Risk, redistribution and retirement

Bonenkamp, J.

Publication date:
2013

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

Risk, redistribution and retirement: the role of pension schemes

Risk, redistribution and retirement: the role of pension schemes

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan Tilburg University op gezag van de rector magnificus, prof. dr. Ph. Eijlander, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit op woensdag 5 juni 2013 om 14.15 uur door

JOANNES PAULUS MARIA BONENKAMP

geboren op 13 september 1979 te Hengelo.

PROMOTIECOMMISSIE

PROMOTORES:	Prof. Dr. C. van Ewijk
	Prof. Dr. A.C. Meijdam
OVERIGE LEDEN:	Prof. Dr. A.L. Bovenberg
	Prof. Dr. E.H.M. Ponds
	Prof. Dr. H.A.A. Verbon
	Dr. E.W.M.T. Westerhout

ACKNOWLEDGEMENTS

In October 2007, I started my PhD research at CPB Netherlands Bureau for Economic Policy Analysis as part of the Netspar theme 'The macroeconomics of pension reform'. At that time, Dutch pension funds were doing well financially. The average funding ratio, i.e., the amount of pension assets divided by the liabilities, was more than 140%. This changed dramatically at the end of 2008 when the funding ratio dropped to 95%, after the bankruptcy of Lehman Brothers and the following world-wide collapse of financial markets. At this moment, when I make the last changes to this thesis, the financial situation of pension funds has improved to some extent but is still far from optimistic. As researcher working on pension issues, these five years were without doubt an interesting and challenging period, not least because the crisis moved pensions to the front page of the newspapers. There were also times that writing this thesis felt more like a lesson in persistence, in particular when urgent policy work threatened to swallow up all my research time. Fortunately, there were several people who helped and supported me during these years of study and I am greatly indebted to them.

First of all, I would like to thank my supervisors Casper van Ewijk and Lex Meijdam for their guidance. Despite their tight schedules, they always had time to discuss my research. Their constructive comments and suggestions have had a huge impact on my thesis. With pleasure I look back to the meetings with Lex in Tilburg, where we not only talked about my research but also about daily issues. I very much appreciated and enjoyed this personal contact. I also want to extend my gratitude to Casper in his position as deputy director of CPB, who allowed me to spend a large part of my working time on this thesis.

At CPB, during this project I greatly enjoyed the support and guidance of Ed Westerhout, in particular with respect to Chapter 3. Also his constructive comments on other part of the manuscript as member of the PhD committee are greatly appreciated. I am also grateful to Andre Nibbelink for his help with all the computational problems I had to overcome. I also want to thank my former department heads Peter Kooiman, Ruud de Mooij and Bas ter Weel for their support.

I am also thankful to Yvonne Adema for her guidance in our joined research projects which form the basis of Chapter 4 and Chapter 5. I appreciate our cooperation which already has its origins from 1998, when we were assigned to the same tutorial group during our study Economics in Groningen.

Furthermore, I am grateful to Lans Bovenberg, Eduard Ponds and Harrie Verbon for the comments and suggestions they gave me during the pre-defense and their willingness to join my PhD committee.

Ook wil ik mijn familie bedanken, in het bijzonder mijn ouders Paul en Anneke, voor jullie oprechte vertrouwen en betrokkenheid met mij, Rebekka en Lucas. Mijn speciale dank gaat ook uit naar mijn broers Daan, Dries en Teun voor de talloze keren dat ik bij jullie terecht kon om reistijd uit te sparen. Ook ben ik veel dank verschuldigd aan mijn schoonfamilie, niet in de laatste plaats mijn schoonouders Numan en Hatun, voor jullie belangstelling maar ook voor de vele uren dat jullie Lucas met liefde hebben opgevangen als ik mijn zorgplicht weer niet kon nakomen.

Mijn grootste dank, tot slot, gaat uit naar mijn lieve vrouw Rebekka. Ik realiseer me goed dat de vrijheid die jij me de afgelopen jaren hebt gegeven om dit proefschrift te kunnen schrijven, verre van vanzelfsprekend is geweest. Dat geldt misschien nog wel meer voor jouw onvoorwaardelijke liefde en steun die ik elke dag opnieuw weer mag ervaren.

Jan Bonenkamp
April 2013

CONTENTS

CHAPTER 1

INTRODUCTION

Amsterdam, 2 October 2004. More than 300,000 supporters of the political opposition and members of labour unions went to Amsterdam to demonstrate against the plans of the government to abolish the favourable tax treatment of early retirement schemes in the Netherlands. In these schemes pension premiums were deductible from the worker's gross salary, while early retirement benefits were being taxed as if they were a regular source of income. Due to the progressive tax system the tax advantage was considerable. With this policy measure, the government aimed to increase the labour force participation of elderly. Despite all the massive protests, the law came into force in 2006 and turned out to be quite successful. The effective age of labour market exit has increased by 2.1 years, from age 61 in 2006 to age 63.6 in 2012.[1]

The Hague, 29 June 2011. Delegates of groups promoting the interests of young workers frustrated the parliamentary debate about a proposal to reform the pension system in the Netherlands. This pension proposal, agreed between social partners and supported by the Dutch government, contains an increase in the pension entitlement age and a conversion of accrued pension entitlement from nominal guarantees to entitlements which are more conditional on stock market performance. The proposal is an answer to population ageing and the global financial crisis which have put the funding ratios to historical

[1]Numbers are obtained from the StatLine database of Statistics Netherlands (www.cbs.nl).

minima. At the end of 2012, the average funding ratio of Dutch pension funds amounted to 102%.[2]

As illustrated by these two events, any reform in collective pension systems evokes large social resistance. The illustrations are referring to recent pension reforms in the Netherlands, but we could equally well take examples from other countries like Greece, France or Italy. The social involvement with pension reforms is a widespread phenomenon in all modern countries. If we just take a look to some statistics, it is not surprising that people are concerned with their pension entitlements. In most countries, the share of first-pillar and second-pillar benefits in total pension income is larger than 80%. In the Netherlands, this share is about 90% (DNB, 2009). The average replacement rate, the ratio between pension income and last-earned labour income, is 57% in the OECD countries, albeit with great variation across countries. In the Netherlands the average replacement rate is much higher and amounts to 88%. Taking the first pillar and second pillar together, the average share of pension expenditures in GDP amounts to 8.4% in the OECD countries. The Netherlands is again a positive outlier; there the percentage of pension expenditures is 10% of GDP (OECD, 2011).

Of course, these kinds of financial statistics do not tell the whole story why people have a strong emotional involvement with pension reforms. Also the way pensions are financed plays an important role. Most of first-pillar and second-pillar programmes, funded and unfunded, are financed on a *collective* basis. This means that in principle all participants contribute the same percentage of their salaries to the pension contract. This type of financing always has distributional implications within generations and across generations. The demonstrators in Amsterdam, in October 2004, were predominantly older workers who more or less viewed their early retirement entitlements as vested rights for which they had paid contributions in the past. The activists who interrupted the parliamentary debate in June 2011 were young workers who feared that their contributions would mainly serve to protect the pension entitlements of the elderly and would not sufficiently result in future pension benefits for themselves.

These types of generational conflicts of interest sometimes mask that collective pension schemes also contain elements which in principle could be beneficial for all participants. There is convincing evidence that large-scale pension

[2]This number is obtained from online statistics of the Dutch Central Bank (www.dnb.nl).

funds benefit from economies of scale. Operating costs of collective pension funds, like administrative and investment costs, are typically much lower than those of individual pension schemes. More importantly, collective funds may provide insurance for all kinds of risks, like longevity shocks or wage shocks, which could not be insured by individual pension arrangements. Last but not least, in contrast to individual schemes, collective pension schemes may also provide intergenerational risk-sharing arrangements enabling pension funds to smooth the effect of shocks beyond the life time of single individuals.

This thesis studies the effects of redistribution and risk sharing in collective pensions on economic behaviour, with particular attention paid to the retirement decision. Throughout the analysis, we make a clear distinction between redistribution and risk sharing. *Risk sharing* comes into play after a person for whatever reason suffers a loss and other people entirely or partially compensate him for that. This form of solidarity is the basic principle of insurance contracts. That means, *ex ante*, i.e., before the occurrence of an event, people do not know whether they will be net receivers or net contributors. To illustrate, people with fire insurance on their house do not know in advance whether they will have to claim or not at the moment the contract is signed. *Redistribution*, though, is a form of solidarity that is independent of a certain event occurring, but in advance, based on information about individual characteristics, leads to a certain transfer between participants. For example, high-income persons know that they will most likely transfer money to low-income persons when both types of individuals participate in a social security system with flat benefits (independent of past earnings).

This distinction between risk sharing and redistribution is not always clear and, in a certain sense, artificial. The reason for this is that, once a shock has occurred, risk sharing turns into redistribution. Generations that enter a pension fund which has a huge funding deficit can reasonably expect that according to current expectations they should pay in more than they will ever receive. Hence, existing deficits (or surpluses) can be viewed as a form of pure redistribution rather than risk sharing. After all, a pension contract contains an implicit agreement with future generations to always communicate all unexpected surpluses and deficits. Reasoning along this line, existing deficits and surpluses are more a type of risk sharing than redistribution.

Our analysis on redistribution and risk sharing will be performed in the light of three important world-wide pension reforms, namely the switch from collective defined-benefit (DB) pension schemes to more individual defined-contribution (DC) pension schemes, the increase in the pension entitlement age and the introduction of more flexibility in this entitlement age. These kinds of reforms are currently also on the policy agenda in the Netherlands and will have important implications for redistribution and risk sharing in the first and second pension pillar. Our research will focus on both pension pillars, dependent on the specific reform considered. With the Dutch policy agenda in mind, the introduction of a flexible pension entitlement age is particularly relevant for the unfunded first pillar, the switch from DB to DC contracts is more tied to the funded second pillar whereas an increase in the pension entitlement age applies to both pillars.

Redistribution, especially from rich to poor in unfunded systems, is often viewed as one of the main objectives of a pension system. According to the well-known proposal of the World Bank (World Bank, 1994), the first pillar should exactly perform this task, while the income smoothing function and the risk-sharing function should be achieved by the second pillar. While redistribution and risk sharing certainly have advantages in terms of preventing old-age poverty and completing incomplete markets, they may also come along with a welfare cost. Pension contributions are usually levied proportional to labour income. Then the transfers (arising from redistribution or risk sharing) break down the link between contributions and benefits and therefore distort the labour-leisure decision. Hence, collective pension schemes always face a trade-off between on the one hand providing redistribution and risk sharing and on the other hand minimizing labour-supply distortions. The aim of this thesis is to get a better understanding how this trade-off will be affected by pension reforms that increase the risk properties of pension benefits or increase the level and flexibility of the pension entitlement age.

When the traditional DB schemes came into prominence in the 50s and 60s of last century, there were relatively much young workers protecting older generations from risks. Nowadays, the ageing of the population has increased the weight of older compared to younger generations. Therefore, to protect the pension benefits of the elderly larger adjustments in the contribution rate are needed. Faced with more uncertainty about their pension contribution, young workers may become less inclined to act as the residual claimant. As a result, pension claims are increasingly being made contingent on shocks in longevity and on developments in financial markets. Besides ageing, also financial considerations contribute to a lower popularity of traditional DB pensions, especially in case of corporate pension funds. The volatility of financial markets in the recent decade has confronted corporations with the risks of their corporate funds. To take control of the overall pension costs, corporations are moving to pension schemes with fixed contributions. *A priori* it is not immediately clear that a movement from collective funded DB pensions towards more individual funded DC pensions will improve social welfare. This depends to a large extent on the pros and cons of risk sharing. Indeed, the added value of collective DB pensions is that these contracts can spread the effects of shocks over more generations. A switch to individual DC implies that this intergenerational risk sharing is no longer possible. In this thesis, we give a theoretical and quantitative analysis of the value of intergenerational risk sharing. This analysis requires a neat comparison between the insurance gains of risk sharing and the welfare losses due to labour-supply distortions.

The increasing burden of ageing also has implications for first-pillar pensions. In an attempt to keep the public pension costs in control, many countries will increase or have already increased the official retirement age. Raising the retirement age certainly has implications for the level and direction of redistribution. The reason is that the policy, if performed in a uniform way, will typically affect all workers while not all individuals will benefit equally from an increase in longevity. Indeed, there is evidence that the ongoing decline in mortality is disproportionately concentrated among the more-educated part of the population. This can make the pension scheme more regressive after an increase in the pension entitlement age. It should be stressed, however, that redistribution effects are not only determined by (rather exogenous) differences in life expectancy but are also the result of behavioural responses. In this thesis, we will investigate the redistribution effects of an increase in the official retirement age (as response to ageing) taking into account that people are different in terms of life expectancy and ability and may have different preferences for retirement.

Apart from increasing the statutory retirement age, many countries have also taken measures to accommodate more choice in pension systems. Early retirement schemes which often contained large implicit taxes on continued activity have been replaced by pension systems that allow for a flexible choice of the retirement age with more or less actuarially-neutral adjustments.[3] Also

[3] In this thesis, we make a distinction between actuarial fairness and actuarial neutrality. *Actuarial fairness* requires that the present value of (life-time) contributions equals the present value of (life-time) benefits. *Actuarial neutrality* is a marginal concept and requires that the present value of accrued benefits for working an additional year is the same as in the year before (Queisser and Whitehouse, 2006).

in the Netherlands there are plans to introduce flexible pension take-up in the first pillar. This move from inflexible to flexible contracts raises several interesting (and policy-relevant) questions, both on the issue of risk sharing and redistribution. For example, to what extent can flexible retirement provide insurance to workers against all type of pension risks? We will deal with this question in this thesis. We shall consider to what extent the hedging function of retirement flexibility depends on individual preferences and on the type of risk factor the individual is exposed to.

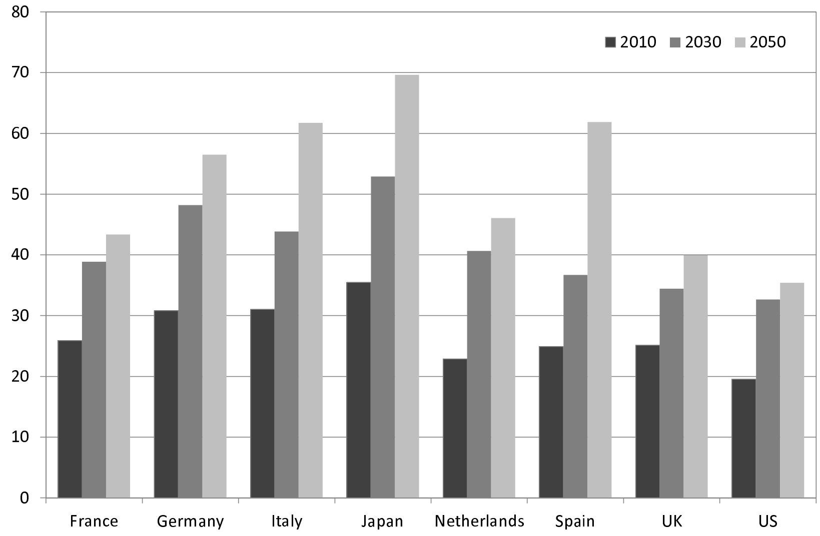
With respect to redistribution, an actual issue is how to determine the actuarial adjustment of benefits when an individual retires earlier or later than the normal retirement age. In most real-world pension schemes this adjustment is based on some average life expectancy index. This leads to redistribution because the adjustment rate is only actuarially neutral for some *average* individual. For workers with long life spans the reward rate of later retirement is typically too high; for individuals with short life spans, however, this rate is too low. In the final part of the thesis we study the redistribution and welfare effects of flexible pension take-up. We will consider alternative ways to adjust benefits. We also investigate whether flexible pension take-up could alleviate the classical trade-off between equity and efficiency. The idea is that individuals, when they are able to choose their retirement age themselves, can avoid to some extent implicit taxation in the redistributive pension scheme. If flexibility would stimulate people with higher life expectancies (typically the more wealthy people) to continue working it may also foster redistribution.

The rest of this introductory chapter is organized as follows. In Section 1.1 we illustrate some stylized facts about the pension reforms that underlie the questions we try to answer in this thesis. Section 1.2 will explore the main concepts used in the thesis. We end the chapter by giving an outline of the thesis in Section 1.3.

1.1 Pension reforms

In this section we give some facts and figures about the switch to more risk-bearing pension benefits, the increase in the pension eligibility age, and finally, the introduction of flexibility in the pension entitlement age. We start this section with a brief overview of some demographic, economic and social trends which have triggered these reforms.

Figure 1.1: Old-age dependency ratio



Source: United Nations, Population Division, World Population Prospects: The 2010 Revision (medium variant). The old-age dependency ratio is defined as the population aged 65+ divided by the population aged 15-64.

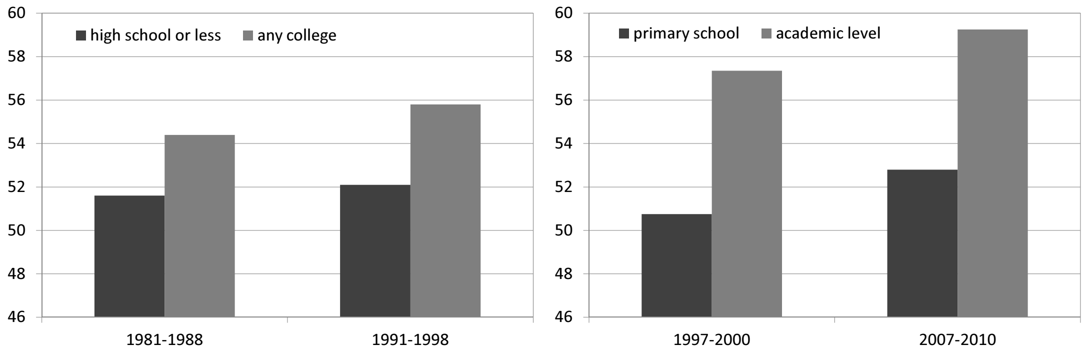
1.1.1 Setting the scene

The shift in the age distribution towards older ages, *population ageing*, is the most prominent global demographic trend of this century and has far-reaching implications for the sustainability and shock resistance of pension schemes. It is a direct consequence of the global fertility decline after World War II and of the ongoing mortality decline at older ages. After 1945 there was a baby boom which continued until the 1960s. In Western Europe, for example, the fertility rate, i.e., the average number of births per woman, was 2.4 in 1950. In the US, the fertility rate was even higher at that time, namely 3.4.[4] From the 1970s, though, fertility rates dropped significantly in the developed countries, mainly because of the wide introduction of contraceptives and the increased labour participation of women. Nowadays, the average births per woman are 1.7 in Europe and 2.4 in the US. For the coming decades, population projections foresee some improvement in the fertility rates, but they will stay at least in Europe below the replacement level of 2.1.

The other important source of population ageing is the decline in mortality

[4] All demographic figures mentioned in this subsection are taken from the World Population Prospects: The 2010 Revision (medium variant) of the United Nations.

Figure 1.2: Total life expectancy at age 25 by education level



(a) United States (b) Netherlands

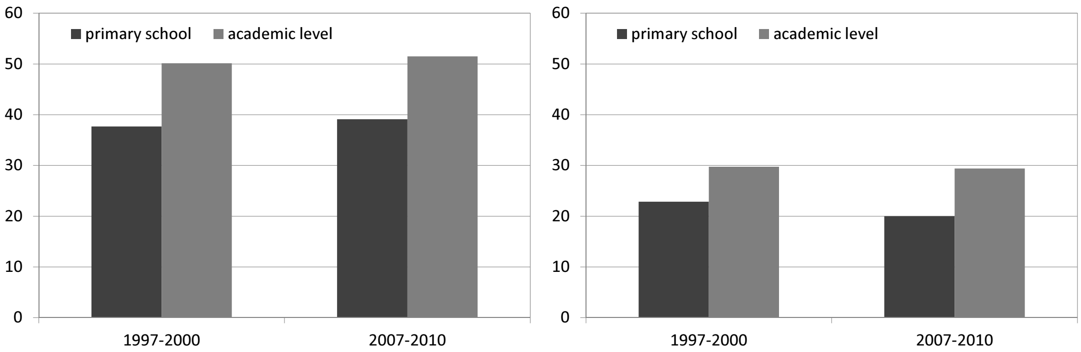
Source: US data are obtained from Meara et al. (2008); Dutch data are obtained from the CBS StatLine database and averaged over men and women.

rates. In 1950, the average life expectancy at birth in the US was 68.6 years. Nowadays, newborn US citizens are expected to live 78.8 years. In Western Europe, the rise in life expectancy was even bigger, from 67.8 in 1950 to 81.0 in 2010. This higher life expectancy is mainly the result of better nutrition and hygiene as well as better medical care. Contrary to the fertility drop, the rising trend in life expectancy seems to be a permanent process. Demographic projections for the US foresee that life expectancy at birth further increases to 83.6 in 2050. For Western Europe, current projections assume that newborns will on average live until 85.7 in 2050.

An often-used indicator for ageing is the old-age dependency ratio: the ratio of the number of people aged 65 and older to the working-age population. Figure 1.1 shows the development of the old-age dependency ratio for some developed countries. For all these countries, the dependency ratio will increase in the coming decades. Notice that the Netherlands take a middle position with a projected dependency ratio of 46% in 2050. In the US and UK the ageing of the population is relatively mild. For these countries the dependency ratio is expected to reach 35% and 40% in 2050, respectively. The absolute outlier on the top is Japan where the dependency ratio will increase to almost 70% in 2050. This high dependency ratio can be explained by both very low fertility and mortality rates. Also the populations in Italy and Spain are rapidly ageing; there the dependency ratio is expected to increase to 62% in the coming decades.

Although average life expectancy is rising, there is still a large educational disparity in life expectancy. More-educated people live on average signifi-

Figure 1.3: Life expectancy in good health at age 25 by education level



(a) Without disabilities (b) No chronic illness

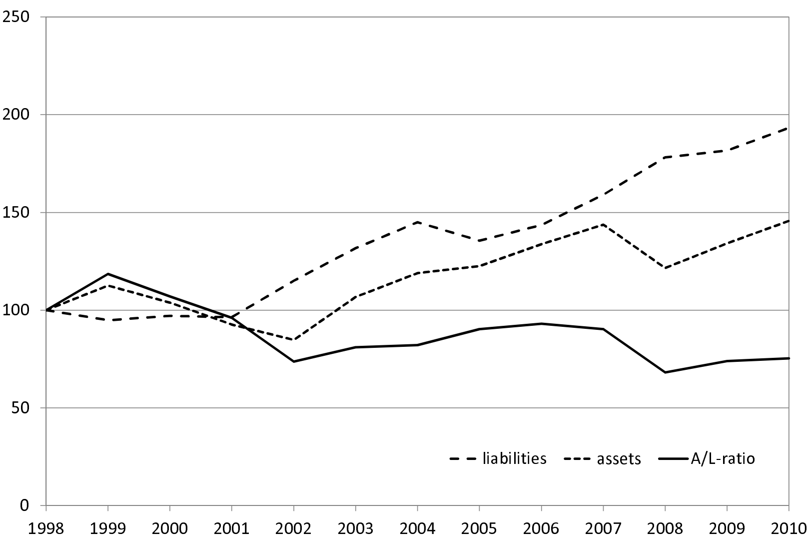
Source: Data are obtained from the CBS StatLine database and these data are averaged over men and women.

cantly longer than less-educated people. This difference can partly be attributed to risk factors like smoking and obesity which are more common among the less educated than among the better educated. In fact, for many countries there is evidence that the educational gap in life expectancy has widened over the last decades (see e.g., Mackenbach et al., 2003; Pappas et al., 1993). In the US, for example, life expectancy at age 25 grew 1.6 years for the high-education groups between 1990 and 2000, but remained unchanged for the low-education groups (see Figure 1.2a).

The Netherlands does not completely fit into this picture. The life expectancy of a high-educated person is much larger than that of a low-educated person, but the educational gap has not changed in the last decade. Over the period 1997-2010, the life expectancy of people with academic schooling remained about 6.5 years higher than the life expectancy of persons with only primary education (see Figure 1.2b). If we focus on life expectancy in good health rather than total life expectancy, the picture becomes more mixed (see Figure 1.3). Statistics Netherlands uses different definitions of good health: defined in terms of no disabilities, the educational gap in life expectancy has also hardly changed over the last decade, although the level of this gap is twice as much as that in total life expectancy. However, if we focus on life expectancy in the absence of chronic illness, there is an increase in the educational gap of 2.5 years over the period 2000-2010.

The financial sustainability problems of social security systems and pension funds due to population ageing have been strengthened by the *global financial*

Figure 1.4: DB asset/liability index (1998=100)

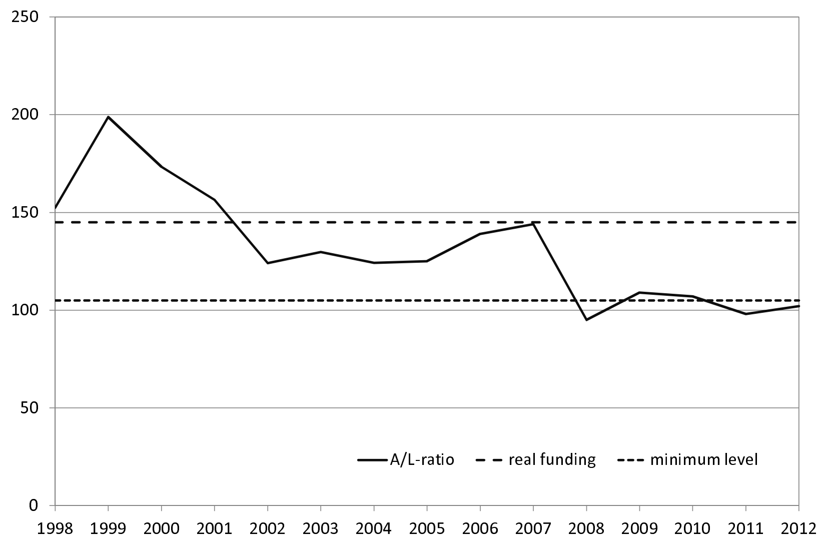


Source: Global Pension Asset Study 2011, Towers Watson. The following countries are included: Australia, Brazil, Canada, France, Germany, Hong Kong, Ireland, Japan, Netherlands, South Africa, Switzerland, UK and US.

crisis which started in 2008. It resulted in the collapse of large financial institutions, the bailout of banks by governments, and downturns in world-wide stock markets. The financial crisis was triggered by a complex interplay of valuation and liquidity problems in the US banking system. The collapse of the US housing bubble caused the values of securities tied to real estate pricing to drop, damaging financial institutions all around the world. Economies went into a recession during this period, as credit tightened and international trade declined. Governments and central banks responded with huge fiscal incentives policies and monetary policy expansion. At the beginning of 2010, the financial crisis reached a second phase with serious concerns about the sustainability of sovereign debt levels, especially among certain countries within the Euro Area. Since then, European leaders have taken several measures to restore financial stability. These measures included a rescue package with public loan guarantees and proposals to create a common fiscal union with enforceable solvency rules.

The financial crisis has had (and still has) a large impact on pension funds as developed countries' stock markets lost about 45% of their value in 2008. Figure 1.4 shows the development of the asset index (dotted line) and liability index (dashed line) of world-wide DB pensions. It also shows the funding ratio (solid line) which is defined as the ratio between the asset and liability

Figure 1.5: Dutch DB asset/liability index



Source: Dutch Central Bank.

index. Countries with typical DB pensions are Canada, the UK, Switzerland, the US and the Netherlands. After a steady increase from 2002 onwards, the asset index dropped by 15% in 2008. Notice that this decline is lower than the downturn in stock markets because not all pension wealth is invested in stocks. Since pension liabilities increased by 10% in 2008, the funding ratio dropped with 25% in one year time, from a level of 90% in 2007 to 68% in 2008.

From an international point of view, the drop in the average funding ratio of Dutch pension funds is rather large. Figure 1.5 shows the nominal funding ratio (solid line) of Dutch pension funds in the period 1998-2012. Before the onset of the financial crisis (end 2007), the nominal funding ratio was 144%. Within one year the funding ratio dropped by 35% to a historical minimum of 95%. After a short recovery in 2009, the funding ratio again started to decline at the beginning of 2010. At the end of 2012 the funding ratio amounted to 102%, 3%-points below the minimum required level (dotted line). In terms of inflation-adjusted pensions, the situation is even worse: to guarantee that pensions can be adjusted with inflation or wage growth, pension funds need a funding ratio of about 145% (dashed line).

Two main factors contributed to this extreme fall in Dutch funding ratios. Obviously, as stressed before, the world-wide collapse of share prices is an important determinant. But interest-rate developments played a significant role as well. Dutch pension funds are obliged to discount nominal pension liabilities with the nominal market yield curve. The sharp decline of nominal interest rates since the start of the financial crisis increased the present value of nominal pension liabilities considerably. This was not compensated by the increase in the value of bonds on the asset side of the balance sheet because the asset side did not match the structure of the nominal liabilities: not only were bond holdings much smaller than nominal pension liabilities, also the duration of the bonds owned by pension funds was much shorter than that of pension liabilities.

The development of pension schemes is also affected by important *social trends*, like the increasing labour mobility, the growing heterogeneity of the labour force and the ongoing emancipation of women. The traditional household model with the husband as breadwinner working his whole career for the same employer disappears more and more. Temporary labour contracts become a growing part of the range of contracts in modern labour markets. In the Netherlands, for example, the share of temporary employment has increased from 11% in 1996 to about 16% in 2008 (Cörvers et al., 2011). At the same time, there is a trend going on from salaried employment to self employment. During the period 1996-2009, the share of self-employed workers has grown from 6.4% to 8.6% in the Netherlands. Another important social trend is the increase in labour participation of women. Over the past few decades, the labour force participation of women has grown strongly in most countries. In the OECD countries, labour force participation of women has increased from 52.7% in 1980 to 61.8% in 2010 on average. In the Netherlands, it has increased from 36.1% to 72.6% during this period.[5] This increase is spectacular in the Netherlands although it must mainly be attributed to an increase in part-time employment.

All these social developments have contributed to a more individualized society in which people stand more as individuals rather than as group. This process of individualization has changed the view on social security and pension schemes (Natali, 2008). As far as social security schemes are concerned, traditional programmes were mainly anchored to the idea of *redistributive* justice, in line with the aim of providing workers with the same level of income before and after retirement. In these schemes old-age risks were shared and individuals were protected against many unpredictable factors (like poverty, ageing or investment risks). By contrast, the wave of pension reforms which started at the end of previous century points more to the idea of so-called *pro-*

[5]Figures are obtained from the database OECD.Stat (stats.oecd.org).

cedural justice. The aim is to link the provision of benefits to the payment of contributions: the more people contribute to the system, the more they receive from the old-age provision. The growing popularity of DC elements in unfunded and funded schemes confirms such a broad tendency. The necessary consequence of these developments is that risk-pooling and redistributive characteristics of pension schemes become less important.

As response to all these economic, demographic and social developments, in many countries pension systems have been reformed or are planning to be restructured in the coming years. Various countries have undertaken far-reaching reforms that have changed the structure of their pension systems, other countries adopted a series of smaller reforms which, taken together, often also have had a substantial impact on future pension entitlements. The OECD (2007) has identified five main categories of pension reform that have been undertaken since the 1990s:

1. introducing defined-contribution schemes;
2. increasing the pension eligibility age;
3. adjustment of work and retirement incentives;
4. changing the number of years in the benefit formula;
5. introducing an adjustment mechanism in the benefit formula for higher life expectancy.

The main motivation for all these reforms has of course been a fiscal one: improving the sustainability, affordability and shock resistance of pension schemes. But improving work incentives (and thus economic efficiency) has also been an important reform goal, especially for the third category of reforms. In this thesis, we will concentrate on the first three types of pension reforms, i.e., the switch to more risk-bearing pension benefits, increasing the pension eligibility age and introducing flexibility in the pension entitlement age. These three types of reform are currently also on the policy agenda in the Netherlands and underlie our main research questions.

One might wonder why these pension reforms are on the policy agenda today, given the fact that many underlying causes are already known for a long period of time. Indeed, we already know for many decades that the population structure will change in the future towards more elderly and less younger people. The same holds for all kinds of social trends, like the shift from the social group as focus point to the individual. Of course, not all recent developments (could) have been foreseen. In the Netherlands, the life expectancy has increased much more in the last decade than forecasted by demographic models. For five years ago, most of the economic models used put too little weight on extreme shocks like the recent financial crisis. Although it is interesting, the question why pension reforms are currently occurring is outside the scope of this thesis. Our starting point is the observation that pension schemes are under revision in many countries and analyse those implications on redistribution, risk sharing and aspects of economic behaviour (in particular retirement).

1.1.2 Towards risk-bearing pension benefits

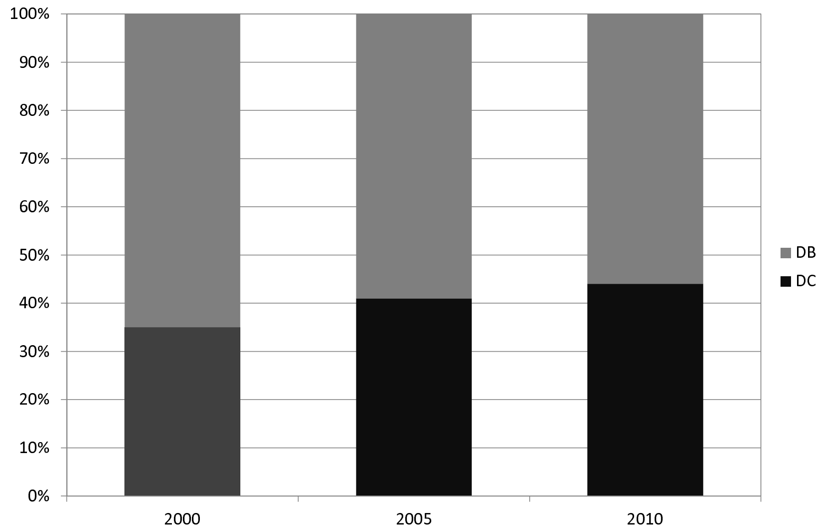
The international trend to more risk-bearing pension benefits is one of the most prominent pension reforms of the last decades. The transfer of macroeconomic and demographic risks from workers to retirees has been implemented in funded schemes as well as unfunded schemes. In Italy, Poland and Sweden, for example, so-called notional DC schemes have been introduced in the pay-as-you-go (PAYG) pillar with adjustments of benefits in line with increases in life expectancy, thus making the system endogenous to demographic risk. In this thesis, however, we will only focus on the trend from DB to DC pension contracts within *funded* pensions.

Figure 1.6 shows the development of the share of DB and DC contracts in total pension assets in the seven countries with the largest pension markets.[6] In the first decade of this century, DC assets have grown at a rate of 7.5% per year while DB assets have grown at a much slower pace of 2.9% per year. Currently, DC assets represent 44% of total pension assets compared to 40% in 2005 and 35% in 2000. The markets with a larger proportion of DC assets are Australia, Switzerland and the US, while Japan essentially remains 100% DB. The Netherlands, also historically DB minded, is now showing sign of a shift to DC contracts: the share of DC assets has been grown from 1% in 2005 to 6% in 2010.

In absolute terms, the DC share is still very low in the Netherlands. In practice, however, the distinction between defined benefit and defined contribution is less strict than the above numbers suggest. Almost all Dutch occu-

[6] These countries are: Australia, Canada, Japan, Netherlands, Switzerland, UK, US.

Figure 1.6: DB/DC asset split

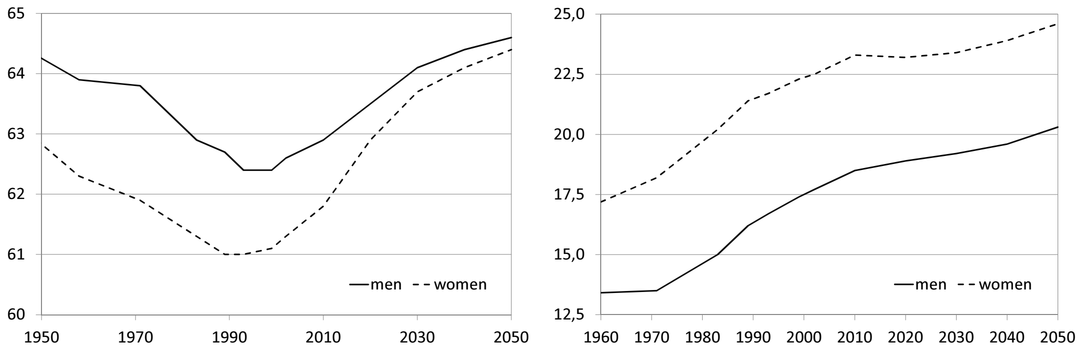


Source: Global Pension Asset Study 2011, Towers Watson. The following countries are included: Australia, Canada, Japan, Netherlands, Switzerland, UK and US.

pational funded pensions include indexation of entitlements which is conditional on the funding ratio of the fund. So-called 'indexation ladders' specify how the degree of indexation depends on the funding ratio. Such indexation ladder typically takes a simple non-linear form: the degree of indexation is zero below some minimum funding ratio and 100% of wage or price inflation above some maximum funding ratio and increases linearly between this minimum and maximum level. Hence, the occupational pension schemes are essentially hybrid contracts that incorporate DC elements in traditional DB contracts which link the benefit level to past labour earnings.

The move from DB to DC contracts has different causes. First of all, it is a response to population ageing which increases the weight of older generations compared to younger generations. This implies that the pension liabilities have expanded compared to the contribution base. Accordingly, in a funded DB scheme the risk-bearing capital of young generations has to be leveraged with an increasing number of elderly. In other words, unanticipated financial or longevity shocks require larger adjustments in pension contributions in order to protect the benefits of the elderly against these shocks. The insurance provided to the retirees has thus become more expensive in the sense that the working generations are exposed to more uncertainty about their contributions. This poses a serious threat to the shock resistance of the traditional

Figure 1.7: Average pensionable age and retirement period in OECD countries



(a) Pensionable age (b) Retirement period

Source: OECD (2011).

DB pension schemes. Besides ageing, the move to DC contracts is further accelerated by the financial crisis in combination with an increased mobility of workers. Indeed, any commitment of young working generations to close a funding gap becomes more difficult to enforce if funding ratios are low and, at the same time, workers are more mobile. Finally, also the diminishing willingness of corporations to guarantee the obligations of the associated pension fund plays a role. New accounting standards (IFRS) force corporations to put risks related to pension promises explicitly on their balance sheet. To avoid these risks, many corporations change their pension plans with the aim to shift more risks to the members. In face of population aging which reduces the risk-bearing capacity of the younger workers, retirees then necessarily end up as residual risk bearers in these funds.

1.1.3 Increase in retirement age

In response to the growing burden of ageing, many countries have increased (like Germany, Japan, Poland and the Slovak Republic) or will increase (like the US, UK, Australia and France) the pension entitlement age in an attempt to put a ceiling on pension contributions. Other countries (like Denmark) have introduced automatic mechanisms to adjust the entitlement age to higher life expectancy. Figure 1.7a shows the evolution of the average pensionable age in OECD countries covering a period of a century, looking back to 1950 and forwards to 2050. The pensionable age is defined as the age at which people can first draw *full* benefits, that is, without actuarial reduction for early retirement. The average pensionable age dropped by nearly two years during the second half of the 20th century to 62.4 for men and 61.1 for women. From the beginning of the 1990s and thereafter, governments started taking action to reverse the trend and put in place legislation that has increased pensionable ages. Already by 2010, the average pension ages have increased by 0.5 years for men and by 0.8 years for women from the low point. Looking forward, current plans will further increase the average pensionable age and reduce the gap between the sexes. Legislation already in place will increase the pension age almost to 65 for both sexes in 2050.

Despite these increases, it is striking that the average pension age for men will only reach the same level as in 1950 by 2040. For women, the pensionable age will reach the level it was in 1950 from 2020 onwards. Life expectancy is projected to grow faster than these increases in the pension age. Figure 1.7b shows the expected period of retirement based on the evolution of the pension age as shown in Figure 1.7a and information on developments in life expectancy. The expected period of retirement is an important determinant of the cost for paying pensions. In this respect, it is striking that the increases in the pension age since the 1990s are not sufficient to reduce the expected retirement period. On the contrary, the expected duration of retirement is projected to increase by about 1.8 years for men and 1.3 years for women between 2010 and 2050.

1.1.4 Flexible pensions

Since 1970, the effective retirement age of labour market exit has declined substantially in almost all Western countries. To reverse this development, in recent years more attention has been given to pension reforms that improve labour-supply incentives and encourage people to work longer. Most pension systems now allow for a flexible choice of the retirement age with more or less actuarially-neutral adjustments. This kind of reforms has the double advantage that it increases the sustainability of pension systems but also reduces the distortions to the labour market caused by incentives to retire early. Another important driver of flexible pensions is the ongoing process of individualization and the resulting acknowledgement that individuals differ in their tastes for leisure, earnings capacities, wealth positions, and therefore have different preferences for retirement. From 2005 onwards, people indeed stay longer in the labour market although effective retirement ages are currently still below

Table 1.1: Examples of reforms affecting timing of retirement

Country	Reform	When
UK	Higher pension for retiring between 65-70: raised from 7.5% to 10.4% per year, with lump-sum option added.	2005
Finland	Flexible retirement age from age 62-68: 7.2% bonus for delaying retirement to age 63 and 4.5% therafter to age 68.	2005
Germany	Pension 3.6% lower if retire at age 63-64; 6% higher for each year after age 65.	1997-04
France	Bonus of 3% for each year the pension is postponed beyond age 60 (for those already at the full rate).	2004
Netherlands	Abolishment of favourable tax treatment of early retirement schemes.	2006
Denmark	Pension reduction of about 10% for retiring at 60-62.	1999
	Lump-sum bonus for working between 62 and 65.	1999
	Higher pensions for deferring after age 65 (e.g. +7% if defer to age 66).	2004
Sweden	Flexible retirement from age 61 with actuarially-based adjustments.	1999

Source: OECD (2006).

the normal statutory pension age in the majority of OECD countries (OECD, 2011).

In many countries, the way pensions were calculated in pre-reformed systems discouraged older workers from staying in the labour market: benefits were not reduced much (or not all) when retirement was taken early, and working beyond the official retirement age did not result in higher pension entitlements (Gruber and Wise, 1999). To avoid such distortions, pension systems with high implicit taxes for continuing in work were changed into more actuarially-neutral schemes with flexible retirement ages. This resulted in penalties for early retirement or increases in the number of years of contributions required to receive a full pension. Similarly, other pension schemes have introduced (or increased the increments for) bonuses paid to people retiring after the normal pension age. These measures aim to reduce early pension benefits by an amount that corresponds both to the lower amount of contributions paid by the worker and to the increase in the period over which the worker will receive pension benefits.

In recent years, penalties for earlier retirement and rewards for later retirement were increased in a number of countries although these adjustments are still not actuarially neutral (Queisser and Whitehouse, 2006). Generally, the penalty rate is not as high as the reward rate. Moreover, countries differ at lot in the incentives provided for working an additional year. Table 1.1 gives some illustrations of recent flexibility measures undertaken by countries in the calculation of their earnings-related pension benefits. The UK, for example, has sharply increased the reward for staying in the labour market between age 65 and 70. At the same time, it has given people the option of taking the reward as lump sum rather than a higher annuity which could make it even more attractive. But also other countries like Finland and France have increased incentives for people to work after the normal retirement age. In the Netherlands, as already mentioned at the start of this chapter, the government abolished the favourable tax treatment of early retirement schemes in 2006.

1.2 Main concepts

This thesis deals with the question how the pension reforms as just discussed will affect two important characteristics of collective pension schemes, redistribution and risk sharing. This section introduces these two concepts: it gives an overview of the main literature and relates this to what will be done in this thesis.

1.2.1 Redistribution

In practice, almost all pension systems have distributional consequences, both within generations (so-called intragenerational redistribution) and across generations (so-called intergenerational redistribution), some intentional and others unintentional. In this thesis we will discuss both types of redistribution but the focus will be on intragenerational redistribution.

Once a pension system includes a PAYG element, it necessarily redistributes wealth across generations. This PAYG part of the scheme implies that the first generations, living at the time of the introduction of the scheme, receive a public pension provision without having made enough contributions to the system in the past. Since any pension contract is a zero-sum game in terms of cash flows, this windfall gain of the initial generations has to be born by other generations. Indeed, it will be passed on to all subsequent generations in the form of higher pension contributions. In other words, a pension scheme with PAYG financing redistributes wealth from all subsequent generations to the first generations (Sinn, 2000). For a given benefit level, these future generations are forced to contribute more to the scheme than when they could save for their pensions at the capital market themselves. In Chapter 2 of this thesis it will be shown that this type of intergenerational redistribution is not confined to pure PAYG systems but also operates (although at a smaller scale) in funded schemes with uniform contribution and accrual rules.

Collective pension systems also have intragenerational redistribution effects. The most traditional channel is the rich-to-poor redistribution. This type of redistribution is often considered as one of the core objectives of public pensions (World Bank, 1994) and has therefore attracted a lot of attention in the literature. This literature mainly focuses on the distinction between a redistributive Beveridgean system offering flat benefits and Bismarckian pensions which do not include intragenerational redistribution by offering earnings-related benefits. Some of this literature asks whether there is a negative relation between the amount of intragenerational redistribution and the size of the PAYG scheme (Casamatta et al., 2000; Koethenburger et al., 2008). Other studies try to explain why real-world pension schemes usually contain Beveridgean and Bismarckian characteristics (Cremer and Pestieau, 2003b; Conde-Ruiz and Profeta, 2007 or Kolmar, 2007).

In addition to the classical redistribution from rich to poor, public pension systems contain common design features that may have other (unintentional) distributive consequences. Traditionally, public DB pensions offer a 'collective' annuity to the retired. The defining character of such an annuity is that it does not depend on individual mortality rates. As a result, individuals with a high life expectancy will receive benefits for a longer period than individuals with a low life expectancy. This makes the pension scheme regressive because it is well known that average life expectancy tends to increase with income.[7] The existing literature indicates that these intragenerational transfers are large in pension systems and create substantial deadweight losses (see e.g., Börsch-Supan and Reil-Held, 2001 or Ter Rele, 2007). Almost all of these studies focus on first-pillar PAYG schemes. However, intragenerational transfers might also play a role in funded pension schemes with collective annuities. In Chapter 2 we will focus on funded pensions and quantify the level of intragenerational

[7] See Adams et al. (2003) for an extensive listing of studies dealing with the association of socioeconomic status and longevity.

redistribution in the Dutch occupational pension schemes.

Given that intragenerational transfers are empirically relevant, it is surprising that the literature has paid little attention to the economic implications of these transfers. There are a few exceptions. Cremer et al. (2010) analyse the desirability of collective annuities in Bismarckian and Beveridgean pension systems assuming that income and longevity are positively correlated. They show (with a utilitarian social welfare function) that collective annuitization implies an undesirable redistribution from low-income earners to high-income earners and a desirable redistribution from short- to long-lived individuals. Therefore, collective annuitization is desirable only when wage differentials are sufficiently small and/or longevity differentials are sufficiently important. Borck (2007) shows that a correlation between income and life expectancy has important implications for voting outcomes about public pensions. In Chapter 5 we also take the position that individual life spans and income are positively correlated. We explore the impact of recent reform towards more flexible pensions and a higher pension entitlement age on intragenerational redistribution and welfare.

1.2.2 Risk sharing

Besides realizing intentional (or unintentional) redistribution, pension funds also engage in risk-sharing arrangements. Risk sharing can increase welfare for the same reason that insurance increases the welfare of insured people: by spreading shocks over a larger group, the individual exposure to shocks can be lower than when everyone would have to form their own provision. Bovenberg and Van Ewijk (2012) distinguish three functions of pension systems related to risk sharing: *i)* facilitating life-cycle financial planning; *ii*) insuring idiosyncratic risks and *iii)* sharing macroeconomic risks across generations. The first function concerns consumption smoothing over the life cycle, taking into account individual circumstances and preferences. The second function concerns pooling of intragenerational risks in the face of imperfect insurance markets. The third function concerns intergenerational risk sharing of macroeconomic shocks in the face of incomplete markets. In this thesis, we only focus on the first and third function and we will not deal with intragenerational risk sharing.[8]

The fundamental question is what kind of market incompleteness a pension

[8] For studies about intragenerational risk sharing and pension systems, see e.g., Nishiyama and Smetters (2007) or Fehr and Habermann (2008).

fund is able to solve by engaging in risk sharing. Actually, two market failures stand out. First, and most importantly in the context of this thesis, current generations are unable to trade risks with generations that are not born yet. This point was first made by Diamond (1977) and Gordon and Varian (1988). A pension fund can solve this market inefficiency by defining pension benefits independently from *ex post* returns on the underlying financial assets (which is the case with DB pension schemes) and imposing the mismatch risk between assets and liabilities upon the workers. In this way, the young generations share in the risks of the financial markets in which they cannot trade in a *laissez-faire* economy.

