point above the reference point. However, habits can also be specified in a way that allows consumption to fall below the habit. In particular, habits can be included multiplicatively rather than additively as in (16), i.e. C/\bar{C} rather than $C - \bar{C}$ enters the utility function. See Gomes and Michaelides (2003) and Davidoff et al. (2005).

60 This can be easily generalized for an exogenously specified minimum growth rate or maximum decline.

7 CONCLUSIONS

This paper has explored the academic literature on optimal saving and investment over the life cycle. We started out with a simple benchmark model in which human capital is riskless, standard preferences apply and equity risk is the only risk factor. Within this benchmark model we derived the optimal individual contract in both the absence and the presence of constraints on borrowing against future labor income. We also considered a number of collective pension contracts and discussed the advantages and disadvantages of collective versus individual pension schemes. The final section explored a number of extensions of the benchmark framework. We considered more elaborate models of human capital, additional risk factors in financial markets as well as non-standard specifications of preferences.

Section 6 indicates a wide variety of questions for future research on this fascinating and highly relevant topic. Among other things, it points to the importance of the nature of human capital for saving and investment behavior. This suggests that pension plans should be tailored to the type of human capital of individuals. Hence, pools in collective pension schemes should be based on that. Another implication of the relevance of human capital is that financial economics and labor economics should be integrated more. An important future research topic concerns saving and investment implications of the career pattern of women, which account for an increasingly important share of labor supply in modern economies. This career pattern differs in a number of important aspects from that of men. Family risks (i.e. elderly parents, young children) are new sources of idiosyncratic risk (see also the survey by Browning and Crossley (2001)).

Whereas endogenous labor supply extends the scope of risk sharing by adding consumption of leisure as an additional way to adjust to shocks, it also allows individuals to evade the risk-sharing contract ex post. In this way, it weakens the commitment of future workers to risk sharing with current generations. An important issue for future research is to endogenize to what extent pension schemes can adopt the continuity perspective by modelling the mobility of labor, capital, and consumers. In this way, we can model the optimal recovery period, i.e. optimal risk sharing among non-overlapping generations. As a short cut, we can impose a maximum tax that workers are willing to pay ex post as part of a risk-sharing contract.

REFERENCES

- Aarssen, Karin and Barthold Kuipers (2006), 'Iedereen wint, maar sommigen meer dan in anderen', in: S.G. van der Lecq and O.W. Steenbeek (eds.), *Cost and Benefits Of Collective Pension Systems*, Dutch, Kluwer.

- Abowd, J.M. and D. Card (1989), 'On the Covariance Structure of Earnings and Hours Changes,' *Econometrica*, 57, pp. 411–445.
- Ameriks, John, Andrew Caplin and John Leahy (2003), 'Wealth Accumulation And The Propensity To Plan,' *Quarterly Journal of Economics*, 118, pp. 1007–1047.
- Ameriks, John and Stephen Zeldes (2004), 'How Do Household Portfolio Shares Vary With Age?' Unpublished Paper, Columbia University.
- Ang, Andrew and Geert Bekaert (2006), 'Stock Return Predictability: Is It There?', *The Review of Financial Studies*, forthcoming.
- Barbaris, N., M. Huang and T. Santos, (2001), 'Prospect Theory and Asset Prices,' *Quarterly Journal of Economics*, 116, pp. 1–53.
- Beetsma, R.M.W.J. and A.L. Bovenberg (2007), 'Pensions, Intergenerational Risk Sharing And Inflation,' *Netspar Discussion Paper*.
- Benartzi, Shlomo and Richard Thaler (2001), 'Naive Diversification Strategies In Defined Contribution Savings Plans,' *American Economic Review*, 91, pp. 79–98.
- Benzoni L., P. Collin-Dufresne and R.S. Goldstein (2007), 'Portfolio Choice over the Life-Cycle when the Stock and Labor Markets are Cointegrated,' *Journal of Finance* (forthcoming).
- Berkelaar, Arjan, Roy Kouwenberg and Thierry Post (2004), 'Optimal Portfolio Choice under Loss Aversion,' *Review of Economics and Statistics*, 86, pp. 973–986.
- Bernatzi, S. and R. Thaler (1995), 'Myopic Loss Aversion and the Equity Premium Puzzle,' *Quarterly Journal of Economics*, 110, pp. 73–92.
- Bikker, Jaap and Jan de Dreu (2006), 'Uitvoeringskosten van pensioenverstrekkers', in: S.G. van der Lecq and O.W. Steenbeek (eds.), *Cost and benefits of collective pension systems*, Dutch, Kluwer.
- Blake, David, Andrew Cairns and Kevin Dowd (2006), 'Living with Mortality: Longevity Bonds and Other Mortality Linked Securities', *British Actuarial Journal*, 45, pp. 153–197.
- Bodie, Zvi, Robert C. Merton and William F. Samuelson (1992), 'Labor Supply Flexibility and Portfolio Choice in a Life-Cycle Model,' *Journal of Economic Dynamics and Control*, 16, pp. 427–449.
- Boeijen, Dick, Corne Jansen, Niels Kortleve and Jan Tamerus (2006), 'Leeftijdssolidariteit in de doorsneepremie', in: S.G. van der Lecq and O.W. Steenbeek (eds.), *Cost and Benefits Of Collective Pension Systems*, Dutch, Kluwer.
- Boender, C.G.E., A.L. Bovenberg, S. van Hoogdalem and Th.E. Nijman (2006), 'Optimale Risicodeling in Individuele en Collectieve Pensioencontracten,' in: S.G. van der Lecq and O.W. Steenbeek (eds.), *Cost and Benefits Of Collective Pension Systems*, in Dutch, Kluwer.
- Boender, C.G.E., S. van Hoogdalem, E. van Lochem and R.M.A. Jansweijer (2000), 'Intergenerationele Solidariteit en Individualiteit in de Tweede Pensioenpijler: Een Scenario-analyse,' Report No. 114, *Dutch Scientific Council for Government Policy (WRR)*, The Hague, The Netherlands.
- Bovenberg, Lans, Ralph Koijen, Theo Nijman and Coen Teulings (2007), 'Saving and investing over the life cycle and the role of collective pension funds', *Netspar Panel Paper*.
- Brav, Alon, George M. Constantinides and Christopher C. Geczy (2002), 'Asset Pricing with Heterogeneous Consumers and Limited Participation: Empirical Evidence,' *Journal of Political Economy*, 110, pp. 793–824.