A second point where pension funds can complement a *laissez-faire* economy concerns capital markets which theoretically could exist but for whatever reason do not or hardly develop in practice. An example of such an incomplete market is the market for wage-indexed bonds (Shiller, 1999). A pension fund can (partly) solve this market incompleteness by linking benefits to wages. In this way, the retirees acquire an implicit claim on the human capital of the workers which is not traded on financial markets. Other examples of undeveloped markets where pension funds can help are markets for longevity bonds or price-indexed bonds.

If designed properly, risk-sharing contracts lead to welfare improvements for all generations from an *ex ante* perspective, i.e., before shocks materialize that determine the size and direction of the transfers between generations. But as always with insurance contracts, some generations may be worse off from an *ex post* perspective, i.e., after shocks materialize. A pension fund with voluntary participation is typically unable to commit future generations to risk sharing because new potential entrants will not join if this is not in their interest from the *ex post* perspective. Risk-sharing pension contract are at odds with free entrance. Hence, a feasible risk-sharing contract requires mandatory participation and can only be implemented by institutions which have sufficient power to enforce intergenerational commitments, like national governments or large-scale pension funds. The question which institution is best be able to employ risk sharing is interesting in its own but beyond the scope of this thesis. We simply take the perspective that our pension institution has a unique power to enforce risk sharing.

The existing literature that deals with intergenerational risk sharing within funded pensions report large welfare gains. Estimates range from a 2.3% increase in certainty equivalent consumption (Cui et al., 2011) to 19.0% (Gollier, 2008). However, these two studies ignore two important aspects related to the labour market which could overstate their gains from risk sharing. First, labour-supply distortions are not taken into account. In practice, pension funds finances funding deficits (*surpluses*) with income-related taxes (*subsidies*) which may distort the labour-supply decision. Second, these studies ignore wage risks. It is well known that financial returns and returns to human capital are at least in the long run positively correlated (see e.g., Benzoni et al., 2007). This comovement decreases the preference of young workers to absorb the financial risks of the old because they are already overexposed to correlated wage risks. In Chapter 3 we explore the merits of intergenerational risk sharing in a funded DB scheme where labour-supply distortions and a positive co-movement between financial and human capital returns are explicitly taken into account.

Besides providing intergenerational risk sharing, pension institutions can also facilitate life-cycle planning which concerns the smoothing of consumption over the individual life cycle. The pension fund can do this is in a direct way by providing an optimal investment policy of the collected contributions in accordance with the individual preference for risk taking over the life cycle (Teulings and De Vries, 2006 or Bovenberg et al., 2007). It can also indirectly contribute to life-cycle planning by e.g., facilitating flexible pensions with a variable starting date. Flexible pensions may contribute to a better employment of the retirement decision as instrument to hedge unexpected shocks during the life cycle (Bodie et al., 1992).

Basically, an individual has three options on how to react to macroeconomic shocks. The individual may adjust consumption over the remaining life time, he can reallocate his asset portfolio or adjust the amount of labour supplied. In practice, individuals will most likely apply a combination of these three options. Recent evidence finds that labour-supply adjustments are mainly participation decisions which are most sensitive to financial incentives concentrating at the end of the career. In addition, it turns out that the design of pension schemes has a large effect on the retirement decision of older workers (French and Jones, 2012). This suggests that the effectiveness of labour-supply adjustments to absorb shocks may be facilitated by introducing flexible pension schemes.

If capital markets would be well functioning, there seems little reason for flexible pension schemes, however. In that case, the age of pension take-up would not matter, as it would be possible to either annuitize revenues, in case of postponing retirement, or borrow against future pension income, in case of advancing retirement. In practice, however, capital markets are far from perfect due to borrowing constraints, moral hazard issues or institutional restrictions. Retirement flexibility and pension schemes with flexible starting dates are therefore closely linked to each other (Van Vuuren, 2011).

An important implication of flexible retirement opportunities is that it allows for more risk taking in pension assets. This point was first made by Bodie et al. (1992) and further worked out by Choi and Shim (2006) and Farhi and Panageas (2007) among others. The basic mechanism behind this result is the negative correlation between financial returns and labour income due to wealth effects in the retirement decision. Indeed, a negative wealth shock causes the marginal utility from leisure to decrease and hence agents increase labour supply which, in turn, raises labour income. However, these studies only focus on capital market shocks and ignore that retirement decisions are also affected by other important shocks like, for example, shocks in wage income. The distinction between different risk factors is important because they may constitute a rather different effect on income and substitution effects in labour supply. Indeed, as will be shown in Chapter 4, the relative strength of income and substitution effects determines whether retirement flexibility indeed serves as a hedge against poor asset returns.

1.2.3 Methodology

Our main analytical tool is the two-period overlapping-generations (OLG) model, first developed by Samuelson (1958) and later extended by Diamond (1965). In this model individuals live for two periods, as a young worker in the first period and as an old retiree in the second period of life. This structure implies that at any given time period both a young generation and an old generation are alive. Due to this simple demographic structure, the model is a very powerful tool for analysing redistribution and risk sharing between young and old generations. Of course, the model has its limitations when it comes down to providing realistic quantitative results and drawing clear policy implications. Our main interest, however, is to shed light on the most important economic mechanisms going on in the issues we address, preferably by obtaining clear analytical insights. We will support these analytical results with numerical illustrations to get an idea about the order of magnitude of the effects at hand.

In the different chapters, we make refinements to the standard model, depending on the relevance for the topic at hand. In particular, when we analyse risk sharing issues we introduce uncertainty in the model. We follow Matsen and Thøgersen (2004), Beetsma and Bovenberg (2009) and Bohn (2009) among others and introduce stochastic productivity and depreciation in the two-period OLG model to allow for imperfect correlation between labour and capital income. In contrast to Bohn (2009), who studies risk-sharing implications of alternative fiscal policies, and Matsen and Thøgersen (2004), who explore optimal division between PAYG and DC funding from a risk-sharing perspective, we mainly focus on funded pensions of the DB type. Beetsma and Bovenberg (2009) also include various funded DB pensions but their model only lasts two periods and does not include endogenous labour supply. We, in contrast, employ an infinite horizon and allow for endogenous labour supply, either in the first or second period of life.

To model retirement in a two-period OLG model, we follow Cremer and Pestieau (2003a) and Casamatta et al. (2005) and assume that individuals can decide which fraction of the second period they spend on working and enjoying retirement. In this way, we basically merge the labour-supply decision on the extensive and intensive margin. In other words, the labour-supply decision can either be interpreted as a retirement decision (extensive margin) or as a decision about the amount of hours worked (intensive margin). In this thesis, we prefer the first interpretation and consider the labour decision in the second period as a retirement decision.

All models used in this thesis are based on the rational agent principle. That means, agents have clear preferences, perfect foresight (or rational expectations in case of uncertainty) and always choose to perform the action with the optimal expected outcome for itself from among all feasible actions. We know from the literature that not all individuals are completely rational. As argued by e.g., Feldstein (1985), a principal rationale for mandatory pension institutions is that individuals often lack the foresight to save enough for their retirement years. However, the *raison d'être* of the pension schemes as considered in this thesis has more to do with redistribution and risk-sharing issues rather than myopic behaviour.

1.3 Outline of the thesis

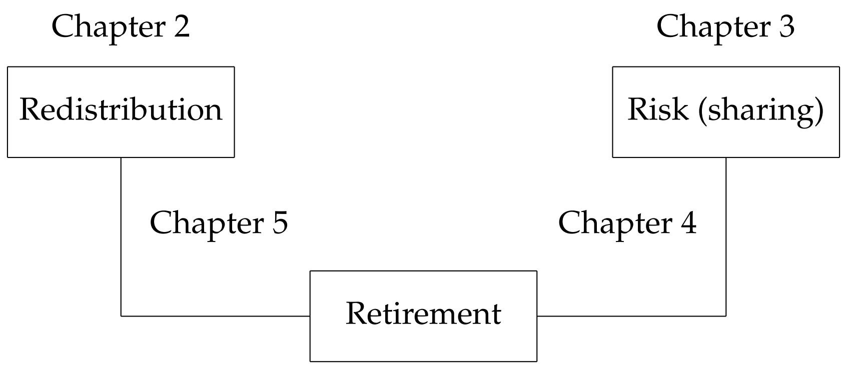
This thesis consists of four conceptual chapters which can all be read independently of each other. All chapters have the same structure: we start with a short abstract, followed by an introductory section containing the explanation and motivation of the research topic. It also discusses the contribution of the analysis to the previous literature and sketches its main policy implications. In subsequent sections, we will present the body of the analysis. All chapters end with a concluding section which summarizes the main findings and explores opportunities for further research. Most analytical derivations are moved to appendices attached to each chapter.

Figure 1.8 gives a schematic overview of the relation between the topics raised in this thesis. The figure should be read clockwise beginning at the top left: in Chapter 2 we analyse the direction and magnitude of the intergenerational and intragenerational redistribution effects in Dutch occupational pension schemes. Chapter 3 considers the value of intergenerational risk sharing in a funded DB scheme assuming that pension contributions are distortionary and wages and asset returns are both risky. In Chapter 4 we explore the interaction between retirement flexibility and portfolio choice and investigate under which conditions this flexibility serves as a hedge against unforeseen events. We then consider the relation between retirement flexibility and intragenerational redistribution in Chapter 5 in the light of recent pension reforms, like the increase in the official retirement age and the introduction of flexible retirement ages.

Chapter 6 summarizes the main findings, discusses the key policy implications and sketches directions for further research. In the remaining of this section we will provide a more detailed outline of the conceptual chapters.

Chapter 2 analyses the direction and magnitude of intergenerational and intragenerational redistribution in Dutch occupational pension schemes. We start this chapter with a description of these two sources of redistribution. We use the level of educational attainment, gender and age to classify the pension fund population. Hence, intragenerational redistribution takes place between males and females on the one hand and between individuals with different levels of education on the other hand. We measure redistribution on a life-time basis, i.e., as the difference between the expected present values of life-time pension contributions and life-time benefits. We proceed with a formal description of the occupational pension scheme used in the calculations. This scheme is characterized by uniform contribution and accrual rules, funding and forced annuitization, all characteristics which are typical for Dutch pension funds. After discussing the calibration, we present the redistribution effects for a baseline scenario with a constant demography. We will perform a

Figure 1.8: Schematic overview of the chapters



sensitivity analysis on most important parameter values used in the baseline. Finally, we present two alternative scenarios with a time-varying demography: the first scenario allows for a general increase in life expectancy; in the second one we allow for a further increase in the labour force participation of women in the coming decades.

In **Chapter 3** we move to intergenerational risk sharing and analyse those benefits in case of funded DB pensions. Like in real-world schemes, we allow for pension contributions proportional to labour income enabling us to distinguish the pure risk-sharing gains from the welfare losses related to labour-supply distortions. We will present in detail the model used which consists of two overlapping generations and represents a small open economy with risky asset returns and risky wages. Based on Cobb Douglas per-period utility, we will analytically derive the optimal behaviour of individuals (in terms of consumption, leisure and portfolio allocation) and the optimal investment policy of the pension fund. By solving the model for distortionary transfers as well as lump-sum transfers, we are able to derive a separate expression for the size of the welfare gain of risk sharing and the size of the welfare loss of labour-supply distortions. These analytical results will be accompanied with numerical simulations based on more general preferences. We will also perform an extensive robustness check of the default results by solving the model for alternative parameter values. Finally, we consider five important model extensions, namely the introduction of a short-sale constraint for consumers, the inclusion of labour income taxation, discounting benefits with the portfolio return (rather than the risk-free rate), putting a cap on the mismatch contribution and, finally, extending the risk-sharing horizon to an infinite amount of periods (rather than two periods).

In the last two chapters of the thesis we will focus on the role of retirement flexibility. We define retirement flexibility as a setting in which the individual optimally chooses his own personal retirement age according to his own preferences and circumstances. In Chapter 4 we investigate the relation between retirement flexibility and individual risk sharing (i.e., consumption smoothing); in Chapter 5 we turn to the relation between retirement flexibility and intragenerational redistribution.

In **Chapter 4** we explore the interaction between retirement flexibility and portfolio choice. The idea of this chapter is to investigate the conditions under which retirement flexibility serves as an efficient hedge against unforeseen shocks. We will present the stochastic model used which, like in the previous chapter, consists of two overlapping generations. Individuals choose upon consumption, the share of firm equity and government bonds in their portfolio and the retirement age. To isolate the effects of retirement flexibility, we compare two different retirement settings: one in which agents choose their retirement age *before* shocks are known (fixed retirement) and one in which the retirement age is chosen *after* shocks are known (flexible retirement). The model contains shocks to equity holdings and wages which originate from stochastic depreciation and total factor productivity in the production function. We solve the model quasi-analytically using a log-linearization around the stochastic steady state. We will show how the interaction between retirement flexibility and portfolio choice differs in partial equilibrium and general equilibrium and how it depends on the source of uncertainty (i.e., depreciation risk versus productivity risk). Finally, we consider how this interaction is affected by important model parameters, like the substitution elasticity between consumption and leisure.

In the final conceptual chapter of this thesis, **Chapter 5**, we explore the impact of some widely-used pension reforms on intragenerational redistribution and welfare. We consider reforms that introduce automatic links between pensions and higher longevity as well as reforms aimed at increasing the flexibility of individual pension take-up. For that purpose, we will present a deterministic two-period overlapping-generations model with a Beveridgean PAYG pension scheme and heterogeneous agents who differ in ability and life expectancy. We first show how intragenerational redistribution effects across high-skilled and low-skilled agents are affected by an increase in longevity. It will be shown that the direction of these effects depends on the adjustment mechanism the government uses to absorb this shock (i.e., lowering benefits versus increasing the entitlement age) and on the individual heterogeneity in life expectancies. Then, we study the redistribution and welfare effects of introducing pension schemes that allow for actuarial adjustment of benefits when retirement is postponed or advanced. We consider three different cases to determine the actuarial adjustment factor. We first assume that the government applies uniform actuarial adjustment based on the average life span of the whole population. Then we assume that the government uses individual-specific adjustment factors based on individual life spans. In the last case, the adjustment factor is made conditional on skill level. For each of these scenarios we explore the welfare effects and, in particular, whether it can result in a Pareto improvement.

CHAPTER 2

MEASURING REDISTRIBUTION IN DUTCH OCCUPATIONAL PENSIONS[1]

This chapter explores how the Dutch system of occupational pensions redistributes across and within generations. The approach in this chapter deviates from the usual approach by incorporating the full life cycle in the measurements, rather than only the annual effects. In order to quantify redistribution, we use the level of educational attainment, gender and age to classify the pension fund population. For all groups distinguished, we measure in present value terms the average net benefit from participating in occupational pensions. The results indicate a sizable redistribution from males to females and from low-educated to higher-educated workers. On a life-time basis, the impact of intergenerational transfers is modest.

2.1 Introduction

Redistribution is an important objective of unfunded first-pillar pensions. According to the well-known proposal of the World Bank, first-pillar pensions should exactly perform this task, while the saving function should be achieved by the second pillar (see World Bank, 1994). Against this background, it is not surprising that most of the existing literature focusing on redistribution

[1]This chapter is an extended version of Bonenkamp (2009).

restricts to first-pillar pensions (see, e.g., Cubeddu, 2000; Börsch-Supan and Reil-Held, 2001; Sommacal, 2006 and Ter Rele, 2007). In practice, however, redistribution may also play a role in second-pillar pensions. In this chapter, we analyse redistribution in the second pillar of the Dutch pension system.

Occupational earnings-related pension schemes in the second pillar are often funded via a uniform contribution and accrual rate, determined as a fraction of the wage earned. In the Netherlands, charging a uniform contribution rate is legally obliged. The uniform contribution and accrual system (henceforth denoted as uniform pricing) drives a wedge between the market price of the annuity contract and the actual contributions charged. The market value depends on individual characteristics, like age and gender, which, by definition, does not hold true for uniform pricing. Differences between the market price of a pension scheme and the costs imply redistribution between groups.

The occupational pension schemes considered in this chapter are supplementary to the unfunded first-pillar pension provision and are characterized by funding, collectivity, mandatory participation, forced annuitization and uniform pricing. The pension schemes have defined benefits (DB) related to the average wage earned. In a DB scheme, the pension rights depend on the labour history of the participant. Pension schemes with flat contribution rates, as studied in this chapter, are common in the Netherlands, but are also important in other countries, like the UK, the US and Canada.

In a society that becomes increasingly more individualistic, solidarity can be under tension if there are no good reasons to justify it, which especially happens when the transfers become too large or tend to flow in one direction. In the Netherlands, there is currently a public debate on the desirability of uniform pricing (see e.g., Bovenberg et al., 2006; Boeijen et al., 2007). The discussion mainly focuses on the systematic transfers from young to old participants. In the absence of uniform pricing, young workers would pay a lower contribution rate than old workers, because the period over which contributions yield returns decreases as people get older. In the years after World War II, the system of uniform pricing ensured that the old people, carried on the shoulders of the young people, could still accrue a reasonable pension income in a short time period. This redistribution from young to old would not be problematic as long as participants enter and leave the pension scheme at the same age, because then the transfers will mostly smooth out over the life time. However, in the current climate of increasing labour mobility, this smoothing will not necessarily occur.

So far in the discussion, a clear picture about the magnitude and direction of the transfers is missing. Moreover, little attention is given to intragenerational transfers which may also be important. The aim of this chapter is to fill these gaps and to measure redistribution in Dutch occupational pensions for various socioeconomic groups. We measure redistribution as the difference between the expected present values of pension contributions and benefits, which is a standard way of measuring life-time redistribution in the literature (Börsch-Supan and Reil-Held, 2001). A natural alternative is to focus on annual redistribution as indicator of redistribution. However, this might be misleading because an individual's net benefit from a funded pension scheme may vary with age.

To quantify redistribution, we distinguish between inter- and intragenerational life-time redistribution. We use the level of educational attainment, gender and age to classify the pension fund population. Hence, intragenerational redistribution takes place between males and females on the one hand (denoted cross-gender redistribution) and between individuals with different levels of education on the other hand (denoted intereducational redistribution). Life-time intergenerational redistribution relates to an implicit tax imposed on future generations to service the gains given away to the generations living at the time uniform pricing was introduced.

Throughout the analysis, we focus solely on *ex ante* (or non-reciprocal) redistribution that is independent of a shock occurring and *ex ante* leads to transfers between groups of participants. Occupational pension schemes often also contain *ex post* (or reciprocal) redistribution that relates to risk sharing and occurs after the pension fund experiences a financial loss or gain. This form of redistribution does not lead to structural transfers from one group to another and will be considered in Chapter 3 of this thesis.[2] Apart from this, the scope of this study is limited to the redistribution within the occupational pension schemes. Redistribution among different socioeconomic groups also plays a role in other parts of the Dutch economy, like the first pension pillar, the health care system and the tax system.[3]

This article complements the study of Börsch-Supan and Reil-Held (2001) and Ter Rele (2007) who measure life-time redistribution in the first pillar

[2]For other examples of studies that focus on sharing reciprocal risks, see Gollier (2008), Hoevenaars and Ponds (2008) or Cui et al. (2011).

[3]For the Netherlands, Bonenkamp and Ter Rele (2013) compare the intragenerational redistribution effects in the first and second pension pillar. They conclude that the redistribution from high-educated to low-educated people in the first pillar is much larger than the reverse redistribution in the second pillar.

of the German and the Dutch pension system, respectively. Regarding redistribution in occupational pensions, related literature includes Aarssen and Kuipers (2007) and Boeijen et al. (2007). Compared to these studies, the novelty of our approach is that we incorporate the full life cycle in the measurements, rather than only the annual effects. Moreover, our approach includes intergenerational and intragenerational redistribution, not only the first. Hári et al. (2006) also analyses the attractiveness of participation in collective pension schemes for different socioeconomic groups, but in contrast to our work, their analysis restricts to a comparison of the benefits of participation. To measure redistribution properly, both benefits and costs should be included.

We find that the Dutch occupational pension schemes contain sizable transfers from males to females and from low-educated to high-educated employees. On a life-time basis, the impact of the intergenerational transfer seems to be modest, although its size is very sensitive to the market interest rate. If we account for the expected convergence of life expectancies between males and females in the coming decades, the cross-gender redistribution will reduce. From a social point of view, the intragenerational transfers, especially those from low-educated to high-educated people, might be considered as unintended solidarity. In this respect, our results are remarkable, because for the legislator uniform pricing has always been considered a necessary element of solidarity to justify the lack of competition caused by the mandatory participation for companies and individuals (Lutjens, 2007).

The rest of this chapter is organized as follows. In Section 2.2 we introduce the different sources of redistribution in a funded pension scheme with uniform pricing. Section 2.3 describes the representative pension fund and shows how to calculate the different sources of redistribution defined in Section 2.2. In Section 2.4 we present the baseline scenario in which the economic and demographic exogenous variables are held constant, whereas Section 2.5 discusses two alternative scenarios in which we allow for two future demographic developments, namely increasing life expectancy and increasing labour force participation of females. Finally, Section 2.6 concludes.

2.2 Defining life-time redistribution

The difference between the contribution that a participant actually pays each year (i.e., the uniform contribution) and the contribution that should be paid according to the actuarial value of the accrued pension entitlements (i.e., the actuarially-fair contribution), defines the yearly redistribution between a participant and other members of the pension fund. Since the actuarial cost price of a given pension benefit increases with age, uniform pricing implies a redistribution from young to old workers. However, since a worker is a net contributor when he is young and a net receiver when he is old, uniform pricing involves to some extent a redistribution of pension contributions over an individual's career. We correct for these intrapersonal payments by defining redistribution on a life-time basis.

Deviations from absolute equivalence is the standard way of defining transfers in the literature.[4] According to this concept, a pension scheme is viewed as a fair insurance (no transfers) if, for every individual, the expected present discounted value of benefits is equal to the present discounted value of contributions. Any difference between the two present discounted values is defined as a transfer to or from an individual. In this chapter, we apply this definition to Dutch occupational pensions to disentangle intragenerational from intergenerational transfers.

In a funded pension scheme with uniform pricing there are two reasons why *ex ante* the present value of life-time pension contributions can differ from that of life-time pension benefits. The first reason is that the pension contributions of the current and future participants partly entail a redistribution to former generations for which they do not get any compensation (Section 2.2.1). Second, since participants of a pension fund generally differ in terms of life expectancy, income perspectives and labour force participation, uniform pricing also redistributes between individuals of the same generation (Section 2.2.2).

2.2.1 Intergenerational redistribution

As already noted, in the first part of one's working life workers subsidize older workers, during the second part they receive a subsidy. Viewed in this way, a pension scheme with uniform pricing has some characteristics of a pay-as-you-go (PAYG) system. This notion allows us to draw a parallel between a PAYG scheme and a funded pension scheme with uniform pricing.

Consider a country that introduces a public PAYG pension scheme. At the time of introduction, the people who are retired benefit because they receive

[4] See Börsch-Supan and Reil-Held (2001) for an overview of the different concepts of life-time redistribution.

a public pension provision without having made any contribution to the system in the past. However, this introductory gain of the first generations cannot be passed on to all subsequent generations without any cost. These generations are forced to participate in a pension scheme with a lower return than the market rate of interest they would earn if the PAYG contributions were invested in the capital market. To see this, note that in a stable economic and demographic environment the implicit return in a PAYG system is equal to the population growth rate plus the growth rate of labour productivity. In a dynamically efficient economy this composite growth rate is lower than the market interest rate in the long run. Since any pension scheme must be a zero-sum game among all participating generations, the present value of the missed returns of all subsequent generations is equal to the introductory gain of the first generation (Sinn, 2000).

A funded pension scheme with uniform pricing can be viewed as a mixture between a pure PAYG scheme and a completely fair funded pension scheme. Therefore, the economic logic of a PAYG scheme also partly applies to a funded scheme with uniform pricing. As long as a pension fund bases its contribution rate on the pension accrual of all participants, uniform pricing creates a gain for the elderly working generations at the time of introduction.[5] These generations benefit from below cost-effective uniform contributions without having made above cost-effective contributions earlier. Like in a pure PAYG scheme, the burden of the introductory gain is necessarily imposed on all subsequent generations. These generations participate in a pension deal with a composite rate of return that, as long as an economy is dynamically efficient, falls somewhere between the market interest rate and the implicit return of a PAYG scheme. This implies that for a given level of pension benefits, a participant has to contribute more than he would contribute in a funded pension scheme with fair pricing. Similar to a PAYG scheme, the additional contributions are equal to the introductory gain in present value terms. Therefore, these additional contributions can be viewed as an intergenerational redistribution from

[5]Theoretically it is possible to construct a contribution rate that is actuarially fair over the life time of a generation (see equation (2.12) for an example). If there is no intragenerational heterogeneity and the economy is in a steady state, this contribution is the same for all individuals at each point in time. Then we have a situation in which the pension fund levies a uniform contribution rate whereas there are no intergenerational transfers. We abstract from this possibility, however, by imposing that a pension fund cannot charge gender-specific contribution rates. Rather, the fund bases the contribution rate on the pension accrual of all participants together (see equation (2.6) for an example), which is a standard practice in real-world pension schemes.

future generations to the generations who received the introductory gain.

2.2.2 Intragenerational redistribution

For Dutch pension schemes, Kuné (2005) gives an extensive overview of the types of intragenerational redistribution we may think of, such as redistribution between males and females, between individuals with a steep and flat career, between low-educated and high-educated people, between workers and disabled persons or between single and married people. Some of these transfers might be desired, others might be undesired, but they all share the property that they are not related to the financial position of a pension fund and, hence, are pure forms of *ex ante* redistribution.

Theoretically, intragenerational transfers mainly originate from differences in life expectancy, income growth and labour force participation growth. The actuarial value of future pension benefits is increasing in life expectancy since in expectation people with low life expectancies will receive pension benefits over a shorter period than people with high life expectancies. Therefore, uniform pricing redistributes from persons with a short life expectancy to persons with a long life expectancy. It also redistributes from persons with a flat income or participation profile to persons with a steep one. People with a steep profile benefit from uniform pricing because they accumulate relatively more pension rights at the end of the career, the period in which pension accrual is subsidized by young workers.[6] As such, in a funded pension scheme it is not so much the *level* of income or labour force participation that determines intragenerational redistribution, but more the individual *change* in income or labour force participation relative to those of other persons.

The extent to which intragenerational redistribution is profitable or harmful for an individual, depends on the distribution of the relevant individual characteristics (i.e., life expectancy, income profile and labour force participation profile) over the total pension fund population. If these characteristics are more or less uniformly distributed, the gains and losses will be of equal size at the individual level. If the distribution is skewed, however, the persons with extremely deviating characteristics will experience large gains or losses while the majority of the persons will hardly be affected. For example, the advantage of uniform pricing is higher for a person with a long life expectancy, if he is the *only* individual with this characteristic, because then the uniform

[6]Note that this channel of intragenerational redistribution is more important in a final-wage scheme than in an average-wage scheme.

contribution rate is low compared to his actuarially-fair contribution rate. At the same time, the disadvantage of the other persons is limited as the burden can be spread over a relatively large group.

2.3 The pension fund

Although the second pillar in the Netherlands consists of hundreds of pension funds, more than 90% of the participants have a DB pension scheme based on the average-wage system (DNB, 2009). In this system, pension benefits are linked to the average wage over a participant's entire career. We represent the second pillar using a model of a single representative pension fund which offers an average-wage DB scheme.

In the Dutch system of occupational pensions, significant reforms have been implemented in recent years.[7] Triggered by the plunging reserves during the creeping stock market collapse at the beginning of this century, many pension funds have introduced schemes which explicitly make the indexing of entitlements conditional on the fund's financial position. For simplicity, we abstract from conditional indexation and assume that the pension fund does not experience funding deficits or surpluses. Pension rights are unconditionally indexed to wages in our model. This simplification is defensible since we are only interested in *structural* transfers that take place regardless of the financial position of the pension fund. As long as risk allocation rules do not change, sharing of reciprocal mismatch risks (as for example reflected in conditional indexation) does not lead to structural transfers from one group to the other.[8]

2.3.1 The participants

It is well known that females have a longer life expectancy than males. Therefore, a pension scheme with uniform pricing will redistribute pension contributions from males to females. Also, there is much evidence that high-educated people have a higher life expectancy than low-educated people. Hoyert et al. (2001), for example, show that mortality rates for Americans aged 25 to 64 who have attended college are less than half the rates for those who

[7] See Van Ewijk (2005) for an overview of the major reforms in the system of occupational pensions.

[8] Hoevenaars and Ponds (2008) show that changes in risk allocation rules, like for example a switch to a more risky asset mix, involve large intergenerational transfers.

stopped education after completing high school. Deboosere and Gadeyne (2002) conclude, using Belgian data, that the difference in life expectancy at age 25 between high-educated and low-educated males is 5.3 years. For females this difference amounts to 3 years. In the Netherlands this difference is 4.9 years for males, while for females it amounts to 2.6 years (Van Herten et al., 2002).

To allow for these socioeconomic differences in life expectancy, we consider a pension fund that consists of participants who differ in age, level of education and gender. There are four educational levels: low education (L), low secondary (LS) education, high secondary (HS) education and high (H) education.[9] Since we also distinguish between male and female workers, there are in total eight socioeconomic groups, each of which has its own survival probabilities, labour force participation and income profile. As a consequence, intragenerational redistribution in this chapter can be split up into redistribution between males and females (cross-gender redistribution) and transfers between agents of the same gender but with different educational level (intereducational redistribution).

2.3.2 The pension scheme

To describe our pension scheme, we need to specify the following ingredients: the population structure, the contribution base, the uniform contribution rate and the accrual and benefit rules. Before discussing these elements, we first introduce some notational convention.

Notation

In our model, an individual-specific variable has four dimensions: gender, educational level, age and time. We use the letter t as time indicator and the letter j to indicate the year of birth. Hence, $x = t - j$ is the age of an individual at time t. In addition, h is the gender indicator which distinguishes males (m) from females (f), that is $h = \{m, f\}$, and i is the indicator with respect to educational attainment, so $i = \{L, LS, HS, H\}$. Hence, the value of a variable v at time t of a representative agent born at time j, with gender h and level of education i, will be indicated as $v(h, i, j, t)$.

[9]Low education means primary school only, low secondary education contains lower vocational training, high secondary education represents secondary and intermediate college level and finally, high education contains higher vocational training and academic level.

Agents enter the pension scheme at age x_w and they retire at age $x_r > x_w$. Nobody becomes older than age $x_e > x_r$. Hence, $x_w \leq x < x_r$ defines the active period in which agents earn wages, pay pension premiums and accumulate pensions rights; $x_r \leq x \leq x_e$ is the inactive period in which agents receive pension benefits.

Average-wage scheme

It is assumed that the development of the population is only determined by death and birth. There is no emigration or immigration. We allow for population growth (decline) and time-varying mortality rates. It is further assumed that deaths and births occur at the end of a period. Let n denote the growth rate of cohort sizes at birth and ϵ the probability that an individual also lives throughout the next period. Then the size of a cohort p born in year j and measured at time t is equal to:

$$p(h,i,j,t) = \begin{cases} (1+n(t-1))\, p(h,i,j-1,t-1) & \text{if} \quad x = 0 \\ \epsilon(h,i,j,t-1)\, p(h,i,j,t-1) & \text{if} \quad 0 < x \leq x_e \end{cases} \tag{2.1}$$

Individual wages (w) grow because of three factors: inflation (φ), productivity growth (g) and an incidental component (γ) reflecting promotion and professional experience. This incidental wage component depends on gender and level of education. So we have,

$$w(h,i,j,t) = \begin{cases} (1+\theta(t))\, w(h,i,j-1,t-1) & \text{if} \quad x = x_w \\ (1+\theta(t))\,(1+\gamma(h,i,j,t))\, w(h,i,j,t-1) & \text{if} \quad x_w < x < x_r \end{cases} \tag{2.2}$$

where $\theta \equiv (1+g)(1+\varphi) - 1$ is the nominal productivity growth. Wages w are expressed in full-time equivalents.

All residents of the Netherlands receive a first-pillar PAYG benefit from the age of 65. Funded pensions in the second pillar are supplementary to this first-pillar benefit. This implies that workers do not need to accumulate future pension benefits over their entire income. Instead, a franchise is deducted to compensate for the first pillar. The franchise (F) and contribution base (y) are then given by:

$$F(t) = (1+\theta(t))\, F(t-1) \tag{2.3}$$

$$y(h,i,j,t) = w(h,i,j,t) - F(t) \tag{2.4}$$

In an average-wage scheme the level of pension benefits depends on the average wage income the participant has earned during his career. Each year the participant accumulates a uniform percentage (α) of his contribution base as pension entitlement. To correct for part-time employment, non-participation and early retirement, the contribution base will be multiplied by the labour force participation rate (λ), expressed in full-time equivalents. So, the individual pension accrual (a) equals,

$$a(h,i,j,t) = \alpha y(h,i,j,t)\,\lambda(h,i,j,t), \quad x_w \leq x < x_r \tag{2.5}$$

The pension scheme is characterized by a uniform contribution rate (π_U). The pension contributions will be fully attributed to the participants. This assumption implies that the employer's part of the contributions are shifted on to the employee. The uniform contribution rate is defined as the present value of the total (i.e., aggregated over all socioeconomic groups) pension accrual divided by the total contribution base. That is,

$$\pi_U(t) = \frac{\sum_h \sum_i \sum_{j=t-x_r+1}^{t-x_w} p(h,i,j,t)\,a(h,i,j,t)\,\delta(h,i,j,t)}{\sum_h \sum_i \sum_{j=t-x_r+1}^{t-x_w} p(h,i,j,t)\,y(h,i,j,t)\,\lambda(h,i,j,t)} \tag{2.6}$$

where δ denotes the unit cost price of a wage-indexed pension benefit, which satisfies the following first-order difference equation:

$$\delta(h,i,j,t) = \begin{cases} \frac{\epsilon(h,i,j,t)(1+\theta(t+1))}{1+r(t+1)}\,\delta(h,i,j,t+1) & \text{if} \quad 0 \leq x < x_r - 1 \\ \frac{\epsilon(h,i,j,t)(1+\theta(t+1))}{1+r(t+1)}\,(1+\delta(h,i,j,t+1)) & \text{if} \quad x_r - 1 \leq x < x_e \end{cases}$$

with r the discount rate and where $\delta(h,i,t-x_e,t) = 0$. The unit cost price increases with age because, when agents get closer to the retirement age, the investment horizon of the contributions decreases while the probability to actually reach this age increases.

The uniform contribution rate is the rate that is actually paid by the participants but it is not equal to the actuarially-fair contribution rate (π_F), which is defined as,[10]

$$\pi_F(h,i,j,t) = \frac{a(h,i,j,t)\,\delta(h,i,j,t)}{y(h,i,j,t)\,\lambda(h,i,j,t)}, \quad x_w \leq x < x_r \tag{2.7}$$

or, by substituting equation (2.5),

$$\pi_F(h,i,j,t) = \alpha\delta(h,i,j,t) \tag{2.8}$$

[10]Note that for the pension fund it does not matter whether the contributions are financed by π_U or π_F. In both cases it collects the present value of the total pension accrual given by the nominator of equation (2.6).

In contrast to the uniform contribution rate π_U, the actuarially-fair rate π_F is increasing in age because this rate is linear in δ.

The individual pension benefit (b) completes the description of the model,

$$b(h,i,j,t) = \begin{cases} \sum_{k=1}^{x_r - x_w} a(h,i,j,t-k) \prod_{l=1}^{k} (1+\theta(t+1-l)) & \text{if} \quad x = x_r \\ (1+\theta(t))\, b(h,i,j,t-1) & \text{if} \quad x_r < x \leq x_e \end{cases} \tag{2.9}$$

2.3.3 Measuring redistribution

Recall that we define redistribution as any difference between the expected present discounted value of benefits (PV_b) minus the expected present discounted value of contributions (PV_c). Seen from the perspective of the worker, this deviation from absolute equivalence represents the net benefit (NB) of participating in the pension scheme. Formally, the net benefit of a worker who is born in year j, evaluated at the age of entrance (x_w), equals:

$$NB(h,i,j) = PV_b(h,i,j) - PV_c(h,i,j) \tag{2.10}$$

in which:

$$PV_b(h,i,j) = \sum_{k=x_w}^{x_r-1} a(h,i,j,j+k)\, \delta(h,i,j,j+k)\, \Psi(h,i,j,k)$$

$$PV_c(h,i,j) = \sum_{k=x_w}^{x_r-1} \pi_U(t+k)\, y(h,i,j,j+k)\, \lambda(h,i,j,j+k)\, \Psi(h,i,j,k)$$

$$\Psi(h,i,j,k) \equiv \prod_{l=x_w+1}^{k} \frac{\epsilon(h,i,j,j+l-1)}{1+r(j+l)}$$

By definition, the discounted value of benefits is equal to the present value of the actuarially-fair contributions. It should be emphasized that the net benefit, as defined in equation (2.10), is a rather narrow concept and only measures the degree of redistribution. It does not capture welfare effects of the pension scheme, derived from e.g., sharing risks among pension members.[11]

In the calculations later on, we break down the uniform contributions into a saving share and a transfer share. This boils down to rewriting equation (2.10) in $PV_c = PV_b + PV_T$, with $PV_T \equiv -NB$, and denoting PV_b as the saving share and PV_T as the transfer share. So defined, a *positive* transfer represents a tax, a *negative* transfer a subsidy.

[11] The welfare effects of risk sharing in pension schemes are the topic of Chapter 3.

Table 2.1: Measures of redistribution (% contribution base)

Total	=	uniform CR (equation (2.6))	−	individual CR (equation (2.11))
Intergenerational	=	uniform CR (equation (2.6))	−	generational CR (equation (2.12))
Cross-gender	=	generational CR (equation (2.12))	−	gender-specific CR (equation (2.14))
Intereducational	=	gender-specific CR (equation (2.14))	−	individual CR (equation (2.11))

Note: CR = contribution rate.

Equation (2.10) takes the discount rate r as given. When we instead interpret the net benefit as a function of the discount rate, set $NB(r)$ to zero and solve for r, we obtain the implicit rate of return of the pension contributions. Differences in the rates of return within a generation can be interpreted as intragenerational redistribution, while differences of the implicit returns between generations represent intergenerational redistribution.

Recall from Section 2.2 that the total transfer (PV_T) can be split up into inter- and intragenerational transfers. The intragenerational transfer, in turn, can be subdivided into cross-gender and intereducational transfers. In order to identify these different sources of redistribution, we define three separate contribution rates: the individual contribution rate, the gender-specific contribution rate and the generational contribution rate. These contribution rates share the common property that they are *constant* and *actuarially fair* over the working life of, respectively, an individual, gender or generation. Appendix 2.A provides formal definitions of these contribution rates.

Having defined these contribution rates, total redistribution (PV_T), expressed as percentage of the contribution base, is equal to the uniform contribution rate minus the individual contribution rate (see Table 2.1 for a schematic representation). To isolate the intergenerational transfer, we have to compare the uniform contribution rate with the generational contribution rate. In the exceptional case that the population growth rate plus the growth rate of productivity (i.e., the implicit return of a PAYG scheme) is equal to the market interest rate, the uniform contribution rate coincides with the generational contribution rate. In this case there is no intergenerational redistribution.[12] However, if the market interest rate exceeds the implicit return, as we have assumed,

[12] Appendix 2.A.2 gives a formal proof of this statement.

there is intergenerational redistribution.

To compute the cross-gender and intereducational transfers out of the intragenerational transfer, we use the gender-specific contribution rate, defined for males and females separately. The difference between the generational and gender-specific contribution rate then measures the cross-gender transfer. Finally, the difference between the gender-specific contribution rate and the individual contribution rate represents intereducational redistribution.

2.4 Baseline scenario

In this section, we quantify life-time redistribution related to uniform pricing. We present a baseline scenario in which demographics and labour force participation rates are held constant. We start by explaining the parameter values and data, then we turn to our main results and finally, we examine how sensitive these results are for the underlying assumptions.

2.4.1 Parameter values

We assume that the accrual rate (α) is 2% of the contribution base. In the base year (2005), the franchise (F) is set at 10 thousand euros. These numbers are commonly used in Dutch occupational pension schemes (DNB, 2009). At this point we abstract from population growth ($n = 0$) and set the real productivity growth (g) and the inflation rate (φ) at 1.7% and 2%, respectively. It is further assumed that the pension fund can only invest in one asset with a certain real rate of return of 3%. Note that this configuration implies that the implicit return on the intergenerational transfer is lower than the explicit rate of return on pension savings (see the discussion in Section 2.2.1). The retirement age is exogenous. All participants start working at age $x_w = 25$ and retire at age $x_r = 65$. Nobody becomes older than age $x_e = 99$.

2.4.2 Data

Participants differ in terms of age, gender and educational level. These differences boil down to three factors: survival probabilities, labour force participation and income. Recall that differences in life expectancy (or survival probabilities), income profile and labour force participation profile determine the direction and magnitude of the intragenerational transfers. In addition, the magnitude of these transfers also depends on the relative size of the socioeconomic groups in the pension fund population. In this subsection, we will discuss the baseline values of these variables together with the population composition.

Demography

Panel A of Table 2.2 shows the composition of the Dutch population between age 25 and age 64 distinguished by gender and level of education in 2005. The distribution of the educational levels over males and females is quite similar. For each gender the fraction of the people with high secondary education is the highest, while the fraction of the people with low education is the lowest. Relative to the female population, the percentage of the people with high education is higher in the male population while the percentage of the people with low secondary education is lower.

Once we know the size of the educational- and gender-specific birth cohorts, the population structure is completely determined by equation (2.1). We have calibrated these birth cohorts in such a way that the relative sizes of the socioeconomic groups in the total population is consistent with the figures in panel A.

While educational-specific life expectancies at birth are publicly available for the Netherlands, this is not the case for the underlying mortality rates. Fortunately, Deboosere and Gadeyne (2002) calculated educational-specific mortality rates for Belgium for the period 1991-1996. We use their results to estimate Dutch mortality rates for each socioeconomic group. To compute these estimates, we have largely followed the procedure described in Hári et al. (2006). The main idea is the following. First, we calculate, using these Belgian mortality data, for each socioeconomic group the ratio between the educational-specific mortality rate and the average mortality rate. The average mortality rate in this case, is the weighted average of Belgian mortality rates where weights are based on the number of persons present in each socioeconomic group in the Netherlands. Second, we apply these ratios to gender-specific mortality rates of the Dutch population, which are publicly available. Finally, we rescale the ratios in such a way that for each socioeconomic group the life expectancy of a 25-year-old individual exactly matches the corresponding life expectancy in actual Dutch data, as published by Van Herten et al. (2002).

In the baseline calculation we keep the mortality rates constant over time.

Table 2.2: Data

	Males				Females			
	L	LS	HS	H	L	LS	HS	H
A. Composition population (%)								
25-65	3.8	9.3	21.5	15.7	4.7	11.5	20.3	13.2
B. Life expectancy in years at age								
25	73.1	76.0	76.0	78.0	79.5	82.0	82.1	82.1
35	73.8	76.4	76.4	78.2	79.9	82.3	82.3	82.3
45	74.5	77.0	76.9	78.6	80.4	82.7	82.7	82.7
55	75.9	78.1	78.0	79.4	81.4	83.6	83.6	83.5
64	78.4	80.1	80.0	81.2	83.0	84.9	84.9	84.8
C. Labour force participation rate (%) per age group								
25-34	65.0	84.5	87.2	90.8	27.6	53.3	73.4	87.6
35-44	65.8	84.2	88.7	92.7	32.8	52.6	67.2	80.8
45-54	64.8	82.9	86.2	90.9	32.5	49.7	65.8	77.3
55-64	37.9	50.5	51.5	60.5	14.3	18.4	32.0	47.8
D. Annual incidental wage growth (%) per age group								
25-34	2.0	2.5	3.0	3.5	1.3	1.8	2.3	2.8
35-44	1.0	1.5	2.0	2.5	0.2	0.7	1.2	1.7
45-54	0.5	1.0	1.5	1.5	0.1	0.6	1.1	1.1
55-64	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

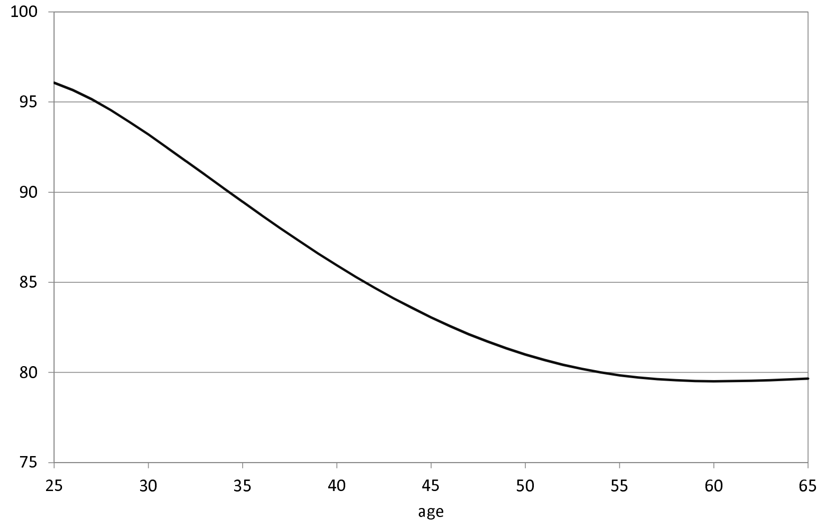
Source: Statistics Netherlands (panel A and C); Deboosere and Gadeyne (2002), Van Herten et al. (2002) and own calculations (panel B); figures in panel D are postulated.

Panel B of Table 2.2 displays for each socioeconomic group and for different ages the life expectancy. The difference in life expectancy at the age of 25 between the low- and high-educated group is 4.9 years for males and 2.6 years for females. Given gender the most pronounced difference in life expectancy is between the low education group and the other groups. There is little difference between the life expectancy of males with low secondary and high secondary education. For females we observe equal life expectancies for the three highest levels of education.

Labour participation and income profiles

Statistics Netherlands provides labour force participation rates per level of education for groups of age cohorts. Panel C of Table 2.2 displays these par-

Figure 2.1: Female income as percentage of male income by age, 2004



Source: Statistics Netherlands.

ticipation rates for the year 2005. From this panel we notice three facts that are relevant for the size of the redistribution effects. First, the labour force participation of females is significantly lower than that of males, in particular for participants with low and low secondary levels of education. Second, for each gender labour force participation positively depends on the level of education. Third, there is a remarkable drop in labour force participation for ages between 55 and 64.

As far as we know, there are no income profiles available by gender and level of education. We might assume however that both the starting wage (at age 25) and the career profile positively depend on the level of education. Panel D of Table 2.2 presents the income profile postulated for male workers.

Income data from Statistics Netherlands reveal that income differences between males and females are increasing in age, see Figure 2.1. The ratio between average full-time income of females and males is around 97% at age 25, which decreases to 80% at age 64. We use this information, together with the postulated incidental wage growth rates of males, to calculate educational-specific career steps for females. These figures are also shown in panel D. By lack of data, we impose that the relative income differential between males and females as shown by Figure 2.1 holds for each level of education. We recognize that the empirical foundation of the income profiles is relatively weak. We therefore shall carefully check how the redistribution results change if we take another assumption at this point.

2.4.3 Results

In the baseline scenario, the uniform contribution rate amounts to 21.7% of the contribution base. We decompose this rate into a saving share and transfer share. The transfer share is subdivided in an intergenerational transfer and an intragenerational transfer which, in turn, consists of a cross-gender and an intereducational transfer. In the baseline scenario, the redistribution effects (in percent of the contribution base) are constant over time. Recall that a positive transfer represents a tax, a negative transfer implies a subsidy.[13]

For each socioeconomic group, the saving share is by far the most important component of the uniform contribution rate (see Table 2.3). Nevertheless, the size of the transfer and saving share differs across groups. The saving share is higher for females than for males and it increases in the level of education, ranging from 15.7% of the contribution base for a male worker with low education to 25.1% for a female worker with high education. For male (*female*) participants the saving part is lower (*higher*) than the uniform contribution paid, implying that they pay (*receive*) a transfer. Notice that $6.0/21.7 = 28\%$ of the contributions paid by a low-educated male is redistributed to other pension members. On the other side of the spectrum, a high-educated female receives a subsidy of $3.3/21.7 = 15\%$ from the other participants.

Table 2.3 also shows the decomposition of the transfer share. Since the real market interest rate (3%) is higher than the real productivity growth rate (1.7%) plus the population growth (0%), the implicit return of a pension scheme with uniform pricing is lower than the explicit return of a pure funded scheme. Therefore, the contribution rate is higher than the generational contribution rate. This difference, defined as the intergenerational transfer, is 0.4% of the contribution base for all groups (which amounts to 2% of the contributions).