- Brennan, Michael and Yihong Xia (2000), 'Stochastic Interest Rates and the Bond-Stock Mix', *European Finance Review*, 4, pp. 197–210.
- Brown, Jeffrey (2002), Differential mortality and the value of individual account retirement annuities, in: Feldstein and Liebman (eds.), *The Distributional Aspects of Social Security and Social Security Reform*, Chicago Press.
- Brown, Jeffrey, Olivia Mitchell and James Poterba (2001), 'The Role of Real Annuities and Indexed Bonds in an Individual Accounts Retirement Program', in: John Campbell and Martin Feldstein (eds.), *Risk Aspects of Investment-Based Social Security Reform*, University of Chicago Press, pp. 321–360.
- Browning, Martin and Thomas Crossley (2001), 'The Life-Cycle Model of Consumption and Saving', *Journal of Economic Perspectives*, 15, pp. 3–22.
- Browning, Martin and Annamaria Lusardi (1996), 'Household Savings: Micro Theories and Micro Facts', *Journal of Economic Literature*, 34, pp. 1797–1855.
- Campbell, John (2006), 'Household Finance', *Journal of Finance*, 61(4), pp. 1553–1603.
- Campbell, J.Y. and J.H. Cocrane (1999), 'By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior', *Journal of Political Economy*, 107, pp. 205–251.
- Campbell, John and Luis Viceira (2001), 'Who should buy: Long term Bonds?', *American Economic Review*, 91, pp. 99–127.
- Campbell, John and Luis Viceira (2005), 'The Term Structure of the Risk-Return Trade-off', *Financial Analysts Journal*, 61(1), pp. 34–44.
- Campbell, John and Motohiro Yogo (2006), 'Efficient tests of Stock Return Predictability', *Journal of Financial Economics*, 81, pp. 27–60.
- Chacko, George and Luis Viceira (2005), 'Dynamic Consumption and Portfolio Choice with Stochastic Volatility in Incomplete Markets', *Review of Financial Studies*, 8(4), pp. 1369–1402.
- Choi, James, David Laibson and Brigitte Madrian (2005), '\$100 Bills on the Sidewalk: Suboptimal Saving in 401(k) Plans', working paper.
- Choi, James, Brigitte Madrian, Andrew Metrick and David Laibson (2002), 'DC Pensions: Plan Rules, Participant Decisions, and the Path of Least Resistance', *Tax Policy and the Economy*, 16, pp. 67–114.
- Clark, Robert, Madeleine d'Ambrosio, Ann McDermed and Kshama Sawant (2006), 'Retirement Plans and Savings Decisions: The Role of Information and Education' *Journal of Pension Economics and Finance*, 5(1), pp. 45–68.
- Clark, Gordon, Emiko Caerlewy and John Marshall (2006), 'Pension Fund Trustee Competence: Decision Making in Problems Relevant to Practice', *Journal of Pension Economics and Finance*, 5(1), pp. 91–110.
- 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, pp. 492–533.
- Cohen, Alma and Liran Einav (2005), 'Estimating Risk Preferences from Deductible Choice', *American Economic Review* (forthcoming).
- Cochrane, John (2006), 'The Dog that did not bark: a Defense of Return Predictability', Working paper, University of Chicago.
- Cochrane, John and Monica Piazzesi (2005), 'Bond risk premia', *American Economic Review*, 95, pp. 138–160.