The cross-gender transfer equals $+1.9\%$ of the contribution base for males and -3.5% for females. This means that $1.9/21.7 = 9\%$ of the pension contributions of all male workers is transferred to female workers. Female workers, in turn, receive a subsidy from the male participants of $3.5/21.7 = 16\%$. The reason for the large difference between the tax and subsidy stems from the fact that there are relatively more male workers in the pension fund population.

[13] To compute the redistribution effects, we do not allow for compensating wage differentials. In a competitive labour market, employees will in principle be compensated for any actuarial unfairness of the pension scheme, reducing the amount of redistribution. However, we study the Dutch pension system in which practically all pension funds use uniform pricing. In this setting, there is less need for employers to provide a compensating wage differential because workers cannot evade these transfers by moving to another pension fund.

Table 2.3: Baseline results

	Males				Females			
	L	LS	HS	H	L	LS	HS	H
Contribution (% y)	21.7	21.7	21.7	21.7	21.7	21.7	21.7	21.7
– Saving part	15.7	18.6	18.6	20.6	22.0	24.4	24.8	25.1
– Transfer part	6.0	3.1	3.2	1.1	−0.3	−2.7	−3.1	−3.3
Inter transfer (% y)	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
Intra transfer (% y)	5.6	2.7	2.7	0.7	−0.7	−3.1	−3.5	−3.8
– Cross-gender	1.9	1.9	1.9	1.9	−3.5	−3.5	−3.5	−3.5
– Intereducational	3.7	0.8	0.8	−1.3	2.8	0.4	0.0	−0.2
Net benefit (% *LTI*)								
– age 25-64	−3.3	−2.1	−2.3	−0.8	0.1	1.6	2.1	2.4
– age 25-44	−4.3	−3.5	−3.9	−2.8	−0.9	0.2	0.2	0.3
– age 45-64	−2.0	−0.2	−0.3	1.4	1.9	4.2	4.9	5.4
Implicit return (%)	1.9	2.5	2.5	2.8	3.0	3.3	3.4	3.4

Notes: The saving and transfer shares are expressed in percent of the contribution base (y). A positive (*negative*) transfer is a tax (*subsidy*). Net benefit is defined as the present value (evaluated at age 25) of life-time benefits minus life-time contributions and expressed in percent of life-time income (*LTI*). The implicit return is defined as the rate of return for which the net benefit is equal to zero.

Irrespective of gender, uniform pricing entails a large redistribution from low-educated workers to higher-educated workers. The intereducational tax of a low-educated male is $3.7/21.7 = 17\%$ of the contribution rate, while a high-educated male receives a subsidy on their contribution of $1.3/21.7 = 6\%$. For females, the tax of a low-educated worker and the subsidy of a high-educated worker are $2.8/21.7 = 13\%$ and $0.2/21.7 = 1\%$, respectively. This high burden imposed on low-educated workers can be explained by two factors. First, the life expectancy of this group is relatively low so that the actuarial value of their pension rights is relatively low. Second, low-educated persons only constitute about 9% of the working population. Consequently, the intereducational transfer is imposed on a relatively small group.

Net Benefit and implicit return

The lower panel of Table 2.3 shows the net benefit (in percent of life-time income) of participating in a pension scheme with uniform pricing. Remember from equation (2.10) that the net benefit is defined as the difference between the present discounted value of benefits and contributions. Also in terms of the net benefit, we observe a sizable redistribution from males to females and from low-educated to high-educated persons. Let us first concentrate on the situation in which an individual contributes from age 25 until age 64 to the pension fund. The net benefit of male workers is then negative, ranging from -3.3% for a low-educated worker to -0.8% for a high-educated worker. For females, the net benefit is positive, varying from 0.1% for a low-educated person to 2.4% for a high-educated person. Surprisingly, a male worker with low secondary education has a higher net benefit (-2.1%) than a worker with high secondary education (-2.3%). The reason for this is that the former has a slightly higher life expectancy than the latter (see Table 2.2). For a female worker with low education, whose net benefit is nearly zero, uniform pricing turns out to be more or less actuarially fair.

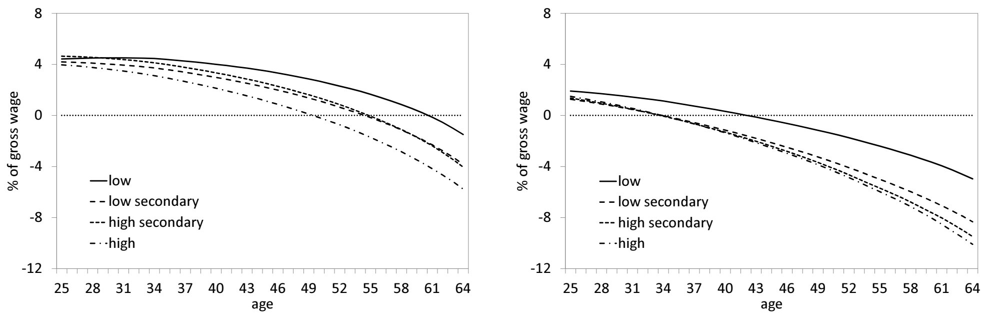
In absolute terms, the life-time taxes and subsidies due to uniform pricing may involve large amounts of money. To illustrate this, in our baseline calculation a male low-educated worker has a total life-time labour income of about 370 thousand euro (2006 price level). This means that this person pays a life-time transfer of $3.3\% * 370 = 12$ thousand euro to other pension fund members. A high-educated female, in turn, earns a life-time labour income of about 800 thousand euro. This person therefore receives a life-time transfer of $2.4\% * 800 = 19$ thousand euro from the other participants.

The net benefit calculations take the discount rate as given. Alternatively, we can solve for the implicit rate of return under the assumption that the net benefit equals zero. Differences in the implicit rates of return within a generation can then be interpreted as intragenerational redistribution. The implicit return of a low-educated male (1.9%) is only slightly higher than the real productivity growth (1.7%). For a low-educated female the implicit return is close to the market interest rate (3%) which confirms the insight already obtained that uniform pricing for this person is close to actuarial fairness. For secondary levels of education, the return of males is roughly 0.5%-points lower than the market interest rate, while the return of females exceeds this rate by about 0.4%-points.

Implicit taxation and incomplete careers

So far, we have assumed that a worker participates his whole career (from age 25 until age 64) in the pension fund. In practice, though, not each indi-

Figure 2.2: Implicit taxes (+) and subsidies (-) of uniform pricing



(a) Males (b) Females

vidual will work for forty years. As shown by Figure 2.2, on an annual basis uniform pricing involves a large redistribution from young to old workers. On average, a young worker of age 25 faces an implicit tax of about 3% of the contribution base, while an old worker of age 64 experiences a subsidy of about 6%. In our model, this intergenerational redistribution is driven by two factors. The first and most important factor is the time value of money. The investment horizon of the pension contribution of a young worker is longer than that of an old worker. In an actuarially-fair scheme, this would result in a higher pension entitlement for the young for a given contribution level. With uniform pricing, though, the young and old person receive the same entitlement, which means that part of the contribution of the young is redistributed to the old. Besides this, young persons also have a higher probability to die before the pension entitlement age than older workers but they do not receive an actuarial discount for this on the contribution rate.

The annual redistribution from young to old workers implies that the net benefit of the pension scheme is sensitive to the participation period. Table 2.3 therefore also shows net benefits for shorter spells of participation, which either occur in the first part of the life cycle (from age 25-44) or in the second part (from age 45-64). For each socioeconomic group we observe that the net benefit is higher if participation to the pension fund is postponed and lower if exit is advanced. By postponing participation people avoid the implicit taxes that occur at the beginning of the working period while people miss the subsidies if they leave the pension fund earlier. Notice that for male workers with low or even secondary education who participate as from age 45 the net benefit is still negative. For these persons the uniform contribution rate also exceeds the actuarially-fair contribution rate during the main part of the career after the age of 45 (see Figure 2.2a). However, for female workers with secondary or high education the net benefit is already positive when participation is confined to the first half of the working period (age 25-44). For these female workers, the uniform contribution rate already falls below the actuarially-fair contribution rate around the age of 35 (see Figure 2.2b).

One can argue that the redistribution from young to old becomes even larger if we would allow for risk (Bovenberg and Mehlkopf, 2013). If part of the pension contributions is invested in risky assets, the redistribution from young to old is larger because the contribution of the young is subject to the excess return over a longer time horizon while this person does not fully receive the compensation for this higher risk exposure in the form of a higher future benefit. Moreover, in practice many pension schemes do include some smoothing mechanism to spread out the effects of a shock over more periods. This further exacerbates the redistribution from young to old because part of the pension risks of the old is shifted to the young who often do not receive an actuarially-fair compensation for this.

Alternative decomposition

To decompose the intragenerational transfer in the baseline scenario, we have first determined the cross-gender transfer and subsequently, conditional on gender, the intereducational transfer (see Table 2.1). However, we could also have chosen an alternative decomposition, by first determining the intereducational transfer and then, conditional on skill level, the cross-gender transfer. Table 2.4 (upper part) shows the intragenerational redistribution effects of this alternative decomposition. These results do not differ substantially from the baseline figures in Table 2.3. The cross-gender transfer is negative for females and positive for males while the intereducational transfer is negative for high-educated workers and positive for all other socioeconomic groups.

Table 2.4 also presents the decomposition of the total net benefit into the part that is related to the intragenerational transfer and the part that is related to the intergenerational transfer. Consistent with earlier findings, the intragenerational part is more important than the intergenerational part. The intergenerational tax reduces the net benefit of all socioeconomic groups with 0.2 to 0.3%-points.[14]

[14] This effect is lower than the presented intergenerational transfer of 0.4% because the net benefit is expressed in percent of life-time income and the transfer in percent of the contribution base (i.e., income exceeding the franchise).

Table 2.4: Alternative decompositions

	Males				Females			
	L	LS	HS	H	L	LS	HS	H
Intra transfer (% y)	5.6	2.7	2.7	0.7	−0.7	−3.1	−3.5	−3.8
– Cross-gender	1.8	2.0	2.2	1.6	−4.5	−3.8	−4.1	−2.8
– Intereducational	3.8	0.7	0.5	−1.0	3.8	0.7	0.5	−1.0
Net benefit (% *LTI*)	−3.3	−2.1	−2.3	−0.8	0.1	1.6	2.1	2.4
– Intergenerational	−0.2	−0.3	−0.3	−0.3	−0.2	−0.3	−0.3	−0.3
– Intragenerational	−3.1	−1.8	−2.0	−0.5	0.4	1.9	2.4	2.8

Notes: The transfers are expressed in percent of the contribution base (y). A positive (*negative*) transfer is a tax (*subsidy*). Net benefit is defined as the present value (evaluated at age 25) of life-time benefits minus life-time contributions and expressed in percent of life-time income (*LTI*).

2.4.4 Sensitivity analysis

To investigate the sensitivity of the results with respect to the assumptions underlying the baseline scenario, we consider four alternative calculations. In the first one we decrease the real market rate of return. In the second and third calculations, we take an alternative view with respect to the gender- and educational-specific income profiles and labour force participation. In the fourth variant the statutory retirement age will be increased. Table 2.5 shows the effects of these alternative calculations on the saving part of the contribution, the intergenerational transfer and the intragenerational transfer (i.e., the sum of the cross-gender and intereducational transfer).

- **Lower market interest rate.** First consider the impact of a lower real market interest rate of 2% instead of 3% (variant I). This increases the present value of future pension benefits, and hence, the uniform contribution rate also increases substantially (with 7.1%-points). The market interest rate is now very close to the productivity growth rate (1.7%), which implies that the intergenerational transfer is almost zero. The saving share of the contribution increases for all socioeconomic groups. The lower interest rate has a relatively larger (positive) effect on the pension entitlements of females than on the entitlements of males. The reason is that the duration of the pension entitlements of females is longer because they generally have a higher life expectancy than males. The intragenerational subsidy provided to females therefore increases as well as the

Table 2.5: Sensitivity analysis

	Baseline	I	II	III	IV
Contribution (% y)	21.7	7.1	0.0	0.4	−1.3
Males, low education					
Saving part	15.7	5.3	0.1	0.0	−1.2
Intergenerational transfer	0.4	−0.4	0.0	0.0	0.0
Intragenerational transfer	5.6	2.2	−0.1	0.3	−0.1
Males, high education					
Saving part	20.6	6.9	−0.2	−0.1	−1.3
Intergenerational transfer	0.4	−0.4	0.0	0.0	0.0
Intragenerational transfer	0.7	0.6	0.2	0.5	0.0
Females, low education					
Saving part	22.0	8.3	0.5	0.1	−1.3
Intergenerational transfer	0.4	−0.4	0.0	0.0	0.0
Intragenerational transfer	−0.7	−0.8	−0.4	0.3	0.0
Females, high education					
Saving part	25.1	9.3	0.0	0.1	−1.3
Intergenerational transfer	0.4	−0.4	0.0	0.0	0.0
Intragenerational transfer	−3.8	−1.8	0.0	0.2	0.0

Notes: The contribution rate, the saving part and the transfers are expressed in percent of the contribution base (y). A positive (*negative*) transfer is a tax (*subsidy*). The variant figures are in absolute differences from the baseline. In variant I the rate of return is decreased by 1%-point. In variant II and III all groups face, respectively, identical wages and labour participation. In variant IV the pension age is increased from age 65 to 66.

intragenerational tax imposed on males.

- **Equalization of wages and labour participation.** Variant II presents the redistribution effects if we assume that all workers, irrespective of gender and level of education, have identical wages. In this scenario, all workers face the wage profile of a low secondary male (see panel D of Table 2.2). Variant III does the same for labour force participation. We observe that in our average-wage scheme both wage equalization and labour force participation equalization have very small effects on redistribution, especially on the intergenerational transfer. Recall that from a theoretical point of view uniform pricing redistributes from persons with a short life expectancy to persons with a long life expectancy and from persons with a flat income or participation profile to persons with a steep one. However, the results from Table 2.5 indicate that in Dutch occupational pensions redistribution from persons with short life expectancies to persons with long life expectancies is more important than the other two factors of intragenerational redistribution.

- **Increase in the retirement age.** In the baseline simulation, the annuity calculations are based on a reference retirement age of 65, which is indeed the current practice in Dutch occupational pension schemes. However, recently, employers and employees have decided to increase the reference pension age to 67 in 2014 and to introduce a direct link between this pension age and life expectancy from 2014 onwards. Variant IV shows the sensitivity of the redistribution effects for an increase in the pension age. In this variant, the retirement age increases by just one year, from age 65 to 66. We assume that people will work half a year longer when the retirement age increases by one year. Notice that a higher retirement age does not affect the transfers. For all socioeconomic groups the lower contribution rate is fully absorbed by a lower saving part.

  As pointed out by Nelissen et al. (2011), raising the retirement age increases the redistribution effects from short-lived agents to long-lived agents in percent of the contributions. The reason for this is that the retirement period of agents with low life expectancy decreases relatively more than that of agents with high life expectancy. We can illustrate this in our model by leaving out intergenerational transfers and by focusing on intragenerational transfers. Table 2.6 shows for all socioeconomic groups the actuarially-fair contribution rate, both for the baseline simulation and the alternative simulation with a pension age of 66. It also shows for each group the percentage difference between this actuarially-fair contribution rate and the generational contribution rate (i.e., the uniform contribution rate without intergenerational transfers). To illustrate, in the baseline the total contribution rate is 21.3% while the actuarially-fair rate of a low-educated male amounts to 15.7%. Hence, this person contributes 26.3% too much. A high-educated female has an actuarially-fair rate of 25.1% and therefore contributes 17.7% too little.

  Increasing the retirement age decreases the contribution rate, which falls from 21.3% to 20.0% of the contribution base. However, in relative terms, it increases the redistribution from agents with lower expected life spans to agents with higher expected life spans. When the pension age is raised

Table 2.6: Intragenerational redistribution effects of an increase in the retirement age from age 65 to 66

	Baseline		Pension age at 66	
	Contribution rate	Relative difference	Contribution rate	Relative difference
Males				
Low education	15.7	26.3	14.5	27.4
Low secondary	18.6	12.7	17.3	13.3
High secondary	18.6	12.8	17.3	13.4
High education	20.6	3.1	19.3	3.3
Females				
Low education	22.0	−3.4	20.7	−3.7
Low secondary	24.4	−14.7	23.1	−15.5
High secondary	24.8	−16.7	23.5	−17.6
High education	25.1	−17.7	23.7	−18.7
Total	21.3	0.0	20.0	0.0

Note: Contribution rates are expressed in percent of the contribution base.

to age 66, a low-educated male contributes 27.4% too much (instead of 26.3%) while a high-educated female contributes 18.7% too little (instead of 17.7%).

2.5 Two alternative scenarios

The population forecast of Statistics Netherlands suggests an increase and convergence in life expectancies of males and females in the coming decades. In addition, it is reasonable to assume that labour force participation of especially females will increase in the Netherlands (Euwals and Van Vuuren, 2005). In this section we will investigate how the redistribution effects of the baseline calculation change if we allow for these two future developments.

2.5.1 Increasing life expectancy

The population forecasts of Statistics Netherlands contains age-specific survival probabilities per gender. We have used these forecasts to calculate the

Table 2.7: Life expectancy in years for a 25-year-old individual

	Males			Females		
	95/05	2025	2050	95/05	2025	2050
Low education	73.1	74.6	75.3	79.5	80.5	80.7
Low secondary education	76.0	77.4	78.1	82.0	82.9	83.1
High secondary education	76.0	77.4	78.1	82.1	83.0	83.2
High education	78.0	79.3	80.0	82.1	83.0	83.1

Source: population forecast of Statistics Netherlands (2006-2050). See appendix 2.B.1 for more details.

future development of the mortality rates per socioeconomic group.[15] There is international evidence that the relative differences in mortality rates between socioeconomic groups has not declined in the last decades (Pappas et al., 1993; Mackenbach et al., 2003 and Meara et al., 2008). We therefore impose that, conditional on gender, the relative differences between the educational-specific survival probabilities will not change in the future. In Appendix 2.B.1 we explain in detail how the survival probabilities have been computed.

Table 2.7 shows the life expectancies at age 25 for three years, the base period (1995/2005), 2025 and 2050. After 2050 the survival probabilities are held constant and hence, life expectancy will not further improve thereafter. The figures reveal a convergence in life expectancy of males and females. Between 2005 and 2050 the life expectancy of males is expected to increase with 2.2 years, which is twice as much as the increase of 1.1 years for females.

Apart from the survival probabilities, we change two other assumptions of the stylized baseline calculation in order to make the demographic environment more realistic. First, we set the population in the base year equal to the actual Dutch population in 2005. Second, the growth rate of the cohort size at birth (n) is no longer set at zero. Instead, the growth rate is calibrated using the population forecasts of Statistics Netherlands. This growth is low and sometimes even negative, reflecting the fact that fertility rates are low in the Netherlands. After 2050 the growth rate is set at zero again.

Variant I in Table 2.8 presents the redistribution effects associated with the increasing life expectancy, expressed in absolute differences from the baseline results. The new demographic assumptions imply that the redistribution ef-

[15] See the online StatLine database of Statistics Netherlands (at www.cbs.nl) for the population forecast 2006-2050.

Table 2.8: Redistribution effects in two alternative scenarios

	Baseline	I		II	
		2007-12	∞	2007-12	∞
Contribution rate (% y)	21.7	1.6	2.1	1.7	2.3
Males, low education					
Saving part	15.7	2.3	2.4	2.3	2.4
Intergenerational transfer	0.4	−0.5	0.0	−0.5	0.1
Intragenerational transfer	5.6	−0.2	−0.3	−0.1	−0.1
Males, high education					
Saving part	20.6	2.3	2.3	2.3	2.3
Intergenerational transfer	0.4	−0.5	0.0	−0.5	0.1
Intragenerational transfer	0.7	−0.2	−0.2	0.0	0.0
Females, low education					
Saving part	22.0	1.1	1.1	1.2	1.2
Intergenerational transfer	0.4	−0.5	0.0	−0.5	0.1
Intragenerational transfer	−0.7	1.0	1.0	1.0	1.0
Females, high education					
Saving part	25.1	1.1	1.1	1.2	1.2
Intergenerational transfer	0.4	−0.5	0.0	−0.5	0.1
Intragenerational transfer	−3.8	1.0	1.0	1.0	1.0

Notes: The contribution rate, the saving rate and the transfers are expressed in percent of the contribution base (y). A positive (*negative*) transfer is a tax (*subsidy*). The figures of the variants are in absolute differences from the baseline. In variant I, life expectancy is assumed to increase in the coming decades. On top of this, in variant II the labour force participation of females increases.

fects are not constant anymore. They will gradually change over time until the population structure becomes stable again. We therefore show the effects for the short term (i.e., averaged over the period 2007-2012) and for the long term (i.e., when the population structure has stabilized) and we focus on a 25-year-old worker new entering the pension fund. As shown, the convergence in life expectancies between males and females reduces the amount of cross-gender redistribution. We observe that, relative to the baseline calculation, the intragenerational tax imposed on males decreases. The same holds for the intragenerational subsidy provided to females.

Interestingly, the intergenerational redistribution is slightly negative for a 25-year-old worker in the short run. Recall that this transfer is defined as the difference between the uniform contribution rate and the generational contribution rate. The generational contribution rate is forward looking and takes the increased life expectancy of a 25-year-old worker fully into account.[16] The uniform contribution rate, instead, reflects the pension accrual of all current workers and will gradually adjust when old workers with relatively low life expectancy are replaced by younger generations of workers with a higher life expectancy. The uniform contribution rate increases by 1.6%-points in the short term (2007-2012) and by 2.1%-points in the long run. The generational contribution rate, however, immediately increases by 2.1%-points. This means that the current generations do not fully pay the additional pension contributions necessary to cover the longer expected retirement period themselves under uniform pricing. A part of this burden is shifted on to future generations who will be confronted with positive intergenerational transfers.[17]

2.5.2 Increasing labour force participation of females

In the baseline calculation, we assumed time-invariant female labour force participation rates. However, due to sociological and cultural considerations, it is reasonable to expect that these participation rates will increase the coming decades (Euwals and Van Vuuren, 2005). Obviously, an increase in the labour force participation rates of females affects the size of the intragenerational redistribution, because it changes the composition of the pension fund population. We therefore extend the previous scenario with increasing female labour force participation. Euwals and Van Vuuren (2005) expect that the labour force participation rates of males will not change very much in the future. We therefore keep these rates constant, as before.

Knoef (2006) decomposes the development of Dutch female labour force participation during the last decade into age, period and cohort effects. The estimated age, period and cohort effects provide a tool to predict future participation rates for different socioeconomic groups. The forecasts of Knoef (2006) are based on the assumption that the relative differences between socioeconomic groups will not change in the future. The forecasts are defined in gross terms (labour force divided by total population) and for each age cohort. We follow the convention used by Statistics Netherlands to define the participation rates for groups of ten age cohorts. In addition, participation rates are

[16] See equation (2.12) in Appendix 2.A.

[17] In the long run, the intergenerational transfer is slightly higher than in the baseline but the difference is too small to observe with numbers rounded to one decimal place.

Table 2.9: Labour force participation, females (% cohort size)

	L	LS	HS	H
2005				
Age 25-35	27.6	53.3	73.4	87.6
Age 35-44	32.8	52.6	67.2	80.8
Age 45-54	32.5	49.7	65.8	77.3
Age 55-64	14.3	18.4	32.0	47.8
2008				
Age 25-35	31.7	56.6	75.4	88.9
Age 35-44	36.4	56.3	70.0	82.6
Age 45-54	35.4	52.6	68.6	79.3
Age 55-64	16.5	21.2	35.3	51.1
2012				
Age 25-35	33.2	58.1	76.7	89.8
Age 35-44	36.6	56.5	70.1	82.6
Age 45-54	36.8	54.0	69.8	80.1
Age 55-64	19.2	25.0	39.6	55.7

Source: Knoef (2006) and own calculations, see Appendix 2.B.2. The participation rates are defined as the active working force in percent of the total population.

defined in net terms (active labour force divided by total population) because only people who are actually working accumulate occupational pension.

Table 2.9 shows the predicted labour force participation rates for two years, 2008 and 2012. The 2005 figures repeat the participation rates of the baseline scenario. After 2012 we keep the labour force participation constant. In the near future, female labour force participation is expected to increase for each cohort group. In particular, the participation rates of female workers of age 55 and older will increase substantially. The same holds for the youngest category female worker with a low level of education.

Variant II of Table 2.8 presents the redistribution effects that include the increase in female labour force participation, on top of the demographic assumptions of Section 2.5.1. So, the differences between variant II and I reflect the pure impact of the change in labour force participation of females. The increase in female labour force participation leads to a small increase (of 0.2%-points) of the uniform contribution rate in the long run. It does not significantly change the redistribution effects. We only observe a small increase in the intragenerational tax for males.

In the sensitivity analysis of Section 2.4.4 we concluded that the impact of labour force participation profiles on the redistribution effects is much lower than that of different life expectancies. In this subsection we observe the same picture. The converging life expectancies of males and females puts upward pressure on the cross-gender transfers, while the increase of female labour force participation induces downward pressure on these transfers. Our results suggest that the effect of converging life expectancies dominates which implies that the cross-gender transfers are likely to decline in the future.

2.6 Concluding remarks

In the Netherlands, there is a public debate on the desirability of the uniform contribution and accrual system. This chapter aims to feed this discussion with relevant information about the magnitude and direction of the redistribution effects associated with uniform pricing. Our approach deviates from the usual one by measuring redistribution on a life-time basis, rather than on an annual basis. In addition, we calculate life-time redistribution in occupational pensions for various socioeconomic groups.

We find that a pension scheme with uniform pricing is not actuarially fair over an individual's entire career. First, analogous to a PAYG system, each participant pays an implicit tax in a dynamically efficient economy. This tax is necessary to service the introductory gain given away to the first generation. This intergenerational transfer is rather small, although its size is very sensitive to the (growth-adjusted) interest rate. Second, uniform pricing leads to intragenerational transfers, which in terms of magnitude are more important than is the intergenerational transfer. We find a large redistribution from males to females and from low-educated to higher-educated people. Since the recent population forecasts predict a convergence in life expectancies of males relative to females, it is likely that the cross-gender transfer will decline in the future. Our analysis reveals that differences in life expectancy are far more important for intragenerational redistribution than differences in income profile or the development of labour force participation.

In the Netherlands, charging a uniform contribution rate is legally obliged. When it was introduced, around the fifties of the last century, the idea was that uniform pricing provides equal opportunities for young and old, because an old and a young employee pay the same amount for a euro of pension. It is not clear whether intragenerational implications of uniform pricing have played a role in the decision process at that time. From a legal point of view, uniform pricing is always seen as a necessary element of solidarity to justify the competitive distortions caused by the mandatory participation for companies (in industry-wide pension funds) and for individuals (in pension schemes related to their collective labour agreement). Viewed in this light, it is remarkable that collective pension schemes contain transfers, especially those from low-educated to high-educated people, which should be characterized as socially unintended or even unnatural solidarity.

Against this background, our findings may provide a starting point to reconsider the desirability of uniform pricing in occupational pension schemes in more detail. This preferably demands a complete welfare analysis of the advantages and disadvantages. Besides undesirable transfers, uniform pricing can also lead to labour market effects since the uniform contribution rate generally differs from the market price of the pension accrual. These effects are not by definition disadvantageous. If the labour-supply elasticity of agents who benefit from uniform pricing (like females and older workers) is higher than the elasticity of those agents who face the negative effects (like males and younger workers), the transfers can stimulate aggregate labour supply and hence, total production or welfare. There is indeed evidence, also for the Netherlands (Evers et al., 2008), that females have a higher labour-supply elasticity than males and there are studies that find a positive relation between age and the labour-supply elasticity (French, 2005; Fenge et al., 2006).

In the chapter we have made some simplifying assumptions. First, the Dutch occupational pension system, which consists of a large amount of industrial pension funds and company pension funds, has been captured in a single representative fund. It is likely that the heterogeneity of the participants in these actual pension funds could be somewhat less pronounced than that in our representative fund. Second, we have only focused on old-age pensions while most of the pension arrangements also include a uniform surcharge to finance surviving dependants' pensions. Since in general the chance that the wife survives her husband is higher rather than the other way around, uniform pricing of surviving dependants' pensions can (partly) mitigate the cross-gender redistribution. On the other hand, in the Netherlands at least, participants can convert the accrued surviving dependants' pension rights into a more generous old-age pension. Further analysis is required to investigate to what extent these simplifying assumptions are decisive for the main conclusions.

APPENDIX TO CHAPTER 2

2.A Appendix to Section 2.3

In this appendix we provide formal definitions of the individual, generational and gender-specific contribution rates used to identify the intergenerational, cross-gender and intereducational transfers.

2.A.1 Individual contribution rate

We define the following two variables for $x_w \leq x < x_r$:

$$\hat{a}(h,i,j,t) \equiv a(h,i,j,t)\,\delta(h,i,j,t)$$
$$\hat{y}(h,i,j,t) \equiv y(h,i,j,t)\,\lambda(h,i,j,t)$$

Then the individual contribution rate (π_I) of an agent born in year j is given by:

$$\pi_I(h,i,j) = \frac{\sum_{k=x_w}^{x_r-1} \hat{a}(h,i,j,j+k)\,\Psi(h,i,j,k)}{\sum_{k=x_w}^{x_r-1} \hat{y}(h,i,j,j+k)\,\Psi(h,i,j,k)} \tag{2.11}$$

with discount factor Ψ already defined in equation (2.10).

2.A.2 Generational contribution rate

The generational contribution rate (π_G) is given by:

$$\pi_G(j) = \frac{\sum_h \sum_i p(h,i,j,j+x_w) \sum_{k=x_w}^{x_r-1} \hat{a}(h,i,j,j+k)\,\Psi(h,i,j,k)}{\sum_h \sum_i p(h,i,j,j+x_w) \sum_{k=x_w}^{x_r-1} \hat{y}(h,i,j,j+k)\,\Psi(h,i,j,k)} \tag{2.12}$$

Now we can stress the relation between the generational contribution rate (π_G) and the uniform contribution rate (π_U). For simplicity, assume that productivity growth (g), incidental wage growth (γ), the nominal interest rate (r), survival probabilities (ϵ) and labour force participation rates (p) are constant over time. Also, let the growth rate of the cohort at birth be zero ($n = 0$). Then, using equation (2.1), (2.2), (2.3) and (2.4), we can rewrite equation (2.12) in,

$$\pi_G(t) = \frac{\sum_h \sum_i \sum_{k=x_w}^{x_r-1} p(h,i,t-k,t)\, \hat{a}(h,i,t-k,t)\, \Phi(k)}{\sum_h \sum_i \sum_{k=x_w}^{x_r-1} p(h,i,t-k,t)\, \hat{y}(h,i,t-k,t)\, \Phi(k)} \tag{2.13}$$

with,

$$\Phi(k) \equiv \left(\frac{1+\theta}{1+r}\right)^{k-x_w}$$

Note that, besides the factor Φ, equation (2.13) is identical to equation (2.6). Since there is no population growth, the nominal implicit return of a PAYG scheme is $1+\theta$. Remember that the rate of return of a pension scheme with uniform pricing is a weighted average of the implicit return of a PAYG scheme and the market interest rate. Hence, if these returns are equal, a pension scheme with uniform pricing offers the same return as the capital market. In this case $\Phi(k) = 1$ for each k, and the generational contribution rate is equal to the uniform contribution rate. However, if $r > \theta$, as we have imposed, the uniform contribution rate exceeds the generational rate and the difference entails intergenerational redistribution.

2.A.3 Gender-specific contribution rate

The gender-specific contribution rate has the same form as equation (2.12), instead that we now have to aggregate over males and females separately. Denoting this contribution rate by π_H we have,

$$\pi_H(h,j) = \frac{\sum_i p(h,i,j,j+x_w) \sum_{k=x_w}^{x_r-1} \hat{a}(h,i,j,j+k)\, \Psi(h,i,j,k)}{\sum_i p(h,i,j,j+x_w) \sum_{k=x_w}^{x_r-1} \hat{y}(h,i,j,j+k)\, \Psi(h,i,j,k)} \tag{2.14}$$

2.B Data computations

2.B.1 Future mortality rates

Since forecasts of the educational-specific mortality rates are not publicly available for the Netherlands, we have to compute these figures ourselves. Starting points are the gender-specific mortality rates of the most recent population forecast of Statistics Netherlands, denoted $\hat{\mu}(h, j, t)$. The forecast horizon of these estimates ranges from 2006 to 2050. The computation of the educational-specific mortality rates involves the following two steps:

1. **Weighting.** The educational-specific mortality rates are generated by the following formula:

$$\hat{\mu}(h, i, j, t) = \omega(h, i, j)\, \hat{\mu}(h, j, t)$$

   in which the adjustment factors $\omega(h, i, j)$ are time-invariant and computed using a procedure described in Hári et al. (2006). In fact, $\omega(h, i, j)$ measures the discrepancy of the mortality rate of an individual of gender h, educational level i and age j relative to the average rate.

2. **Scaling.** We have not used the levels $\hat{\mu}(h, i, j, t)$ in our calculations directly. Instead, we have applied the following scaling to get rid of the discrepancy between the population forecast, $\hat{\mu}(h, j, t)$, and the mortality rates we use in our baseline calculation:

$$\mu(h, i, j, t) = \begin{cases} \mu(h, i, j, \text{base year}) & \text{if} \quad t = \text{base year} \\ \mu(h, i, j, t-1) + \Delta\hat{\mu}(h, i, j, t) & \text{if} \quad t > \text{base year} \end{cases}$$

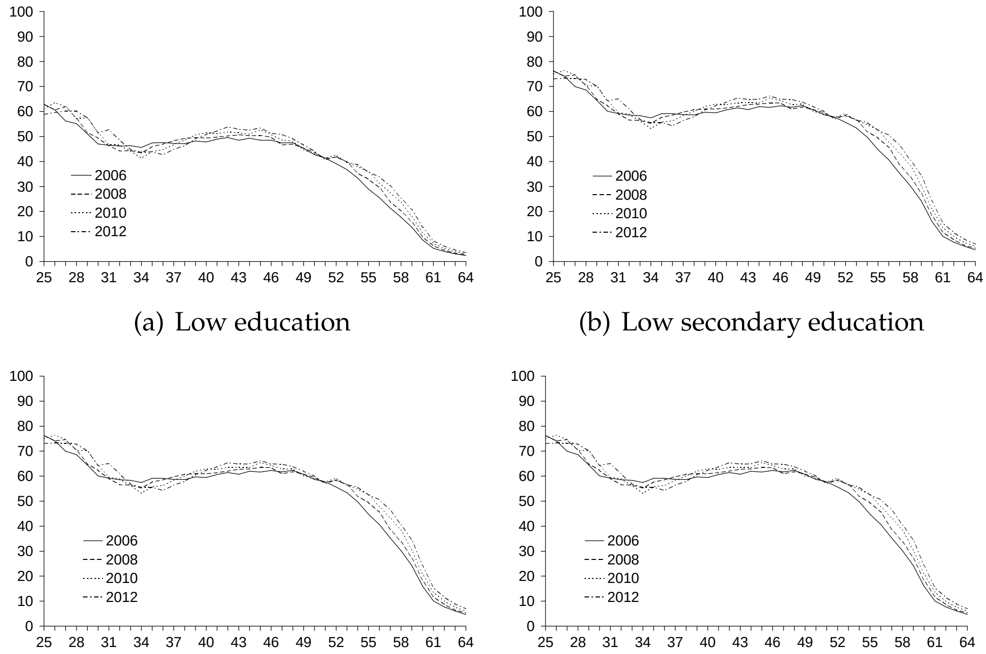
   where 2005 is our base year. Obviously, the survival rates $\epsilon(h, i, j, t)$ used in the formulas in the text are equal to $1 - \mu(h, i, j, t)$.

2.B.2 Labour participation of females

Knoef (2006) predicts female labour force participation rates for each age cohort (see figure 2.3). These rates are in gross terms, i.e., they are defined as the total labour force (employed and unemployed people) divided by the total population. In this study, labour force participation rates are defined in net terms, i.e., as the active working force divided by the total population. In addition, Statistics Netherlands reports educational-specific participation rates for groups of age cohorts only. The transformation from gross participation rates, defined for each age cohort, to net participation rates, defined for groups of age cohorts, involves the following three steps:

1. **Grouping.** Let z denote the group indicator, i.e., $z = 1, 2, 3, 4$, and $\underline{\nu}_z$ and $\overline{\nu}_z$ the lower- and upper-bound of z, expressed in age. The lower-bounds are $\underline{\nu}_1 = 25$, $\underline{\nu}_2 = 35$, $\underline{\nu}_3 = 45$, $\underline{\nu}_4 = 55$ and for the upper-bounds

Figure 2.3: Predicted gross labour force participation rates of females



(a) Low education

(b) Low secondary education

(c) High secondary education

(d) High education

Source: Knoef (2006).

we have $\overline{v}_1 = 34$, $\overline{v}_2 = 44$, $\overline{v}_3 = 54$, $\overline{v}_4 = 64$. Denoting the predicted gross participation by $\hat{\lambda}_B$, the grouped participation rates are defined as:

$$\hat{\lambda}_B(f,i,z,t) = \frac{\sum_{j=\underline{v}_z}^{\overline{v}_z} \hat{\lambda}_B(f,i,j,t)\,\hat{p}(f,j,t)}{\sum_{j=\underline{v}_z}^{\overline{v}_z} \hat{p}(f,j,t)} \tag{2.15}$$

where the age-dependent participation rates are weighted with the corresponding female population obtained from the population forecasts of Statistics Netherlands.

2. **Deflating.** The predicted net participation rates $\hat{\lambda}$ are derived from the formula:

$$\hat{\lambda}(f,i,z,t) = \hat{\lambda}_B(f,i,z,t)\,(1 - \hat{u}(f,i,z,t)) \tag{2.16}$$

with $\hat{u}$ the predicted unemployment rate (i.e., the total number of unemployed as percentage of the labour force). Predictions of educational-specific unemployment rates are not available. Hence we have to rely on some approximation rule. The CPB Netherlands Bureau for Economic Policy Analysis publishes projections of the macro unemployment rate.

We use the following approximation rule:

$$\hat{u}(f,i,z,t) = u(f,i,z,2005)\,\frac{\hat{u}(t)}{\hat{u}(t-1)} \tag{2.17}$$

3. **Scaling.** Confronting the predicted values $\hat{\lambda}$ with the realized values λ for the year 2005 reveal some small differences. Therefore, we will not use these predicted participation rates directly. Instead, we take first differences and relate these differences with the realizations in the base year (2005).

$$\lambda(f,i,z,t) = \begin{cases} \lambda(f,i,z,\text{base year}) & \text{if} \quad t = \text{base year} \\ \lambda(f,i,z,t-1) + \Delta\hat{\lambda}(f,i,z,t) & \text{if} \quad t > \text{base year} \end{cases} \tag{2.18}$$

CHAPTER 3

Risk sharing, endogenous labour supply and funded pensions[1]

Funded defined-benefit (DB) pensions add to welfare on account of providing intergenerational risk sharing, but lower it on account of inducing labour-supply distortions. This chapter shows that a properly designed funded DB pension scheme involves a welfare improvement even if contributions are distortionary and even if individuals face positively correlated wage and equity risks. Numerical calculations indicate that the diversification gains from risk sharing are large compared to the losses related to labour-supply distortions. This result withstands a number of extensions, like the introduction of a short-sale constraint for individuals, the inclusion of a labour income tax or the application of an alternative discount rate for pension benefits.

3.1 Introduction

In the industrialized world, population ageing jeopardizes the fiscal sustainability of public pay-as-you-go (PAYG) pension systems. The recent large government deficits due to the financial crisis load further pressure on the feasibility of these systems. As a result, in various countries the PAYG systems

[1]This chapter is based on the article 'Intergenerational risk sharing and endogenous labour supply within funded pension schemes' which will be published in *Economica*.

are gradually being reduced in favour of more investment-based individual retirement accounts. Examples of countries that have recently moved into this direction are Hungary and the Slovak Republic. At the same time, countries with traditionally large funded systems, such as the US, the UK and Switzerland, have replaced collective defined-benefit (DB) arrangements with individual defined-contribution (DC) arrangements in which benefits are subject to various risks. To illustrate, in the last decade the world-wide share of DC assets has grown from 35% in 2000 to 44% in 2010 (Towers Watson, 2011).

From a welfare perspective, it is *a priori* not immediately clear that a move towards individual DC schemes is the best way to implement (more) funding when compared to collective schemes. Indeed, in financial markets human capital is non-tradable and currently living generations are not able to share risks with those who are not born yet (Gordon and Varian, 1988; Shiller, 1999). As a result, young generations typically face a disproportionately high wage risk which they cannot shift to the old generations. The elderly, in turn, are overexposed to equity risk which cannot be shared with the young. This market incompleteness has recently been used to argue in favour of collective funded pension schemes instead of individual saving accounts.[2] The main advantage of a collective funded scheme is that it smoothes shocks over and beyond the life time of a single generation by disconnecting individual contributions and future benefits. However, in most real-world pension plans, contributions are related to labour income. A disconnection between contributions and benefits then implies that the contribution rate contains an implicit tax (or subsidy) which distorts labour supply. Moreover, it is well known that wages and stock returns are positively correlated, at least in the long run (Benzoni et al., 2007). This decreases diversification possibilities and, hence, may further reduce the attractiveness of risk sharing by the pension fund.

In this chapter, we explore the welfare aspects of collective funded pension schemes by comparing the risk-sharing gains with the labour-supply distortions. For this purpose, we develop a stylized two-period model with a young and an old generation. The economy is subject to two, potentially correlated, risk factors, equity risk and wage risk. The two generations cannot trade risks because the young are not able to participate in the capital market before shocks occur. As a result, young people are overexposed to wage risk because they own human capital while the elderly are overexposed to equity risk. The pension fund alleviates this market inefficiency. By providing safe benefits to the elderly, investing savings partly in equity and imposing the mismatch risk

[2]See Gollier (2008); Beetsma and Bovenberg (2009); Cui et al. (2011) or Beetsma et al. (2013).

between assets and liabilities upon the young, the pension fund lets young generations share in equity risk. Taking into account the optimal decisions of individuals, the pension fund determines the optimal degree of risk sharing. Mismatch contributions are levied on the labour income of the young generation. In this way, our model captures both the advantages (risk-sharing gains) and disadvantages (distortions on the intensive margin of labour supply) of collective funded schemes.

This chapter provides three main results. First, using Cobb-Douglas utility we analytically show that the introduction of a collective funded scheme with defined benefits and distortionary contributions involves an *ex ante* Pareto improvement. Using numerical simulations, we show that this result also holds for more general preferences and alternative model settings. We show that labour-supply distortions are always overcompensated by risk-sharing gains, even if the correlation between equity returns and wages is positive. The key intuition behind this result is the optimal investment policy of the pension fund. With this instrument the fund is able to control both the welfare gains from risk sharing and the size of the labour-supply distortions. Therefore, the investment strategy of the pension fund rules out the possibility that labour-supply distortions exceed the welfare gain from insurance.

Second, the benefits of risk sharing in this chapter do not imply a lower level of risk, but show up in a different guise. Intergenerational risk sharing increases the risk-bearing capacity of the economy and hence increases the demand for risky investments. Individuals are therefore able to exploit the positive equity premium already early in life, resulting in higher expected life-time consumption and welfare. This point was also made by Gollier (2008) and Cui et al. (2011) among others. This chapter adds to this that the benefits of risk sharing may also arise in a different combination of labour and leisure. Risk sharing shifts the labour-leisure choice towards labour supply as individuals can only capture the equity premium if they work.

Finally, we find that endogenous labour supply can reduce individual demand for risky assets if contributions are distortionary. This result contrasts with existing studies on the interaction between labour supply and portfolio choice (see e.g., Bodie et al., 1992 and Farhi and Panageas, 2007). These studies show that labour-supply flexibility offers insurance against adverse shocks which justifies more risky asset portfolios. The idea is that income effects in labour-supply behaviour cause a negative correlation between asset returns and labour income allowing individuals to take more risk. This chapter, however, shows that income-related transfers also introduce substitution effects.

These substitution effects work in the opposite direction and generate a positive correlation between labour income and asset returns. Hence, labour supply is subject to pro-cyclical pressure which reduces the risk-bearing capacity of individuals. For the pension fund investing in risky assets is therefore less attractive the more important are substitution effects.

The risk-sharing characteristics of pension systems have attracted much attention in the literature, especially with respect to PAYG systems.[3] Some recent studies also look at the role of funded pension systems in facilitating intergenerational risk sharing. Most of these studies assume exogenous labour supply and only focus on equity risk (Teulings and De Vries, 2006; Gollier, 2008; Cui et al., 2011), thereby overstating the risk-sharing gains. There are a few exceptions: Matsen and Thøgersen (2004) also include wage risk and explore the optimal split between a PAYG system and a funded DC system. They do not include endogenous labour supply, however. Beetsma et al. (2013) do allow for wage risk, equity risk and endogenous labour by exploring whether the combination of a PAYG pillar and a funded pillar with defined benefits can provide for optimal risk sharing. However, this study excludes any distortionary effects of pension contributions in the funded pillar.

Our result that a collective funded DB pension scheme increases welfare, no matter how distortionary are the contributions that are levied on the labour income of the young, may seem strong. Indeed, it conflicts with a number of studies on risk sharing and distortionary side effects that find that the indirect welfare losses on account of distortions may dominate the direct gains from risk sharing. In particular, Krueger and Kubler (2006) find this is the case for a more or less realistic calibration of the risk aversion of households. In a study of sharing of demographic risks, Sánchez-Marcos and Sánchez-Martín (2006) draw a similar conclusion. On the other hand, Nishiyama and Smetters (2007) find the risk-sharing gains of the US social security scheme to dominate its distortionary effects. Besides the fact that these studies focus on PAYG systems rather than funded systems, they differ from ours on a fundamental point: they explore the welfare effects of a typical pension scheme of which the size is not necessarily optimally chosen. Our study, in contrast, lets the pension fund choose optimally the amount of risk sharing, thereby preventing that the welfare effect of the pension scheme is negative.

At least four other studies are similar to ours on this particular point, although they differ in the risks or distortions on which they focus. Varian

[3] See, for example, Enders and Lapan (1982); Demange (2002); Ball and Mankiw (2007) and Gottardi and Kubler (2011).

(1980) and Persson (1983) find the gains from insurance to dominate distortionary effects if the size of the insurance scheme is optimally chosen. Furthermore, Fehr and Habermann (2008) study an optimizing unfunded pension scheme with basic allowances for contributions and a flat benefit fraction. They find that the gains from sharing idiosyncratic wage risks dominate the losses from a distortion of the labour-supply decision. Finally, Mehlkopf (2011) also concludes that the gains from risk sharing are much larger than the losses from labour-supply distortions, like in our study. However, his study does not allow for wage risk in the analysis of this trade-off.

The remainder of this chapter is structured as follows. In Section 3.2 we discuss the model. Section 3.3 derives an analytical solution for the *ex ante* welfare gain of the pension scheme in case of a Cobb-Douglas felicity function in consumption and leisure. Section 3.4 presents numerical simulation results for the more general version of the model that allows the intratemporal substitution elasticity to be lower or higher than one. Section 3.5 explores the welfare consequences of some relevant model extensions. Finally, Section 3.6 concludes.

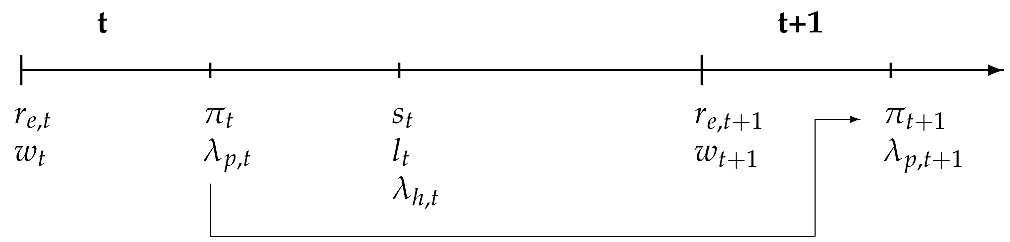
3.2 The model

We consider an economy populated with overlapping generations that cannot share risks through direct trade in financial markets. Each generation consists of a large number of individuals (normalized to unity) who live for two periods, such that in each period both a young and an old generation are alive. The model includes a collective pension fund that defines benefits independently from realized returns and lets the young bear the associated mismatch risk. In this way, the pension fund provides new opportunities for risk sharing, which agents cannot replicate in private markets. The model represents a small open economy in which, as usual, capital is assumed to be perfectly mobile and labour is perfectly immobile.