- Constantinides, George M. (1990), 'Habit formation: A resolution of the equity-premium puzzle,' *Journal of Political Economy*, 48, pp. 519–543.
- Constantinides, George M., John B. Donaldson and Rajnish Mehra (2002), 'Junior can't Borrow; A New Perspective on the Equity Premium Puzzle,' *Quarterly Journal of Economics*, 117, pp. 269–296.
- Constantinides, George M. and Durrell Duffie (1996), 'Asset Pricing with Heterogeneous Consumers,' *Journal of Political Economy*, 105, pp. 219–240.
- Cui, J., F.C.J.M. de Jong, and E.M.H. Ponds (2006), 'Intergenerational Risk Sharing within Funded Collective Pension Schemes', SSRN Working Paper.
- Davidoff, Thomas, Jeffrey A. Brown and Peter Diamond, (2005), 'Annuities and Individual Welfare,' *American Economic Review*, 95, pp. 1573–1590.
- Davis, S.J., F. Kubler and P. Willen (2005), 'Borrowing Costs and the Demand for Equity over the Life Cycle,' Federal Reserve Bank of Boston, Working Paper No. 05-7.
- Dufló, Ester and Emanuel Saez (2003), 'The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment', *Quarterly Journal of Economics*, 68, pp. 815–842.
- Dybvig, P.H. (1995), 'Dusenberry's Ratcheting of Consumption: Optimal Dynamic Consumption and Investment Given Intolerance for Any Decline in Standard of Living,' *Review of Economic Studies*, 62, pp. 287–313.
- Farhi, E. and S. Panageas (2007), Saving and Investing for Early Retirement: A Theoretical Analysis. *Journal of Financial Economics*, 83, pp. 87–121.
- Finkelstein, Amy, James Poterba and Casey Rothschild (2005), 'Redistribution by Insurance Market Regulation: Analyzing a Ban on Gender-Based Retirement Annuities', *NBER Working Paper W12205*.
- Gollier, Christian (2005), 'Understanding Saving and Portfolio Choices with Predictable Changes in Asset Returns', Working paper, University of Toulouse.
- Gollier, C. and R.J. Zeckhauser (2002), 'Horizon Length and Portfolio Risk,' *Journal of Risk and Uncertainty*, 24 (3), pp. 195–212.
- Gollier, C. (2006), 'Intergenerational risk sharing and risk taking in a pension fund,' Working paper, University of Toulouse.
- Gollier, C. and J.W. Pratt (1996), 'Risk Vulnerability and the Tempering Effect of Background Risk', *Econometrica*, 64, pp. 1109–1124.
- Gomes, Francisco and Alexander Michaelides (2003), 'Portfolio Choice with Internal Habit Formation A Life-Cycle Model with Uninsurable Labor income Risk,' *Review of Economic Dynamics*, 6, pp. 729–766.
- Goyal, Amid and Ivo Welch (2006), 'A comprehensive look at the empirical performance of equity premium prediction', *The Review of Financial Studies*, (forthcoming).
- Guiso, Luigi and Tullio Japelli (2005), 'Awareness and stock market participation', *Review of Finance*, 9, pp. 537–567.
- Gul, F. (1991), 'A theory of disappointment aversion', *Econometrica* 59, pp. 667–686.
- Haliassos, Michael and Carol Bertaut (1995), 'Why do so few hold stocks', *Economic Journal*, 105, pp. 1110–1129.
- Hoevenaars, R. and E. Ponds, 'Valuation of Intergenerational Transfers in Funded Collective Pension Schemes,' Netspar Discussion Paper 2006–19.