3.2.1 Timing

The sequence of events is shown in Figure 3.1. At the beginning of period t, a shock occurs in the equity rate of return ($r_{e,t}$) and the wage rate (w_t). After these shock have revealed, first the pension fund decides on the contribution

Figure 3.1: Timing of events



rate (π_t) and the portfolio share ($\lambda_{p,t}$) to be invested in equity. The pension fund acts as a benevolent Stackelberg leader, taking into account the future reactions of households to its decisions. A key property of the model is that the portfolio choice of the pension fund at time t only affects life-time utility of the next generation, born in $t+1$. As visually emphasized with an arrow in Figure 3.1, this dependency is driven by the direct impact of the fund's equity investment on next period's contribution rate. After the actions of the pension fund, the consumers decide upon private savings (s_t), the amount of leisure (l_t) and the equity share ($\lambda_{h,t}$), taking the pension contribution rate as given. The decisions of the consumers are based on the distribution of the future asset return. Since future pension benefits are only linked to current wages, consumers only face uncertainty about the return on their private savings.

3.2.2 Stochastic environment

There are two risk factors in the model economy, which are the return to equity holdings and the return to human capital. We assume stationary processes for these risk factors. The log return on the risky asset in excess of the log risk-free return, i.e., $\log(1+r_e) - \log(1+r_f)$, is an independently and identically distributed normal variable with mean μ_r and variance σ_r^2. In addition, the log wage rate, i.e., $\log w$, is also normally, independently and identically distributed over time with mean μ_w and variance σ_w^2. The covariance between the two factor prices is denoted by $\sigma_{wr} \equiv \rho_{wr}\sigma_w\sigma_r$, with ρ_{wr} the correlation coefficient between the log return on equity and human capital.

3.2.3 Individuals

Agents derive utility from consumption and leisure. The preference structure is represented by a time-separable, nested constant-elasticity-of-substitution

(CES) utility function that separates the aversion to risk and to intertemporal variation (Epstein and Zin, 1991). This separation is important for an analysis of pension contracts that alter both the risk properties and the timing of individual consumption flows. The utility function of an agent born at time t is thus defined as:

$$U_t = \left\{ u(c_{y,t}, l_t)^{1-\gamma} + \beta \left[\mathrm{E}_t\, u(c_{o,t+1}, 1)^{1-\theta} \right]^{\frac{1-\gamma}{1-\theta}} \right\}^{\frac{1}{1-\gamma}}, \gamma > 0, \theta > 0 \tag{3.1}$$

where $c_{y,t}$ and l_t denote consumption and leisure when young at time t, $c_{o,t+1}$ denotes consumption when old at time $t+1$ and β is the time discount factor. The parameters θ and γ define the coefficient of relative risk aversion and the inverse of the intertemporal substitution elasticity with respect to the consumption bundle of commodities and leisure. When $\gamma = \theta$, equation (3.1) reduces to a standard expected utility function where no distinction is made between risk aversion and intertemporal substitution. The per-period utility function $u(\cdot)$ is defined over commodities and leisure consumption, assuming a CES specification:

$$u(c,l) = \begin{cases} \left[(1-\eta)c^{1-\rho} + \eta l^{1-\rho}\right]^{\frac{1}{1-\rho}} & \text{for} \quad \rho > 0, \rho \neq 1 \\ c^{1-\eta} l^{\eta} & \text{for} \quad \rho = 1 \end{cases} \tag{3.2}$$

with $0 < \eta < 1$. The inverse of the intratemporal substitution elasticity is given by ρ and the utility parameter η governs the relative preference for leisure. In the following, we will use the following shorthand notation for per-period utility: $u_{y,t} \equiv u(c_{y,t}, l_t)$ and $u_{o,t+1} \equiv u(c_{o,t+1}, 1)$.

When young, an agent spends a fraction l_t of his total time endowment (normalized to unity) on leisure. A fraction π_t of labour income is contributed to the pension fund; the rest is devoted to consumption and private saving s_t. During the second period, the agent is retired. He then receives a labour-related pension benefit $(1-l_t)b_{t+1}$, where b_{t+1} denotes the maximum attainable level and the factor $1-l_t$ reflects the accumulation of pension benefits. The individual budget constraints are thus equal to:

$$c_{y,t} + s_t = (1-\pi_t)(1-l_t)w_t \tag{3.3}$$

$$c_{o,t+1} = (1 + r_{h,t+1})s_t + (1-l_t)b_{t+1} \tag{3.4}$$

with r_h the return on the household portfolio. Agents can either invest in a risk-free asset with return r_f or in equity with return r_e. The part of private savings that is invested in equity is denoted by λ_h, so that the return on the portfolio can be defined as:

$$r_{h,t+1} \equiv (1-\lambda_{h,t})r_f + \lambda_{h,t} r_{e,t+1} \tag{3.5}$$

Maximizing the objective function, equation (3.1), subject to the budget constraints, equation (3.3) and (3.4), gives the following set of first-order conditions with respect to $c_{y,t}$, l_t and $\lambda_{h,t}$:

$$u_{y,t}^{\rho-\gamma} c_{y,t}^{-\rho} = \beta \left(\mathrm{E}_t\, u_{o,t+1}^{1-\theta}\right)^{\frac{\theta-\gamma}{1-\theta}} \mathrm{E}_t \left[(1+r_{h,t+1}) u_{o,t+1}^{\rho-\theta} c_{o,t+1}^{-\rho}\right] \tag{3.6}$$

$$\begin{aligned}\eta u_{y,t}^{\rho-\gamma} l_t^{-\rho} = {} & (1-\eta)\beta \left(\mathrm{E}_t\, u_{o,t+1}^{1-\theta}\right)^{\frac{\theta-\gamma}{1-\theta}} \\ & \mathrm{E}_t \left\{\left[(1+r_{h,t+1})(1-\pi_t)w_t + b_{t+1}\right] u_{o,t+1}^{\rho-\theta} c_{o,t+1}^{-\rho}\right\}\end{aligned} \tag{3.7}$$

$$0 = \mathrm{E}_t \left[u_{o,t+1}^{\rho-\theta} c_{o,t+1}^{-\rho} (r_{e,t+1} - r_f)\right] \tag{3.8}$$

Equation (3.6) is the standard Euler equation which equalizes the marginal utility of first-period consumption to the discounted expected marginal utility of second-period consumption. Equation (3.7) is the first-order condition with respect to leisure and equation (3.8) is the condition for an optimal portfolio allocation.

The Euler equation specifies a relation between the marginal utility of consumption and the rate of return on assets. This relation becomes more clear by rewriting equation (3.6) in the form $\mathrm{E}_t\left[m_{t+1}(1+r_{h,t+1})\right] = 1$, where

$$m_{t+1} = \left[\frac{\left(\mathrm{E}_t\, u_{o,t+1}^{1-\theta}\right)^{\frac{1}{1-\theta}}}{u_{o,t+1}}\right]^{\theta-\gamma} \left(\frac{c_{o,t+1}}{c_{y,t}}\right)^{-\rho} \left(\frac{u_{o,t+1}}{u_{y,t}}\right)^{\rho-\gamma} \beta \tag{3.9}$$

defines the stochastic discount factor (SDF). The SDF measures the marginal value of a unit of consumption next period per unit of current consumption. The term in square brackets enters because of non-expected utility and compares next-period utility with its certainty-equivalent counterpart. A consumer that is relatively risk averse ($\theta > \gamma$) has a certainty-equivalent utility that is lower than expected utility.[4] That is, the consumer applies a correction factor to next period's marginal utility which is less than one for most states, implying that he discounts the future more heavily on average than an

[4]By Jensen's inequality, we have that $\left(\mathrm{E}_t\, u_{o,t+1}^{1-\theta}\right)^{\frac{1}{1-\theta}} < \mathrm{E}_t\, u_{o,t+1}$.

expected-utility consumer. The SDF can be used to discount expected pay-offs on any asset to find their prices. To see this, from equation (3.8), using (3.5), it follows that:

$$\mathrm{E}_t\left[u_{o,t+1}^{\rho-\theta}c_{o,t+1}^{-\rho}(1+r_{h,t+1})\right]=\mathrm{E}_t\left[u_{o,t+1}^{\rho-\theta}c_{o,t+1}^{-\rho}(1+r_{i,t+1})\right],\quad i=f,e \tag{3.10}$$

Using equation (3.6), we then have:

$$\mathrm{E}_t\left[m_{t+1}(1+r_{i,t+1})\right]=1,\qquad i=f,h,e \tag{3.11}$$

3.2.4 Pension fund

We consider a collective pension fund in which households are obliged to participate. Like in many real-world pension schemes, contributions are levied proportional to labour income.[5] The pension fund collects contributions from the young generation, invests these contributions in the capital market and pays out benefits to the same generation in the second period of life. The maximal attainable pension benefit b_{t+1} is risk free and defined as:

$$b_{t+1}=\alpha w_t \tag{3.12}$$

with α the replacement rate. Recall that the pension contract is related to labour history so that the actual pension benefit paid to the old generation at time $t+1$ equals $(1-l_t)b_{t+1}$.

As the pension fund on the one hand provides *safe* benefits to the elderly but on the other hand invests part of the contributions in *risky* capital markets, the mismatch risk between assets and liabilities has to be absorbed by the young in the form of additional contributions. The total contribution rate (π_t) therefore consists of two parts: a pure funded component ($\pi_{f,t}$) reflecting the own pension accrual of the young and a pure PAYG component ($\pi_{m,t}$) reflecting the mismatch risk between liabilities and assets, i.e., $\pi_t\equiv\pi_{f,t}+\pi_{m,t}$. The funded contributions $\pi_{f,t}(1-l_t)w_t$ are invested in the risk-free asset and the risky asset and earn a return:

$$r_{p,t+1}\equiv(1-\lambda_{p,t})r_f+\lambda_{p,t}r_{e,t+1} \tag{3.13}$$

[5]Since individual abilities are unobservable, policy makers (or pension funds) necessarily use observable wages to distribute shocks. Wage-related contributions can also be justified from constant relative risk aversion. In that case, optimal risk sharing implies that shocks should be distributed proportionally over pension members, based on total wealth (Bovenberg et al., 2007). One way to implement this is to use income-related contributions. In practice, funded pension contributions are often linked to wages exceeding some franchise level (see e.g., Chapter 2). In this chapter, however, we abstract from this and simply relate the contribution rate to wages.

with λ_p the share of pension contributions invested in equity.

From an *ex ante* point of view, the pension scheme is an actuarially fair deal if the funded component of the contribution rate is equal to the value the participant attaches to the future pension benefit. That is,

$$\pi_{f,t}(1-l_t)w_t = \mathrm{E}_t\left[m_{t+1}(1-l_t)b_{t+1}\right] \tag{3.14}$$

where we have used the SDF to discount future benefits to determine their current price. Equation (3.14) is the funding condition which ensures that the pension contract does not contain *ex ante* redistribution. Solving for the cost-effective component of the contribution rate, using (3.11), then gives:

$$\pi_{f,t} = \frac{\alpha}{1+r_f} \tag{3.15}$$

Notice that the risk-free return shows up as discount rate. This makes sense because for a young agent the accrued pension entitlement is equivalent to an investment of $\pi_{f,t}(1-l_t)w_t$ in the risk-free asset.

The mismatch part of the contribution rate, π_m, has to ensure that the solvency constraint of the pension fund will satisfy. This solvency constraint equals,

$$(1+r_{p,t+1})\pi_{f,t}(1-l_t)w_t + \pi_{m,t+1}(1-l_{t+1})w_{t+1} = (1-l_t)b_{t+1} \tag{3.16}$$

which states that the pension fund finances benefits in period $t+1$ (reflecting entitlements accumulated in period t) with actuarially-fair contributions levied in period t, the portfolio return earned on this in period $t+1$ and an intergenerational mismatch transfer levied in period $t+1$ on the basis of period $t+1$ labour supply. Hence, risk sharing is confined to two overlapping generations. As one model period represents roughly twenty years, the potential for risk sharing is maximized at forty years. This is not unrealistic if we look at risk-sharing mechanisms in actual pension schemes, which are often restricted by rigid solvency regimes.[6] Solving for the recovery rate, using equations (3.12), (3.13) and (3.15), gives:

$$\pi_{m,t+1} = -\frac{n_t(r_{e,t+1}-r_f)}{(1-l_{t+1})w_{t+1}} \tag{3.17}$$

with $n_t \equiv \pi_{f,t}\lambda_{p,t}(1-l_t)w_t$ the absolute amount of contributions invested in equity. For the scheme to be sustainable we need $\pi_{m,t+1} < 1$, otherwise

[6] In Section 3.5.5, though, we will relax this assumption and see how welfare changes when the risk-sharing horizon is extended.

the young generation is not always able to provide safe benefits to the old generation.

Equation (3.17) shows that the pension fund in effect issues risk-free bonds of size n_t to the old generation and invests the resources on behalf of the young generation in equity. In this way, young people are taking short positions in the risk-free asset and long positions in equity and, hence, they are sharing equity risk with the elderly which is not possible in private markets. Obviously, the mismatch contribution is zero if the pension fund does not invest in equity ($n_t = 0$). If it invests in equity ($n_t > 0$), the mismatch contribution is positive (*negative*) in case of a funding deficit (*surplus*) and the young effectively makes (*receives*) a transfer to (from) the old generation. On average the transfer is negative: if the rate of return on equity happens to be equal to its mean, the young generation receives a transfer which reflects the compensation for risk taking.

The pension fund uses its investment policy (n_t) to maximize expected utility of all currently living and future generations. This optimization problem does not depend on the current state of the economy because shocks are independently distributed and the only intergenerational link in the model is the mismatch rate (recall Figure 3.1). Indeed, the portfolio decision of the pension fund at time t only affects life-time utility of the generation born at $t+1$, through its direct impact on the intergenerational transfer. This property is a consequence of the utility-based valuation of the actuarially-fair contribution rate which is based on the risk-free return (and not the portfolio return). Hence, both the benefit of a risky investment strategy at time t (i.e., a higher expected portfolio return) and the mismatch risk between assets and liabilities show up only in the transfer of one generation which is young at time $t+1$. For the benevolent pension fund it is therefore sufficient to maximize *ex ante* life-time utility of one representative generation (i.e., life-time utility evaluated *before* the occurrence of shocks to wages and stock returns in the first period of the life of the agent and based upon the distribution of shocks in the two periods of his life),

$$\mathcal{W}_t = \left(\mathrm{E}_t\, U_{t+1}^{1-\theta}\right)^{\frac{1}{1-\theta}} \tag{3.18}$$

subject to constraint (3.17) and the individual first-order conditions, equations (3.6)-(3.8). Consequently, if the investment strategy of the pension fund improves welfare for this single generation, the policy is automatically a Pareto improvement.

3.3 Analytical solution

In this section we show analytically for Cobb-Douglas intratemporal utility ($\rho = 1$) that the funded DB scheme involves an *ex ante* Pareto improvement. In the simulations later on, we will present results for the general model based on CES utility ($\rho \neq 1$). In case of Cobb-Douglas utility, the first-order conditions (3.6)-(3.8) simplify to:

$$c_{y,t}^{-\varphi} l_t^{1-\omega} = \beta(1+r_f)\left(\mathrm{E}_t\, c_{o,t+1}^{1-\zeta}\right)^{\nu} \mathrm{E}_t\, c_{o,t+1}^{-\zeta} \tag{3.19}$$

$$\eta c_{y,t}^{1-\varphi} l_t^{-\omega} = (1-\eta)\beta(1+r_f)(1-\pi_{m,t})w_t\left(\mathrm{E}_t\, c_{o,t+1}^{1-\zeta}\right)^{\nu} \mathrm{E}_t\, c_{o,t+1}^{-\zeta} \tag{3.20}$$

$$0 = \mathrm{E}_t\left[c_{o,t+1}^{-\zeta}(r_{e,t+1} - r_f)\right] \tag{3.21}$$

where we used equation (3.10) to substitute out the stochastic portfolio return $r_{h,t+1}$ for the risk-free return r_f. Equations (3.19) to (3.21) include the parameters $\varphi \equiv 1-(1-\eta)(1-\gamma) > 0$, $\omega \equiv 1-\eta(1-\gamma) > 0$, $\zeta \equiv 1-(1-\eta)(1-\theta) > 0$ and $\nu \equiv (\theta-\gamma)/(1-\theta)$. The parameters φ and ω are the inverse of the intertemporal substitution elasticity with respect to consumption and leisure, respectively; ζ denotes the coefficient of relative risk aversion with respect to consumption and ν reflects the importance of non-expected utility.

3.3.1 Consumer problem

We start to solve for the individual portfolio allocation. We first consolidate the individual budget constraints (3.3) and (3.4) into:

$$c_{o,t+1} = (1+r_{T,t+1})\left[(1-l_t)w_{T,t} - c_{y,t}\right] \tag{3.22}$$

with,

$$w_{T,t} \equiv (1-\pi_{m,t})w_t \tag{3.23}$$

$$r_{T,t+1} \equiv (1-a_t)r_f + a_t r_{e,t+1} \tag{3.24}$$

$$a_t \equiv \frac{\lambda_{h,t}(1+r_f)s_t}{(1+r_f)s_t + (1-l_t)b_{t+1}} \tag{3.25}$$

Note that $w_{T,t}$ is full life-time income of individuals and that the portfolio share a_t relates the household's investment in equity to its *total* wealth, which is defined as the sum of financial wealth and pension wealth. As a result, $r_{T,t+1}$ is the effective return on the individual's total portfolio. Substituting equation (3.22) in first-order condition (3.21) gives:

$$\mathrm{E}_t\left[(1+r_{T,t+1})^{-\zeta}(r_{e,t+1} - r_f)\right] = 0 \tag{3.26}$$

Since shocks in $r_{e,t+1}$ are independently and identically distributed, there is a unique time-invariant solution a to equation (3.26). In Appendix 3.A.1 we show that this solution is approximately equal to:

$$a = \frac{\bar{\mu}_r}{\zeta \sigma_r^2} \tag{3.27}$$

where $\bar{\mu}_r = \mu_r + \frac{1}{2}\sigma_r^2$ is the expectation of the excess return on the risky asset. Equation (3.27) is similar to the result obtained by Merton (1969) and Samuelson (1969). The portfolio fraction invested in the risky asset is increasing in its expected excess return and decreasing in the variance of the excess return and the preference from consumption smoothing as measured by relative risk aversion ζ.

To solve for consumption and leisure demand, we first substitute the budget constraint (3.22) in first-order condition (3.19) to obtain:

$$c_{y,t} = \frac{(1 - l_t) w_{T,t}}{1 + z_t} \tag{3.28}$$

with z_t and the certainty-equivalent (CE) rate of return $r_{ce,t}$ defined as:[7]

$$z_t \equiv \left[\beta l_t^{\omega - 1} (1 + r_{ce})^{1-\varphi} \right]^{\frac{1}{\varphi}} \tag{3.29}$$

$$1 + r_{ce,t} \equiv \left[\mathrm{E}_t (1 + r_{T,t+1})^{1-\zeta} \right]^{\frac{1+\nu}{1-\varphi}} \tag{3.30}$$

The CE rate of return is the return on a hypothetical risk-free investment strategy that provides individuals the same expected utility level as they receive from optimally investing their wealth into the tradable risk-free and risky asset. Since it is assumed that the equity return is independently and identically distributed, the CE rate of return can be treated as an unconditional expectation. In Appendix 3.A.2 we show that this return is approximately equal to:

$$r_{ce} = r_f + \frac{1}{2} \frac{\bar{\mu}_r{}^2}{\zeta \sigma_r^2} \tag{3.31}$$

If the risky asset offers no excess return ($\bar{\mu}_r = 0$), agents will not invest in the risky asset so that the CE rate of return is equal to the risk-free return. In the

[7]To derive the definition of r_{ce} we use:

$$\begin{aligned} \mathrm{E}_t (1 + r_{T,t+1})^{1-\zeta} &= \mathrm{E}_t \left[(1 + r_{T,t+1})^{-\zeta} (1 + r_{T,t+1}) \right] \\ &= \mathrm{E}_t (1 + r_{T,t+1})^{-\zeta} (1 + r_f) + a \, \mathrm{E}_t \left[(1 + r_{T,t+1})^{-\zeta} (r_{e,t+1} - r_f) \right] \\ &= \mathrm{E}_t (1 + r_{T,t+1})^{-\zeta} (1 + r_f) \end{aligned}$$

where the last step follows from first-order condition (3.26).

more realistic case in which $\bar{\mu}_r > 0$, the CE rate of return exceeds the risk-free return and positively depends on the reward-to-variability ratio $(\bar{\mu}_r/\sigma_r)$.

Dividing (3.20) by equation (3.19) shows that the marginal rate of substitution between leisure and consumption is equal to the price of leisure,

$$\frac{\eta}{1-\eta}\frac{c_{y,t}}{l_t} = w_{T,t} \tag{3.32}$$

Recall from equation (3.23) that only the recovery rate shows up in the price of leisure, because, by construction, the cost-effective contribution rate is equal to the utility-based value of the accrued pension entitlement (that is, the cost-effective contribution rate is actuarially fair). Using equations (3.28) and (3.32), we can solve for consumption and leisure demand:

$$c_{y,t} = \frac{1-\eta}{1+(1-\eta)z_t} w_{T,t} \tag{3.33}$$

$$c_{o,t+1} = \frac{(1-\eta)z_t(1+r_{T,t+1})}{1+(1-\eta)z_t} w_{T,t} \tag{3.34}$$

$$l_t = \frac{\eta}{1+(1-\eta)z_t} \tag{3.35}$$

Inserting equation (3.35) in (3.29), yields for z_t:

$$z_t = \beta^{\frac{1}{\varphi}} \left[\frac{(1-\eta)z_t + 1}{\eta}\right]^{\frac{1-\omega}{\varphi}} (1+r_{ce})^{\frac{1-\varphi}{\varphi}} \tag{3.36}$$

We assume that this equation has a unique and positive solution for z. This solution is constant over time because, apart from z, only exogenous variables and structural parameters show up. Since z is a constant, labour supply is a constant as well. This is a well-known property of the Cobb-Douglas specification combined with a budget constraint without non-labour income. Then the income and substitution effects exactly cancel out implying a zero uncompensated labour-supply elasticity. Hence, we omit the time index of labour supply in the rest of the section.

Equation (3.36) can be used to derive the effect of uncertainty on the consumption and leisure decision. Taking the total differential of this equation and rearranging terms gives:

$$\frac{\mathrm{d}z}{\mathrm{d}r_{ce}} = \frac{\frac{1-\varphi}{\varphi}\frac{z}{1+r_{ce}}}{1-\frac{1-\omega}{\varphi}\frac{(1-\eta)z}{(1-\eta)z+1}} \tag{3.37}$$

Using the definition of φ, it can be shown that this derivative is unambiguously negative (*positive*) for $\gamma > 1$ ($\gamma < 1$) and zero for $\gamma = 1$. An increase in the degree of risk has two opposite effects on saving (see Sandmo, 1970). On the one hand, it makes the consumer less inclined to expose his resources to the possibility of loss (positive substitution effect on consumption). On the other hand, higher riskiness makes it necessary to save more in order to protect oneself against very bad states of low future consumption (negative income effect on consumption). If $\gamma > 1$, the income effect dominates the substitution effect implying that an increase in uncertainty (represented by a decrease in r_{ce}) leads to lower first-period consumption and leisure. Of course, if $\gamma < 1$, the opposite holds.

The solution for consumption enables us to solve for the fraction of private savings invested in the risky asset:

$$\lambda_{h,t} = \frac{a(1-\pi_{m,t})z}{(1-\pi_{m,t})z-(1+z)\pi_{f,t}} \tag{3.38}$$

If the pension fund increases the actuarially-fair contribution rate (i.e., higher π_m) to fund a more generous pension benefit, workers respond by increasing the equity share in private savings (i.e., higher λ_h). Since the pension benefit is safe, the actuarially-fair contribution is equivalent to an investment in the risk-free asset. Agents therefore counteract the actions of the pension fund with their private savings in such a way that in terms of total wealth the investment in the risky asset is constant. This offsetting response will be reinforced when agents are confronted with a positive surcharge ($\pi_{m,t} > 0$), because in that case the share of financial wealth in total wealth would decline even further. Note that $\lambda_{h,t}$ can in principle be larger than unity. In that case, the worker goes short in bonds to buy equity.

3.3.2 Pension fund problem

To derive the optimal policy of the pension fund we first need to specify the indirect utility function of households. Substituting the solutions for consumption, equation (3.33) and equation (3.34), and leisure, equation (3.35), in equation (3.1), gives:

$$V_t = \frac{(1+z)^{\frac{1}{1-\gamma}}\eta^{\eta}(1-\eta)^{1-\eta}}{1+(1-\eta)z}\, w_{T,t}^{1-\eta} \tag{3.39}$$

Inserting this function together with the definitions of life-time income and the recovery rate, equations (3.23) and (3.17), in the objective function of the pension fund, equation (3.18), and taking the derivative with respect to n_t, we obtain the following first-order condition:

$$\mathrm{E}_t \left[w_{T,t+1}^{-\zeta} (r_{e,t+1} - r_f) \right] = 0 \tag{3.40}$$

Note that condition (3.40) has a similar structure as the optimality condition of the portfolio choice of households, see equation (3.26). The difference is that the optimality condition of the pension fund contains two stochastic variables, the equity return and the wage rate.

Since labour supply is constant (because of Cobb-Douglas preferences) and shocks are independently and identically distributed, equation (3.40) has a unique time-invariant solution for the (absolute) equity investment n. This solution, based on a normalization of the expected wage level, can be approximated by the following equation (see Appendix 3.A.3):

$$n = \frac{\bar{\mu}_r - \zeta \sigma_{wr}}{\zeta \sigma_r^2} (1 - l) \tag{3.41}$$

The optimal policy indeed involves a constant equity investment proportional to labour supply. The term $\bar{\mu}_r / (\zeta \sigma_r^2)$ reflects the standard mean-variance optimal portfolio. The other term, σ_{wr} / σ_r^2, is the human capital hedge term: if the correlation between equity returns and wages is positive, the human capital hedge reduces the demand for equity. If the correlation is negative, the opposite holds. Like individuals, the pension fund invests a smaller amount in the risky asset if risk aversion increases (higher ζ), reflecting the higher preference for consumption smoothing across states of nature. As the *absolute* amount invested in equity is constant, the *relative* equity investment, i.e., equity investment as percentage of the actuarially-fair contributions $\pi_f (1 - l) w_t$, necessarily becomes time-dependent. That is,

$$\lambda_{p,t} = \frac{\bar{\mu}_r - \zeta \sigma_{wr}}{\zeta \sigma_r^2} \frac{1}{\pi_f w_t} \tag{3.42}$$

For lower (higher) values of π_f or w_t, the fund collects relatively less (much) contributions. In these cases, the pension fund needs to invest a larger (smaller) share of the contributions in the risky asset, so that its risk exposure in absolute terms is left unchanged.

3.3.3 Welfare measure

We use the income-equivalent variation as the welfare measure to determine the performance of the collective funded DB pension scheme compared to the individual DC scheme: it indicates how much the wage rate of the representative agent should be increased in the situation without a pension scheme (individual DC) in order to make this person indifferent between participating in the funded DB scheme or not. We take an *ex ante* perspective, i.e., *before* the agent knows the state of nature in the first period of life. Let $\mathcal{W}^{DC}$ denote *ex ante* indirect utility in the individual DC scheme and $\mathcal{W}^{DB}$ *ex ante* indirect utility in case there is a collective DB scheme. Then the income-equivalent variation (denoted by x_t) must be solved from the following equality:

$$\mathcal{W}_t^{DB} = \mathcal{W}_t^{DC}(x_t) \tag{3.43}$$

If $x_t > 0$ ($x_t < 0$) the funded DB pension scheme involves a welfare gain (loss) compared to the individual DC scheme.

Proposition 3.1 *Under the assumptions that pension contributions are proportional to income, wages and equity returns are jointly lognormally distributed and the per-period utility function has a Cobb-Douglas specification ($\rho = 1$), the collective funded DB pension scheme involves an ex ante Pareto improvement. This welfare gain is (approximately) equal to:*

$$x = \frac{1}{2\zeta}\left(\frac{\bar{\mu}_r - \zeta\sigma_{wr}}{\sigma_r}\right)^2 \tag{3.44}$$

Proof See Appendix 3.B.1. ∎

As long as the mean-variance term is larger than the human capital hedge term, i.e., $\bar{\mu}_r > \zeta\sigma_{wr}$, the pension will invest in the risky asset so that agents can capture the equity premium which is welfare enhancing ($x > 0$). Note that, if $\bar{\mu}_r < \zeta\sigma_{wr}$, the pension fund takes a short position in equity. However, also in this case the pension scheme allows for a welfare gain.[8] Notice further that the welfare gain is increasing with the reward for risk taking ($\bar{\mu}_r/\sigma_r$), decreasing with the covariance between human capital and equity returns (σ_{wr}) and decreasing with the coefficient of relative risk aversion (ζ). A higher reward for risk taking induces the pension fund to invest more in the risky asset. This raises the scope for intergenerational risk sharing resulting in higher welfare gains. For higher correlations between wages and equity returns or higher

[8]The case $\bar{\mu}_r < \zeta\sigma_{wr}$ seems empirically not very plausible though. Given our empirically based choices regarding the mean and standard deviation of equity returns and wages (see Section 3.4.1), we have $\bar{\mu}_r > \zeta\sigma_{wr}$, even if $\rho_{wr} = 1$. Hence, in the rest of this paper we restrict to the case $\bar{\mu}_r > \zeta\sigma_{wr}$.

degrees of risk aversion, however, the pension fund invests a smaller amount in the risky asset which decreases the scope for risk sharing.

To show that this pension scheme is *ex ante* Pareto improving, suppose that the scheme is introduced at time t^*. Then the generations born before time t^* are obviously indifferent between the case with and without a pension fund. The young generation at time t^* only contributes the actuarially-fair rate π_{f,t^*} of disposable income to the pension fund. Since there are no intergenerational mismatch transfers in t^* (i.e., $\pi_{m,t^*} = 0$) and the full-funding requirement is satisfied, we have that the utility of this generation also remains unaffected. The generations born at the beginning of time $t^* + 1$ and beyond, benefit from the introduction of the collective funded pension scheme. The reform is thus Pareto improving.

This Pareto-improving nature of the pension scheme implies that the welfare gain from risk sharing exceeds the welfare loss due to labour-supply distortions. The driving force behind this result is the optimal investment policy of the pension fund. With this instrument the fund is able to control both the insurance gains as well as the labour-supply distortions. Indeed, as *ultimum remedium* the fund can always replicate the private economy in which intergenerational risk sharing and labour-supply distortions are absent by investing all contributions in the risk-free asset. This investment strategy rules out the possibility that labour-supply distortions exceed the gain from insurance.

Interestingly, the welfare gain and the optimal equity investment of the pension fund do not depend on the replacement rate, α. Since the young generation values the defined pension benefit as an investment in the risk-free asset, the saving function (as represented by π_f) can be completely separated from the risk-sharing function (as represented by π_m). In the most extreme case with $\alpha = 0$, the fund does not collect actuarially-fair contributions at all (i.e., $\pi_f = 0$) but still invests an amount n in equity thereby providing a welfare gain of x. In this extreme situation, the pension fund explicitly takes short positions in safe assets to buy equity on behalf of future generations. However, once we introduce short-sale constraints on investment behaviour of individuals or the pension fund, this independency between the welfare gain and the pension fund size breaks down (see Section 3.5.1).

3.3.4 Lump-sum transfers

To disentangle the welfare gains from intergenerational risk sharing from the labour-supply distortions associated with income-related transfers, we also solve the model with lump-sum transfers (i.e., transfers that are unrelated to individual labour supply). We continue to assume Cobb-Douglas utility ($\rho = 1$) and we will use the same symbols for the endogenous variables as employed before, in case of proportional contributions. With lump-sum transfers, the budget constraint becomes:

$$c_{o,t+1} = (1 + r_{T,t+1}) \left[(1 - l_t)w_t - T_t - c_{y,t}\right] \tag{3.45}$$

with $T_t \equiv -n_{t-1}(r_{e,t} - r_f)$. Similar to the case with proportional transfers, a *positive* lump-sum transfer ($T_t > 0$) is a tax imposed on the young generation and a *negative* transfer ($T_t < 0$) is a subsidy. With lump-sum transfers, the first-order condition with respect to leisure changes into:

$$\eta c_{y,t}^{1-\varphi} l_t^{-\omega} = (1-\eta)\beta(1+r_f)w_t \left(\mathrm{E}_t\, c_{o,t+1}^{1-\zeta}\right)^{\nu} \mathrm{E}_t\, c_{o,t+1}^{-\zeta} \tag{3.46}$$

Combining (3.46) with (3.19) gives the result that the marginal rate of substitution between leisure and consumption equals the gross wage rate rather than the wage rate after the mismatch pension contributions as in the previous section,

$$\frac{\eta}{1-\eta}\frac{c_{y,t}}{l_t} = w_t \tag{3.47}$$

First-period consumption, second-period consumption and leisure then satisfy:

$$c_{y,t} = \frac{1-\eta}{1+(1-\eta)z_t}\,(w_t - T_t) \tag{3.48}$$

$$c_{o,t+1} = \frac{(1-\eta)z_t\,(1+r_{T,t+1})}{1+(1-\eta)z_t}\,(w_t - T_t) \tag{3.49}$$

$$l_t = \frac{\eta}{1+(1-\eta)z_t}\frac{w_t - T_t}{w_t} \tag{3.50}$$

with z_t already defined in equation (3.29).

Equations (3.48)-(3.50) are no closed-form solutions in the sense that their right-hand sides contain endogenous variables. Indeed, z_t is a function of leisure which is not constant anymore. As a consequence, it is not possible in general to solve for the optimal investment policy of the pension fund analytically. Only for a particular case, when life-time utility is log-linear in first-period and second-period consumption and leisure (i.e., $\gamma = \theta = \rho = 1$), is it possible to get closed-form expressions and to derive the optimal pension fund policy. In that case, z_t is constant and does not depend on leisure.

For this particular case the optimal pension fund policy is given by (see Appendix 3.B.2):

$$n = \frac{\bar{\mu}_r - \sigma_{wr}}{\sigma_r^2} \frac{1}{1 + r_f} \tag{3.51}$$

Comparing solution (3.51) with the solution in case of distortionary transfers, see equation (3.41) with $\gamma = \theta = \rho = 1$, it follows that in the latter case the pension fund invests a *smaller* amount in the risky asset. In the numerical simulations, later on, it will be shown that this result also holds if the assumption of log-linear utility is relaxed.

The lower demand for risky investments when contributions are proportional to income is driven by two effects. First, the intergenerational transfers are accompanied by distortions in the labour-supply decision which makes risky investments less attractive. Second, if contributions relate to labour income, the intergenerational payments introduce a substitution effect in the labour-supply decision. This substitution effect creates a positive correlation between labour income and asset returns, because if stock returns drop, the pension fund has to increase the contribution rate which reduces the price of leisure, depresses labour supply and hence, reduces labour income. This pro-cyclical pressure on labour-supply behaviour has a destabilizing effect on consumption of the young, making it less attractive to use their future human capital as collateral when investing in the capital market. Consequently, the portfolio holdings of the pension fund need to be shifted towards safe assets.[9] This result differs from the existing literature on the interaction between labour supply and portfolio selection that only focus on income effects (see e.g., Bodie et al., 1992). Income effects in labour-supply behaviour cause a negative correlation between asset returns and labour income. This has a stabilizing effect on consumption allowing individuals to take more risk.

For the case of log-linear life-time utility, the welfare gain can be decomposed into the welfare gain from risk sharing and the welfare loss from the labour-supply distortion associated with the mismatch rate.

Proposition 3.2 *Under the assumptions that contributions are financed lump sum, wages and equity returns are jointly lognormally distributed and expected life-time utility is log-linear ($\gamma = \theta = \rho = 1$), the collective funded DB pension scheme involves an ex ante Pareto improvement. This welfare gain is (approximately) equal to:*

$$x = \frac{1}{2} \left(\frac{\bar{\mu}_r - \sigma_{wr}}{\sigma_r} \right)^2 \left[1 + \frac{\eta}{(1 - \eta)(1 + \beta)} \right] \tag{3.52}$$

[9] Mehlkopf (2011) numerically derives this result using a multi-period model.

Proof See Appendix 3.B.2. ∎

Comparing equations (3.52) and (3.44), it follows that the income-equivalent variation in case of lump-sum transfers is *higher* than with proportional transfers. The difference between these two equations represents the size of the labour-supply distortion and when we express this difference relative to the income-equivalent variation with lump-sum transfers, we obtain $\eta/[1+\beta(1-\eta)]$. This term measures the fraction of the maximum welfare potential of risk sharing that is eroded by labour-supply distortions. This 'erosion factor' is increasing in the utility parameter η: if the share of leisure expenditures in total expenditures increases (higher η), the distortionary effects of contributions increase as well which leads to a larger erosion of the gains from risk sharing. Furthermore, the erosion factor is decreasing in the time discount rate β: if more weight is given to future consumption (higher β), agents choose to work more during the first period which reduces the distortionary impact of the mismatch contributions.

3.4 Numerical illustrations

This section provides numerical illustrations of the welfare gain of intergenerational risk sharing. We use numerical integration methods like Gaussian quadrature to solve the model. There are two reasons for switching to numerical simulation. First, the model version with CES preferences ($\rho \neq 1$) cannot be solved explicitly for the optimal decisions of individuals and the pension fund. Second, in the case of Cobb-Douglas preferences ($\rho = 1$), the analytical solution relies on an approximation of the log portfolio return and life-time income.[10]

3.4.1 Model parameters

Table 3.1 reports the values for the parameters used in the default scenario. In our model economy, agents live for two periods. Therefore, we interpret one model period to last twenty years (as in Matsen and Thøgersen, 2004). We set the time discount factor at $\beta = 0.8$, which corresponds to an annual time discount rate of 1.1%. We choose as benchmark an intertemporal elasticity of

[10]To derive equations (3.27) and (3.41), we assume that $1 + r_T$ and w_T are lognormally distributed. This is an approximation because both variables are linear combinations of the lognormally distributed equity return and, respectively, the risk-free return or the wage rate.

Table 3.1: Benchmark parameterization

Parameter	θ	γ	η	ρ	β	α	r_f	$\bar{\mu}_r$	σ_r	μ_w	σ_w	ρ_{wr}
Value	5	2	0.5	1	0.8	0.4	0.02	0.03	0.2	0	0.2	0.4

Note: the risk-free rate of return (r_f) and the mean ($\bar{\mu}_r$) and standard deviation (σ_r) of the equity return are annualized figures.

substitution of 0.5 ($\gamma = 2$) and an intratemporal substitution of unity ($\rho = 1$). An intertemporal substitution elasticity of one half is commonly used in the macro and public finance literature and it lies well within the range of available estimates (see e.g., Attanasio and Weber, 1995). We set the coefficient of relative risk aversion at $\theta = 5$ which is also standard in macro and finance literature (see e.g., Gollier, 2008). The calibrated share parameter η is set at 0.5 and the risk-free return (r_f) is set at 2% per year. The replacement rate α offered by the pension scheme amounts to 40% of labour income.[11]

Without loss of generality (in the power utility setting), we normalize the expected wage rate to unity ($\mu_w = 0$). Consistent with Constantinides et al. (2002), we assume that the 20-year standard deviation of the log wage rate (σ_w) is 20%. We further assume that the annual mean of the equity return is 5%, implying a risk premium ($\bar{\mu}$) of 3% per year. The standard deviation of the annual equity return (σ_r) is set at 20%. To construct 20-year shocks, we transform the annual mean and variance of the lognormal distribution (of equity returns) to the corresponding moments of the normal distribution (of log equity returns).[12] Since the log equity returns add up, the mean of the 20-year log return is just 20 times the yearly mean while the 20-year standard deviation is $\sqrt{20}$ times the yearly standard deviation. To reduce the sample variation, these 20-year log returns are scaled in a uniform way to ensure that the simulated mean and standard deviation are equal to their theoretical counterparts.

[11]The average (gross) replacement rate of mandatory and voluntary funded DB schemes in the OECD countries roughly amounts to 40%, albeit with large variation across countries (see OECD, 2009).

[12]If $\mathrm{E}(1+r_e)$ and $\mathrm{Var}(1+r_e)$ are the mean and variance of the lognormally distributed gross equity return, then μ and σ^2 satisfy:

$$\mu = 2\log \mathrm{E}(1+r_e) - \log(1+r_f) - \frac{1}{2}\log\left\{\mathrm{Var}(1+r_e) + [\mathrm{E}(1+r_e)]^2\right\}$$
$$\sigma^2 = \log\left\{\mathrm{Var}(1+r_e)\,[\mathrm{E}(1+r_e)]^{-2} + 1\right\}$$

Many studies, like e.g., Baxter and Jermann (1997) or Benzoni et al. (2007), have shown that capital and labour income are positively correlated at long horizons. A positive correlation between wages and equity returns is also consistent with a standard neoclassical production technology which is subject to productivity and depreciation shocks. Bohn (2009) provides estimates for the long-run correlation coefficient between the log equity return and log labour income, ranging from 0.17 to 0.69. In our baseline, we take a middle position and assume $\rho_{wr} = 0.4$. However, as for all the other parameters, we will perform an extensive sensitivity check to this value.

3.4.2 Baseline results

We start by exploring the simulation results for the default parameterization. Table 3.2 shows results for the private economy with we denote as the individual DC scheme and for the economy with a collective funded pension scheme, whereby it distinguishes between the defined-benefit with proportional transfers (DBP) scheme and the defined-benefit with lump-sum transfers (DBL) scheme. The table reports expected values, medians and the 10th and 90th percentiles. The reported statistics are based on 10,000 simulated paths for the equity return and the wage rate.[13]

Compared to the DC scheme, the introduction of the DBP scheme does not change labour-supply behaviour due to the assumption of a unitary elasticity of intratemporal substitution (i.e., income and substitution effects cancel). In both settings, individuals spend about 40% of available time on leisure. However, the pension scheme does increase the expected consumption levels in the first and second period. At the same time, it also raises the risk born by each generation as reflected by the wider 80%-confidence intervals of consumption. Hence, in this small open economy, the benefit of risk sharing is not to reduce risk but rather to increase the expected payoff from risky investments by generations who have not entered the labour market yet. The collective pension scheme alleviates the constraint that prevents generations to trade in financial markets before they are born. By imposing the mismatch risk between assets and liabilities upon the young, this generation in fact has a claim on the equity stock of the pension fund before entering the economy, resulting in higher expected life-time wealth. Gollier (2008) and Cui et al. (2011) also point to this

[13]With 10,000 simulations the distributions of the stochastic variables turn out to be converged. In addition, the 80% confidence interval of the computed welfare gains is not larger than 0.15%-points which we consider as sufficiently small.

Table 3.2: Baseline results

	c_y	c_o	l	a	$\lambda_h s$	n	x
DC							
expectation	0.401	0.386	0.394	0.249	0.054	–	–
10th percentile	0.304	0.243	0.394	0.249	0.041	–	–
median	0.394	0.345	0.394	0.249	0.053	–	–
90th percentile	0.508	0.566	0.394	0.249	0.069	–	–
DBP							
expectation	0.444	0.427	0.394	0.249	0.060	0.057	2.9%
10th percentile	0.293	0.239	0.394	0.249	0.040	–	–
median	0.418	0.372	0.394	0.249	0.056	–	–
90th percentile	0.621	0.671	0.394	0.249	0.084	–	–
DBL							
expectation	0.428	0.422	0.416	0.249	0.059	0.063	3.3%
10th percentile	0.300	0.241	0.371	0.249	0.040	–	–
median	0.411	0.368	0.402	0.249	0.056	–	–
90th percentile	0.577	0.654	0.477	0.249	0.081	–	–

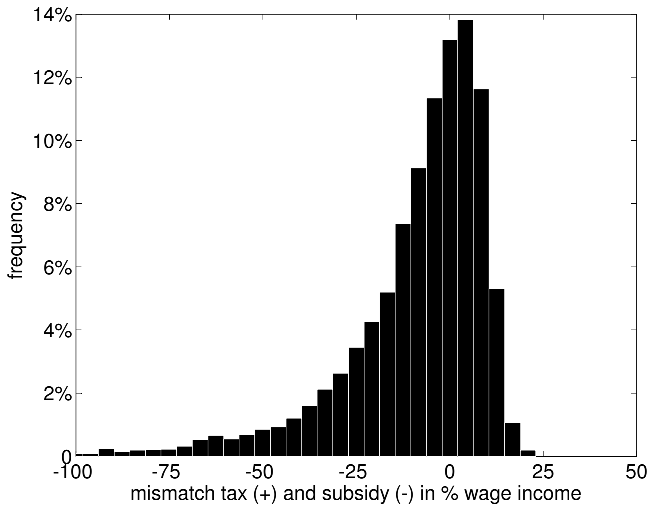
Notes: results are based on 10,000 simulations. Income-equivalent variations (x) are expressed in percent of the wage rate.

effect of intergenerational risk sharing.

When young individuals enter the economy and shocks are realized, the collective pension fund does no longer add anything to the transaction possibilities in financial markets. Indeed, at that time, the claim young people have to the pension fund is equivalent to an investment in the risk-free asset. Consequently, in all three economies considered individuals invest the same share (equal to 25%) of total wealth (financial plus pension wealth) in equity holdings. However, since intergenerational risk sharing increases expected life-time wealth, the absolute equity investment ($\lambda_h s$) is higher in the DBP and DBL economy as compared to the DC economy.

Of course, individuals can only capture the positive equity premium in the first period if they choose to work. In this way, the distortionary effect of the mismatch contribution rate acts as a subsidy on labour supply. Indeed, in the DBP scheme leisure time is lower than in the DBL scheme, while average consumption is higher. Besides this, the distortionary effect of the mismatch contribution rate also affects the investment policy of the pension fund. Proportional transfers reduce the risk-bearing capacity of individuals as they

Figure 3.2: Distribution of the mismatch contribution



distort the labour-leisure decision and simultaneously induce a pro-cyclical impact on consumption. Consequently, the optimal response of the pension fund is to downsize the investment in equity as compared to the lump-sum case.

From an *ex ante* perspective (i.e., *before* economic shocks materialize), the pension scheme with intergenerational risk sharing leads to a welfare improvement for all generations. In the default setting, the income-equivalent variation amounts to 2.9%.[14] That means, before knowing the state of the world at the time of entrance, individuals are willing to maximally give up 2.9% of their wage income to participate in the DBP pension scheme. The income-equivalent variation in the DBL scheme is 3.3%. This means that the distortionary effect of income-related transfers is about 0.4%-points, which implies that $(3.250 - 2.904)/3.250 = 10.7\%$ of the pure welfare gain from risk sharing vanishes due to labour-supply distortions.

From an *ex post* perspective (i.e., *after* shocks materialize), however, not all generations will benefit from participating. In states in which realized equity returns are lower than expected, the DBP scheme forces the young generation to transfer part of its wealth to the currently living old generation. Those agents would be better off by not participating. Figure 3.2 shows the distribution of the mismatch contribution rate (π_m) as percentage of wage income.

[14] If we compute the income-equivalent variation from the analytical solution, equation (3.44), we obtain a welfare gain of 3.3%.

Positive (*negative*) numbers here indicate that the young generation pays (*receives*) a transfer to (*from*) the elderly. The expected value is negative and equal to −9%. There is a probability of 61% (39%) that agents enter the pension contract with a subsidy (*tax*). Once shocks are known, the pension fund does not provide additional diversification gains anymore as compared to private savings. Therefore, from an *ex post* perspective, agents are only indifferent between participating or not if the transfer is occasionally equal to zero; if they experience a positive (*negative*) transfer, participation makes them worse (*better*) off. The average loss agents incur is 6% of wage income, while the average profit amounts to 20%. Although profits are generally much larger than the losses, voluntary participation is unable to commit young generations to risk sharing in all states. Consequently, the risk-sharing arrangement is only feasible if participation in the pension fund is mandatory.

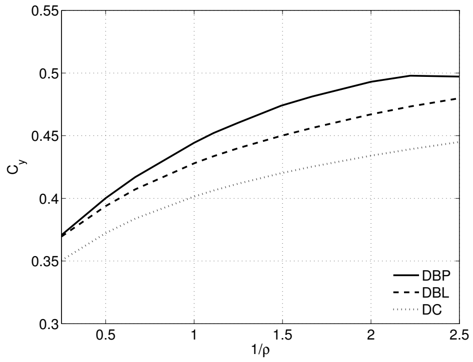
3.4.3 Simulations with CES utility

Until now, we have assumed a unitary intratemporal substitution elasticity between consumption and leisure, implying an uncompensated labour-supply elasticity of zero (income and substitution effects cancel). Actually, there is a lot of evidence suggesting a non-zero labour-supply elasticity (see e.g., Blundell and MaCurdy, 1999; Evers et al., 2008). Based on thirty different studies, Evers et al. (2008) construct a data set of empirical estimates of the uncompensated labour-supply elasticity. They show that the mean elasticity of men equals 0.07, while for women it equals 0.34. Mean elasticities for men range between −0.08 and 0.18. For women, mean elasticities range between 0.03 and 2.79. In our model, an intratemporal substitution elasticity larger (smaller) than unity coincides with a positive (negative) labour-supply elasticity, which means that the substitution effect dominates (is dominated by) the income effect.[15]

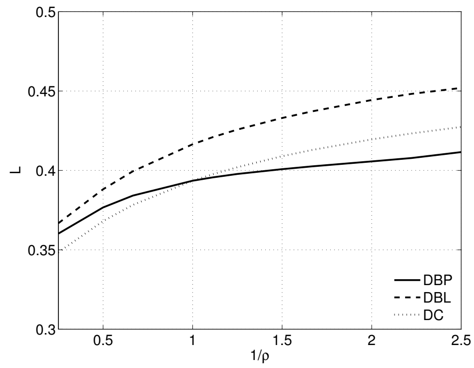
Figure 3.3 shows how optimal consumption and leisure of the young generation as well as the equity investment of the pension fund and the welfare gain vary with intratemporal substitution. We show results for the DBP scheme (solid lines) and the DBL scheme (dashed lines) and, if applicable, also for the individual DC scheme (dotted lines). To capture the empirical evidence regarding the labour-supply elasticity, intratemporal substitution ranges from 0.25 to 2.5 which coincides with an interval for the labour-supply elasticity

[15] See Appendix 3.C for a formal proof of the relation between the intratemporal substitution elasticity and the uncompensated labour-supply elasticity.

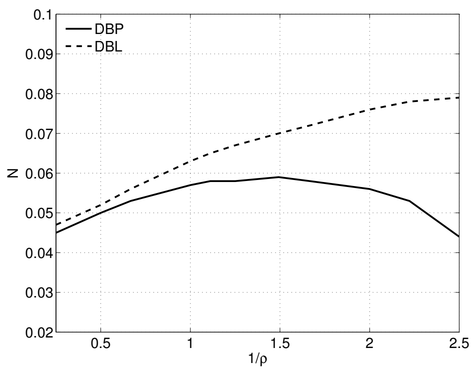
Figure 3.3: Simulation results with CES preferences



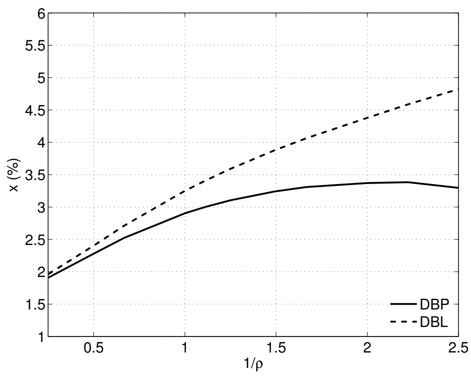
(a) Consumption



(b) Leisure



(c) Equity investment



(d) Welfare

Notes: consumption (c_y) and leisure (l) are averaged over 10,000 simulations. The income-equivalent variation (x) is expressed in percent of the wage rate.

that ranges from -0.3 to 0.5. Observe that risk sharing in the two collective pension schemes on average results in higher consumption as compared to the DC scheme. In the DBL economy, risk sharing also unambiguously results in more leisure consumption. Whether risk sharing also leads to higher leisure in the DBP economy, depends on the sign of the labour-supply elasticity. Recall that the collective pension scheme enables agents to capture the positive equity premium already in their first period of life. In this way, the mismatch contribution rate acts on average as a subsidy on labour supply tilting the labour-leisure choice in the direction of more working. Indeed, if the labour-supply elasticity is positive (intratemporal substitution larger than unity) individuals will actually enjoy less leisure than in the private economy. If the elasticity is negative (intratemporal substitution smaller than unity), we have the opposite: then the welfare gain is split between higher consumption and more leisure.

The equity investments of the pension fund in the DBP and DBL scheme are diverging if the substitution elasticity increases. In fact, in the DBP scheme the equity investment actually starts declining for higher values of the intratemporal substitution elasticity. If intratemporal substitution increases, the distortionary effect of the mismatch transfers becomes larger, which reduces the risk-bearing capacity of individuals. In addition, the higher substitution elasticity causes labour supply to become pro-cyclical: the effective wage rate decreases after a negative equity shock and increases after a positive shock. Labour supply thus becomes more positively correlated to equity returns, which further reduces the demand for risk taking in the investment portfolio of the pension fund. Consequently, the optimal response of the pension fund is to downsize the equity investment for higher degrees of intratemporal substitution in the DBP scheme, thereby preventing that the welfare effect becomes negative. However, it cannot prevent that an increasing share of the potential diversification gains are eroded by labour-supply distortions. To illustrate, for $1/\rho = 0.25$ the percentage of the gain eroded by distortions only amounts to $(1.962 - 1.908)/1.962 = 2.8\%$; for $1/\rho = 2.5$ this percentage has increased to $(4.822 - 3.297)/4.822 = 31.6\%$.

3.4.4 Sensitivity analysis

To explore the robustness of the baseline results we solve the model also for alternative parameter values. We compute the income-equivalent variation in the presence of proportional transfers (DBP scheme) as well as lump-sum transfers (DBL scheme) and calculate the fraction of the welfare gain eroded by labour-supply distortion. We consider alternative values for the preference parameters and the statistical moments of the risk factors. From our analytical exposition we know that these parameters are important determinants of the welfare gain. The findings obtained are presented in Table 3.3 and summarized as follows:

- **Preference parameters.** The welfare effects and the size of the labour-supply distortion are very sensitive to changes in the leisure parameter (η) and the coefficient of relative risk aversion (θ). Notice that the welfare gain is decreasing in the degree of relative risk aversion, a result already derived analytically. For lower degrees of risk aversion, the pension fund takes more risk which raises the scope for risk sharing (and hence,

Table 3.3: Income-equivalent variation and equity investment: sensitivity analysis

	DBP		DBL		Distortion
	x	n	x	n	
Baseline	2.9%	0.057	3.3%	0.063	10.7%
$\eta = 0.3$	1.6%	0.047	1.7%	0.049	4.7%
$\eta = 0.7$	5.4%	0.063	7.0%	0.080	22.7%
$\theta = 3$	6.5%	0.109	7.8%	0.131	17.1%
$\theta = 7$	1.4%	0.033	1.5%	0.035	7.9%
$\bar{\mu}_r = 0.02$	0.6%	0.026	0.7%	0.029	11.1%
$\bar{\mu}_r = 0.04$	7.2%	0.089	8.0%	0.099	10.6%
$\sigma_r = 0.15$	7.4%	0.119	8.3%	0.134	11.4%
$\sigma_r = 0.25$	1.2%	0.028	1.3%	0.031	10.5%
$\sigma_w = 0.1$	4.8%	0.080	5.3%	0.088	10.3%
$\sigma_w = 0.3$	1.5%	0.036	1.7%	0.040	11.0%
$\rho_{wr} = 0.0$	6.8%	0.089	7.5%	0.098	10.0%
$\rho_{wr} = 0.8$	0.6%	0.027	0.7%	0.030	11.4%

Notes: Results are based on 10,000 simulations. Baseline parameters are shown in Table 3.1. Income-equivalent variations (x) are expressed in percent of the wage rate. The labour-supply distortion is defined as the absolute difference between the income-equivalent variation of the DBL and the DBP scheme divided by the income-equivalent variation of the DBL scheme.

welfare). On the contrary, the income-equivalent variation is increasing in the leisure parameter η. If this parameter increases (*decreases*), the welfare gain increases (*decreases*) because the average subsidy the pension scheme provides increases (*decreases*) as a percentage of labour income. At the same time, an increase in η also increases the volatility of the mismatch payments, which widens the gap between the income-equivalent variation in the DBP and DBL scheme.

- **Mean and standard deviation.** The income-equivalent variations are also very sensitive to the mean and volatility of wages and equity returns because changes in these parameters directly affect the risk-return trade-off. An increase in the equity premium ($\bar{\mu}_r$), for a given degree of uncertainty, makes equity investments more attractive resulting in higher welfare gains from risk sharing. Obviously, an increase in the standard deviation of equity (σ_r) or wages (σ_w) leads to the opposite outcome. As these first and second statistical moments do not influence individual preferences (and, hence, distortions), the fraction of the diversification gain eroded by labour-supply distortions remains more or less constant in all these cases and roughly equal to 10%.

- **Correlation between equity returns and wages.** In the baseline, we have assumed a correlation coefficient between wages and equity returns (ρ_{wr}) of 40%. However, as mentioned earlier, the empirical estimates show a wide interval for this correlation coefficient. Note that the welfare gain is decreasing in the correlation coefficient. For high correlation between wages and equity returns, young generations are less inclined to bear the equity risk of the old generations because they are already exposed to a correlated risk factor via their human capital. The pension fund therefore invests less in equity which decreases the scope for risk sharing and, hence, leads to lower welfare gains. For low (and negative) correlation we obviously have the opposite: then young individuals are better equipped to act as the residual claimant of the assets of the fund since losses at the capital market may be compensated by gains at the labour market and *vice versa*. Accordingly, the pension fund invests more in equity which increases the diversification gains from risk sharing.

To summarize, calculated for a wide range of realistic parameter values, the reported welfare gains from risk sharing are large compared to the labour-supply distortion associated with the intergenerational payments. The welfare costs of labour-supply distortions are not negligible, however. For some cases, the size of the labour-supply distortion can increase to roughly one quarter of the pure welfare gain of risk sharing.

3.5 Extensions

This section considers the welfare implications of five extensions to the preceding analysis: the introduction of a short-sale constraint for individuals (Section 3.5.1), the inclusion of a labour income tax (Section 3.5.2), discounting of the pension benefit with the portfolio return rather than the risk-free rate (Section 3.5.3), putting a cap on the intergenerational transfers (Section 3.5.4) and, finally, extending the risk-sharing horizon (Section 3.5.5).

3.5.1 Short-sale constraint consumers

So far, we have assumed that agents do not face a borrowing or liquidity constraint. That means, agents can take short asset positions. In addition, if that would be optimal, agents can choose to borrow in the first period to smooth consumption over the first and second period. In practice, though, it is often difficult or even impossible for young people to take short positions in an asset, because human capital alone does not collateralize major loans in modern economies for reasons of moral hazard and adverse selection problems (Shiller, 1999). To overcome this objection, this section solves the model with a short-sale constraint for individuals, i.e., we impose $0 \leq \lambda_{h,t} \leq 1$. We do not have to consider a non-negativity constraint on private savings because in the baseline scenario the young's optimal savings turn out to be positive in any scenario.

When there is no short-sale constraint, we have seen before that the size of the collective pension scheme in terms of the exogenous replacement rate (α) does not play a major role in the model. Indeed, any actions by the pension fund can be undone by the individual. However, if there is a short-sale constraint, the size of the pension sector will matter because there is a possibility that agents cannot offset the decision of the pension fund. In our DBP pension scheme it is indeed the case that the implementation of a short-sale constraint restricts individual behaviour in some states. We already discussed that for the benchmark parameterization agents invest 25% of *total* wealth in the risky asset (see Table 3.2). As a percentage of *financial* wealth, however, this equity investment is much larger in the DBP scheme. Table 3.4 shows that the equity investment is 105.7% of financial wealth in the expected path, which means that agents take short positions in the risk-free asset on average.

The introduction of the short-sale constraint reduces the risk exposure considerably: the expected value of λ_h declines from 105.7% to 86.6% and the 90th percentile decreases from 159.1% to the imposed upper bound of 100%. The pension fund does not take short positions. In the baseline scenario the pension fund invests 35.6% of the collected contributions in the risky asset. In case of a short-sale constraint, it is optimal to decrease this share to 33.1%. As discussed in Section 3.3, agents are inclined to take short positions if there are funding deficits ($\pi_m > 0$) to ensure that the fraction of their risky asset holdings in total wealth will not change. However, when agents are not allowed to do this, a welfare-maximizing pension fund takes care of this by investing a smaller amount in equity. This investment strategy decreases the

Table 3.4: Equity investment and income-equivalent variation: effect short-sale constraint

	Baseline			No short selling		
	λ_h	λ_p	x	λ_h	λ_p	x
expectation	105.7	35.6	2.9	86.6	33.1	2.7
10th percentile	58.2	27.0		59.7	25.1	
median	95.9	34.9		96.6	32.5	
90th percentile	159.1	45.2		100.0	42.0	

Notes: results are based on 10,000 simulations. The equity share of individuals (λ_h) and the pension fund (λ_p) are expressed in percent of financial wealth and the income-equivalent variation (x) is expressed in percent of the wage rate.

probability of funding deficits, and hence, the probability that the short-sale constraint will bind an individual. The welfare consequence of the inability to take short positions is modest: the income-equivalent variation decreases with only 0.2%-points.

3.5.2 Labour income tax

In general, the labour-supply decision is determined by the total marginal tax burden which is not only affected by implicit taxes or subsidies in collective pension schemes but also by explicit labour income taxes. In the model analysed so far, we have ignored labour income taxation. This section investigates how the introduction of a funded DB pension scheme affects individual welfare if there is already an initial labour income tax in the economy. We consider two different cases: in the first case the government spends the tax revenues on services from which the consumer does not derive utility while in the second case tax revenues are redistributed back in the form of lump-sum income transfers. Hence, in the first case the labour income tax has a substitution and an income effect, in the second case it only has a substitution effect. We analyse the welfare implications for a labour income tax rate (τ) of 10%, 20% and 30% and for different values of the intratemporal substitution elasticity.

Table 3.5 shows the results. Consider first the case without lump-sum redistribution of the collected tax revenues, the left panel of the table. Note that in the benchmark parameterization with a substitution elasticity of unity the welfare gain is independent of labour income taxation: if $\rho = 1$ the income and substitution effects cancel against each other so that the labour-supply

Table 3.5: Income-equivalent variation: effect labour income taxation

		Without lump-sum transfer			With lump-sum transfer		
	$\tau = 0$	$\tau = 0.1$	$\tau = 0.2$	$\tau = 0.3$	$\tau = 0.1$	$\tau = 0.2$	$\tau = 0.3$
$1/\rho = 0.5$	2.28	2.18	2.08	1.98	2.27	2.25	2.23
$1/\rho = 1$	2.90	2.90	2.90	2.90	2.84	2.74	2.60
$1/\rho = 1.5$	3.24	3.35	3.47	3.62	3.07	2.85	2.58

Notes: results are based on 10,000 simulations. Income-equivalent variations (x) are expressed in percent of the wage rate.

decision is not affected.[16] For substitution elasticities below unity, the welfare gain of the funded pension scheme is lower if there is an initial tax distortion. In this case, the income effect dominates the substitution effect implying that labour supply in the economy is higher if $\tau > 0$ compared to $\tau = 0$. Consequently, if there is an initial tax distortion, the welfare-improving intergenerational transfers associated with the risk-sharing contract decrease as percentage of net disposable labour income, resulting in lower welfare gains. If the substitution elasticity is higher than unity, instead, the opposite holds. Then income taxation reduces labour-supply incentives which relatively increases the scope of welfare-improving intergenerational transfers.

Concentrating on the case in which the government redistributes the collected tax receipts using lump-sum payments, the right panel of Table 3.5, it follows that the welfare gain of the pension scheme is unambiguously lower if $\tau > 0$ compared to $\tau = 0$. Labour income taxation introduces an additional (negative) substitution effect on labour supply which reduces the risk-bearing capacity of individuals. The optimizing pension fund responds to this by reducing its equity exposure, resulting in a lower welfare gain.

3.5.3 Alternative discount rate

To impose the requirement that the pension scheme is fully funded, we have used the stochastic discount factor derived from the first-order conditions of the intertemporal optimization problem to value the pension benefit. The funding requirement, as given by equation (3.14), ensures that at the mar-

[16]This can also be proved formally: if $\rho = 1$, we have $\tilde{\lambda}_{p,t} = (1-\tau)\lambda_{p,t}$, where $\tilde{\lambda}_{p,t}$ denotes the optimal equity investment of the pension fund in case of an initial income tax rate τ. Substituting this expression in the condition for x_t, i.e., $\mathrm{E}_t[(1+x_t)w_{t+1}]^{1-\zeta} = \mathrm{E}_t[w_{t+1} + \pi_{f,t}\tilde{\lambda}_{p,t}w_t(1-\tau)^{-1}(r_{e,t+1} - r_f)]^{1-\zeta}$, the tax rate τ cancels out.

gin the young generation is indifferent about contributing or not contributing one additional unit to the pension fund. As the pension benefit is safe from the perspective of the young, the funding requirement implies that the risk-free return is the appropriate discount rate in the actuarially-fair contribution, i.e., recall $\pi_{f,t} = \alpha/(1+r_f)$. This contribution rate is exactly equal to the value young agents attach to the benefit and, hence, ensures that the pension scheme does not redistribute between generations *ex ante*.

In practice, though, pension funds often discount their liabilities using the return they expect to receive on their investments. If pension funds invest part of their wealth in high-yielding risky assets, this expected return will be higher than the risk-free rate of return. Following this strategy, our actuarially-fair contribution rate changes into $\pi_{f,t} = \alpha/\,\mathrm{E}_t(1+r_{p,t+1})$. Since $\mathrm{E}_t\, r_{p,t+1} > r_f$, this contribution rate is lower than the market price which is based on the risk-free return. In our model, any discount rate *higher* than the risk-free rate thus necessarily implies that the introduction of the pension scheme redistributes *ex ante*, i.e., from future generations to the first generation that enters the scheme. Indeed, the introduction of the scheme favours the first participating generation because these people receive pensions (when old) without having made enough contributions (when young). Like in a PAYG system, the burden of this windfall gain is shifted to all subsequent generations (in the form of higher mismatch contributions π_m). Given our default parameter values, discounting with the expected portfolio return decreases welfare of the steady-state generation from 2.9% to 1.3%. Nevertheless, since the transitory generation experience a welfare gain, the introduction of the scheme still involves an *ex ante* Pareto improvement. In fact, discounting with the expected return makes it possible to strictly improve welfare of all generations, also that of the first contributory generation.

3.5.4 Cap on contribution rate

At an earlier stage, we concluded that a risk-sharing contract is only feasible if participation is mandatory. However, even with mandatory participation high mismatch taxes rates imposed on the labour income of the young generation can still undermine the viability of the pension scheme. People can decide to reduce the number of hours working or to seek employment elsewhere, in another company for instance (in case of a company pension fund) or in another business sector (in case of an industry-wide pension fund). Also, the possibility to become a self-employed or moving to a foreign country should

Table 3.6: Income-equivalent variation: effect of cap on mismatch contribution rate

	Baseline	$-\pi_m^{max} \leq \pi_m \leq +\pi_m^{max}$		
		$\pi_m^{max} = 0.2$	$\pi_m^{max} = 0.15$	$\pi_m^{max} = 0.1$
x	2.9%	2.0%	1.7%	1.3%
n	0.057	0.038	0.033	0.026
$\lambda_h s$	0.056	0.051	0.050	0.049

Note: Income-equivalent variations (x) are expressed in percent of the wage rate.

be kept in mind. In the opposite case, if the fund is overfunded, there are also discontinuity risks. The older generations, for example, can try via the political process to share in the investment gains of the pension fund.

To avoid or at least reduce these discontinuity problems, the pension can decide to restrict intergenerational transfers by putting a cap on the mismatch rate. As shown by Figure 3.2, without such a cap the mismatch tax or subsidy can constitute a large part of labour income. The 95%-confidence interval ranges between a subsidy of 67% and a tax of 13% of labour income. Although a contribution cap might reduce the risk of discontinuity in the pension scheme, it reduces the *ex ante* welfare gain of risk sharing. Tabel 3.6 shows the welfare implications if the pension fund introduces a contribution cap (π_m^{max}) at, respectively, 20%, 15% and 10%.[17] As shown, the benefits of risk sharing decrease when the cap is set more tight, although the pension scheme still offers a gain of 1.3% if the maximum and minimum cap amounts to 10%. A lower cap implies that more (tail) risk is shifted from the young to the elderly. As the old are less equipped to bear this risk, the pension fund reduces its risk exposure (n). Moreover, as agents get a larger stake in equity through the pension system in the second period of life, they lower their own risky investments to some extent ($\lambda_h s$).

3.5.5 First-best risk sharing

Up to now, we have assumed that the pension fund does not share funding shortfalls or surpluses over more than two overlapping generations. Of

[17]To reduce complexity, labour supply is exogenous in these simulations and equal to the baseline level. Recall from equation (3.35) that labour supply is deterministic in the baseline because income and substitution effects cancel. Hence, exogenous labour supply will probably not that restrictive in this case.

course, this is a realistic point of departure because in practice pension funds are often legally restricted in shifting risk to the future. Restricting the horizon of risk sharing is also defensible from a continuity perspective. Contribution payers or sponsoring companies could withdraw from the system if they are exposed to huge funding surpluses.

Even if there are good reasons for it, the question how much welfare potential is given up by restricting the scope of risk sharing is still relevant. To answer this question, we expand the instrument set of the pension fund in this section. We impose that the pension fund is no longer restricted to the per-period solvency constraint, equation (3.16). That means, the pension fund does not only decide upon the optimal portfolio allocation of pension contributions, it also decides upon the optimal level of the contribution rate itself. To solve this problem, we make three simplifying assumptions: first, wages are assumed to be non-stochastic ($\sigma_w^2 = 0$); second, the felicity function has a Cobb-Douglas specification ($\rho = 1$) which implies that labour supply is deterministic and, finally, life-time preferences are represented by a CRRA utility function ($\theta = \gamma$). If these assumption are satisfied, we know from Gollier (2008) that the problem has a well-defined analytical solution.

If the pension fund is no longer obliged to balance the budget each period, it can also involve human capital of future generations into risk sharing. Total wealth of the pension fund then consists of the sum of financial wealth and the discounted value of the claim on labour earnings of all future generations, denoted H and which is equal to:

$$\begin{aligned} H &= \sum_{t=0}^{\infty}(1+r_f)^{-t}(1-l)w \\ &= \frac{(1+r_f)(1-l)w}{r_f} \end{aligned} \tag{3.53}$$

Consider now the decision problem of the pension fund at the beginning of period $t = 0$. The pension fund then maximizes the discounted sum of expected utilities. We look for the dynamic policy of the pension fund in terms of the contribution rate π_m and the portfolio allocation n that maximizes welfare of all generations:

$$\max_{n,w_T} W_t = \mathrm{E}\sum_{t=0}^{\infty}\beta^t V(w_{T,t})^{1-\theta} \tag{3.54}$$

subject to:

$$\begin{aligned} A_{t+1} &= (1+r_f)\left[A_t + (1-l)(w - w_{T,t})\right] + n_t(r_{e,t+1} - r_f) \\ A_t + H &\geq 0,\ \forall t \geq 1 \end{aligned}$$

and some initial wealth position A_0. Recall that the indirect utility function of households $V(\cdot)$ is already defined by equation (3.39). For analytical convenience, we have written problem (3.54) in terms of full life-time income $w_T \equiv (1-\pi_m)w$. In addition, without loss of generality, we have put the replacement rate of the defined-benefit scheme to zero (i.e., $\alpha = 0$).[18]

Proposition 3.3 *The first-best contribution rate and portfolio asset allocation are determined in such a way that*

$$(1-l)w_T^{fb} = q(A+H) \tag{3.55}$$

$$n^{fb} = d(A+H) \tag{3.56}$$

with,

$$\begin{aligned} q &\equiv 1 - \left\{\beta \, \mathrm{E}\left[1 + r_T(a^*)\right]^{1-\zeta}\right\}^{1/\zeta} \\ d &\equiv (1-q)a^* \end{aligned}$$

where a^* is the unique solution to equation (3.26). *If all future generations receive the same expected utility, this welfare gain equals:*

$$1 + x^{fb} = \frac{q(A_0 + H)}{(1-l)w} \tag{3.57}$$

Proof See Appendix 3.B.3. ∎

The optimal strategy of the pension fund is characterized by two decision rules that determine the degree of risk taking and the way these risk will be allocated over generations. Equation (3.55) shows that the pension fund pins down the contribution rate π_m in such a way that life-time income of agents is proportional to total wealth. The same holds for the money investment in the risky asset, see equation (3.56).

Notice that the *ex ante* welfare gain of the pension scheme depends on the initial wealth position of the fund. The young enters the model with zero financial capital, hence $A_0 = 0$ in our case. Given the default parameter choices,

[18]Alternatively, we can also assume $\alpha > 0$. As long as the initial wealth position A_0 is sufficient to cover the defined benefit to the old generation at $t = 0$ both approaches are equivalent.

it turns out that the fund destributes each period a constant share of $q = 37.1\%$ of the aggregate fund's wealth to the working young generation. In addition, the ratio between the total pension wealth and gross labour income, $\frac{A_0+H}{(1-l)w}$, is equal to 3.06. Therefore, the first-best welfare gain is $x^{fb} = 0.371 * 3.06 - 1 = 13.5\%$. The corresponding welfare gain if the pension fund does not share risks among more than two overlapping generations is 7.0%. This implies that extending the risk sharing horizon maximally doubles the welfare gain.

This result has to be interpreted with some caution though. Recall that we have abstracted from wage risk in this section. We already know from previous sections that there is empirical evidence that shows that stock and labour markets move together in the long run. If this is indeed the case, then it becomes less attractive for future generations to share in current risks, because they are already affected by current risks via their human wealth. Hence, in the presence of labour income risk, there could be limited scope to shift risks into the future (Mehlkopf, 2011).

3.6 Concluding remarks

In this paper, we have investigated the welfare aspects of intergenerational risk sharing in the context of funded pensions with defined benefits. We have developed a stylized two-period overlapping-generations model in which generations cannot trade risks privately, thereby creating a rationale for pension funds. The analysis allows for endogenous labour supply (in relation to income-related contributions) and for a positive correlation between stock returns and wages, two factors which *a priori* disadvantage the added value of intergenerational risk sharing.

We show that a properly defined funded pension scheme is welfare improving, even if pension contributions are distortionary and even if capital and labour markets are subject to correlated risks. In this scheme, efficient risk sharing implies that the pension fund invests a constant fraction of the human capital of the young in equity, in accordance with household preferences for risk. With this investment strategy, the fund controls both the diversification gains from risk sharing and the welfare losses due to labour-supply distortions, thereby preventing that the gains are dominated by the losses. In addition, we show that efficient risk sharing increases the risk-bearing capacity of the economy and, hence, shifts the portfolio choice towards risky assets. It also shifts the labour-leisure choice towards labour supply as individuals can only exploit the positive equity premium if they work. As a result, efficient risk sharing shows up in an increase in expected consumption rather than a decrease in volatility. Finally, our analysis contributes to the understanding of the role of labour-supply flexibility and portfolio choice. Endogenous labour supply makes it less attractive to use human capital of the young as collateral when investing in the capital market and, hence, reduces risk-sharing possibilities.

There are reasons to presume that the welfare gains from intergenerational risk sharing in this chapter serve as a lower bound. First, our analysis only focuses on the sharing of capital market risks. There are other sources of risk sharing that can further strengthen the welfare-enhancing role of collective funded pensions as compared to individual defined-contribution plans. Most importantly, pension benefits can be linked to current wages enabling older people to acquire a claim on the human capital of the young. In this way, the pension scheme alleviates another market inefficiency related to the non-tradability of human capital.

Second, due to the two-cohort structure, the model used in this chapter is only appropriate to analyse *intertemporal* risk sharing between generations that are not both alive prior to the occurrence of a shock. This model cannot be used for *contemporaneous* risk sharing which involves generations that are already alive before a shock is realized. This type of risk sharing calls for a multi-cohort model. In reality, there are risk factors (like demographic risk, inflation risk or wage risk) which in principle can be traded between contemporaneous generations but the markets for these risks do not easily arise for reasons of moral hazard and adverse selection. Instead of using private trade as an instrument for intergenerational risk sharing, this kind of risk sharing is often organized via pension institutions. An important question left for future research is how the welfare gains of contemporaneous risk sharing relate to the gains of intertemporal risk sharing.

The results of this paper are relevant for the debate on pension reform. In the past fifteen years, at least half the OECD countries have undergone pension reforms that all point into the direction of individual defined contribution (OECD, 2009). A number of OECD countries (such as Hungary, Mexico and the Slovak Republic) introduced funded DC plans as a substitute for part of the public pension scheme. Other countries (such as the US, the UK and Ireland) replaced their funded DB plans with funded DC plans. There are also countries (like Germany, Denmark, Italy) that reduced the generosity of public pensions, for example, by increasing the pension eligibility age or applying a less generous indexation base. Making pensions less generous implies that individuals have to rely more on private savings if they want to maintain their pensions at their original levels. As private savings are similar to savings in an individual DC scheme, reducing the generosity of collective pensions in fact implies a switch towards defined contribution. Our results emphasize that this development is not optimal with regard to the scope of intergenerational risk sharing. This paper argues that collective funded pension schemes with well-structured risk sharing are better candidates to introduce more funding than individual pension schemes that do not share risks among generations, even if we allow for the welfare losses from labour market distortions induced by these collective schemes.

Appendix to Chapter 3

In this appendix, we derive the analytical solution of the consumer and pension fund problem in case of Cobb-Douglas utility (Section 3.A). We also provide proofs of the propositions (Section 3.B) and derive the labour-supply elasticity (Section 3.C). Throughout, we use the notation $\hat{r} \equiv \log(1+r)$ for returns and $\hat{y}_t \equiv \log y_t$ for any other variable y_t.

3.A Consumer and pension fund optimization

3.A.1 Portfolio allocation consumer

Following Campbell and Viceira (2002), we assume that the joint distribution of consumption and returns is lognormal. If a variable y_t is lognormally distributed, then there holds:

$$\log \mathrm{E}_t\, y_{t+1} = \mathrm{E}_t\, \hat{y}_{t+1} + \frac{1}{2} \mathrm{Var}_t\, \hat{y}_{t+1} \tag{3.58}$$

Taking logs of equation (3.26) and using equation (3.58), we obtain:

$$\begin{aligned} \mathrm{E}_t(-\zeta \hat{r}_{T,t+1} + \hat{r}_{e,t+1}) + \frac{1}{2} \mathrm{Var}_t(-\zeta \hat{r}_{T,t+1} + \hat{r}_{e,t+1}) = \\ \mathrm{E}_t(-\zeta \hat{r}_{T,t+1} + \hat{r}_f) + \frac{1}{2} \mathrm{Var}_t(-\zeta \hat{r}_{T,t+1} + \hat{r}_f) \end{aligned} \tag{3.59}$$

Simplifying this expression gives:

$$\mathrm{E}_t\, \hat{r}_{e,t+1} + \frac{1}{2} \mathrm{Var}_t\, \hat{r}_{e,t+1} - \hat{r}_f = \zeta\, \mathrm{Cov}_t(\hat{r}_{T,t+1}, \hat{r}_{e,t+1}) \tag{3.60}$$

As the return on the portfolio is a linear combination of the return on stocks and the return on bonds, see equation (3.5), and the log of a linear combination is not the same as a linear combination of logs, we follow Campbell and Viceira (2002) and use a Taylor approximation of the non-linear function relating log individual-asset returns to log portfolio returns. First note that equation (3.24) can be written as:

$$1 + r_{T,t+1} = 1 + r_f + a_t \left[(1 + r_{e,t+1}) - (1 - r_f) \right] \tag{3.61}$$

Dividing this expression by $1 + r_f$ and then taking logs gives:

$$\hat{r}_{T,t+1} - \hat{r}_f = \underbrace{\log \left\{ 1 + a_t \left[\exp(\hat{r}_{e,t+1} - \hat{r}_f) - 1 \right] \right\}}_{f(\hat{r}_{e,t+1} - \hat{r}_f)} \tag{3.62}$$

Now we take a second-order Taylor expansion of $f(\cdot)$ around $\hat{r}_{e,t+1} - \hat{r}_f = 0$, which gives:[19]

$$\hat{r}_{T,t+1} \approx \hat{r}_f + a_t(\hat{r}_{e,t+1} - \hat{r}_f) + \frac{1}{2} a_t (1 - a_t) \operatorname{Var}_t \hat{r}_{e,t+1} \tag{3.63}$$

From equation (3.63) it follows:

$$\operatorname{Cov}_t(\hat{r}_{T,t+1}, \hat{r}_{e,t+1}) = a_t \operatorname{Var}_t \hat{r}_{e,t+1} \tag{3.64}$$

Substituting equation (3.64) into (3.60) then gives:

$$a_t = \frac{\mathrm{E}_t \, \hat{r}_{e,t+1} - \hat{r}_f + \frac{1}{2} \operatorname{Var}_t \hat{r}_{e,t+1}}{\zeta \operatorname{Var}_t \hat{r}_{e,t+1}} \tag{3.65}$$

Recall our statistical assumptions:

$$\mathrm{E}_t(\hat{r}_{e,t+1} - \hat{r}_f) = \mu_r \tag{3.66}$$

$$\operatorname{Var}_t \hat{r}_{e,t+1} = \sigma_r^2 \tag{3.67}$$

Note from equation (3.58) that:

$$\bar{\mu}_r \equiv \log \mathrm{E}_t \left(\frac{1 + r_{e,t+1}}{1 + r_f} \right) = \mu_r + \frac{1}{2} \sigma_r^2 \tag{3.68}$$

Inserting equations (3.66)-(3.68) in equation (3.65) gives equation (3.27).

[19] This approximation holds exactly in continuous time and turns out to be still satisfactory for longer holding periods (see Barberis, 2000).

3.A.2 Certainty-equivalent rate of return

Taking logs of equation (3.30) and using (3.58), we obtain:

$$\hat{r}_{ce,t} = \frac{1+\nu}{1-\varphi}\Big[\underbrace{(1-\zeta)\,\mathrm{E}_t\,\hat{r}_{T,t+1} + \frac{1}{2}(1-\zeta)^2\,\mathrm{Var}_t\,\hat{r}_{T,t+1}}_{g_t}\Big] \tag{3.69}$$

Using equation (3.63), the term g_t can be rewritten to:

$$\begin{aligned} g_t = {} & (1-\zeta)\hat{r}_f + (1-\zeta)a_t\left(\mathrm{E}_t\,\hat{r}_{e,t+1} - \hat{r}_f\right) \\ & + \frac{1}{2}(1-\zeta)a_t(1-a_t)\,\mathrm{Var}_t\,\hat{r}_{e,t+1} + \frac{1}{2}(1-\zeta)^2 a_t^2\,\mathrm{Var}_t\,\hat{r}_{e,t+1} \end{aligned} \tag{3.70}$$

Inserting equation (3.65) into (3.70) and rearranging gives:

$$g_t = (1-\zeta)\hat{r}_f + \frac{1-\zeta}{2\zeta}\frac{\left(\mathrm{E}_t\,\hat{r}_{e,t+1} - \hat{r}_f + \frac{1}{2}\,\mathrm{Var}_t\,\hat{r}_{e,t+1}\right)^2}{\mathrm{Var}_t\,\hat{r}_{e,t+1}} \tag{3.71}$$

Using equations (3.66)-(3.67) together with equation (3.68), we obtain:

$$g_t = (1-\zeta)\left(\hat{r}_f + \frac{1}{2}\frac{\bar{\mu}_r^2}{\zeta\sigma_r^2}\right) \tag{3.72}$$

Inserting equation (3.72) in equation (3.69), and using the fact that $\log(1+y) \approx y$ (for small y), we obtain equation (3.31).

3.A.3 Optimal policy pension fund

Using the property that labour supply is deterministic in case of Cobb-Douglas preferences, we can write life-time income as:

$$w_{T,t+1} \equiv (1-\pi_{m,t+1})w_{t+1} = w_{t+1} + \pi_{f,t}\lambda_{p,t}w_t(r_{e,t+1} - r_f) \tag{3.73}$$

Assume now that $w_{T,t+1}$ is lognormally distributed. Then taking logs on both sides of equation (3.40) and rearranging terms, we obtain:

$$\bar{\mu}_r = \zeta\,\mathrm{Cov}_t(\hat{w}_{T,t+1}, \hat{r}_{e,t+1}) \tag{3.74}$$

Dividing equation (3.73) by $1 + r_f$ and taking logs gives:

$$\hat{w}_{T,t+1} - \hat{r}_f = \underbrace{\log\left\{\exp(\hat{w}_{t+1} - \hat{r}_f) + \pi_{f,t}\lambda_{p,t}w_t\left[\exp(\hat{r}_{e,t+1} - \hat{r}_f) - 1\right]\right\}}_{f(\hat{w}_{t+1}, \hat{r}_{e,t+1})} \tag{3.75}$$

Now we take a second-order Taylor expansion of $f(\cdot)$ around $\hat{w}_{t+1} - \hat{r}_f = 0$ and $\hat{r}_{e,t+1} - \hat{r}_f = 0$, which gives:

$$\begin{aligned}\hat{w}_{T,t+1} \approx\ & \pi_{f,t}\lambda_{p,t}w_t(\hat{r}_{e,t+1} - \hat{r}_f) + \frac{1}{2}\pi_{f,t}\lambda_{p,t}w_t(1 - \pi_{f,t}\lambda_{p,t}w_t)\sigma_r^2 + \\ & \hat{w}_{t+1} - \pi_{f,t}\lambda_{p,t}w_t\sigma_{wr}\end{aligned} \tag{3.76}$$

From equation (3.76), it follows:

$$\mathrm{Cov}_t(\hat{w}_{T,t+1}, \hat{r}_{e,t+1}) = \pi_{f,t}\lambda_{p,t}w_t\sigma_r^2 + \sigma_{wr} \tag{3.77}$$

Substituting equation (3.77) in equation (3.74), we obtain:

$$\pi_{f,t}\lambda_{p,t}w_t = \frac{\bar{\mu}_r - \zeta\sigma_{wr}}{\zeta\sigma_r^2} \tag{3.78}$$

From this equation together with definition $n_t \equiv \pi_{f,t}\lambda_{p,t}(1 - l)w_t$, we obtain equation (3.41).

3.B Proofs

3.B.1 Proof of Proposition 3.1

Proof Substituting the indirect utility function (3.39) in the objective function of the pension fund, equation (3.18), gives the following expression for $\mathcal{W}_t^{DB}$:

$$\mathcal{W}_t^{DB} = \frac{(1+z)^{\frac{1}{1-\gamma}}\eta^\eta(1-\eta)^{1-\eta}}{1+(1-\eta)z}\left(\mathrm{E}_t\, w_{T,t+1}^{1-\zeta}\right)^{\frac{1}{1-\theta}} \tag{3.79}$$

Note that $\mathcal{W}_t^{DC}$ is simply equal to:

$$\mathcal{W}_t^{DC} = \frac{(1+z)^{\frac{1}{1-\gamma}}\eta^\eta(1-\eta)^{1-\eta}}{1+(1-\eta)z}\left(\mathrm{E}_t\, w_{t+1}^{1-\zeta}\right)^{\frac{1}{1-\theta}} \tag{3.80}$$

Then equation (3.43) implies:

$$\mathrm{E}_t\left[(1+x_t)w_{t+1}\right]^{1-\zeta} = \mathrm{E}_t\, w_{T,t+1}^{1-\zeta} \tag{3.81}$$

Taking logs on both sides of equation (3.81) gives:

$$\log(1+x_t) + \mu_w + \frac{1}{2}(1-\zeta)\sigma_w^2 = \mathrm{E}_t\,\hat{w}_{T,t+1} + \frac{1}{2}(1-\zeta)\,\mathrm{Var}_t\,\hat{w}_{T,t+1} \tag{3.82}$$

From equation (3.76) we know that:

$$\mathrm{E}_t\, \hat{w}_{T,t+1} = \mu_w + \alpha\lambda_{p,t} w_t(\bar{\mu}_r - \sigma_{wr}) - \frac{1}{2}(\alpha\lambda_{p,t} w_t)^2\sigma_r^2 \tag{3.83}$$

$$\mathrm{Var}_t\, \hat{w}_{T,t+1} = \sigma_w^2 + (\alpha\lambda_{p,t} w_t)^2\sigma_r^2 + 2\alpha\lambda_{p,t} w_t\sigma_{wr} \tag{3.84}$$

Substituting equations (3.83) and (3.84) into equation (3.82), we find:

$$\log(1 + x_t) = \alpha\lambda_{p,t} w_t(\bar{\mu}_r - \zeta\sigma_{wr}) - \frac{1}{2}\zeta(\alpha\lambda_{p,t} w_t)^2\sigma_r^2 \tag{3.85}$$

Inserting equation (3.78) in this expression, we ultimately have:

$$\log(1 + x_t) = \frac{1}{2\zeta}\left(\frac{\bar{\mu}_r - \zeta\sigma_{wr}}{\sigma_r}\right)^2 \tag{3.86}$$

Using that $\log(1 + y) \approx y$, we obtain equation (3.44). ∎

3.B.2 Proof of Proposition 3.2

Proof With lump-sum transfers, life-time income is defined as:

$$w_{T,t} = w_t + n_{t-1}(r_{e,t} - r_f) \tag{3.87}$$

Similar to equation (3.23), equation (3.87) can be approximated by:

$$\hat{w}_{T,t+1} \approx \hat{w}_{t+1} + n_t(\hat{r}_{e,t+1} - \hat{r}_f) + \frac{1}{2}n_t(1 - n_t)\sigma_r^2 - n_t\sigma_{wr} \tag{3.88}$$

Suppose that life-time utility is log linear (i.e., $\gamma = \theta = \rho = 1$). In that case, we are able to solve the model with lump-sum transfers analytically. With log-linear utility, it follows from equation (3.29) that $z = \beta$ and, hence, does not depend on leisure anymore. Life-time utility equals,

$$U_t = (1 - \eta)\hat{c}_{y,t} + \eta\hat{l}_t + \beta(1 - \eta)\,\mathrm{E}_t\,\hat{c}_{o,t+1} \tag{3.89}$$

Substituting equations (3.48)-(3.50) in this expression gives the following indirect utility function:

$$V_t = \Phi - \eta\hat{w}_t + [1 + \beta(1 - \eta)]\,\hat{w}_{T,t} \tag{3.90}$$

with:

$$\begin{aligned}\Phi \equiv\ & (1 - \eta)(1 + \beta)\log(1 - \eta) + \eta\log\eta + \beta(1 - \eta)\log\beta \\ & + \beta(1 - \eta)\,\mathrm{E}\,\hat{r}_T - [1 + \beta(1 - \eta)]\log[1 + \beta(1 - \eta)]\end{aligned}$$

Notice that Φ is a constant term. Social welfare is maximized if:

$$\frac{\partial W_t}{\partial n_t} = \frac{\partial \, \mathrm{E}_t \, V_{t+1}}{\partial n_t} = 0 \quad \Rightarrow \quad \mathrm{E}_t \left[w_{T,t+1}^{-1} (r_{e,t+1} - r_f) \right] = 0 \tag{3.91}$$

Taking logs on both sides of this expression, ultimately gives:

$$\bar{\mu}_r = \mathrm{Cov}_t(\hat{w}_{T,t+1}, \hat{r}_{e,t+1}) \tag{3.92}$$

Using equation (3.88), we have:

$$\mathrm{Cov}_t(\hat{w}_{T,t+1}, \hat{r}_{e,t+1}) = n_t \sigma_r^2 + \sigma_{wr} \tag{3.93}$$

Inserting equation (3.93) in equation (3.92), we obtain equation (3.51).

With lump-sum transfers, the *ex ante* utility functions in case there is no collective pension scheme $(\mathcal{W}^{DC})$ and in case there is a pension scheme $(\mathcal{W}^{DB})$ are, respectively:

$$\mathcal{W}_t^{DC} = (1-\eta)(1+\beta)\mu_w + \Phi \tag{3.94}$$

$$\mathcal{W}_t^{DB} = [1 + \beta(1-\eta)] \, \mathrm{E}_t \, \hat{w}_{T,t+1} - \eta \mu_w + \Phi \tag{3.95}$$

Equation (3.43) then implies:

$$\begin{aligned} [1 + \beta(1-\eta)] \, \mathrm{E}_t \, \hat{w}_{T,t+1} = {} & (1+\beta)(1-\eta) \log(1+x_t) \\ & + [1 + \beta(1-\eta)] \, \mu_w \end{aligned} \tag{3.96}$$

Combining equation (3.41) and equation (3.88), we have that:

$$\mathrm{E}_t \, \hat{w}_{T,t+1} = \frac{1}{2} \left(\frac{\bar{\mu}_r - \sigma_{wr}}{\sigma_r^2} \right) + \mu_w \tag{3.97}$$

Inserting this expression in equation (3.96) gives:

$$\log(1+x_t) = \frac{1}{2} \left(\frac{\bar{\mu}_r - \sigma_{wr}}{\sigma_r^2} \right) \left[1 + \frac{\eta}{(1-\eta)(1+\beta)} \right] \tag{3.98}$$

Again using the approximation $\log(1+y) \approx y$, we obtain the welfare gain with lump-sum transfers, equation (3.52). ∎

3.B.3 Proof of Proposition 3.3

Proof To save on notation, we first define the following variables:[20]

$$p \equiv (1-l) w_T \tag{3.99}$$

$$k \equiv (1-l) w \tag{3.100}$$

[20]This proof is inspired by Gollier (2008).

Further, rewrite the indirect utility function of the household as follows:

$$V = \Omega p^{1-\eta} \tag{3.101}$$

with,

$$\Omega \equiv \frac{(1+z)^{\frac{1}{1-\gamma}} \eta^{\eta}(1-\eta)^{1-\eta}(1-l)^{\eta-1}}{(1-\eta)z+1}$$

Now we have the following Bellman equation of problem (3.54):

$$v(A) = \max_{p,n} \; V(p)^{1-\theta} + \beta \,\mathrm{E}\, v\left[(1+r_f)(A-p+k) + n(r_e - r_f)\right] \tag{3.102}$$

Suppose the value function has solution $v(A) = j\Omega(A+H)^{1-\zeta}$, for some positive scalar j. This implies,

$$\begin{aligned} & v\left[(1+r_f)(A-p+k) + n(r_e-r_f)\right] = \\ & j\Omega\left[(1+r_f)(A-p+H) + n(r_e-r_f)\right]^{1-\zeta} \end{aligned} \tag{3.103}$$

where we have used equation (3.53). The first-order conditions are equal to:

$$p^{-\zeta} = j\beta(1+r_f)\,\mathrm{E}\left[(1+r_f)(A-p+H) + n(r_e-r_f)\right]^{-\zeta} \tag{3.104}$$

$$0 = \mathrm{E}\left\{\left[(1+r_f)(A-p+H) + n(r_e-r_f)\right]^{-\zeta}(r_e-r_f)\right\} \tag{3.105}$$

We can infer that $p(A) = \hat{q}(A+H)$ and $n(A) = (1-\hat{q})a^*(A+H)$ are solutions, with $\hat{q} \equiv \left(hj^{1/\zeta}+1\right)^{-1}$ and $h \equiv \left\{\beta\,\mathrm{E}\left[1+r_T(a^*)\right]^{1-\zeta}\right\}^{1/\zeta}$. We can retrace j by substituting the solutions in the Bellman equation, which gives:

$$j = \hat{q}^{1-\zeta} + j\beta(1-\hat{q})^{1-\zeta}\,\mathrm{E}\left[1+r_T(a^*)\right]^{1-\zeta} \tag{3.106}$$

Using the expressions for $\hat{q}$ and h, we can rewrite the expression for j as:

$$\begin{aligned} j &= \left(hj^{1/\zeta}+1\right)^{\zeta-1} + jh^{\zeta}\left(\frac{hj^{1/\zeta}}{hj^{1/\zeta}+1}\right)^{1-\zeta} \\ &= \left(hj^{1/\zeta}+1\right)^{\zeta-1}\left(hj^{1/\zeta}+1\right) = \left(hj^{1/\zeta}+1\right)^{\zeta} \end{aligned} \tag{3.107}$$

Hence, $j = q^{-\zeta}$ with $q = \hat{q}$.