- Huberman, Gur (2001), 'Familiarity breeds investment', *Review of Financial Studies*, 14, pp. 659–680.
- Jong, F.C.J.M. de (2005), 'Valuation of Pension Liabilities in Incomplete Markets,' *Netspar Discussion Paper* 2005.
- Koijen, Ralph, Theo Nijman and Bas Werker (2006a), 'Optimal portfolio choice with annuitization', *Netspar Discussion Paper*.
- Koijen, Ralph, Lans Bovenberg and Theo Nijman (2007), 'Optimal retirement provision: comparing individual and collective pension schemes', Working paper, in progress.
- Koijen, Ralph, Theo Nijman and Bas Werker (2006b), 'When can Life-cycle Investors Benefit from Time-Varying Bond Risk Premia?', *Netspar Discussion Paper*.
- Kortleve, Niels and Eduard Ponds (2006), 'Pension deals and value-based ALM', in: Niels Kortleve, Theo Nijman and Eduard Ponds (eds.), *Fair Value and Pension Fund Management*, Elsevier Publishers.
- Laibson, David, Andrea Repeto and Jeremy Tobacman (1998), 'Self-Control and Saving for Retirement', *Brookings Papers on Economic Activity*, 1998(1) pp. 91–196.
- Leland, H.E. (1968), Savings and Uncertainty: The Precautionary Demand for Savings, *Quarterly Journal of Economics*, 45, pp. 621–636.
- Lettau, Martin and Stijn van Nieuwerburgh (2006), 'Reconciling the Return Predictability Evidence', the *Review of Financial Studies*, forthcoming.
- Liu, Jun (2007), Portfolio Selection in Stochastic Environments, *Review of Financial Studies*, 20(1), 39.
- Lusardi, Annamaria (1999), 'Information, expectations and savings for retirement' in: Henry Aron (ed.), *Behavioral Dimensions of Retirement Economics*, Washington DC, Brookings Institution, pp. 81–115.
- Merton, R. C. (1969), 'Lifetime portfolio selection under uncertainty: The continuous-time case,' *Review of Economics and Statistics*, 51, pp. 247–257.
- Merton, R. C. (1971), 'Optimum consumption and portfolio rules in a continuous-time model', *Journal of Economic Theory*, 3, pp. 373–413.
- Merton, R. C. and P.A. Samuelson (1974), 'Fallacy of the log-normal approximation of optimal portfolio decision-making over many periods', *Journal of Financial Economics* 1, pp. 67–94.
- Mitchell, Olivia and Stephen Utkus (2003), 'The Role of Company Stock in Defined Contribution Plans' in: Olivia Mitchell and Kent Smetters (eds.), *The Pension Challenge: Risk Transfers and Retirement Income Security*, Oxford University Press.
- Mitchell, Olivia, James Poterba, Mark Warschavsky and Jeffrey Brown (1999), 'New evidence on the Money's Worth of Individual Annuities', *American Economic Review*, 89, pp. 1299–1318.
- Mossin, J. (1968), 'Optimal multiperiod portfolio policies,' *Journal of Business*, 41, pp. 215–229.
- Munnell, Alicia and Annika Sunden (2004), 'Coming Up Short: The Challenge of 401(k) plans', The Brookings Institution, Washington.
- Munk, Claus and Carsten Sorensen (2005), 'Dynamic Asset Allocation with Stochastic Income Interest Rates', Working paper.

- Nordahl, Helge and Trond Dokseland (2006), 'Intergenerational Effects of Guaranteed Pension Contracts', Working Paper, Norwegian School of Economics.
- Poterba, James (2006), 'Annuity markets', in: Gordon Clark, Alicia Munnell and Michael Orszag, (eds.), *The Oxford Handbook of Pensions and Retirement Income*, Oxford University Press.
- Poterba, James, Joshua Rauh, Steven Venti and David Wise (2005), 'Utility evaluation of risk in retirement savings accounts', in: D. Wise (ed.), *Analyses in the Economics of Aging*, Chicago University Press, Chicago.
- Kooreman, Peter and Henriette Prast (2007), 'Policy implications of behavioral economics', *Netspar Reference Paper*.
- Samuelson, P. A. (1969), 'Lifetime portfolio selection by dynamic stochastic programming,' *Review of Economics and Statistics*, 51, pp. 239–246.
- Sangvinatsos, Antonios and Jessica Wachter (2005) 'Does the Failure of the Expectations Hypothesis matter for Long Term Investors?', *The Journal of Finance*, 60, pp. 179–230.
- Teulings, Coen and Casper de Vries (2006), 'Generational Accounting, Solidarity and Pension Losses,' *De Economist*, 146, pp. 63–83.
- Thaler, Richard and Shlomo Benartzi (2004), 'Save More Tomorrow', *Journal of Political Economy*, 112/1, pp. 164–187.
- Topel, R.H. and M.P. Ward (1992), 'Job Mobility and the Careers of Young Men,' *Quarterly Journal of Economics*, 107, pp. 439–479.
- Tversky, Amos and Daniel Kahnemann (1992), 'Advances in Prospect Theory: Cumulative Representation of Uncertainty,' *Journal of Risk and Uncertainty*, 5, pp. 297–323.
- Van Rooij, Maarten, Henriette Prast and Clemens Kool (2007), 'Risk return preferences in the pension domain: are people able to choose?', *Journal of Public Economics*, Vol. 91, p. 701–722.
- Venti, Steven (2006), 'Choice, Behaviour and Retirement Savings', in: Gordon Clark, Alicia Munnell and Michael Orszag, (eds.), *The Oxford Handbook of Pensions and Retirement Income*, Oxford University Press.
- Viceira, L.M. (2001), 'Optimal Portfolio Choice for Long-Horizon Investors with Nontradable Labor Income,' *Journal of Finance*, 56, pp. 433–470.