Inserting solutions (3.55) and (3.56) together with equation (3.53) in the budget constraint, gives:

$$A_{t+1} + H = (1-q)\left[1+r_{T,t+1}(a^*)\right](A_t + H) \tag{3.108}$$

Because shocks are assumed to be serially independent, income p_t at any time $t > 0$ equals:

$$p_t = q(A_0 + H)(1-q)^t \prod_{i=1}^{t} [1 + r_{T,i+1}(a^*)] \tag{3.109}$$

Therefore, the expected utility of a generation entering the economy at t years after the introduction is equal to:

$$\mathrm{E}\left[V(p_t)^{1-\theta} \middle| A_0 \right] = q^{1-\zeta}\Omega(A_0 + H)^{1-\zeta}\Psi^t \tag{3.110}$$

with,

$$\Psi \equiv (1-q)^{1-\zeta}\, \mathrm{E}\,[1 + r_T(a^*)]^{1-\zeta}$$

Note that expected utility of generations increases over time if $\Psi > 1$. Following Gollier (2008), we impose that expected utility is the same for each generation, i.e., $\Psi = 1$. For this to be the case, the discount factor should be equal to:

$$\beta = \left\{ \mathrm{E}\,[1 + r_T(a^*)]^{1-\zeta} \right\}^{\frac{1}{\zeta-1}} \tag{3.111}$$

Then equation (3.43) implies:

$$q^{1-\zeta}\Omega(A_0 + H)^{1-\zeta} = \Omega(1-l)^{1-\zeta}W^{1-\zeta} \;\Rightarrow\; 1 + x = \frac{q(A_0 + H)}{(1-l)W} \tag{3.112}$$

This completes the proof. ∎

3.C Labour-supply elasticity

This section derives the uncompensated labour-supply elasticity ϵ. To that end, we first need to approximate $u_{y,t}$ and $u_{o,t+1}$. Taking logs of equation (3.2) for the first period gives:

$$\hat{u}_{y,t} = \frac{1}{1-\rho} \underbrace{\log\left\{ \exp\left[\log(1-\eta) + (1-\rho)\hat{c}_{y,t}\right] + \exp\left[\log\eta + (1-\rho)\hat{l}_t\right] \right\}}_{f(c_{y,t},l_t)} \tag{3.113}$$

Approximating $f(\cdot)$ with a first-order Taylor expansion around $(\hat{c}_{y,t}, \hat{l}_t) = (0, 0)$ gives:

$$f(\hat{c}_{y,t}, \hat{l}_t) \approx (1-\rho)\left[(1-\eta)\hat{c}_{y,t} + \eta\hat{l}_t\right] \tag{3.114}$$

So, we have:

$$u_{y,t} \approx c_{y,t}^{1-\eta} l_t^{\eta} \tag{3.115}$$

Along the same lines, we can approximate equation (3.2) for the second period, resulting in:

$$u_{o,t+1} \approx c_{o,t+1}^{1-\eta} \tag{3.116}$$

Substituting equation (3.115) and (3.116) in (3.6), gives:

$$c_{y,t} = \frac{(1-l_t)w_{T,t}}{1+z_t} \tag{3.117}$$

In case of CES utility, z_t is defined as:

$$z_t = \left\{\beta(1+r_f)l_t^{\omega-1}\,\mathrm{E}_t(1+r_{T,t+1})^{-\zeta}\left[\mathrm{E}_t(1+r_{T,t+1})^{(1-\eta)(1-\theta)}\right]^{\nu}\right\}^{\frac{1}{\varphi}} \tag{3.118}$$

with $\varphi \equiv \rho-(1-\eta)(\rho-\gamma)$, $\zeta \equiv \rho-(1-\eta)(\rho-\theta)$, $\omega \equiv 1-\eta(\rho-\gamma)$ and ν and r_T as defined in the main text. Dividing equation (3.7) by equation (3.6), we obtain:

$$\frac{c_{y,t}}{l_t} = \left(\frac{1-\eta}{\eta}\,w_{T,t}\right)^{\frac{1}{\rho}} \tag{3.119}$$

Using this condition, we have:

$$c_{y,t} = \frac{\left(\frac{1-\eta}{\eta}\,w_{T,t}\right)^{\frac{1}{\rho}}}{(1+z_t)\left(\frac{1-\eta}{\eta}\right)^{\frac{1}{\rho}}w_{T,t}^{\frac{1}{\rho}-1}+1} \tag{3.120}$$

$$l_t = \frac{1}{(1+z_t)\left(\frac{1-\eta}{\eta}\right)^{\frac{1}{\rho}}w_{T,t}^{\frac{1}{\rho}-1}+1} \tag{3.121}$$

From equation (3.121) we derive the labour-supply elasticity ϵ, which is equal to:

$$\epsilon \equiv \frac{\mathrm{d}(1-l_t)}{\mathrm{d}w_{T,t}}\frac{w_{T,t}}{1-l_t} = \frac{l_t}{1-l_t}\frac{1-\rho}{\rho}\frac{1+z_t}{1+\frac{\gamma}{\varphi}z_t+\left(\frac{\eta}{1-\eta}\right)^{\frac{1}{\rho}}w_{T,t}^{1-\frac{1}{\rho}}} \tag{3.122}$$

so we have that $\epsilon > 0$ if $\rho < 1$, $\epsilon < 0$ if $\rho > 1$ and $\epsilon = 0$ if $\rho = 1$.

CHAPTER 4

Retirement flexibility and portfolio choice

This chapter explores the interaction between retirement flexibility and portfolio choice in an overlapping-generations model. We analyse this interaction both in a partial equilibrium and general equilibrium setting. Retirement flexibility is often seen as a hedge against capital market risks which justifies more risky asset portfolios. We show, however, that this positive relationship between risk taking and retirement flexibility is weakened and under some conditions even turned around if not only capital market risks but also productivity risks are considered. Productivity risk in combination with a high elasticity of substitution between consumption and leisure creates a positive correlation between asset returns and labour income, reducing the willingness of consumers to bear risk. Moreover, it turns out that general equilibrium effects can either increase or decrease the equity exposure, depending on the degree of substitutability between consumption and leisure.

4.1 Introduction

In Western countries, pension schemes typically move from contracts with high implicit taxation and predominantly inflexible payout periods towards more actuarially-neutral arrangements with flexible payout periods (OECD, 2007). This move to flexible pension schemes is partly forced by population ageing and the financial crisis which put the traditional social security schemes under financial pressure. Another important factor is the ongoing process of individualization and the resulting acknowledgement that individuals differ in their tastes for leisure, earnings capacities, wealth positions, and therefore have different preferences for retirement. As such, flexible pension schemes have contributed to create new opportunities in which individuals can optimally choose their retirement age (Van Vuuren, 2011).

In this chapter, we raise the question how this increasing importance of retirement flexibility will affect consumption and portfolio decisions during working life. As stressed in the literature, the important advantage of retirement flexibility is that it provides insurance against all types of risks, like disability risk (Diamond and Mirrlees, 1978) or stock market risk (Pestieau and Possen, 2010). It gives individuals the ability to adjust working life to their own preferences and to avoid abrupt changes in consumption. Viewed in this way, flexible retirement serves as a hedge against adverse investment outcomes which allows for more risk taking in pension assets (Bodie et al., 1992). The basic mechanism behind this result is the negative correlation between asset returns and labour income due to wealth effects in the retirement decision. Indeed, a negative wealth shock causes marginal utility from leisure to decrease and hence agents increase labour supply which, in turn, raises labour income. The aim of this chapter is to investigate the robustness of this result if we generalize the analysis by incorporating more risk factors, general equilibrium effects and more general preferences.

The number of studies that focus on the interaction between portfolio, consumption and retirement decisions is rather limited. Starting point is the paper of Bodie et al. (1992) which analyses this interaction assuming that labour supply can be adjusted continuously. Subsequent studies, like Lachance (2004), Choi and Shim (2006), Farhi and Panageas (2007) and Choi et al. (2008), model optimal retirement as a discretionary stopping problem. Although all these studies differ in many respects, they have in common that they use partial equilibrium models and mainly stick to capital market risks. In addition, they all find that more flexibility in the retirement decision increases the portfolio share invested in stocks.

Compared to the existing literature in general and the work of Bodie et al. (1992) in particular, we add three important elements to the analysis on portfolio choice and retirement. First, we complement the partial equilibrium approach with a general equilibrium one. A general equilibrium perspective seems the most natural road to take because the move to flexibility in the retirement date clearly is an international phenomenon. With general equilibrium, we explicitly recognize that consumption and labour-supply decisions affect factor prices which, in turn, influence the insurance effect of retirement flexibility. To illustrate, if every older worker decides to work longer after an adverse shock, wages will decline making the insurance of retirement flexibility less effective. Second, we distinguish between productivity and depreciation (or financial) risk and these risk factors are directly linked to production. This distinction is important because both risk factors constitute a rather different effect on income and substitution effects in labour supply. As will be shown, the relative strength of income and substitution effects determines whether retirement flexibility indeed serves as a hedge against poor asset returns. Third, following Choi et al. (2008), we allow for more general preferences which are characterized by a constant elasticity of substitution (CES) function of consumption and leisure. This specification allows the elasticity of substitution between labour and leisure to take any positive number.

To analyse the interaction between portfolio choice, consumption and retirement decisions, we develop a two-period overlapping-generations (OLG) model of a closed economy in the spirit of Samuelson (1958) and Diamond (1965). The model includes government debt and incorporates endogenous retirement. In our framework, the young working generation decides upon his consumption and the allocation of his asset portfolio. Agents can either invest in risk-free government bonds or in risky firm stocks. Our model is related to the model of Adema (2008) which is also a stochastic two-period OLG model of a closed economy with government debt. There, however, the return on bonds is subject to inflation risk while retirement is exogenous. In our model, retirement is endogenous and we compare two different retirement settings: under *flexible* retirement, the old generation can freely postpone or advance retirement in the second period after a realization of shocks; under *fixed* retirement, this generation has to make this decision already before shocks are revealed. Once set, this decision cannot be subsequently changed when new information becomes available.

We use log-linearization techniques to characterize the main insights of the model. This method is widely applied in the real business cycle literature (see e.g., Campbell, 1994; Uhlig, 1999 or King et al., 2002), but it is also often used in stochastic overlapping-generations models (see Matsen and Thøgersen, 2004; Hougaard Jensen and Jørgensen, 2008 or Bohn, 2009). The standard procedure used in these studies is to first derive the non-stochastic steady state and then to take first-order Taylor approximations around this steady state. The resulting system of log-linear difference equations can then be solved either numerically or analytically. To study macroeconomic dynamics, as most of the aforementioned studies do, this procedure is sufficient. It is less suitable, however, for an analysis involving asset-pricing issues, as we do here. We therefore log-linearize the model around a stochastic steady state which explicitly takes the second-order risk terms into account. This method has already been used by Bovenberg and Uhlig (2008) and Beetsma and Bovenberg (2009), who both study risk-sharing issues in relation to social security, but until now it has never been applied to portfolio allocation in relation to endogenous retirement.

This chapter shows that existing views from the literature may change if we incorporate more realistic elements into the analysis. First, the positive relation between retirement flexibility and a higher demand for risky assets is weakened and under some conditions even turned around if not only depreciation shocks but also productivity shocks are considered. Depreciation shocks mainly affect the return on capital and through the income effect these shocks contribute to the traditional view that retirement flexibility increases risk-taking behaviour. Productivity shocks, in contrast, do not only affect capital returns but also influence wages. Consequently, productivity shocks also induce substitution effects in labour supply which work in the opposite direction. These substitution effects generate a positive correlation between asset returns and labour income, thereby reducing the risk-bearing capacity of consumers.

Second, confining the analysis to Cobb-Douglas utility, as most of the existing studies do, ignores the essential role of the elasticity of substitution between consumption and leisure in studying retirement flexibility. This elasticity of substitution governs the relative strength of income and substitution effects in labour supply and, hence, determines the insurance provided by retirement flexibility. Our analysis clearly shows that flexible retirement amplifies consumption volatility if substitution effects are important, a notion also put forward by Basak (1999).

Finally, we find that general equilibrium effects play an important role in the interaction between portfolio choice and retirement. Ignoring these effects by sticking to a partial equilibrium framework can either overstate or understate the hedging effect of retirement flexibility, dependent on the willingness of consumers to substitute between consumption and leisure. If the elasticity of substitution is high, agents choose to supply less labour after a negative productivity shock. In general equilibrium, this labour-supply response exacerbates the direct fall in the return on capital due to the productivity contraction.

Compared to partial equilibrium, this higher sensitivity of the capital return for productivity risk results in lower portfolio shares invested in equity. Of course, for low elasticities of substitution the opposite holds: then the insurance effect is more effective in general equilibrium than in partial equilibrium, leading to higher equity shares.

The results we derive in this chapter are relevant for private or public pension institutions, like corporate pension funds, trust funds or life-insurance companies, to which individuals have dedicated or will dedicate their saving and investment decisions in the future. As the development towards tailor-made pension products is still an ongoing process in many countries, the acknowledgement that investment policy should be based on individual preferences for retirement will become increasingly important. Even if individuals are able to make the retirement decision conditional on future states, our analysis shows that risky investment strategies are not always in their interest. This is in particular the case if shocks to pension wealth and wages are positively correlated or if consumers view leisure and consumption as close substitutes. Indeed, many authors argue that the stock market and human capital are highly positively correlated (Baxter and Jermann, 1997; Benzoni et al., 2007). Moreover, many empirical studies exploring the impact of financial incentives on the retirement decision typically find modest wealth effects (Krueger and Pischke, 1992; French, 2005 and Bloemen, 2011) but large substitution effects (Gruber and Wise, 1999; Coile and Gruber, 2001 and Asch et al., 2005).

This chapter is organized as follows. Section 4.2 sets out the basics of the stochastic OLG model. In Section 4.3, we explain how to solve this model using a log-linearization technique around the stochastic steady state. Section 4.4 presents analytical results for a simplified version of the model that reproduces the main findings of the current literature. In Section 4.5 we present and compare numerical results for the partial equilibrium model and for the general equilibrium model. Finally, Section 4.6 concludes.

4.2 The model

In this section, we develop a two-period OLG model of a closed economy. In order to analyse the interaction between retirement and portfolio choice, we include government debt in the model as an alternative investment vehicle for future consumption and introduce endogenous retirement in the second period of life. The economy is subject to productivity risk and depreciation risk.

At each point in time, the young individual determines consumption of a single good and the proportion of financial wealth to invest in firm stocks. The old generation decides which fraction of the second period it will spend on working and enjoying retirement. Following Bodie et al. (1992), we consider two retirement settings: *i*) under *flexible* retirement, the old generation can freely postpone or advance retirement in the second period after a realization of shocks; *ii*) under *fixed* retirement, the retirement decision has to be made before shocks are revealed. Once set, the retirement age cannot be subsequently changed after new information has become available. Whatever the retirement setting (flexible or fixed), an individual sets his decision variables optimally, conditional on his information to date: his current financial wealth, the future dynamics of the asset returns and his uncertain future wage.[1]

4.2.1 Production

The young and old generation are composed of the same large number of individuals and this number is normalized to unity. Production per young worker is described by a standard neoclassical constant-returns-to-scale Cobb-Douglas production function:

$$f(k_t, z_t) = A_t k_t^{\alpha} (1 + z_t)^{1-\alpha} \tag{4.1}$$

with A_t the stochastic total productivity parameter, α the capital share in production and k_t the capital stock per young worker. Total labour supply, $1 + z_t$, consists of young workers inelastically supplying one unit of labour and old workers, each spending a fraction $0 \leq z_t \leq 1$ of time on working. Profit maximization and perfect competition among producers results in the standard equilibrium conditions:

$$w_t = (1 - \alpha) A_t k_t^{\alpha} (1 + z_t)^{-\alpha} \tag{4.2}$$

$$r_{k,t} + \delta_t = \alpha A_t k_t^{\alpha-1} (1 + z_t)^{1-\alpha} \tag{4.3}$$

[1]Note that our interpretation of fixed retirement differs from the usual interpretation of a fixed statutory retirement age, which is of course not a decision variable at the individual level. In this chapter, though, we are primarily interested in the *isolated impact* of retirement flexibility (i.e. the ability to condition the retirement age on the state of the economy) on portfolio choice and therefore compare two model settings in which retirement is optimally chosen, given the amount of information available.

where w_t is the real wage, $r_{k,t}$ the return on capital and δ_t can be interpreted as the stochastic depreciation rate of capital.

Production and capital investment are important in this context because they endogenize the (positive) correlation between capital and labour income.[2] Productivity risk directly affects the capital return and the wage rate, while depreciation risk only directly affects the return on capital. Of course, there is an indirect link between the wage rate and depreciation risk, to the extent that labour-supply behaviour affects factor prices in general equilibrium. Stochastic depreciation not only breaks down the (perfect) correlation between wages and capital returns, it also increases return volatility and may give capital returns a higher one-period-ahead variance than wages. The two stochastic processes for total factor productivity and capital depreciation are:

$$\log A_t = \log A + \omega_{A,t} \tag{4.4}$$

$$\log \delta_t = \log \delta + \omega_{\delta,t} \tag{4.5}$$

with $\omega_{A,t}$ and $\omega_{\delta,t}$ independently and identically distributed with mean zero and variance σ_A^2 and σ_δ^2.

4.2.2 Consumers

Individuals derive utility from consumption and leisure. Expected life-time utility of a representative individual born at t is given by the following constant-relative-risk-aversion (CRRA) utility function:

$$U_t = \frac{c_{y,t}^{1-\gamma} - 1}{1-\gamma} + \beta \frac{\mathrm{E}_t\, u(c_{o,t+1}, 1 - z_{t+1})^{1-\gamma} - 1}{1-\gamma} \tag{4.6}$$

where $c_{y,t}$ is consumption when young at time t, $c_{o,t+1}$ is consumption when old at $t+1$, β is the time discount factor and γ is the coefficient of relative risk aversion which is identical to the inverse of the intertemporal elasticity of substitution. The per-period utility function $u(\cdot)$ has a CES specification and is defined as:[3]

$$u(c_o, 1-z) = \left[(1-\eta)c_o^{1-\rho} + \eta(1-z)^{1-\rho}\right]^{\frac{1}{(1-\rho)(1-\eta)}} \tag{4.7}$$

[2]Recent empirical evidence indeed suggests that aggregate labour income and stock returns are positively correlated, at least in the long run (see e.g., Baxter and Jermann, 1997; Bohn, 2009 and Benzoni et al., 2007). As one model period represents about 30 years, shocks in our model should be interpreted as long-term shocks.

[3]Defining the per-period function in this way implies that the coefficient of relative risk aversion with respect to consumption is equal to γ if $\rho = 1$.

where η defines the relative preference for leisure and ρ represents the inverse of the elasticity of substitution between consumption and leisure in the second period. This specification includes the familiar Cobb-Douglas period utility function $u(c_o, 1-z) = c_o(1-z)^{\eta/(1-\eta)}$ if $\rho = 1$.

People can either invest in firm stocks which yield the stochastic return $r_{k,t+1}$ or in government bonds with the risk-free return $r_{b,t+1}$. The share of savings that is invested in stocks is denoted by λ_t, so that the return on the asset portfolio can be defined as:

$$r_{t+1} \equiv (1-\lambda_t)r_{b,t+1} + \lambda_t r_{k,t+1} \tag{4.8}$$

Consumption in the first and second period of life are respectively given by:

$$c_{y,t} + s_t = w_t - \tau_t \tag{4.9}$$

$$c_{o,t+1} = (1 + r_{t+1})s_t + z_{t+1}w_{t+1} \tag{4.10}$$

where τ_t are lump-sum taxes to finance the interest obligations on the government debt.

Maximizing life-time utility with respect to consumption ($c_{y,t}$ and $c_{o,t+1}$) and the portfolio allocation (λ_t) subject to the budget constraints gives the following Euler condition:

$$c_{y,t}^{-\gamma} = \beta\, \mathrm{E}_t \left[(1 + r_{j,t+1}) c_{o,t+1}^{-\rho} u(c_{o,t+1}, z_{t+1})^{\rho-\varphi} \right] \tag{4.11}$$

for $j = b, k$ and with $\varphi \equiv \gamma - \eta\,(1-\rho)$.

The first-order condition with respect to labour supply (z_{t+1}) differs between flexible and inflexible retirement.[4] In the first case, the optimality condition is:

$$\left(\frac{c_{o,t+1}}{1 - z_{t+1}} \right)^{\rho} = \frac{w_{t+1}}{\theta} \tag{4.12}$$

with $\theta \equiv \eta/(1-\eta)$. In the optimum, the marginal rate of substitution between leisure and consumption is equal to the wage rate. Since agents can freely adjust labour supply in period $t+1$, this decision is conditional on the shocks that affect consumption and the wage rate in that period, i.e., $\omega_{A,t+1}$ and $\omega_{\delta,t+1}$. With inflexible retirement, though, the first-order condition is:

$$\begin{aligned} \theta(1 - z_{t+1})^{-\rho}\, \mathrm{E}_t \left[u(c_{o,t+1}, z_{t+1})^{\rho-\varphi} \right] = \\ \mathrm{E}_t \left[w_{t+1} c_{o,t+1}^{-\rho} u(c_{o,t+1}, z_{t+1})^{\rho-\varphi} \right] \end{aligned} \tag{4.13}$$

[4]Throughout the analysis, z_{t+1} indicates labour supply in the second period of life. Under fixed retirement, however, z_{t+1} is chosen in the first period and therefore known at time t.

Since agents are not able to condition the retirement decision at the state of the economy in $t+1$, they have to form (rational) expectations. Obviously, z_{t+1} is known at time t.

4.2.3 Government

The government debt per young worker, b_{t+1}, is equal to the amount of debt in the previous period plus the interest obligations on the outstanding debt minus the collected tax receipts. That is,

$$b_{t+1} = (1 + r_{b,t})b_t - \tau_t \tag{4.14}$$

The government can accumulate debt for a certain amount of time, but at some point in time it has to raise additional taxes in order to keep debt per young worker constant, i.e., $b_{t+1} = b_t = b$. These lump-sum taxes are denoted by τ and are equal to:

$$\tau_t = r_{b,t}b \tag{4.15}$$

Like the capital stock and labour supply (in case of fixed retirement), the bond return $r_{b,t}$ is a predetermined variable: it denotes the interest that is paid at time t on the government debt that is issued one period before, in $t-1$.

4.2.4 Equilibrium

The capital market (and the goods market as well) is in equilibrium when savings at time t finance the capital stock and the government debt in the next period:

$$s_t = k_{t+1} + b_{t+1} \tag{4.16}$$

Moreover, the portfolio allocation has to be such that the right amount of private savings goes to the capital stock and the government debt:

$$\lambda_t s_t = k_{t+1} \tag{4.17}$$

This implies that there are two equilibrium conditions and k_{t+1} and $r_{b,t+1}$ adjust to make sure that these equilibrium conditions are satisfied. Obviously, equation (4.16) and equation (4.17) imply that $(1 - \lambda_t)\, s_t = b_{t+1}$.

The whole model is summarized in Table 4.1. To construct equation (T4.1.1) we have substituted equations (4.15) and (4.16) in (4.9). Equation (T4.1.2) is

Table 4.1: The model

$$w_t - c_{y,t} - r_{b,t} b = b + k_{t+1} \tag{T4.1.1}$$

$$c_{o,t} = (1 + r_{b,t})b + (1 + r_{k,t})k_t + z_t w_t \tag{T4.1.2}$$

$$c_{y,t}^{-\gamma} = \beta \, \mathrm{E}_t \left[(1 + r_{k,t+1}) c_{o,t+1}^{-\rho} u(c_{o,t+1}, z_{t+1})^{\rho-\varphi} \right] \tag{T4.1.3}$$

$$c_{y,t}^{-\gamma} = \beta (1 + r_{b,t+1}) \, \mathrm{E}_t \left[c_{o,t+1}^{-\rho} u(c_{o,t+1}, z_{t+1})^{\rho-\varphi} \right] \tag{T4.1.4}$$

$$w_t = (1 - \alpha) A_t k_t^{\alpha} (1 + z_t)^{-\alpha} \tag{T4.1.5}$$

$$r_{k,t} + \delta_t = \alpha A_t k_t^{\alpha-1} (1 + z_t)^{1-\alpha} \tag{T4.1.6}$$

$$\left(\frac{c_{o,t+1}}{1 - z_{t+1}} \right)^{\rho} = \frac{w_{t+1}}{\theta} \tag{T4.1.7a}$$

$$\mathrm{E}_t \left[w_{t+1} c_{o,t+1}^{-\rho} u(c_{o,t+1}, z_{t+1})^{\rho-\varphi} \right] = \theta (1 - z_{t+1})^{-\rho} \, \mathrm{E}_t \left[u(c_{o,t+1}, z_{t+1})^{\rho-\varphi} \right] \tag{T4.1.7b}$$

the result of inserting the portfolio rate of return, equation (4.8), and the equilibrium conditions (4.16) and (4.17) into equation (4.10). The remaining equations, equation (T4.1.4)-(T4.1.7b), just repeat equation (4.11) (for $j = k$ and $j = b$) and equations (4.2), (4.3), (4.12) and equation (4.13).

4.3 Solving the model

There are various ways to solve this model. One way is to solve the model numerically using dynamic programming methods or using perturbation methods around the deterministic steady state (see, for instance, Collard and Juillard, 2001 or Schmitt-Grohé and Uribe, 2004). Another possibility is to approximate the model using log-linearization around the steady state. The latter gives a bit more insight into the working of the model, and it is the road we will take. It should be understood that log-linearization is a small-shock approximation or an approximation to shocks with bounded support (Samuelson, 1970). Despite these limitations of log-linear approximations, this method clearly helps to explore the most important economic factors that affect the interaction between retirement behaviour and portfolio choice. As such, it provides a useful starting point for further qualitative explorations with higher-order numerical techniques.[5]

4.3.1 The steady state

A linearization around a deterministic steady state is sufficient for understanding macroeconomic dynamics, but it is not necessarily sufficient for an economic analysis involving uncertainty, such as questions about precautionary savings and asset-pricing issues. Following Juillard and Kamenik (2005), Bovenberg and Uhlig (2008) and Beetsma and Bovenberg (2009), we therefore use the concept of a *stochastic* steady state. This concept is defined as a situation in which each period shocks are equal to their expectations but agents are not aware of this (i.e., conditional variances are not zero). This point is solved from a non-linear system, and hence the solution does not generally correspond to the expected values of the variables involved.[6]

This non-linear system of steady-state equations is described in Table 4.2 where variables without time index denote steady-state values. Notice that equations (T4.2.1), (T4.2.2), (T4.2.5), (T4.2.6) and (T4.2.7a) have the same form as the original model equations of Table 4.1. The remaining expectational equations, i.e., equations (T4.2.3), (T4.2.4) and (T4.2.7b), are derived using second-order Taylor approximations of the original first-order conditions.[7] The use of a stochastic steady state implies that risk terms $\sigma^2_{r_k-u}$, σ^2_u, $\sigma^2_{w-c_o}$ and $\sigma^2_{c_o}$ show up in the first-order conditions reflecting a precautionary motive for saving and postponing retirement. These conditional (co)variances are defined as:

$$\sigma^2_{r_k-u} \equiv \text{Var}_t\left[\log(1+r_{k,t+1}) - \varphi \log c_{o,t+1} + \theta(\rho-\varphi)\log(1-z_{t+1})\right] \quad (4.18)$$

$$\sigma^2_u \equiv \text{Var}_t\left[-\varphi \log c_{o,t+1} + \theta(\rho-\varphi)\log(1-z_{t+1})\right] \quad (4.19)$$

$$\sigma^2_{w-c_o} \equiv \text{Var}_t\left(\log w_{t+1} - \varphi \log c_{o,t+1}\right) \quad (4.20)$$

$$\sigma^2_{c_o} \equiv \text{Var}_t\left[(\rho-\varphi)\log c_{o,t+1}\right] \quad (4.21)$$

[5]We also checked our results with higher order approximations using Dynare++. Although quantitatively the results give some small differences, the qualitative observations are exactly the same.

[6]Since the solution is not necessarily equal to expected values of the variables, Beetsma and Bovenberg (2009) label this solution as the *median solution*. We prefer to use the term stochastic steady state to indicate that the steady state is adjusted for risk.

[7]See Appendix 4.A.1 for more details. See also Viceira (2001).

Table 4.2: The steady state

$$w - c_y - r_b b = b + k \tag{T4.2.1}$$

$$c_o = (1 + r_b)b + (1 + r_k)k + zw \tag{T4.2.2}$$

$$c_y^{-\gamma} = \beta(1 + r_k)c_o^{-\varphi}(1 - z)^{\theta(\rho-\varphi)} \exp\left(\frac{1}{2}\sigma^2_{r_k - u}\right) \tag{T4.2.3}$$

$$c_y^{-\gamma} = \beta(1 + r_b)c_o^{-\varphi}(1 - z)^{\theta(\rho-\varphi)} \exp\left(\frac{1}{2}\sigma^2_{u}\right) \tag{T4.2.4}$$

$$w = (1 - \alpha)Ak^{\alpha}(1 + z)^{-\alpha} \tag{T4.2.5}$$

$$r_k + \delta = \alpha A k^{\alpha-1}(1 + z)^{1-\alpha} \tag{T4.2.6}$$

$$\left(\frac{c_o}{1 - z}\right)^{\rho} = \frac{w}{\theta} \tag{T4.2.7a}$$

$$\left(\frac{c_o}{1 - z}\right)^{\rho} = \frac{w}{\theta} \exp\left[\frac{1}{2}\left(\sigma^2_{w-c_o} - \sigma^2_{c_o}\right)\right] \tag{T4.2.7b}$$

At this point, we implicitly assume that these variances are constant over time. This will be justified later on, when solving for the linear recursive law of motion of the log-linearized system (see Section 4.3.2).

In general, the system in Table 4.2 cannot be solved analytically. Only for a particular case we are able to obtain explicit solutions, namely if: *i*) life-time utility is log-linear in consumption and leisure ($\gamma = \rho = 1$); *ii*) there is full depreciation ($\delta = 1$) and *iii*) all conditional covariances are perceived to be zero (deterministic steady state). In that case, we obtain the following analytical expressions for retirement z and the capital-labour ratio $k/(1+z)$:[8]

$$z = \frac{\lambda(1 - \alpha) - \alpha\theta}{\lambda(1 + \theta - \alpha) + (1 - \lambda)\alpha\theta} \tag{4.22}$$

$$\frac{k}{1 + z} = \left[\frac{\alpha\beta A(1 + \theta + \alpha)\lambda - 2\alpha^2\beta A}{(1 - \alpha)\lambda + \alpha\beta(2 + \theta)\lambda + 2\alpha}\right]^{\frac{1}{1-\alpha}} \tag{4.23}$$

Notice from these expressions that both labour supply and the capital-labour ratio positively depend on the portfolio share λ invested in firm stocks: if λ

[8]See Appendix 4.A.2 for the formal derivation.

decreases, for example because of a higher government debt, this leads to a crowding out of firm stocks which reduces the capital-labour ratio. In general equilibrium, a lower capital-labour ratio reduces the wage rate and, hence, labour-supply incentives. Simulations confirm that this property of the model also holds under more general assumptions for which analytical results are not available. Equations (4.22) and (4.23) pin down the steady-state solution of all remaining variables.

4.3.2 The log-linearized model

In the usual situation of no uncertainty, the steady state can be computed separately from the recursive laws of motion. With a stochastic steady state, though, this procedure does no longer apply. In this case, deriving the recursive laws involves the calculation of a fixed point: note from equations (T4.2.3), (T4.2.4) and (T4.2.7b) that the steady state requires knowledge of the conditional variances, which can be calculated, given the log-linear recursive laws of motion. But the latter are solutions to a system of equations of which the coefficients depend on the steady state. Hence, we are forced to simultaneously solve for the steady state and the log-linear model. Throughout this chapter, we use the following notation for log-linearized variables: $\tilde{x}_t \equiv \log x_t - \log x$. The complete log-linearized model is reported in Table 4.3.

Solving for the steady state and the log-linearized equilibrium laws involves a three-step procedure. For the mathematical details we refer to Appendix 4.B, here we stick to a description of these three steps.

- **Step 1: rewriting the linear system.** The first step is to write the log-linearized endogenous variables as function of the endogenous and exogenous state variables. Our model contains two exogenous state variables, productivity shocks ($\omega_{A,t}$) and depreciation shocks (ω_δ) and one endogenous state variable, which is the capital stock ($\tilde{k}_t$). Recall that the return on government bonds ($\tilde{r}_{b,t}$) and labour supply in case of retirement inflexibility ($\tilde{z}_t$) are predetermined variables at time t. It turns out, however, that both variables are proportional to the capital stock so that they can be eliminated from the state space.

  The proportional (and negative) relation between the return on bonds and the capital stock follows from capital market equilibrium: a higher capital stock combined with a constant level of government debt has to

Table 4.3: The log-linearized model

$$w\tilde{w}_t - c_y\tilde{c}_{y,t} = k\tilde{k}_{t+1} + r_b b\tilde{r}_{b,t} \tag{T4.3.1}$$

$$c_o\tilde{c}_{o,t} = r_k k\tilde{r}_{k,t} + (1+r_k)k\tilde{k}_t + r_b b\tilde{r}_{b,t} + zw(\tilde{z}_t + \tilde{w}_t) \tag{T4.3.2}$$

$$\varphi\,\mathrm{E}_t\,\tilde{c}_{o,t+1} - \gamma\tilde{c}_{y,t} = \frac{r_k}{1+r_k}\,\mathrm{E}_t\,\tilde{r}_{k,t+1} - \theta(\rho-\varphi)\frac{z}{1-z}\,\mathrm{E}_t\,\tilde{z}_{t+1} \tag{T4.3.3}$$

$$\varphi\,\mathrm{E}_t\,\tilde{c}_{o,t+1} - \gamma\tilde{c}_{y,t} = \frac{r_b}{1+r_b}\,\tilde{r}_{b,t+1} - \theta(\rho-\varphi)\frac{z}{1-z}\,\mathrm{E}_t\,\tilde{z}_{t+1} \tag{T4.3.4}$$

$$\tilde{w}_t = \alpha\tilde{k}_t - \alpha\frac{z}{1+z}\,\tilde{z}_t + \omega_{A,t} \tag{T4.3.5}$$

$$\tilde{r}_{k,t} + \frac{\delta}{r_k}\tilde{\delta}_t = \frac{r_k+\delta}{r_k}\left[(1-\alpha)\frac{z}{1+z}\,\tilde{z}_t - (1-\alpha)\tilde{k}_t + \omega_{A,t}\right] \tag{T4.3.6}$$

$$\tilde{z}_{t+1} = \frac{1-z}{\rho z}\,\tilde{w}_{t+1} - \frac{1-z}{z}\,\tilde{c}_{o,t+1} \tag{T4.3.7a}$$

$$\tilde{z}_{t+1} = \frac{1-z}{\rho z}\,\mathrm{E}_t\,\tilde{w}_{t+1} - \frac{1-z}{z}\,\mathrm{E}_t\,\tilde{c}_{o,t+1} \tag{T4.3.7b}$$

result in a more aggressive asset portfolio. To make this happen, the risk-free return on bonds will fall. The proportional relation between labour supply and the capital stock in case of retirement inflexibility can either be positive or negative, depending on the relative strength of income and substitution effects: a higher next-period capital stock leads to higher future wage expectations. Hence, rational agents, who plan the retirement decision before shocks are revealed under retirement inflexibility, will postpone retirement if the substitution effect dominates and will advance retirement if the income effect dominates.

Accordingly, the capital stock is the only endogenous state variable in the model. For any endogenous variable $\tilde{x}_t$ we are looking for the following recursive equilibrium law:

$$\tilde{x}_t = \pi_{x,k}\,\tilde{k}_t + \pi_{x,A}\,\omega_{A,t} + \pi_{x,\delta}\,\omega_{\delta,t} \tag{4.24}$$

where $\pi_{x,k}$ is the partial elasticity of $\tilde{x}_t$ with respect to $\tilde{k}_t$, $\pi_{x,A}$ is the partial elasticity of $\tilde{x}_t$ with respect to $\omega_{A,t}$ and $\pi_{x,\delta}$ is the partial elasticity of $\tilde{x}_t$ with respect to $\omega_{\delta,t}$.

- **Step 2: determining the conditional variances.** As a second step, we use the derived recursive law to write the conditional variances in terms of the steady-state values and the exogenous shock terms. Then we obtain for the variance terms of the consumption Euler equations:

$$\sigma^2_{r_k-u} = \sum_{i=A,\delta} \left[\frac{r_k}{1+r_k} \pi_{r_k,i} - \varphi \pi_{c_o,i} - \frac{\theta(\rho-\varphi)z}{1-z} \pi_{z,i} \right]^2 \sigma^2_i \tag{4.25}$$

$$\sigma^2_u = \sum_{i=A,\delta} \left[-\varphi \pi_{c_o,i} - \frac{\theta(\rho-\varphi)z}{1-z} \pi_{z,i} \right]^2 \sigma^2_i \tag{4.26}$$

Note that these variances are indeed constant over time, as assumed in the previous subsection. Equations (4.25) and (4.26) apply to the flexible retirement setting as well as to the inflexible retirement setting, although the partial elasticities differ in both cases. With retirement inflexibility, we also have to derive the covariance terms that show up in the optimality condition of labour supply. These covariances are equal to:

$$\sigma^2_{w-c_o} = \sum_{i=A,\delta} \left(\pi_{w,i} - \varphi \pi_{c_o,i} \right)^2 \sigma^2_i \tag{4.27}$$

$$\sigma^2_{c_o} = \sum_{i=A,\delta} \left[(\rho-\varphi) \pi_{c_o,i} \right]^2 \sigma^2_i \tag{4.28}$$

- **Step 3: solving the linear system.** In the final step, we numerically solve for the steady-state variables, given the derived expressions for the conditional variances. In case of retirement flexibility, this boils down to solving equations (T4.2.1)-(T4.2.7a), equation (4.25) and equation (4.26). For retirement inflexibility, the complete system of equations is described by equations (T4.2.1)-(T4.2.6), (T4.2.7b) and (4.25)-(4.28). Once solved for the steady state, the computed formulas in Appendix 4.B.1 (for flexible retirement) and Appendix 4.B.2 (for flexible retirement) retrieve the partial derivatives, and hence, the linear recursive system.

4.4 Retirement as hedge: benchmark result

As stressed in the introduction of this chapter, the current literature on retirement flexibility and portfolio choice only focuses on partial equilibrium models and mainly sticks to capital market risks. The main result that can be derived from this literature is that flexibility in the retirement decision increases the fraction of wealth invested in equity.[9] Viewed in this way, labour-supply flexibility creates a kind of insurance against adverse investment outcomes. In this section, we will illustrate this benchmark result in the context of our model. With reference to the literature, we take a partial equilibrium perspective (factor prices are exogenous) and assume that there is only capital market risk implying that wages are non-stochastic. To keep the analysis as simple as possible, we impose that expected life-time utility is log-linear in first-period consumption, second-period consumption and leisure (i.e., $\rho = \gamma = 1$).

To derive an explicit solution for the portfolio choice λ_t, we follow the approach of Hansen and Singleton (1983) and Campbell and Viceira (2002) and assume that the joint distribution of consumption and asset returns is lognormal. Then the optimal solution for portfolio choice in case of flexible retirement is given by (see Appendix 4.C.1):

$$\lambda_t^F = \left[1 + \frac{w_{t+1}}{(1+r_{b,t+1})s_t}\right] \frac{\log \mathrm{E}_t(1+r_{k,t+1}) - \log(1+r_{b,t+1})}{\mathrm{Var}_t \log(1+r_{k,t+1})} \tag{4.29}$$

The optimal investment share in the risky asset is increasing in the expected excess return of the risky asset and decreasing in its variance. In case of inflexible retirement, the optimal equity share equals (see Appendix 4.C.2):

$$\lambda_t^I = \left[1 + \frac{w_{t+1}z_{t+1}}{(1+r_{b,t+1})s_t}\right] \frac{\log \mathrm{E}_t(1+r_{k,t+1}) - \log(1+r_{b,t+1})}{\mathrm{Var}_t \log(1+r_{k,t+1})} \tag{4.30}$$

Note that equation (4.29) and equation (4.30) are identical except for one factor: λ^F contains maximum potential human capital while λ^I contains actual labour income which is scaled by $z_{t+1} < 1$.[10] Hence, it is straightforward to derive the following result:

Proposition 4.1 *The investment allocation to the risky asset is larger in the case of flexible retirement compared to the inflexible retirement case, i.e.,* $\lambda_t^F > \lambda_t^I$.

Proposition 4.1 is well known from the literature, and was first derived by Bodie et al. (1992). If agents have the possibility to postpone retirement after an adverse shock, they can afford to take more investment risk during working life. As shown by equations (4.29) and (4.30), this higher risk taking stems from a wealth effect. The demand for the risky asset depends positively on the

[9] See, e.g., Bodie et al. (1992), Choi and Shim (2006), Farhi and Panageas (2007) or Choi et al. (2008).

[10] In principle, private savings may not be equal in the flexible and fixed retirement case. However, in Appendix 4.C we show that $s_t^F = s_t^I$.

amount of human wealth of an individual. With flexible retirement, the individual has in effect a larger store of human capital upon which to draw. Since human capital is risk free (at least until now), the individual rebalances his total wealth holdings by investing a larger share of financial wealth in the risky asset. By contrast, with fixed retirement an individual has a smaller amount of potential human capital from which to invest and therefore requires less rebalancing.

Obviously, these differences in portfolio allocation have consequences for the retirement decision. With flexible labour supply, the optimal solution for retirement is equal to (see again Appendix 4.C.1):

$$z_{t+1}^F = 1 - \frac{\theta\beta(1+r_{T,t+1})}{1+\beta(1+\theta)} \left(\frac{w_t - \tau_t}{w_{t+1}} + \frac{1}{1+r_{b,t+1}} \right) \tag{4.31}$$

with,

$$r_{T,t+1} \equiv (1-a_t)r_{b,t+1} + a_t r_{k,t+1} \tag{4.32}$$

$$a_t \equiv \frac{\lambda_t s_t}{s_t + \frac{w_{t+1}}{1+r_{b,t+1}}} \tag{4.33}$$

Note that a_t is the fraction of an individual's total wealth (financial wealth plus human wealth) invested in the risky asset. Hence, $r_{T,t+1}$ is the effective return on the individual's total portfolio when human wealth (i.e., the discounted value of future labour income) is also taken into account. In case of a positive equity shock, i.e., r_T is high, agents will retire earlier due to a positive wealth effect, and *vice versa*. With inflexible retirement, the optimal retirement decision equals (see again Appendix 4.C.2):

$$z_{t+1}^I = 1 - \frac{\theta\beta(1+r_{b,t+1})}{1+\beta(1+\theta)} \left(\frac{w_t - \tau_t}{w_{t+1}} + \frac{1}{1+r_{b,t+1}} \right) \tag{4.34}$$

Note that the risk-free return $r_{b,t+1}$ now enters the retirement function rather than the stochastic return $r_{T,t+1}$. Accordingly, it is possible to derive the following result:

Proposition 4.2 *The expected retirement age in the flexible retirement case is lower than in the inflexible case, i.e.,* $\mathrm{E}_t\, z_{t+1}^F < z_{t+1}^I$.

Proof Using the optimal solution for s_t (derived in Appendix 4.C), it follows from equation (4.29) that $\lambda_t s_t > 0$. Using equation (4.33), this implies that $a_t > 0$ and, hence, $\mathrm{E}_t\, r_{T,t+1} > r_{b,t+1}$. ∎

In summary, when people can adjust their retirement decision, they will invest more in the risky asset. Since the risky asset has a higher expected return, these people can on average afford to retire earlier.

4.5 Numerical results

This section explores the quantitative properties of the model and solves for the steady state and the reaction of the various variables to productivity and depreciation shocks.[11] We first use the model to gain insight in the partial equilibrium effects of retirement (in)flexibility. Then we turn to the general equilibrium effects and relate these to the partial equilibrium results.

4.5.1 Parameterization

In order to quantify the interaction between portfolio choice and the retirement decision, we first have to parameterize the model. We normalize the average productivity parameter at $A = 1$. The capital share in the Cobb-Douglas production function is taken to be $\alpha = 0.3$, as in Krueger and Kubler (2006) and Olovsson (2010). We set δ, the average depreciation rate, to 0.75. Assuming that one model period lasts about 30 years, this corresponds with a depreciation rate of 5% per year, like in Olovsson (2010). We choose as benchmark an intertemporal elasticity of substitution of one half, i.e., $\gamma = 2$, and an intratemporal substitution of $\rho = 1$. An intertemporal elasticity of substitution of one half lies well within the range of available estimates (see e.g., Attanasio and Weber, 1995) and is commonly used in the macro and public finance literature (it implies a coefficient of relative risk aversion of 2). We choose as time discount factor $\beta = 0.71$, or a time discount rate of 1.1% per year. The leisure parameter is set at $\eta = 0.5$ and the supply of government debt is set at $b = 0.016$, a combination which provides plausible values for the retirement age and the risk-free return on government bonds.

Since productivity risk directly affects all factor prices in the economy (wages and asset returns) and depreciation risk only influences capital returns, the two risk factors have a different effect on retirement and portfolio decisions. We will therefore analyse the model for depreciation and productivity risk separately. In order to make the results comparable, we calibrate the standard

[11] In this section, we restrict to a discussion of the steady state and the model dynamics. In Appendix 4.B.3, we also present simulation results (first and second moments) of the model.

Table 4.4: Benchmark parameterization

Parameter	β	η	ρ	γ	α	A	δ	b	σ_A	σ_δ
Values	0.71	0.5	1	2	0.3	1	0.75	0.016	0.31	1.26

deviation of the exogenous shock (i.e., σ_A in case of productivity risk and σ_δ in case of depreciation risk) in such a way that the annualized standard deviation of the return on capital is the same in both cases and equal to 8.2%.[12] This leads to $\sigma_A = 0.31$ and $\sigma_\delta = 1.26$.

All parameter values used in the benchmark model are summarized in Table 4.4.

4.5.2 Partial equilibrium

For flexible labour supply, the partial equilibrium solution is determined by equations (4.18) and (4.19), equations (T4.2.1)-(T4.2.4) and equation (T4.2.7a). In case of fixed labour supply, we have to solve for equations (4.18), (4.19), (4.20), (4.21), equations (T4.2.1)-(T4.2.4) and equation (T4.2.7b).

By definition, in the partial equilibrium model factor prices are exogenous and only influenced by the exogenous shock terms $\omega_{A,t}$ and $\omega_{\delta,t}$. The log-linearized equations for wages and capital returns are then:

$$\tilde{w}_t = \omega_{A,t} \tag{4.35}$$

$$\tilde{r}_{k,t} = \frac{r_k + \delta}{r_k}\,\omega_{A,t} - \frac{\delta}{r_k}\,\omega_{\delta,t} \tag{4.36}$$

The partial elasticities of the wage rate and the return on capital with respect to productivity and depreciation shocks (i.e., $\pi_{w,A}$, $\pi_{w,\delta}$, $\pi_{r_k,A}$ and $\pi_{r_k,\delta}$) are the same as those derived for the general equilibrium model with fixed retirement.[13] This makes sense because with fixed labour supply both the capital stock and labour supply are predetermined variables. Conditional on information at time t, the only source of variation in future factor prices comes from the exogenous shocks. Consequently, if the exogenous factor prices are set at the corresponding general equilibrium values, the partial equilibrium model gives exactly the same results.

[12]Because log-linearization is a small-shock approximation, the standard deviation of the return on capital (but also the excess return) is lower than in Chapter 3. Here we follow Campbell and Viceira (2005) who show that returns on stocks are significantly less volatile when the investment horizon is long.

[13]See Appendix 4.B.2.

Table 4.5: Steady state of partial equilibrium models

	Depreciation risk		Productivity risk	
	Fixed	Flexible	Fixed	Flexible
c_y/y	37.74	37.61	37.47	37.44
c_o/y	51.03	49.70	49.63	49.76
s/w	32.79	33.00	33.57	33.62
z	20.75	17.00	16.27	16.83
λ	60.60	78.17	97.56	91.28

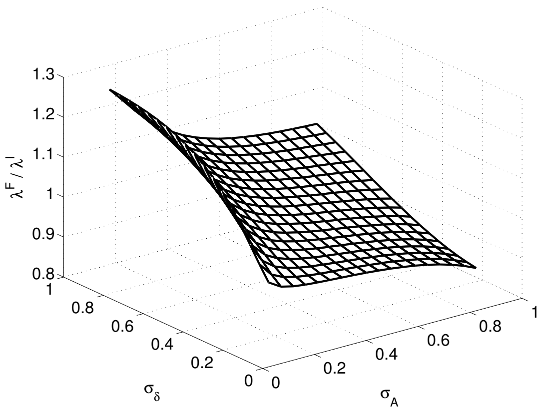
Note: all figures are expressed in percentages.

Table 4.5 compares the steady-state results for fixed and flexible labour supply. The table distinguishes between depreciation and productivity risk. The capital return, the return on bonds and the wage rate are exogenous and obtained from the general equilibrium model with flexible labour supply. Note that, in case of depreciation risk, our model reproduces the traditional view that retirement flexibility increases risk exposure, the first result analytically derived in the previous section. From equation (4.35) and (4.36) we see that wages and capital returns are not correlated when depreciation risk is the only source of uncertainty. A positive depreciation shock (i.e., a negative wealth shock) causes marginal utility from working to increase and, hence, agents increase labour supply (or postpone retirement). Consequently, income effects generate a negative correlation between asset returns and labour income, enabling investors to take greater advantage of the equity premium. The result of this investment strategy is that retirement flexibility induces agents to retire earlier on average compared to retirement inflexibility, the second result derived in Section 4.4. Given our parameterization, agents choose to retire after 66.2 years in case of inflexible retirement while they retire on average after 65.1 years in case of flexible retirement, a difference of about 14 months.[14]

If productivity risk is the sole risk factor, however, the results will turn around. In that case, retirement flexibility may instead be used to amplify the productivity shocks absorbed into consumption, leading to less risk exposure and a higher retirement age compared to fixed retirement. The reason is that productivity shocks do not only induce an income effect in labour supply but

[14]We assume that each model period lasts 30 years. Life time then consists of one period of childhood and schooling that are not accounted for, another period of full activity and a last period the first part of which is devoted to working and lasts 30z years. The retirement age is thus $60 + 30z$.

Figure 4.1: Equity share: fixed versus flexible retirement



also a substitution effect which works in the opposite direction. This substitution effect exacerbates the positive correlation between labour income and capital returns, making equity investment relative unattractive under retirement flexibility. When productivity goes down, both the return on capital and the wage rate decrease. When people can freely adjust retirement, they will respond to this lower wage rate by reducing labour supply, which decreases labour income even further. Hence, under retirement flexibility labour-supply behaviour is subject to procyclical pressure which reduces the risk bearing capacity of consumers.[15] As a result, people are forced to work longer on average. Given our parameterization, this additional work span amounts to almost 2 months.

Figure 4.1 shows the change of the relative equity share (i.e., the equity share in case of flexible retirement divided by the equity share in case of inflexible retirement) for different values for σ_A and σ_δ. The two standard deviations are varied between 0.1 at the lower end and 0.9 at the upper end. When the retirement decision is flexible in the second period of life, agents invest relatively much in equity if depreciation risk is high and productivity risk low and *vice versa*.

4.5.3 General equilibrium

Now we turn to the general equilibrium solution. Table 4.6 shows the steady-state results in case of general equilibrium and again distinguishes between

[15]Notice that this result is similar to what we had in Chapter 3 where the substitution effect related to the mismatch contributions reduced the risk-bearing power of the working generation.

depreciation and productivity risk. The table also shows the results for the deterministic steady state, i.e., when the conditional variances are zero.

Comparing the deterministic steady state with the stochastic steady states illustrates the role of uncertainty in the model. Obviously, if there is no uncertainty, the equity premium (denoted by μ) is equal to zero since capital investments and government bonds are perfect substitutes. In the stochastic steady state, the equity premia are positive reflecting the higher riskiness of capital investments.[16] Including the risk terms in the optimality conditions introduces a precautionary motive for more savings and later retirement. Note that the saving rate and labour supply are higher in the stochastic steady state than in the deterministic steady state.

In general equilibrium, exactly the same risk features appear as in partial equilibrium but they are now operating through price adjustments rather than quantity adjustments. With exogenous factor prices, we saw that agents invest more in equity under flexible labour supply than under fixed labour supply if depreciation risk is the dominant source of uncertainty. When productivity risk is the dominant source, we found the opposite result, namely that agents invest less in equity under retirement flexibility than under retirement inflexibility. With endogenous factor prices and a fixed supply of government bonds, though, different risk attitudes affect the price of risk taking, i.e., the equity premium. If productivity risk is the sole risk factor, the equity premium is *higher* in case of flexible retirement than in case of inflexible retirement. The intuition for this lower demand for risk taking under flexible retirement is the same as before: the substitution effect related to labour market flexibility exacerbates the positive correlation between asset returns and labour income which decreases the risk-bearing capacity of consumers. Hence, people are only willing to invest in the domestic capital stock if they receive a higher expected compensation. If there is only depreciation risk, however, the insurance mechanism related to the income effect dominates, resulting in a *lower* equity premium under labour market flexibility.

Like in the partial equilibrium model, steady-state labour supply is lower with flexible retirement than with inflexible retirement if there is only depreciation risk. In the former case, people on average choose to retire after 65.1 years while in the latter case they retire after 66.3 years, a difference of about 15 months. When agents have no retirement flexibility and only face depre-

[16] Note that the reported risk premia are on the low side, which is a manifestation of the equity premium puzzle.

Table 4.6: Steady state of general equilibrium models

	No risk	Depreciation risk		Productivity risk	
		Fixed	Flexible	Fixed	Flexible
c_y/y	37.61	36.72	37.67	37.56	37.42
c_o/y	50.23	51.55	49.76	49.49	49.66
s/w	31.96	32.11	32.89	33.58	33.65
μ	0.00	0.52	0.32	0.33	0.37
r_b	2.47	2.07	2.06	1.96	1.92
z	16.44	21.12	16.90	16.52	17.00
k/y	16.22	15.63	16.75	17.26	17.23
λ	84.41	84.23	85.06	85.55	85.57

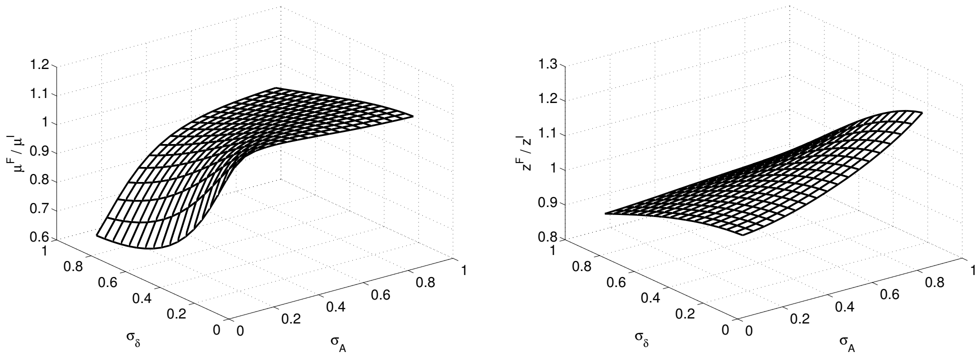
Notes: the equity premium (i.e., $\mu \equiv r_k - r_b$) and the return on government debt are annualized figures. All figures are expressed in percentages.

ciation risk, labour supply is an attractive way to finance future consumption compared to private savings, because wages are not uncertain while the proceeds of savings are uncertain. On the contrary, with retirement flexibility equity savings are attractive because people will probably earn the equity premium while they always have the option to postpone retirement if things go wrong. Hence, compared to the inflexible setting, agents save more and a higher fraction of these savings is allocated to firm equity. Since the supply of government debt is given in general equilibrium, the equity premium has to decline to make sure that enough savings are allocated to this debt. It turns out that the wealth effect (more savings) dominates the price effect (lower equity premium), resulting in lower labour supply under retirement flexibility.

If there is only productivity risk, instead, retirement flexibility is less interesting from an insurance perspective because capital returns are low in states in which wages are also low. Therefore, agents have a relative high demand for risk-free bonds which drives down the interest rate on government debt. This negative wealth effect implies that agents on average retire about 2 months later with flexible labour supply.

Figure 4.2 shows the dependence of portfolio and retirement decisions on the two risk factors in a more general way. These figures compare the equity premium (left panel) and labour supply (right panel) in case of retirement flexibility with those in case of retirement inflexibility. If depreciation risk is high and productivity risk low, the risk premium is lower under flexible

Figure 4.2: Equity premium and labour supply: fixed retirement versus flexible retirement



(a) Equity premium (b) Labour supply

retirement, reflecting the self-insurance role of voluntary retirement. When productivity risk becomes more important, the equity premium increases and ultimately passes the levels of the fixed retirement setting. A comparable pattern emerges for labour-supply behaviour. For higher degrees of productivity risk, the hedging effect of retirement flexibility decreases which leads to a higher demand for risk-free government bonds and, given the fixed level of government debt, to lower risk-free interest rates. This negative wealth effect induces agents to postpone retirement.

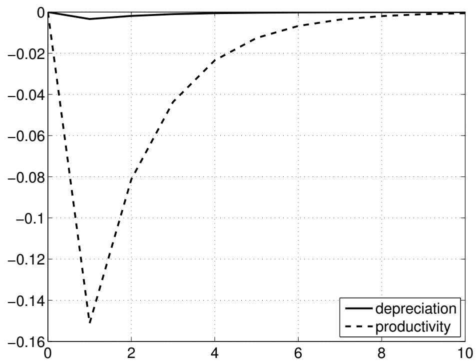
It should be stressed that from a welfare perspective flexibility is always preferable to inflexibility because the model does not include any distortion or externality. With retirement flexibility, expected life-time utility is unambiguously higher, both in case of depreciation risk and productivity risk.[17]

4.5.4 Dynamics

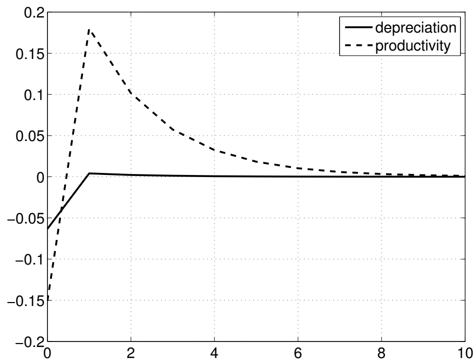
The different roles in the interaction between retirement flexibility and portfolio allocation played by productivity and depreciation shocks can best be illustrated using impulse response functions. Figure 4.3 shows the response of the capital stock, the return on capital and bonds, the wage rate, labour supply and old-age consumption to a 10% positive depreciation shock (solid lines) and to a 10% negative productivity shock (dashed lines). The responses

[17]By simulating the derived recursive laws, we have calculated the unconditional means of most important model variables. It turns out that the unconditional mean of life-time utility in case of flexible retirement is always higher than that in case of inflexible retirement, see Table 4.7 in Appendix 4.B.3.

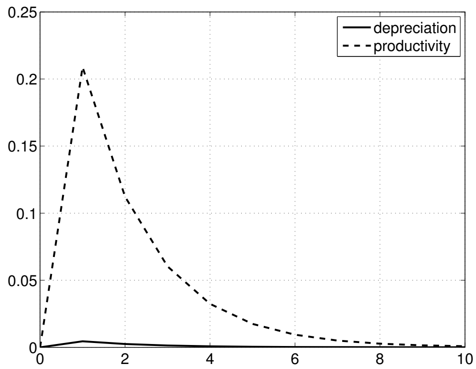
Figure 4.3: Impulse response functions



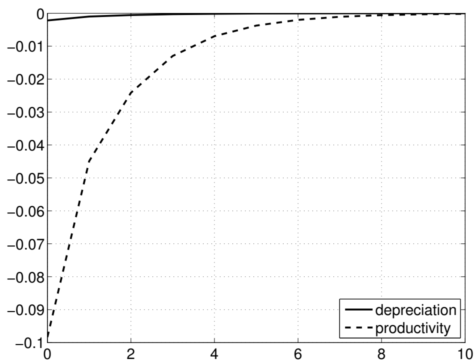
(a) Capital stock



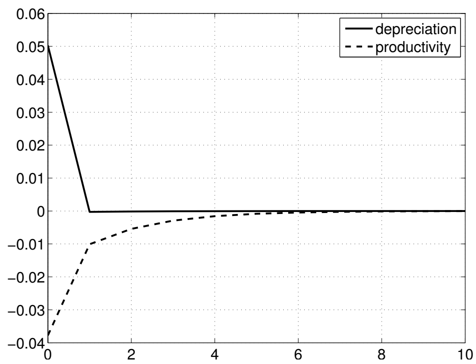
(b) Capital return



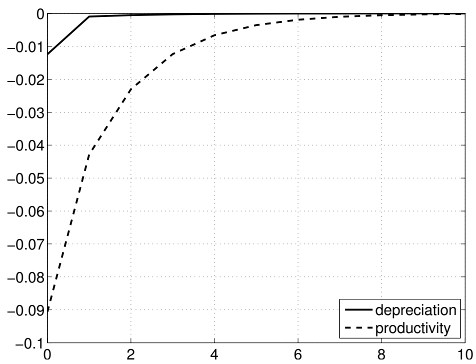
(c) Bond return



(d) Wage rate



(e) Labour supply



(f) Old-age consumption

are expressed in percent deviation from the steady state.

Note first that depreciation shocks lead to relative small responses compared to productivity shocks. After a depreciation shock of 10%, the capital return immediately decreases and, due to the income effect, labour supply increases. This negative correlation between the capital return and labour supply moderates consumption volatility and that is why flexibility provides insurance against adverse shocks. At impact, the decline of old-age consumption is small compared to the decline of the capital return. The capital stock is a predetermined variable and falls one period later. This lower level of the capital stock increases its marginal product so that labour supply declines and, hence, wages and consumption gradually return to their pre-shock levels. The return on bonds moves in the opposite direction of the capital stock: a lower capital stock increases its marginal product leading to a higher demand for capital investment and a lower demand for bond investments. As a result, the return on bonds should increase in order to ensure that the fixed supply of government debt will be financed each period.

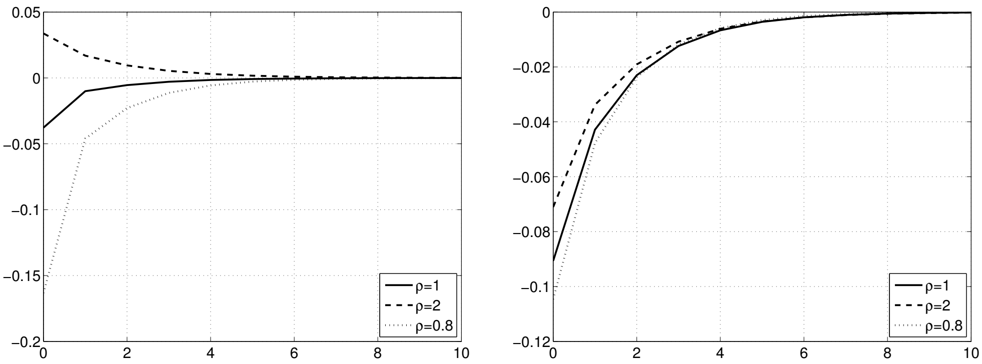
The economic responses after a productivity shock are much larger. In this case, the decrease in the capital return is even larger than the initial decline in productivity itself. Compared to a depreciation shock, a productivity shock does not only directly affect the return on capital but also the wage rate which falls at impact. This shock induces income and substitution effects in labour supply. Indeed, given the benchmark parameterization, the substitution effect dominates the income effect and that is why labour supply slightly decreases. Hence, productivity shocks result in pro-cyclical labour-supply behaviour which exacerbates the volatility of consumption. Note that the initial decline in old-age consumption is almost as high as the relative decrease in productivity. From an investment point of view, the positive co-movement between capital returns and labour income reduces the demand for risk taking. Consequently, the equity premium will be relatively higher under retirement flexibility.

4.5.5 Sensitivity analysis

The previous analysis has shown that the insurance effect of retirement flexibility very much depends on income and substitution effects in labour supply. In our benchmark parameterization, the substitution effect slightly dominates the income effect so that old-age consumption becomes more sensitive to productivity risk in case of retirement flexibility. As a result, agents ask for a higher risk compensation (in general equilibrium) or decrease the equity share in the total asset portfolio (in partial equilibrium).

The relative strength of income and substitution effects is governed by the elasticity of substitution between consumption and leisure. In this section, we study the role of the substitution elasticity in the hedging effect of retirement flexibility into more detail. Figure 4.4 shows the responses of labour supply

Figure 4.4: Impulse response functions with CES utility



(a) Labour supply (b) Consumption

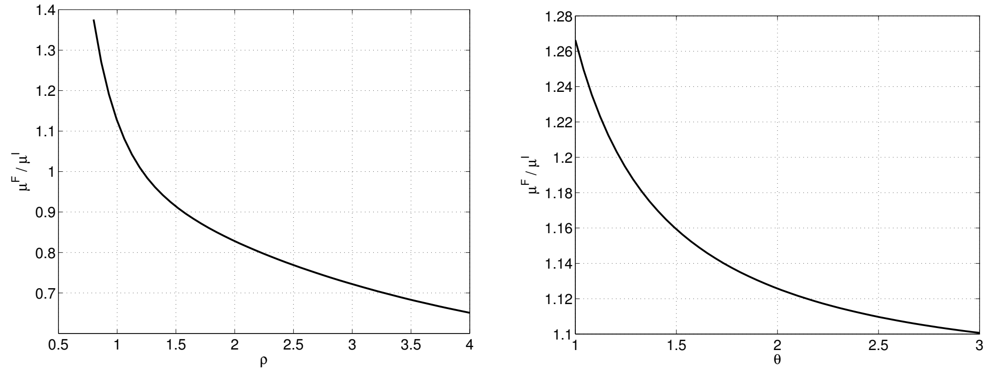
and consumption to a negative productivity shock of (again) 10% for various degrees of substitutability between consumption and leisure. As shown, the labour-supply choice may be used to amplify the productivity shocks absorbed into consumption if the elasticity of substitution of leisure for consumption is high. If this substitution elasticity is low, however, consumers may use their labour/leisure choice to mitigate the effect of the shock on consumption. It is a well-known macroeconomic finding that consumption and labour move positively (Blanchard and Fisher, 1989). Therefore, for high substitutability our model is consistent with the data.

When retirement is flexible, the positive comovement of consumption and labour leads to higher equity premia if the elasticity of substitution increases. Figure 4.5a shows the reaction of the equity premium in case of retirement flexibility relative to the equity premium in case of inflexibility for different degrees of substitution between consumption and leisure.[18] For low values of ρ (high elasticity of substitution), the equity premium under flexible retirement exceeds the equity premium under inflexible retirement. For higher values of ρ (lower elasticity of substitution), the income effect becomes gradually more important and, hence, also the insurance effect of retirement flexibility increases. So when the elasticity of substitution is high, flexible retirement acts in the direction of resolving the equity risk premium puzzle (Basak, 1999).

Figure 4.5b illustrates the sensitivity of the relative equity premium, now for different degrees of risk aversion (or intertemporal substitution). As one can see, for all values of γ considered, the ratio is decreasing in relative risk

[18] In Figure 4.5, it is assumed that productivity risk is the sole risk factor, because substitution effects in labour supply are not relevant in case of depreciation risk.

Figure 4.5: Relative equity premium for various degrees of intratemporal substitution and risk aversion



(a) Intratemporal substitution (b) Risk aversion

aversion but it never falls below unity.[19] This means that, contrary to the elasticity of intratemporal substitution, the coefficient of relative risk aversion does not alter the order of the equity premium: the equity premium is higher with flexible retirement than with fixed retirement.

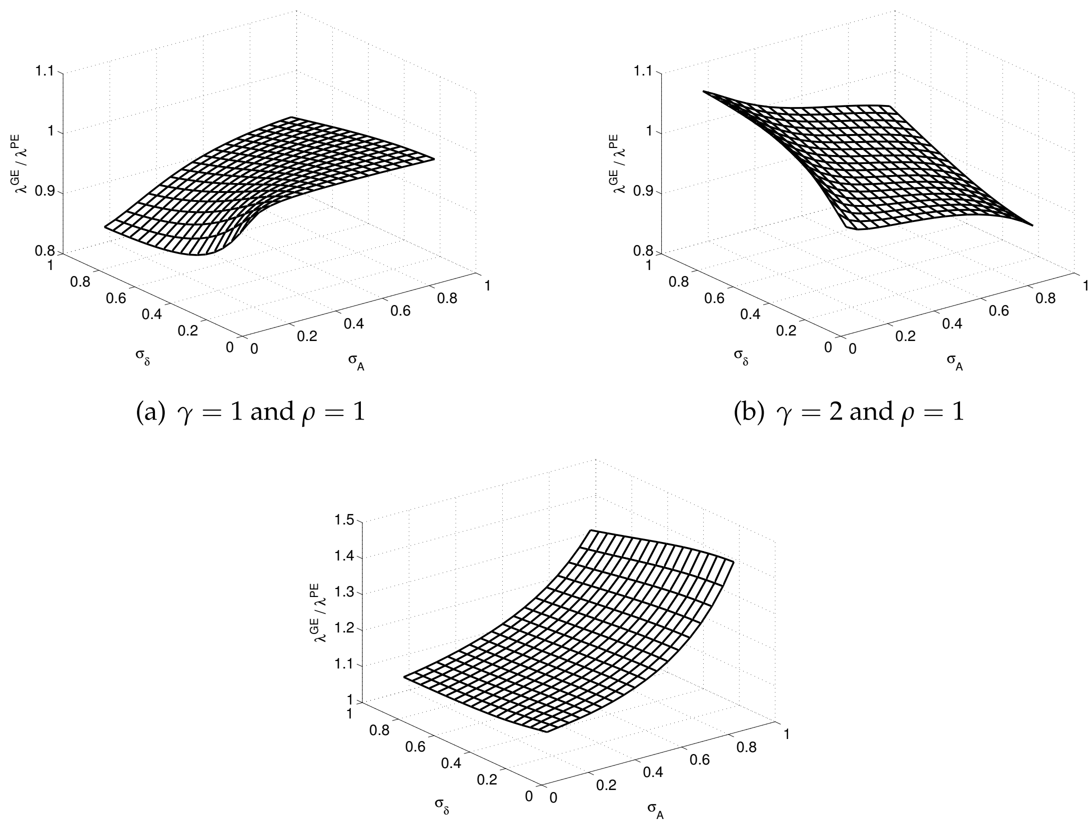
4.5.6 Importance of general equilibrium effects

An interesting question is whether the general equilibrium effects increase or decrease the demand for risky assets compared to a partial equilibrium approach.[20] In this final section we use our model to isolate the general equilibrium effects of retirement flexibility and to identify the main factors that determine the direction of these effects. Existing studies in the field of retirement and portfolio choice only focus on partial equilibrium models thereby ignoring the potentially important general equilibrium effects. As will be discussed, the differences between general equilibrium and partial equilibrium results can be reduced to differences in the partial elasticities of the capital re-

[19] Figure 4.5b shows that in *relative* terms (i.e., in percentage of the equity premium under fixed retirement) the equity premium under flexible retirement is decreasing in risk aversion. In *absolute* terms, however, the difference between the equity premia under flexible and fixed retirement is increasing in risk aversion, as would be expected: for higher degrees of risk aversion, agents ask for a higher expected return to compensate for the positive correlation between factor prices under retirement flexibility.

[20] Remember that under fixed retirement the partial equilibrium solution coincides with the general equilibrium solution. Hence, in this section, the comparison between partial and general equilibrium only points to flexible retirement.

Figure 4.6: Equity share: partial versus general equilibrium



(a) $\gamma = 1$ and $\rho = 1$

(b) $\gamma = 2$ and $\rho = 1$

(c) $\gamma = 2$ and $\rho = 2$

turn and labour supply with respect to the exogenous shocks (i.e., $\pi_{r_k,A}$, $\pi_{z,A}$, $\pi_{r_k,\delta}$ and $\pi_{z,\delta}$). Recall from equations (4.25) and (4.26) that these elasticities determine the conditional variances $\sigma^2_{r_k-u}$ and σ^2_u under flexible retirement.

Figure 4.6 shows the portfolio share of equity in general equilibrium compared to that in partial equilibrium, again plotted for various degrees of productivity and depreciation risk. In order to make a comparison possible, for each combination of standard deviations, the exogenous factor prices in partial equilibrium are imposed to be the same as the calculated factor prices in general equilibrium. Figure 4.6a is based on log-linear utility ($\gamma = 1$ and $\rho = 1$). On the whole grid of standard deviations, the relative equity exposure is below unity meaning that in general equilibrium agents invest less in equity than in partial equilibrium. Note that this difference in risk exposure is particularly large if depreciation risk is high. Since everyone decides to work longer (or to postpone retirement) after an adverse depreciation shock, wages will decline in general equilibrium. Consequently, the positive elasticity of labour supply with respect to depreciation shocks ($\pi_{z,\delta}$) is lower in general equilibrium which makes the insurance of retirement flexibility less effective. Optimizing agents respond to this by lowering their risk exposure. At the same time, the higher supply of labour will also moderate the decline of the capital return in general equilibrium. In other words, the elasticity of the capital return with respect to depreciation shocks ($\pi_{r_k,\delta}$) is less negative than in partial equilibrium. This improves the effectiveness of the insurance and, hence, tends to boost risky investments. With this parameterization, though, the negative effect on risky investments (due to a lower $\pi_{z,\delta}$) dominates the positive effect (due to a less negative $\pi_{r_k,\delta}$).

Why is the relative equity share still below unity for higher degrees of productivity risk? As seen before, with an elasticity of substitution equal to one, agents choose to advance retirement after a negative productivity shock (see Figure 4.3e). In other words, the substitution effect dominates the income effect in labour supply (i.e., $\pi_{z,A} > 0$). In general equilibrium, this reduction in labour supply exacerbates the direct fall of the capital return on account of the productivity contraction. Hence, the capital return is more sensitive to productivity risk than in partial equilibrium (i.e., $\pi_{r_k,A}$ higher) which decreases the effectiveness of the hedging effect of retirement flexibility.

If we increase risk aversion (see Figure 4.6b), the insurance effect is still less effective in general equilibrium for higher levels of productivity risk. However, it becomes more effective for lower degrees of productivity risk and higher degrees of depreciation risk. If risk aversion is higher, the relatively low sensitivity of the capital return with respect to depreciation risk in general equilibrium (which improves the effectiveness of the insurance effect) now dominates the relatively low response in labour supply (which worsens the effectiveness). For higher degrees of risk aversion, agents invest less in firm equity, especially in case of partial equilibrium in which there are no supply restrictions of government bonds. As a consequence, the elasticity of the retirement choice with respect to depreciation shocks drops relatively more in partial equilibrium than in general equilibrium. Hence, the disadvantage of a less effective hedging effect in general equilibrium (associated with this lower sensitivity of the retirement decision) becomes smaller, compared to partial equilibrium.

In the previous section, we have seen that the elasticity of substitution between consumption and leisure plays a crucial role in whether retirement flexibility increases or decreases the demand for stocks. From Figure 4.6c, it can be seen that this parameter is also decisive in the direction of the general equilibrium effects. This figure is based on an elasticity of substitution of one half, implying that income effects now dominate substitution effects (i.e., $\pi_{z,A} < 0$). That means, a negative productivity shock induces people to retire later in time. In general equilibrium, this retirement shift moderates the direct drop in the capital return due to the negative productivity shock. In other words, when income effects are dominating, the sensitivity of the capital return to productivity risk ($\pi_{r_k,A}$) is lower in general equilibrium than in partial equilibrium. Because this lower sensitivity increases the insurance effect of retirement flexibility, the relative equity share is now increasing in the degree of productivity risk.

To summarize, the equity exposure can either be higher or lower in general equilibrium than in partial equilibrium. This is true both for productivity and depreciation risk. With depreciation risk, the labour-supply elasticity with respect to shocks is lower in general equilibrium (which depresses equity investments) but, at the same time, the capital return is less sensitive to these shocks (which stimulates equity investments). We have shown that for low (high) levels of risk aversion the first (second) effect is dominating. In case of productivity risk, the elasticity of intratemporal substitution determines whether agents invest more or less in equity in general equilibrium compared to partial equilibrium. For high intratemporal substitution (substitution effect dominates), the capital return is relatively more sensitive to productivity shocks in general equilibrium resulting in lower equity exposures. For low substitution (income effect dominates), the opposite holds, meaning that agents invest relatively more in equity in general equilibrium.

4.6 Conclusion

In this chapter, we have developed a stochastic general equilibrium model with two overlapping generations. The model is used to analyse the interaction between consumption, portfolio choice and retirement decisions. In the literature, retirement flexibility is often viewed as an insurance against bad investment outcomes. This chapter reviews this benchmark result in a more general model. In particular, our model includes risk factors (productivity risk and depreciation risk) that are directly linked to the production structure of the economy. Second, and more importantly, we combine a partial equilibrium approach with a general equilibrium approach thereby explicitly recognizing that correlations between productivity and depreciation shocks are endogenous. Finally, we allow for more general preferences which are characterized by a CES function of consumption and leisure.

Our main findings are as follows. First, the relevance of retirement flexibility as a hedging instrument strongly depends on the type of risk agents are subject to. Productivity risk affects wages and asset returns in the same direction. Under retirement flexibility, this positive correlation between wages and asset returns is reinforced by the substitution effect on labour supply resulting in a *lower* preference for risk taking. In partial equilibrium this lower demand leads to lower equity shares in the total investment portfolio while in general equilibrium it leads to higher equity premia as the supply of assets is (partly) fixed. With depreciation risk, though, wages are only indirectly affected by general equilibrium effects. In this case, the income effect dominates implying that labour income and capital returns are negatively correlated which leads to a *higher* preference for risk taking. In partial equilibrium, this higher demand leads to higher portfolio shares invested in equity and, in general equilibrium, it leads to lower equity premia.

Second, our analysis reveals that the elasticity of substitution between consumption and leisure is of crucial importance in determining to which extent retirement flexibility protects retirees against bad investment returns. Indeed, this elasticity governs the relative strength of income and substitution effects in labour supply and therefore determines the hedging effect of retirement flexibility. Our analysis shows that the advantage of flexible retirement as hedging instrument is smaller if substitution effects are relatively important.

Third, we find that general equilibrium effects play an important role in the interaction between portfolio choice and retirement. Ignoring these effects by sticking to a partial equilibrium framework can either overstate or understate the insurance benefits of retirement flexibility. It is mainly the degree of substitution between consumption and leisure that determines the direction of the general equilibrium effects. For high substitution elasticities labour-supply behaviour amplifies the sensitivity of capital returns to productivity risk making retirement flexibility less effective as hedging tool in general equilibrium than in partial equilibrium.

This chapter has shown that the main results of existing studies in the field of retirement flexibility and portfolio choice (like e.g., Bodie et al., 1992; Choi and Shim, 2006) may not be robust to alternative (i.e., more realistic) model settings. These studies argue that with labour flexibility much of the uncertainty is absorbed by the labour-supply decision, leaving consumption relatively smooth. This suggests a negative covariability between human capital and the equity market and between consumption and labour. These findings are consistent with our model as long as shocks do not directly affect wages (like depreciation shocks). However, if shocks *do* affect wages (like productivity shocks) and if consumers have a high elasticity of substitution of leisure for consumption, the comovement between consumption and labour becomes positive, resulting in a lower demand for risk. Empirical studies typically find that the stock market and human capital are highly correlated, while it is well known that observed consumption and labour move in the same direction. Interpreted to our analysis, these empirical results imply that more weight should be given to productivity risk than to depreciation risk. Interpreted to the existing literature, these findings suggest that the extent of absorption of financial shocks by total wealth or consumption is underestimated in a model with no labour, an observation already raised by Basak (1999).

Appendix to Chapter 4

In this appendix, we first derive the steady-state relations (Section 4.A). Then we solve the log-linearized model, both for the fixed and flexible retirement setting (Section 4.B). Finally, we will solve for optimal consumption and portfolio allocation in case of log-linear utility and deterministic wages, again both for the fixed and flexible retirement case (Section 4.C). Throughout, we use the notation $\hat{r} \equiv \log(1+r)$ for returns and $\hat{y}_t \equiv \log y_t$ for any other variable y_t.

4.A The steady state

4.A.1 Derivation first-order conditions

We can write equation (T4.1.3) as,

$$1 = \mathrm{E}_t \Big\{ \exp \big[\underbrace{\log \beta + \log \hat{r}_k + \gamma \hat{c}_{y,t} - \rho \hat{c}_{o,t+1} + (\rho - \varphi) \hat{u}_{t+1}}_{x_{t+1}} \big] \Big\} \tag{4.37}$$

Taking a second-order Taylor expansion of $\exp(x_{t+1})$ around $\mathrm{E}_t\, x_{t+1} \equiv \bar{x}_t$, we obtain,

$$\begin{aligned} 1 &\approx \mathrm{E}_t \left\{ \exp(\bar{x}_t) \left[1 + x_{t+1} - \bar{x}_t + \frac{1}{2} (x_{t+1} - \bar{x}_t)^2 \right] \right\} \\ &= \exp(\bar{x}_t) \left(1 + \frac{1}{2} \mathrm{Var}_t\, x_{t+1} \right) \end{aligned} \tag{4.38}$$

Then, a first-order Taylor expansion around zero gives the result,

$$\begin{aligned} 1 &\approx 1 + \bar{x}_t + \frac{1}{2}\operatorname{Var}_t x_{t+1} \\ &= \exp\left(\bar{x}_t + \frac{1}{2}\operatorname{Var}_t x_{t+1}\right) \end{aligned} \tag{4.39}$$

Note that we can write equation (4.7) as,

$$\hat{u} = \frac{\log\left\{\exp\left[\log(1-\eta) + (1-\rho)\hat{c}_y\right] + \exp\left[\log\eta + (1-\rho)\log(1-z)\right]\right\}}{(1-\rho)(1-\eta)} \tag{4.40}$$

Taking a first-order Taylor expansion around zero then gives:

$$\hat{u} \approx \hat{c}_y + \theta\log(1-z) \tag{4.41}$$

with (again) $\theta \equiv \eta/(1-\eta)$. Combining equations (4.39) and (4.41), we obtain the steady-state Euler equation of capital investment, equation (T4.2.3):

$$c_y^{-\gamma} = \beta(1+r_k)c_o^{-\varphi}(1-z)^{\theta(\rho-\varphi)}\exp\left(\frac{1}{2}\sigma^2_{r_k-u}\right) \tag{4.42}$$

with $\sigma^2_{r_k-u}$ defined in equation (4.18).

The derivation of the second Euler equation regarding government bonds investments, equation (T4.2.4), and that of the optimality condition with respect to fixed retirement, equation (T4.2.7b), are similar to the one above.

4.A.2 Deterministic steady state

Suppose that $\gamma = \rho \to 1$ and $\delta = 1$. Ignoring the risk terms or assuming a non-stochastic steady state implies that $r_k = r_b \equiv r$. Then inserting equation (T4.2.1) and equation (T4.2.2) in the Euler equation (T4.2.3) (or equation (T4.2.4)) gives:

$$\frac{1+\beta}{\beta}k = w - rb - \frac{1+\beta}{\beta}b - \frac{w}{(1+r)\beta}z \tag{4.43}$$

From the optimality condition with respect to leisure, equation (T4.2.7a) (or equation (T4.2.7b)), we derive:

$$k = \frac{w}{(1+r)\theta}(1-z) - \frac{w}{1+r}z - b \tag{4.44}$$

Substituting equation (4.44) in (4.43) and solving for z gives:

$$z = \frac{1+\beta-\beta\theta(1+r)\left(1-\frac{rb}{w}\right)}{1+\beta+\beta\theta} \tag{4.45}$$

Inserting equation (4.44) in equation (4.43) and solving for k leads to:

$$k = \frac{\beta(1+\theta)w\left(1-\frac{rb}{w}\right) - \frac{w}{1+r} - (1+\beta+\beta\theta)b}{1+\beta+\beta\theta} \tag{4.46}$$

Using the factor prices, equation (T4.2.5) and equation (T4.2.6), we can rewrite equation (4.46) into:

$$1+z = \frac{\beta(1+\theta)\left(1-\frac{rb}{w}\right)(1-\alpha)\left(\frac{k}{1+z}\right)^{\alpha-1} - \frac{1-\alpha}{\alpha}}{(1+\beta+\beta\theta)\left(1+\frac{b}{k}\right)} \tag{4.47}$$

In the same way, we can rewrite (4.45) into:

$$1+z = \frac{2(1+\beta)+\beta\theta-\beta\theta\left(1-\frac{rb}{w}\right)\alpha A\left(\frac{k}{1+z}\right)^{\alpha-1}}{1+\beta+\beta\theta} \tag{4.48}$$

Equations (4.47) and (4.48) form a closed system in k and z. Solving these equations gives for the capital-labour ratio,

$$\frac{k}{1+z} = \left[\frac{\left(1-\alpha+\theta+\theta\alpha\frac{b}{k}\right)\alpha\beta\left(1-\frac{rb}{w}\right)}{1-\alpha+\left(1+\frac{b}{k}\right)\alpha(2+2\beta+\beta\theta)}\right]^{\frac{1}{1-\alpha}} \tag{4.49}$$

and for labour supply:

$$z = \frac{1-\alpha-\alpha\theta-\alpha\theta\frac{b}{k}}{1+\theta-\alpha+\alpha\theta\frac{b}{k}} \tag{4.50}$$

Using the definition $\lambda \equiv k/(b+k)$ in equation (4.50), gives the labour-supply decision as function of the portfolio choice, equation (4.22). Notice that equation (4.49) still depends on w and r, which are functions of the capital-labour ratio. Again using equations (T4.2.5) and (T4.2.6), we derive:

$$\frac{rb}{w} = \frac{\alpha A\left(\frac{k}{1+z}\right)^{\alpha-1} - 1}{(1-\alpha)\left(\frac{k}{1+z}\right)^{\alpha-1}}\frac{b}{k}(1+z) \tag{4.51}$$

Finally, substituting this expression in equation (4.49) and using equation (4.50), we obtain:

$$\frac{k}{1+z} = \left[\frac{\alpha\beta A\left(1+\theta-\alpha-2\alpha\frac{b}{k}\right)}{1+\alpha+\alpha\beta(2+\theta)+2\alpha\frac{b}{k}}\right]^{\frac{1}{1-\alpha}} \tag{4.52}$$

Using the definition λ in equation (4.52), gives the capital-labour ratio as function of the portfolio choice, equation (4.23).

4.B Solution of the model

In this appendix, we solve the (log-linearized) model. To do this, we first write the model in the following form:

$$\tilde{k}_{t+1} = \pi_{k,k}\,\tilde{k}_t + \pi_{k,A}\,\omega_{A,t} + \pi_{k,\delta}\,\omega_{\delta,t} \tag{4.53}$$

and:

$$\begin{bmatrix} \tilde{c}_{y,t} \\ \tilde{c}_{o,t} \\ \tilde{r}_{k,t} \\ \tilde{w}_t \\ \tilde{r}_{b,t+1} \\ \tilde{z}_t \text{ or } \tilde{z}_{t+1} \end{bmatrix} = \begin{bmatrix} \pi_{c_y,k} \\ \pi_{c_o,k} \\ \pi_{r_k,k} \\ \pi_{w,k} \\ \pi_{r_b,k} \\ \pi_{z,k} \end{bmatrix} \tilde{k}_t + \begin{bmatrix} \pi_{c_y,A} & \pi_{c_y,\delta} \\ \pi_{c_o,A} & \pi_{c_o,\delta} \\ \pi_{r_k,A} & \pi_{r_k,\delta} \\ \pi_{w,A} & \pi_{w,\delta} \\ \pi_{r_b,A} & \pi_{r_b,\delta} \\ \pi_{z,A} & \pi_{z,\delta} \end{bmatrix} \begin{bmatrix} \omega_{A,t} \\ \omega_{\delta,t} \end{bmatrix} \tag{4.54}$$

where $\pi_{x,y}$ denotes the partial elasticity of endogenous variable x with respect to state variable y. With retirement flexibility, the recursive law for labour supply is based on $\tilde{z}_t$. With retirement inflexibility, it is based on $\tilde{z}_{t+1}$ because retirement is predetermined at time t. Solving this system of equations algebraically needs a lot of tedious (but rather straightforward) calculations. In this appendix, we choose to restrict to a description of the main steps.

4.B.1 Flexible retirement

Note that equations (T4.3.2), (T4.3.5), (T4.3.6) and (T4.3.7a) form an independent system of the endogenous variables $\tilde{c}_{o,t}$, $\tilde{w}_t$, $\tilde{r}_{k,t}$ and $\tilde{z}_t$ in the predetermined variables $\tilde{k}_t$ and $\tilde{r_{b,t}}$ and the exogenous shocks $\omega_{A,t}$ and $\omega_{\delta,t}$. From this system we can infer the partial elasticities with respect to productivity and depreciation shocks. To save on notation, we define the following two variables:

$$\Gamma \equiv w^{1-\frac{1}{\rho}}\,\theta^{\frac{1}{\rho}} \tag{4.55}$$

$$\Delta \equiv (1-z)\alpha + (1+z)\rho(1+\Gamma) + \rho\alpha\Gamma \tag{4.56}$$

Then the partial elasticities with respect to productivity shocks are:

$$\pi_{c_o,A} = \frac{(1-z+\rho z+\alpha\rho)y}{c_o\Delta} > 0 \tag{4.57}$$

$$\pi_{r_k,A} = \frac{(r_k+\delta)(\rho+\rho z+\rho\Gamma+1-z)}{r_k\Delta} > 0 \tag{4.58}$$

$$\pi_{w,A} = \frac{\rho(1+z)(1+\Gamma-\alpha)}{(1-\alpha)\Delta} > 0 \tag{4.59}$$

$$\pi_{z,A} = \frac{(1+z)\left[(1-z)(1-\alpha)-\rho\Gamma(\alpha+z)\right]}{z(1-\alpha)\Delta} \tag{4.60}$$

Note that the sign of $\pi_{z,A}$ is ambiguous; it can either be positive or negative, depending on the substitution between consumption and leisure. For the partial elasticities with respect to depreciation shocks we have:

$$\pi_{c_o,\delta} = -\frac{\delta k(\rho+\alpha-\alpha z+\rho z)}{c_o\Delta} < 0 \tag{4.61}$$

$$\pi_{r_k,\delta} = -\frac{\delta\left[\rho(1+z)+(1-z)\alpha+\rho\Gamma(1+z-\alpha z)\right]}{r_k\Delta} < 0 \tag{4.62}$$

$$\pi_{w,\delta} = -\frac{\rho(1-z)\delta k\alpha}{c_o\Delta} < 0 \tag{4.63}$$

$$\pi_{z,\delta} = \frac{(1+z)(1-z)\rho\delta k}{c_o z\Delta} > 0 \tag{4.64}$$

Noting that $\mathrm{E}_t\,\omega_{A,t+1} = \mathrm{E}_t\,\omega_{\delta,t+1} = 0$ and using the Euler equations (T4.3.3) and (T4.3.4), we now can express the bond return $\tilde{r}_{b,t+1}$, the conditional expectations $\mathrm{E}_t\,\tilde{c}_{o,t+1}$ and $\mathrm{E}_t\,\tilde{c}_{r_k,t+1}$ together with first-period consumption $\tilde{c}_{y,t}$ as functions of the next-period capital stock $\tilde{k}_{t+1}$. This ultimately gives:

$$\tilde{r}_{b,t+1} = \Psi_{r_b}\,\tilde{k}_{t+1} \tag{4.65}$$

$$\mathrm{E}_t\,\tilde{c}_{o,t+1} = \Psi_{c_o}\,\tilde{k}_{t+1} \tag{4.66}$$

$$\tilde{c}_{y,t} = \Psi_{c_y}\,\tilde{k}_{t+1} \tag{4.67}$$

$$\mathrm{E}_t\,\tilde{z}_{t+1} = \Psi_z\,\tilde{k}_{t+1} \tag{4.68}$$

where the partial elasticities are equal to,

$$\Psi_{r_b} \equiv -\frac{(1+r_b)\rho(1+z)y\left[(r_k+\delta)(1+\Gamma-\alpha)+\alpha(1-\delta)\Gamma\right]}{r_b y\Delta(1+r_k)+r_b(1+r_b)\rho(r_k+\delta)\Gamma(1+z)b}$$

$$\Psi_{c_o} \equiv \frac{\left[\rho+\alpha+z(\rho-\alpha)\right]\left[(1-\delta)k+r_b b\Psi_{r_b}\right]+\alpha\left[1-z+\rho(z+\alpha)\right]y}{c_o\Delta}$$

$$\Psi_{c_y} \equiv \frac{1}{\gamma}\left[\varphi\Psi_{c_o}-\frac{r_b\Psi_{r_b}}{1+r_b}+\frac{\theta(\rho-\varphi)z\Psi_z}{1-z}\right]$$

$$\Psi_z \equiv \frac{(1-z)(1+z)\left[\alpha c_o-\alpha\rho(y-w)-\rho(1-\delta)k-\rho r_b b\Psi_{r_b}\right]}{c_o z\Delta}$$

Notice from equation (4.65) that $\tilde{r}_{b,t}$ and $\tilde{k}_t$, the two predetermined variables, move proportionally. Therefore, we can substitute out $\tilde{r}_{b,t}$ from the state space.

To obtain the equilibrium law of the capital stock, (4.53), we substitute equation (4.67) in the budget restriction, equation (T4.3.1). This gives the following partial elasticities for the capital stock:

$$\pi_{k,k} = \frac{w\pi_{w,k} - r_b b \Psi_{r_b}}{c_y \Psi_{c_y} + k} \tag{4.69}$$

$$\pi_{k,A} = \frac{w\pi_{w,A}}{c_y \Psi_{c_y} + k} \tag{4.70}$$

$$\pi_{k,\delta} = \frac{w\pi_{w,\delta}}{c_y \Psi_{c_y} + k} \tag{4.71}$$

The system is stable if and only if $\pi_{k,k} < 1$.

The equilibrium law for $\tilde{k}_t$ pins down the solutions of the remaining endogenous variables in the model as shown in equation (4.54). Notice that the equilibrium laws of $\tilde{r}_{b,t+1}$ and $\tilde{c}_{y,t}$ follow from equations (4.65) and (4.67), respectively. This implies that $\pi_{r_b,i} = \Psi_{r_b}\pi_{k,i}$ and $\pi_{c_y,i} = \Psi_{c_y}\pi_{k,i}$ with $i = \{k, A, \delta\}$. The solutions for $\tilde{c}_{o,t}$, $\tilde{w}_t$, $\tilde{r}_{k,t}$ and $\tilde{z}_t$ then follow from equations (T4.3.2), (T4.3.5), (T4.3.6) and (T4.3.7a). This gives the remaining partial elasticities with respect to the capital stock:

$$\pi_{c_o,k} = \Psi_{c_o} \tag{4.72}$$

$$\pi_{r_k,k} = \frac{r_k+\delta}{r_k}\left[\frac{\alpha(\rho+\rho z+\rho\Gamma+1-z)}{\Delta} - \frac{\Gamma\rho(1+z)(k-\delta k+r_b b\Psi_{r_b})}{y\Delta} - 1\right] \tag{4.73}$$

$$\pi_{w,k} = \frac{\alpha\rho(1+z)(1+\Gamma-\alpha)}{(1-\alpha)\Delta} + \frac{\alpha\rho(1-z)(k-\delta k+r_b b\Psi_{r_b})}{c_o\Delta} \tag{4.74}$$

$$\pi_{z,k} = \Psi_z \tag{4.75}$$

4.B.2 Fixed retirement

The derivation of the solution with fixed retirement mainly follows the same steps as that of the flexible retirement setting. When retirement is fixed, equations (T4.3.2), (T4.3.5) and (T4.3.6) form an independent system of the endogenous variables $\tilde{c}_{o,t}$, $\tilde{w}_t$ and $\tilde{r}_{k,t}$ in terms of the predetermined variables $\tilde{k}_t$, $\tilde{r}_{b,t}$ and $\tilde{z}_t$ and the exogenous shocks $\omega_{A,t}$ and $\omega_{\delta,t}$. From this system, we can directly solve for the partial elasticities with respect to the shock terms. For productivity shocks we have:

$$\pi_{c_o,A} = \frac{y-w}{c_o} > 0 \tag{4.76}$$

$$\pi_{r_k,A} = \frac{r_k+\delta}{r_k} > 0 \tag{4.77}$$

$$\pi_{w,A} = 1 \tag{4.78}$$

and for depreciation shocks:

$$\pi_{c_o,\delta} = -\frac{\delta k}{c_o} < 0 \tag{4.79}$$

$$\pi_{r_k,\delta} = -\frac{\delta}{r_k} < 0 \tag{4.80}$$

$$\pi_{w,\delta} = 0 \tag{4.81}$$

With inflexible retirement, equations (4.65)-(4.67) do not change except that the definition of Γ (used in the Ψ-terms) now becomes,

$$\Gamma \equiv w^{1-\frac{1}{\rho}}\,\theta^{\frac{1}{\rho}}\exp\left[\frac{1}{2\rho}\left(\sigma_{c_o}^2 - \sigma_{w-c_o}^2\right)\right] \tag{4.82}$$

Consequently, the dynamic solution of the capital stock is still given by equations (4.69)-(4.71). Therefore, we retain the solution $\pi_{r_b,i} = \Psi_{r_b}\pi_{k,i}$ and $\pi_{c_y,i} = \Psi_{c_y}\pi_{k,i}$ with $i = \{k, A, \delta\}$. In addition, the elasticities of $\tilde{c}_{o,t}$, $\tilde{w}_t$ and $\tilde{r}_{k,t}$ with respect to the capital stock, as given by equations (4.72)-(4.74), are also still satisfied. Note from equation (4.78) that $\pi_{w,\delta} = 0$. Equation (4.71) then implies $\pi_{k,\delta} = 0$ which also means that $\pi_{r_b,\delta} = 0$ and $\pi_{c_y,\delta} = 0$.

Equation (4.68) is no longer satisfied, though, and becomes $\tilde{z}_t = \Psi_z\tilde{k}_t$. For the partial elasticities this means: $\pi_{z,k} = \Psi_z\pi_{k,k}$, $\pi_{z,A} = \Psi_z\pi_{k,A}$ and $\pi_{z,\delta} = 0$.

4.B.3 Simulation results

Table 4.7 shows the unconditional mean and standard deviation of the most important endogenous variables. These moments are calculated by simulating the derived recursive laws. From this table we draw the same conclusions as from the steady-state results, discussed in Section 4.5.3. With depreciation risk, retirement flexibility indeed offers a kind of insurance against adverse investment outcomes as stressed by Bodie et al. (1992). In this situation, the equity premium is lower than in case of inflexible retirement and agents are able to retire earlier on average. With productivity risk, however, we again have the opposite result. Then the equity premium under flexible retirement is higher than under inflexible retirement and agents choose to retire later on average. From a welfare perspective, though, flexibility is preferable to inflexibility. Note that expected life-time utility is unambiguously higher in the first case, irrespective of whether depreciation risk or productivity risk is the sole risk factor.

Table 4.7: Statistical moments of general equilibrium models

	Depreciation risk				Productivity risk			
	Fixed		Flexible		Fixed		Flexible	
	Mean	Stdev	Mean	Stdev	Mean	Stdev	Mean	Stdev
c_y/y	36.72	0.00	37.82	3.36	37.59	1.27	37.47	1.88
c_o/y	53.71	15.71	50.98	11.37	49.59	3.04	49.70	2.13
r_k	2.94	8.20	2.69	7.66	2.64	8.20	2.65	8.22
r_b	2.09	0.00	2.08	0.58	2.18	6.40	2.14	6.27
z	21.12	0.00	20.65	14.50	16.54	0.67	17.13	2.11
k/y	15.63	0.00	16.80	1.23	19.55	10.40	19.53	10.43
U	−6.60	0.97	−6.46	0.98	−6.68	2.37	−6.67	2.34

Notes: the return on capital and the return on government debt are annualized figures. All figures are expressed in percentages.

4.C Appendix to Section 4.4

Suppose that we have log-linear life-time utility in consumption and leisure (i.e., $\rho = \gamma = 1$). Assume further that wages are non-stochastic.

4.C.1 Flexible retirement

Portfolio choice

Inserting equation (4.12) in equation (4.10), and using equation (4.8), we obtain:

$$c_{o,t+1} = \frac{1}{1+\theta}(1 + r_{T,t+1})\left(s_t + \frac{w_{t+1}}{1 + r_{b,t+1}}\right) \tag{4.83}$$

where $r_{T,t+1}$ is defined in equation (4.32). Note that $c_{o,t+1}$ is decomposed in non-stochastic terms (the first and third term) and a stochastic term (the second one). Substituting (4.83) in the two Euler equations (for $j = r_b$ and $j = r_k$) and subtracting both, we have:

$$\mathrm{E}_t\left[(1 + r_{T,t+1})^{-1}(r_{k,t+1} - r_{b,t+1})\right] = 0 \tag{4.84}$$

Taking logs of equation (4.84), we obtain:

$$\mathrm{E}_t\,\hat{r}_{k,t+1} + \frac{1}{2}\mathrm{Var}_t\,\hat{r}_{k,t+1} - \hat{r}_{b,t+1} = \mathrm{Cov}_t(\hat{r}_{T,t+1}, \hat{r}_{k,t+1}) \tag{4.85}$$

where we have used Jensen's inequality condition for a lognormal variable, $\log \mathrm{E}_t\, x_{t+1} = \mathrm{E}_t \log x_{t+1} + 1/2 \operatorname{Var}_t \log x_{t+1}$. To derive the term on the left-hand side of equation (4.85), we follow Campbell and Viceira (2002) and use a second-order Taylor approximation of the portfolio return, equation (4.32). This gives,

$$\hat{r}_{T,t+1} \approx \hat{r}_{b,t+1} + a_t(\hat{r}_{k,t+1} - \hat{r}_{b,t+1}) + \frac{1}{2} a_t (1 - a_t) \operatorname{Var}_t \hat{r}_{k,t+1} \tag{4.86}$$

Hence,

$$\operatorname{Cov}_t(\hat{r}_{T,t+1}, \hat{r}_{k,t+1}) = a_t \operatorname{Var}_t \hat{r}_{k,t+1} \tag{4.87}$$

Substituting equation (4.87) into (4.85) then gives:

$$a_t = \frac{\mathrm{E}_t\, \hat{r}_{k,t+1} - \hat{r}_{b,t+1} + \frac{1}{2} \operatorname{Var}_t \hat{r}_{k,t+1}}{\operatorname{Var}_t \hat{r}_{k,t+1}} \tag{4.88}$$

Finally, inserting (4.88) in (4.33), we end up with the portfolio allocation in terms of financial wealth, i.e., equation (4.29).

Consumption and leisure

Substituting equation (4.83) in equation (4.11) (for $j = r_b$) and rearranging gives:

$$c_{y,t}^{-1} = \beta(1+\theta)(1 + r_{b,t+1})\, \mathrm{E}_t(1 + r_{T,t+1})^{-1} \left(w_t - \tau_t - c_{y,t} + \frac{w_{t+1}}{1 + r_{b,t+1}} \right)^{-1} \tag{4.89}$$

Notice that:

$$\begin{aligned} (1 + r_{b,t+1})\, \mathrm{E}_t(1 + r_{T,t+1})^{-1} &= (1 + r_{b,t+1})\, \mathrm{E}_t(1 + r_{T,t+1})^{-1} \\ &\quad + a_t\, \mathrm{E}_t \left[(1 + r_{T,t+1})^{-1} (r_{k,t+1} - r_{b,t+1}) \right] \\ &= 1 \end{aligned} \tag{4.90}$$

Hence, first-period consumption satisfies:

$$c_{y,t} = \frac{1}{1 + \beta(1+\theta)} \left(w_t - \tau_t + \frac{w_{t+1}}{1 + r_{b,t+1}} \right) \tag{4.91}$$

Note that the propensity to consume is the same as under certainty. Hence, there is no precautionary saving motive, which is a direct implication of the log-utility specification (see Sandmo, 1970). Combining (4.91) and (4.83), we obtain for second-period consumption:

$$c_{o,t+1} = \frac{\beta(1 + r_{T,t+1})}{1 + \beta(1+\theta)} \left(w_t - \tau_t + \frac{w_{t+1}}{1 + r_{b,t+1}} \right) \tag{4.92}$$

Substituting (4.92) in (4.12), we obtain the expression for labour supply, equation (4.31).

4.C.2 Fixed retirement

Portfolio choice

Consider now the fixed retirement setting. Then the intertemporal budget constraint becomes:

$$c_{o,t+1} = (1 + r_{T,t+1}) \left(s_t + \frac{w_{t+1} z_{t+1}}{1 + r_{b,t+1}} \right) \tag{4.93}$$

with $r_{T,t+1}$ again defined as in (4.32) but where a_t now satisfies:

$$a_t = \frac{\lambda_t s_t}{s_t + \frac{w_{t+1} z_{t+1}}{1 + r_{b,t+1}}} \tag{4.94}$$

Inserting (4.93) in the two Euler equations (for $j = r_b$ and $j = r_k$) again gives condition (4.84). Hence, a_t is still given by equation (4.88). Inserting (4.88) into (4.33) we end up with the portfolio share in terms of financial wealth, equation (4.30).

Consumption and leisure

The fact that wages are non-stochastic implies that the first-order condition with respect to leisure consumption, equation (4.13), becomes:

$$\frac{\theta}{1 - z_{t+1}} = w_{t+1} \, \mathrm{E}_t \, c_{o,t+1}^{-1} \tag{4.95}$$

Combining (4.95) and (4.11) (for $j = r_b$), gives:

$$(1 - z_{t+1}) w_{t+1} = \theta \beta (1 + r_{b,t+1}) c_{y,t} \tag{4.96}$$

Substituting (4.93) in (4.11) (again for $j = r_b$) and rearranging gives:

$$c_{y,t}^{-1} = \beta \left(w_t - \tau_t - c_{y,t} + \frac{w_{t+1} z_{t+1}}{1 + r_{b,t+1}} \right)^{-1} \tag{4.97}$$

where we (again) used equality (4.90). Substitution of (4.96) in (4.97) gives:

$$c_{y,t}^{-1} = \beta \left[w_t - \tau_t + \frac{w_{t+1}}{1 + r_{b,t+1}} - (1 + \theta\beta) c_{y,t} \right]^{-1} \tag{4.98}$$

Hence,

$$c_{y,t} = \frac{1}{1+\beta(1+\theta)} \left(w_t - \tau_t + \frac{w_{t+1}}{1+\tau_{b,t+1}} \right) \tag{4.99}$$

Note that consumption (and thus savings) under fixed labour supply is exactly equal to consumption under flexible labour supply. Substituting (4.99) in (4.96) and solving for z_{t+1}, we ultimately obtain the optimal retirement decision, equation (4.34).

CHAPTER 5

REDISTRIBUTION EFFECTS OF PENSION REFORM

This chapter studies the redistribution and welfare effects of pension reforms that aim to improve fiscal sustainability as a response to ageing. We consider reforms that link pension benefits or the pension entitlement age to longevity. We also look at a more fundamental reform that introduces a flexible take-up of pension benefits. To analyse the economic implications of these reforms, we develop a two-period overlapping-generations model with a pay-as-you-go social security scheme and with individuals who differ in ability and life span. We show that linking the entitlement age to life expectancy is beneficial for low-skilled agents at the expense of high-skilled agents, unless the heterogeneity in individual longevity is high. Introducing a flexible pension take-up can induce a Pareto improvement. To obtain this result, we argue that the initial pension scheme must operate within-cohort redistribution and induce early retirement.

5.1 Introduction

Everywhere in industrialized countries, population ageing has put collective pension schemes into financial strain. To improve fiscal sustainability, policy makers have to make the unavoidable choice to increase taxes, to reduce the generosity of pensions or to increase the official retirement age. For a number of reasons, like tax competition, raising taxes is not the solution most countries have adopted. In the past decade, at least half of the OECD countries have undertaken far-reaching pension reforms that point either to a reduction in the generosity of benefits or to an increase in the official retirement age (OECD, 2007). Various countries (like Finland and Portugal) have introduced mechanisms to adjust benefits to increasing life expectancy. Other countries (like Norway and Denmark) have raised the official retirement age or have link it to life expectancy. Some countries (like the UK and Australia) have combined this with measures to increase work incentives, for example by increasing the reward for continuing in work or by introducing a flexible retirement age.

Although the design of these reforms differs among countries, they share one common property in that they are typically implemented in a *uniform* way, applied to all participants. However, since individuals have heterogeneous characteristics (for example in terms of life expectancy or productivity), uniformly implemented reforms may affect individual welfare in a different way. Indeed, it is well known that pension schemes based on uniform policy rules contain large redistribution effects within and across generations, some intentional and others unintentional (see e.g., Börsch-Supan and Reil-Held, 2001; Ter Rele, 2007 and Bonenkamp, 2009). In practice, pay-as-you-go (PAYG) pension schemes, especially those of the Beveridgean type, contain redistribution from high to low incomes. Apart from this, these pension schemes may also contain redistribution from short-lived to long-lived agents because they are typically based on collective annuities which do not depend on individual life expectancy. This makes collective annuities open to the objection that they lead to more regressive pension schemes because it is well known that average longevity tends to increase with income (see e.g., Pappas et al., 1993; Adams et al., 2003 or Meara et al., 2008).

This chapter explores the redistribution and welfare effects of pension reforms that aim to improve fiscal sustainability as a response to ageing. We first analyse reforms that adjust benefits or the pension age to increasing longevity. Then we focus on a more fundamental reform, which entails a change from a payout scheme in which benefits start at the (fixed) statutory retirement age to a scheme where benefits start at the (variable) effective retirement age. This flexible pension take-up is combined with actuarially-based adjustments of pension benefits for early and late retirement.[1] To analyse the economic impli-

[1]Countries that have recently moved into the direction of more individual choice with respect to retirement date are the UK, Finland and Denmark. See Chapter 1 for more empirical evidence. In the Netherlands, there is also a debate to introduce a flexible pension take-up in the first pillar.

cations of these reforms, we use a two-period overlapping-generations model populated with agents who differ in ability and life span. It is assumed that the life span of an individual is positively linked to his productivity. The PAYG social security system is of the Beveridgean type and characterized by life-time annuities and a fixed tax rate which is proportional to wage income. In this way, the pension scheme includes two types of intragenerational redistribution, from high-income earners to low-income earners and from short-lived to long-lived agents.

In this chapter, we follow the literature that considers a PAYG social security scheme as a redistribution device (see e.g., Galasso and Profeta, 2002; Cremer and Pestieau, 2003a; Casamatta et al., 2005). Of course, one might argue about the desirability to use pension schemes for redistribution purposes as compared to the tax system. There may be good reasons though to operate income redistribution using pensions. Protection against myopia is one argument to redistribute income later in the life cycle. Another argument is that individual differences in income could manifest themselves only later in the career. However, the normative question *why* pension schemes should redistribute within cohorts remains outside the scope of this chapter. We take a positive perspective and observe that pension schemes *do* redistribute in practice.

Implementing pension contracts with a variable starting date for benefits, as analysed in this chapter, might be important for various reasons. It may help individuals to adjust the timing of pension income according to their own preferences and circumstances. This is particularly relevant for people who have a preference to retire early (due to disability problems for example) but who are prevented to do that because of liquidity or borrowing constraints. Flexible pensions may also function as a hedge against all types of risks, like disability risks (Diamond and Mirrlees, 1978) or financial risks (Pestieau and Possen, 2010).[2] This chapter adds some other arguments. We will illustrate that flexible pensions can stimulate people voluntarily to postpone retirement, which helps to bear the increasing fiscal burden of ageing. We also show that flexible pension take-up could be used to reduce the element of regressive redistribution in social security schemes. Dependent on the information publicly available, the government can apply different actuarial adjustment factors to low-skilled and high-skilled agents as a way to get rid of the unintended transfers from short-lived to long-lived agents.

This chapter provides some interesting results. We find that the intragen-

[2]See also Chapter 4.

erational redistribution effects of an automatic link between the pension entitlement age and life expectancy depend on the degree of individual life-span heterogeneity. Higher longevity partly induces people to retire later and therefore leads to an increase in the redistribution from high to low incomes. As long as individual differences in life span are not too high, this increased income redistribution dominates the opposite redistribution effects from short-lived to long-lived agents, implying that the unskilled agents (with short life spans and low income levels) benefit from this reform. Simulating this reform in the model, we find a threshold value for life-span heterogeneity between high-skilled and low-skilled agents of about seven years.

More importantly, we find that introducing a flexible pension take-up can be a Pareto-improving reform if actuarial adjustment of benefits occurs in a uniform way (i.e., based on the average life expectancy). Uniform benefit adjustment leads to selection effects in the retirement decision which may reduce initial tax distortions. Indeed, for the high-skilled individuals the uniform reward rate for later retirement is too high from an actuarial point of view, which reduces their implicit tax and stimulates them to continue working. If the payroll tax is sufficiently high, the low-skilled also gain because they receive more pensions, enabled by the additional tax payments of the high-skilled. If the government would use non-uniform benefit adjustment instead, for example by conditioning the adjustment factor on individual or skill-group life expectancies, flexible pension take-up cannot be Pareto improving. With such a more actuarially-neutral approach, selection effects are less important and therefore also the opportunities to reduce existing distortions. Non-uniform actuarial adjustment then only eliminates the unintended transfers from short-lived to long-lived agents, which is beneficial (*harmful*) for the low-skilled (*high-skilled*).

This chapter relates to different strands of literature. It is closely connected to studies that analyse the interaction between pension schemes and retirement decisions (see e.g., Hougaard Jensen et al., 2003; Lau and Poutvaara, 2006) and to a growing literature that focuses on the role of alternative pension systems when income and life expectancy are correlated (see e.g., Borck, 2007; Gorski et al., 2007; Hachon, 2008 and Cremer et al., 2010). In addition, our work is also inspired by Fisher and Keuschnigg (2010) and Jaag et al. (2010) who investigate the labour market impact of pension reforms towards more actuarial neutrality. Most of all these aforementioned studies focus on pension reforms that strengthen the link between contributions and benefits. Our study, in contrast, deals with other types of reforms, like the introduction of an automatic linkage between the generosity of pensions and higher life expectancy and the implementation of a flexible pension take-up.

This chapter is most closely related to Cremer and Pestieau (2003a). They consider a pension reform that generates the same 'double dividend' as the flexibility reform considered in this study: an increase in economic efficiency and an increase in redistribution from people with high wages to people with low wages. To obtain this outcome, both studies need that the benefit rule of the social security scheme must redistribute within generations and must include an initial retirement distortion, the removal of which brings additional resources. However, the studies differ in the reforms they focus on. Cremer and Pestieau (2003a) analyse an increase in the effective retirement age and the driving force behind their efficiency improvement is the implementation of age-dependent tax rates, which are higher for young than for old agents. Our study, in contrast, focuses on a more commonly reform, the introduction of a flexible pension take-up. In this setting, the efficiency improvement stems from selection effects in the retirement decision, induced by uniform actuarial adjustment. As such, actuarial benefit adjustment provides an additional instrument to the government to specifically reduce distortions on the extensive margin of labour supply.

This chapter is structured as follows. In Section 5.2 we introduce the basic model. This model contains a PAYG social security scheme with inflexible pension take-up and life-time annuities and serves as benchmark for the pension reforms considered. The redistribution effects of linking the benefit level and the statutory retirement age to higher longevity are discussed in Section 5.3. Section 5.4 analyses the redistribution effects of reforms aimed at increasing the flexibility of individual pension take-up. In Section 5.5 we elaborate on these flexibility reforms by introducing non-neutral actuarial adjustment of benefits. Section 5.6 concludes the chapter.

5.2 The benchmark model

We consider a two-period overlapping-generations model of a small open economy populated with heterogeneous agents who differ in terms of ability and life span. Agents decide upon the amount of savings in the first period and upon the length of the working period in the second period. The individual ability level determines whether an agent supplies labour as a low-skilled worker or as a high-skilled worker. High-skilled workers earn a higher wage rate than low-skilled workers. The model includes a Beveridgean social security scheme which offers a life-time annuity that starts paying out from the statutory retirement age until the end of life. Agents are allowed to continue working after the statutory retirement age or to stop working before the statutory retirement age. As a consequence, the statutory retirement is related to the date agents receive their pension benefit which is not necessarily equal to the effective retirement date. Since we are primarily interested in the redistribution effects of social security, we abstract from uncertainty in the model.

5.2.1 Preferences

Preferences over first-period and second-period consumption are represented by the following utility function:

$$U(c, x) = u(c) + \pi u(x) \tag{5.1}$$

with $u' > 0$ and $u'' < 0$ and where c denotes first-period consumption, x is second-period consumption and $\pi \leq 1$ is the length of the second period. The interest rate and the discount rate are zero.[3] Second-period consumption is defined net of the (monetary) disutility of labour:

$$x = \frac{d}{\pi} - \frac{\gamma}{2}\left(\frac{z}{\pi}\right)^2 \tag{5.2}$$

where d denotes total consumption of goods when old yielding a consumption stream of d/π, $z \leq 1$ denotes the working period and γ is the preference parameter for leisure. Following Cremer and Pestieau (2003a) or Casamatta et al. (2005), we assume a quadratic specification for the disutility of work. This specification makes the problem more tractable at the cost that there are no income effects in labour supply. However, income effects in the retirement decision are found to be small when compared to substitution effects, see e.g., Krueger and Pischke (1992) or French (2005). Observe that the disutility of working is related to the fraction of the second period spent on working (i.e., z/π). This implies that for given retirement age an agent with a short life span experiences a higher disutility to work than an agent with a long life span because this agent has to work a larger share of his remaining life time.

[3]Zero interest rate and zero discount rate are assumed for clarity sake. We also abstract from population and productivity growth, which implies that the internal rate of return of the PAYG scheme equals the interest rate so that we can concentrate on the intragenerational redistribution effects of the PAYG scheme. These assumptions could easily be relaxed.

5.2.2 Innate ability and skill level

There are two levels of work skill, denoted by 'low' (L) and 'high' (H). Born low-skilled, an agent can acquire extra skills and become a high-skilled worker by investing $1 - a$ units of time in schooling in the first period. The rest of the time, a, is devoted to working as a high-skilled worker.

The individual-specific parameter a reflects the ability of individuals to acquire work skills. The higher is a, the more able is the individual, and the less time a worker needs for acquiring a work skill. The parameter a ranges between 0 and 1 and its cumulative distribution function is denoted by $G(\cdot)$, that is $G(a)$ is the number of individuals with an innate ability parameter below or equal to a. We henceforth refer to an individual with an innate ability parameter of a as an a-individual. For the sake of simplicity, we normalize the number of individuals born in each period to be one, that is: $G(1) = 1$.

It is assumed that a high-skilled worker provides an effective labour supply of one unit per unit of working time. A low-skilled worker provides only $q < 1$ units of effective labour for each unit of working time. This difference between effective labour supply also applies to the second period. Let w denote the wage rate per unit of effective labour. Then the maximum amount of income agents can earn in the first period, denoted $W_y(a)$, depends on the skill level and is defined as:

$$W_y(a) \equiv \begin{cases} qw & \text{for} \quad a \leq a^* \\ aw & \text{for} \quad a \geq a^* \end{cases} \tag{5.3}$$

with a^* the cut-off ability level so that every agent with an ability parameter above a^* will acquire skill and become a high-skilled worker, while all agents with an ability below a^* will not acquire education and remain low-skilled. At this stage, we assume that a^* is exogenous.[4] For the second period of life the maximum labour income, $W_o(a)$, equals:

$$W_o(a) \equiv \begin{cases} qw & \text{for} \quad a \leq a^* \\ w & \text{for} \quad a \geq a^* \end{cases} \tag{5.4}$$

5.2.3 Individual life span

Each individual lives completely his first period of life (with a length normalized to unity) but only a fraction $\pi(a) \leq 1$ of his second period. We assume

[4] In Appendix 5.B we work out the model with an endogenous schooling choice like in Razin and Sadka (1999). As shown, endogenizing the skill level does not change the main results derived in the body of this chapter.

that $\pi'(a) \geq 0$: the higher the innate ability of an agent, the longer the length of life. As a consequence, our model contains a positive association between longevity and skill level. Since high-skilled agents earn a higher wage rate than low-skilled workers, the model is consistent with the empirical evidence that wages positively co-move with life expectancy.[5]

Whenever necessary to parameterize the function $\pi(a)$, we will use the following specification:

$$\pi(a) = \bar{\pi}\left[1 + \lambda(a - \bar{a})\right], \quad \lambda \geq 0 \tag{5.5}$$

where $\bar{a} \equiv \int_0^1 a \, \mathrm{d}G$ denotes the average ability level. This function has the following attractive properties. First, $\bar{\pi}$ represents the average duration of the second phase of life. Second, there is a positive link between ability and the length of life as long as $\lambda > 0$. Indeed, $\mathrm{Cov}(\pi, a) = \lambda \mathrm{Var}(a) \geq 0$. Third, consistent with international evidence (Pappas et al., 1993; Mackenbach et al., 2003; Meara et al., 2008), the relative differences in individual life spans remain constant if the average life span increases. In absolute terms this means that the socioeconomic gap in longevity then increases, i.e., $\pi(a = 1) - \pi(a = 0) = \lambda\bar{\pi}$.

5.2.4 Consumption and retirement

An individual faces the following intertemporal budget constraint:

$$c + d = (1 - \tau)W_y + (1 - \tau)zW_o + P \tag{5.6}$$

where τ is the social security contribution tax rate and P denotes total pension entitlements to be received during old age.[6]

Maximizing life-time utility (5.1) over c, d and z, subject to the life-time budget constraint (5.6) yields the following first-order conditions:

$$u'(c) = u'(x) \tag{5.7}$$

$$(1 - \tau)W_o = \frac{\gamma z}{\pi} \tag{5.8}$$

[5]See Adams et al. (2003) for an extensive listing of studies dealing with the association of socioeconomic status and longevity. See also Section 1.1 of Chapter 1.

[6]Throughout this chapter, it is assumed that non-uniform lump-sum transfers are not available (because individual abilities and life spans are not publicly observable). The skill levels and wages are discrete while ability is continuous. Hence, individual abilities cannot be inferred from skill level or wages.

Equation (5.7) is the standard consumption Euler equation. Equation (5.8) is the optimality condition regarding retirement and states that the marginal benefit of working (net wage rate) should be equal to the marginal cost of working (disutility of labour). From these first-order conditions, we obtain the following expressions for c and z for the benchmark model:

$$c = \frac{1}{1+\pi}\left[(1-\tau)W_y + \frac{(1-\tau)^2 W_o^2 \pi}{2\gamma} + P\right] \tag{5.9}$$

$$z = \frac{(1-\tau)W_o\pi}{\gamma} \tag{5.10}$$

Note that the social security tax distorts the retirement decision: the larger the tax rate τ is, the faster agents leave the labour market, i.e., the lower z, because it reduces the price of leisure. Notice further that our disutility specification ensures that the retirement period is proportional to longevity, i.e., $\pi - z = [1-(1-\tau)W_o/\gamma]\pi$, like in Andersen (2005). Hence, a longer life span is split between later retirement and a longer retirement period. Compared to high-skilled workers, low-skilled workers retire earlier for two reasons. First, since $q < 1$, low-skilled people have a lower wage rate (substitution effect). Second, low-skilled workers will generally have a shorter life span which induces them to leave the labour force earlier (disutility of labour effect).

Equation (5.10) also implies that the labour-supply response of high-skilled agents is higher than that of low-skilled agents after an increase in the average life span.[7] Although the model does not explicitly differentiate between total life spans and life spans without disabilities, we can link this result to actual developments in life expectancy. There is some evidence that in terms of life expectancy *in good health* the socioeconomic gap has widened in the last decade (see Chapter 1). This means that high-skilled agents may be better able to continue working than low-skilled workers in response to an increase in *total* life expectancy.

5.2.5 Social security

The PAYG social security scheme is of the Beveridgean type with defined contributions.[8] Agents receive a flat pension benefit b which starts at the statutory retirement age h and lasts until the end of the individual old-age period

[7] Note that $\partial z(a)/\partial\bar{\pi}$ is increasing in the ability parameter a.

[8] At this point, we deviate from the Dutch first pension pillar (AOW) which is characterized by defined benefits.

π. Hence total pension entitlements P are equal to:

$$P = (\pi - h)b \tag{5.11}$$

A feasible social security pension scheme must satisfy the following resource constraint:[9]

$$\int_0^1 P \,\mathrm{d}G = \tau q w \int_0^{a^*} (1 + z_L)\,\mathrm{d}G + \tau w \int_{a^*}^1 (a + z_H)\,\mathrm{d}G \tag{5.12}$$

Using equations (5.5) and (5.11), we can rewrite this equation as:

$$b(\bar{\pi} - h) = \tau q w \int_0^{a^*} (1 + z_L)\,\mathrm{d}G + \tau w \int_{a^*}^1 (a + z_H)\,\mathrm{d}G \tag{5.13}$$

This condition states that the total amount of pension benefits paid out (left-hand side) has to be equal to the total amount of tax contributions received (right-hand side).[10] Notice that the first term on the right-hand side are the tax payments of the low-skilled workers and the second term are the payments of the high-skilled workers. Throughout, we assume that the contribution tax rate (τ) is fixed, which implies that the implications of pension reforms for the budget constraint are absorbed by the benefit level (b).

As a measure for redistribution, we calculate the net benefit of participating in the pension scheme. The net benefit is the difference between the present value of pension benefits and tax contributions:

$$NB \equiv (\pi - h)b - \tau(W_y + zW_o) \tag{5.14}$$

If the present value of pension benefits exceeds contributions (positive net benefit), an agent is a net beneficiary. Otherwise, the agent is a net contributor. *A priori* it is not immediately clear that the low-skilled agents are the net beneficiaries of this Beveridgean pension scheme. On the one hand, low-skilled agents have a lower wage rate and generally retire earlier than high-skilled agents. These factors imply that the low-skilled benefit from the Beveridgean scheme. On the other hand, low-skilled agents also die earlier than high-skilled agents which works in the opposite direction and implies that low-skilled agents are negatively affected by the pension scheme.

[9] In this chapter, subscript 'L' refers to low-skilled workers and subscript 'H' refers to high-skilled workers.

[10] We impose that $\pi - h > 0$ for any a-individual. In other words, nobody passes away before the statutory retirement age.

Using the definition of net benefit, equation (5.14), the budget constraint of the pension scheme can be rewritten in the following way:

$$\int_0^{a^*} NB_L \, \mathrm{d}G + \int_{a^*}^1 NB_H \, \mathrm{d}G = 0 \tag{5.15}$$

This equation states that the sum of the net benefits of all (young) individuals is equal to zero which reflects the zero-sum game nature of the pension scheme.[11]

5.3 Linking pensions to longevity

In this section, we analyse the redistribution effects of balanced-budget reforms that automatically adjust benefits or the pension entitlement age to increasing longevity. There is some controversy in the literature whether socioeconomic differences in longevity are constant over time in absolute or relative terms. We have assumed constant relative differences in longevity (see Section 5.2.3), but at the end of this section we will explain how the redistribution effects change if we instead assume constant absolute differences.

In the model, changes in longevity affect welfare of agents via two different channels. First, there is a direct utility effect that runs through the preference specification itself. Second, there is an indirect utility effect that is related to changes in the redistribution effects. At this point, we are only interested in this second channel and we will not take into account that a longer life span as such is something that raises utility. We will take the perspective of a young agent who enters the scheme, thereby focusing on the *structural* intragenerational redistribution effects.

5.3.1 Adjusting benefits to increasing longevity

We start to analyse the redistribution effects of an automatic link between the level of pension benefit and longevity. Countries like Finland and Portugal have implemented this kind of financial-sustainability adjustment (OECD, 2007).

[11] With a positive interest rate, the sum of net benefits would be negative because then all future generations have to pay for the gain given to the old generation at the time the pension scheme has been introduced (see also Section 2.2.1 of Chapter 2).

Proposition 5.1 *For any degree of life-span heterogeneity, an increase in average longevity absorbed by a budget-neutral reduction in benefits increases the redistribution from high-skilled agents to low-skilled agents. That is:*

$$\int_0^{a^*} \frac{\partial NB_L}{\partial \bar{\pi}}\, dG > 0 \quad \Leftrightarrow \quad \int_{a^*}^1 \frac{\partial NB_H}{\partial \bar{\pi}}\, dG < 0$$

Proof See Appendix 5.A.1. ∎

Why do low-skilled agents benefit from an increase in longevity at the expense of high-skilled agents? Higher longevity leads to two opposite effects on pension entitlements. On the one hand, it increases pension entitlements because agents receive benefits over a longer time period. On the other hand, a longer life span automatically reduces the per-period benefit level, which lowers pension entitlements. For agents with short life spans (i.e., the low-skilled), the relative increase in the payout period is larger than the relative decrease in the pension benefit because this is based on the *average* payout period. For high-skilled agents the opposite holds, meaning that the relative increase in the payout period is lower than the relative decline in the benefit.

In addition to this, an increase in longevity also affects retirement behaviour. Recall that the retirement period is proportional to the individual life span. Hence, when longevity increases, both the high-skilled and low-skilled agents will extend the working period. Since the high-skilled agents earn a higher wage rate, this longer working period increases the redistribution from high-income earners to low-income earners. This additional redistribution exacerbates the higher net benefits of the low-skilled agents.[12]

Although the analysis of this policy is rather stylized, it clearly shows that uniform rules as we observe in practice may affect individuals very differently. If the government would avoid this, it has to use non-uniform policies. However, there are practical problems to do this because it requires that socioeconomic characteristics of people are publicly observable. With regard to the policy considered in this section, a benefit-neutral approach would imply that the pension benefits of the low-skilled have to be reduced more than those of the high-skilled.

[12]Recall that the model does not take wealth effects on retirement into account, but empirical studies typically find modest wealth effects (Krueger and Pischke, 1992; French, 2005).

5.3.2 Adjusting the pension age to increasing longevity

Instead of linking pension benefits to longevity, the government can also adjust the pension age. Denmark, for example, has introduced a direct link between increasing life expectancy and the pension entitlement age.

Proposition 5.2 *An increase in average longevity absorbed by a budget-neutral adjustment of the pension entitlement age decreases the redistribution from high-skilled to low-skilled agents if heterogeneity in individual life spans is relatively high ($\lambda > \lambda^*$). If heterogeneity in life spans is low ($\lambda < \lambda^*$), the opposite holds. That is:*

$$\int_0^{a^*} \frac{\partial NB_L}{\partial \bar{\pi}} \, \mathrm{d}G \gtreqless 0 \quad \Leftrightarrow \quad \int_{a^*}^1 \frac{\partial NB_H}{\partial \bar{\pi}} \, \mathrm{d}G \lesseqgtr 0 \quad \textit{iff} \quad \lambda \lesseqgtr \lambda^*$$

with the threshold value λ^* defined in equation (5.51).

Proof See Appendix 5.A.2. ∎

It turns out that an increase in the statutory retirement age can be disadvantageous for high-skilled as well as low-skilled agents. The crucial factor is the degree of life-span heterogeneity. If the heterogeneity in life spans is relatively small (i.e., $\lambda < \lambda^*$), the reform is beneficial for the low-skilled. If heterogeneity is relatively large (i.e., $\lambda > \lambda^*$), the reform is beneficial for the high-skilled.

What is the intuition behind this result? The net benefit of agents is again subject to two opposite forces. On the one hand, the increase in life span is in absolute terms higher for high-skilled agents than for low-skilled agents as long as there is life-span heterogeneity ($\lambda > 0$). This implies that the payout period of the high-skilled increases more than that of the low-skilled. On the other hand, like discussed before, an increase in longevity induces people to retire later (and, hence, to pay more taxes) which increases the redistribution effects from high-income earners to low-income earners. If heterogeneity in life spans is low, this last effect is dominating meaning that high-skilled agents suffer from the reform, and *vice versa*.[13]

As the model provides ambiguous theoretical results, a numerical analysis is needed to determine which case is most realistic, $\lambda < \lambda^*$ or $\lambda > \lambda^*$. It should be emphasized that the stylized model set-up with only two generations and

[13] If agents would not spend part of the additional life time at the labour market, this second effect is not there implying that the high-skilled agents benefit for any value of $\lambda > 0$.

two skill levels cannot provide more than a rough indication of the quantitative impact. That being said, suppose that the average longevity increases and the government adjusts the official retirement age in such a way that the social security scheme remains sustainable. Figure 5.1a shows the redistribution effects of this policy in terms of the absolute change in the net benefit of the young agents. This graph compares the model solution with $\bar{\pi} = 0.8$ and $\bar{\pi} = 0.7$ (default value) which corresponds to an increase in the average life span from 81 to about 84 years.[14] The rest of the parameterization is as follows. The tax rate τ is 30%, $w = 1$ and $\gamma = 2$. We assume $h = 1/6$ which implies an initial retirement age of 65. The parameter λ is calibrated such that the difference between the life span of high-skilled and low-skilled agents is at most 3.5 years which is consistent with recent Dutch population estimates. This gives $\lambda = 1/6$.[15] According to recent figures of Statistics Netherlands, about two third of the Dutch population is low-skilled ($a^* = 2/3$) and these people earn about sixty percent less than high-skilled agents ($q = 0.6$). Finally, we assume that ability a follows a uniform distribution, i.e., $G(a) = a$, and that the utility function is logarithmic, i.e., $u(\cdot) = \ln(\cdot)$.

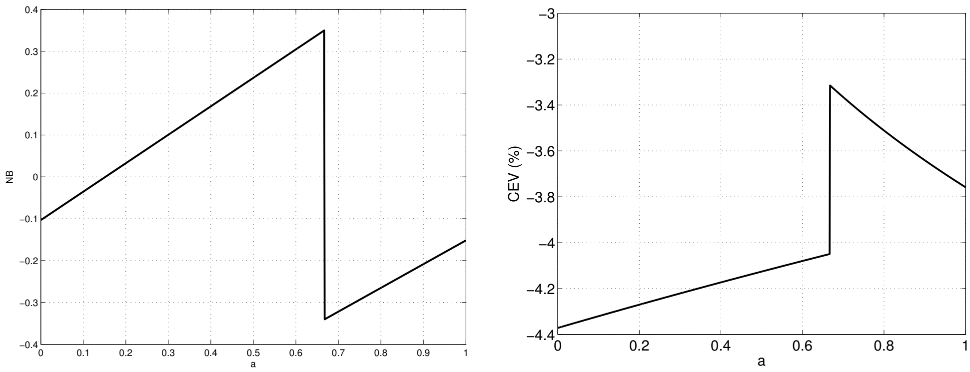
Figure 5.1a shows that an increase in the retirement age has positive effects on the net benefit of (most of the) low-skilled agents and negative effects on that of the high-skilled agents, which means that we are in the situation $\lambda < \lambda^*$. Given the parameter values used, the threshold value is $\lambda^* = 0.37$, which implies that the redistribution results turn around if the difference in life span between the high-skilled and low-skilled group is more than seven years.

The structural utility effects of the reform are presented in Figure 5.1b. Welfare is measured in consumption-equivalent variation (CEV): we ask what percentage of extra consumption an agent would require in the benchmark situation to be as well off as in the flexibility reform. Positive (*negative*) numbers thus indicate welfare gains (*losses*) from the reform. On the one hand, an increase in longevity leads to positive utility effects because it increases life-time income due to the fact that people postpone retirement and receive benefits over a longer time horizon. On the other hand, there are also negative utility effects because the necessary increase in the statutory retirement

[14] Life time consists of 30 years childhood that are not accounted for, 30 years of full potential working time (which can partly be used for tertiary education), and a last period of 30 years. Hence, the average life span is $60 + 30\bar{\pi}$.

[15] We interpret the high skill level as the highest attainable educational levels in the Netherlands (i.e., higher vocational training and academic level) and the low skill level as the collective term of all remaining educational levels.

Figure 5.1: Increase in retirement age in response to ageing



(a) Net benefit (b) Welfare

Notes: Net benefit is expressed in absolute difference from the benchmark and certainty-equivalent variation is expressed in percent of benchmark consumption.

age reduces life-time pension income and, more importantly, individuals have to share total life-time consumption over more periods.[16] As shown by Figure 5.1b, this last effect is dominating for all agents.

5.3.3 Life-span heterogeneity

So far, we have assumed that an increase in the average life span by a certain percentage corresponds to an increase in all individual life spans by the same percentage. This implies that in absolute terms the life span of high-ability agents increases more than that of low-ability agents when the average life span increases. This observation is consistent with many empirical studies, see e.g., Pappas et al. (1993), Mackenbach et al. (2003) or Meara et al. (2008). However, there is also recent evidence based on Dutch data which suggests that (at least for the Netherlands) longevity differences are more or less constant in absolute terms (see Chapter 1). In terms of our model, this property would be supported by the specification $\pi(a) = \bar{\pi} + \lambda(a - \bar{a})$.

Using this alternative specification, it turns out that an increase in longevity

[16] With log utility, $u(\cdot) = \ln(\cdot)$, the utility effect of the reform is given by the following derivative:

$$c\frac{\partial U}{\partial \bar{\pi}} = \frac{(1-\tau)^2 W_o^2}{2\gamma}\frac{\partial \pi}{\partial \bar{\pi}} + b\frac{\partial \pi}{\partial \bar{\pi}} - b\frac{\partial h}{\partial \bar{\pi}} - c\frac{\partial \pi}{\partial \bar{\pi}}$$

In order to compare life-time utility before and after ageing properly, we take π constant in equation (5.1). So we do not take into account that a higher life span is something that is nice for people.

is still beneficial (*harmful*) for the low-skilled (*high-skilled*) agents when this increase is absorbed by a decrease in the pension benefit. However, we now also get this result if this increase is absorbed by raising the statutory retirement age. That means, irrespective of the degree of life-span heterogeneity, raising the statutory retirement age as an automatic response to higher longevity benefits (*harms*) the low-skilled (*high-skilled*) agents.[17] The reason is that the absolute increase in the payout period now is the same for each individual and low-skilled agents benefit from the additional income redistribution due to the longer working period of the high-skilled.

5.4 Pension flexibility reforms

Besides implementing direct links between pensions and life expectancy to restore fiscal sustainability, countries have also taken measures to increase work incentives and to stimulate people voluntarily to continue working. In this section, we consider the welfare and redistribution effects of a more fundamental reform that allows for a flexible starting date of social security benefits, as recently implemented in e.g., the UK, Finland and Denmark. Introducing a variable starting date for benefits may help individuals to adjust the timing of pension income according to their own preferences. We will show that flexible pensions can also help to bear the costs of ageing or to reduce unintended transfers from low-skilled to high-skilled individuals.

In the benchmark model, we have assumed that social security benefits start at the statutory retirement date, irrespective of the individual's effective retirement date. Now we change this and impose that the benefits start at the time the individual actually leaves the labour market. If a person then retires later than the statutory retirement age, he receives an increment to his benefits for later retirement. If this person retires earlier, he receives a decrement. The imposed coincidence of pension take-up and retirement is not an unrealistic assumption because in practice flexible pension schemes often contain legal restrictions to continue work after a person has opted for benefits.[18]

[17] When $\pi(a) = \bar{\pi} + \lambda(a - \bar{a})$, the terms including parameter λ disappear from equation (5.48) and equation (5.49). Then it immediately follows that $\int_0^{a^*} \partial NB_L / \partial \bar{\pi} \, \mathrm{d}G > 0$ and $\int_{a^*}^1 \partial NB_H / \partial \bar{\pi} \, \mathrm{d}G < 0$ for any value of λ.

[18] In countries like Portugal, Spain and France the coincidence of pension take-up and retirement is regulated by law (Van Vuuren, 2011). In Dutch flexible second-pillar schemes the access to pension benefits is also conditional on dismissal.

5.4.1 Actuarial adjustment of benefits

To make this flexible pension take-up concrete, suppose that the government pays benefits p to the individual over its effective retirement period. Hence, total pension entitlements are equal to $P = (\pi - z)p$. Pension earnings per retirement period p are then:

$$p = m(z, \hat{\pi})b \tag{5.16}$$

where b is the reference flat pension benefit independent of contributions and labour history. The conversion factor m is the actuarial adjustment factor which determines to what extent the reference benefit b will be adjusted when agents retire later (or earlier) than the statutory retirement age. That is,

$$m(z, \hat{\pi}) = \frac{\hat{\pi} - h}{\hat{\pi} - z} \tag{5.17}$$

where we impose $\hat{\pi} - z > 0$ to make sure that $m(\cdot) > 0$ to rule out negative pension benefits. The adjustment factor is equal to the ratio between the average statutory retirement period and the individual effective retirement period measured by the reference life-span parameter $\hat{\pi}$ (to be specified below). At the individual level, actuarial non-neutrality arises if $\hat{\pi}$ differs from π. The function $m(\cdot)$ is increasing in the individual retirement decision z: if an agent decides to continue work after the statutory retirement age, the pension benefit in the remaining retirement periods will be adjusted upward.

We consider three different cases for the life span to be used in the adjustment factor which differ with respect to the information set available to the government. We first assume in Section 5.4.2 that the government applies uniform actuarial adjustment based on the average life span of the population ($\hat{\pi} = \bar{\pi}$). Second, in Section 5.4.3 the government instead uses individual-specific adjustment factors based on individual life spans ($\hat{\pi} = \pi$). Finally, in Section 5.4.4 the adjustment factor is made conditional on skill level with the following life span indicator:

$$\hat{\pi} = \begin{cases} \bar{\pi}_L \equiv \int_0^{a^*} \frac{\pi}{G(a^*)} \mathrm{d}G & \text{if} \quad a < a^* \\ \bar{\pi}_H \equiv \int_{a^*}^1 \frac{\pi}{1-G(a^*)} \mathrm{d}G & \text{if} \quad a > a^* \end{cases} \tag{5.18}$$

The adjustment factor for the low-skilled agents will be based on the average life span of the low-skilled group ($\hat{\pi} = \bar{\pi}_L$) and that for the high-skilled agents will be based on the average life span of the high-skilled people ($\hat{\pi} = \bar{\pi}_H$).

5.4.2 Uniform actuarial adjustment of benefits

Individual life spans are difficult to observe in practice. Therefore, real-world pension schemes with a flexible starting date for benefits always rely on uniform actuarial adjustment factors based on some average life expectancy index. In this section we show that this *uniform* adjustment of benefits can increase welfare of all individuals, i.e., induce a Pareto improvement, although individuals are *heterogeneous*.

Actuarial adjustment factor

With uniform adjustment, the reference life-span index is the same for each agent, $\hat{\pi} = \bar{\pi}$, so that the adjustment factor and pension entitlements are:

$$m = \frac{\bar{\pi} - h}{\bar{\pi} - z} \tag{5.19}$$

$$P = \frac{(\pi - z)(\bar{\pi} - h)b}{\bar{\pi} - z} \tag{5.20}$$

Now $m = 1$ for each individual who retires at the statutory retirement age h, implying that $p = b$; agents who retire later than h receive a higher benefit, $p > b$, and agents who retire earlier receive less, $p < b$.

From equation (5.20) we observe that, *ceteris paribus*, total pension entitlements of agents with long life spans are higher than the entitlements of agents with short life spans. This redistribution implies that the pension scheme is not actuarially neutral at the individual level. As the amount of pension entitlements depends on the individual retirement age, uniform actuarial adjustment introduces selection effects in the retirement decision. To show this, we derive from equation (5.20):

$$\Psi(z) \equiv \frac{\partial P(z)}{\partial z} = \frac{(\pi - \bar{\pi})p}{\bar{\pi} - z} \tag{5.21}$$

For agents with above-average life spans ($\pi > \bar{\pi}$), $\Psi > 0$, implying that these agents have an incentive to postpone retirement as this will increase their lifetime pension income. From an actuarial point of view, the conversion factor of these agents is too high. For short-lived people (with $\pi < \bar{\pi}$) it is just the opposite; for these agents the conversion factor of continued activity is too low which stimulates early retirement. For these people postponing retirement would simply mean that total pension entitlements decrease ($\Psi < 0$).

Consumption and retirement

With flexible pension take-up and uniform actuarial adjustment, the life-time budget constraint of the a-individual is still equal to equation (5.6), but now P is defined as in equation (5.20). Only the first-order condition regarding retirement changes:

$$(1-\tau)W_o + \Psi(z) = \gamma \frac{z}{\pi} \tag{5.22}$$

with $\Psi(z)$ given by equation (5.21). Consumption and retirement are then equal to:

$$c_{uni} = c_{ben} + \frac{1}{1+\pi}\left[P_{uni} - P_{ben} - \frac{[\Psi(z)]^2\pi}{2\gamma}\right] \tag{5.23}$$

$$z_{uni} = z_{ben} + \frac{\Psi(z_{uni})\pi}{\gamma} \tag{5.24}$$

where benchmark consumption (c_{ben}) and retirement (z_{ben}) are defined by equations (5.9) and (5.10).[19] Equation (5.24) shows that there is an extra distortion in retirement behaviour. Like before, we have that the contribution rate induces early retirement (through its impact on z_{ben}). The redistribution effects, represented by Ψ, imply an additional distortion in the retirement decision. This redistribution distortion can either stimulate retirement or depress retirement, depending on the individual life span π. For individuals with below-average life spans ($\pi < \bar{\pi}$), $\Psi < 0$, which implies that these people advance retirement as a result of uniform actuarial adjustment. If individuals have above-average life spans ($\pi > \bar{\pi}$), then $\Psi > 0$, and these people will postpone retirement.

Consumption can either be higher or lower compared to consumption in the benchmark. The last term in equation (5.23) is negative and reflects the utility loss resulting from the redistribution distortion in the retirement decision. Of course, flexibility can also induce a utility gain because an agent can choose the retirement age which gives him the highest entitlements. This potential gain is captured by the term $P_{uni} - P_{ben}$. Note from equations (5.11) and (5.20) that total pension benefits are generally not the same in the benchmark scheme and in the flexibility reform with uniform adjustment.[20]

[19] In the remaining of this chapter, subscript 'uni' refers to *uniform* actuarial adjustment, subscript 'ind' to *individual* adjustment, 'edu' to *educational* adjustment and, finally, 'nan' to *not actuarially-neutral* adjustment. We only use these subscripts if it is strictly necessary, i.e., in equations in which we compare one of the flexibility reforms with the benchmark case.

[20] This difference is not only due to the direct effect of a different adjustment factor, but also

Welfare effects

The welfare effects are not trivial because, compared to the benchmark model, uniform adjustment introduces another distortion in the retirement decision which can work into the opposite direction of the existing distortion related to the contribution tax. We will show, however, that under certain conditions this reform can lead to a Pareto improvement.

Suppose that the reform takes place unexpectedly. How will this affect utility of the currently old generation? If the reform would not take place, consumption of this generation would be equal to:

$$\pi x_{ben} = s_{ben} + \frac{(1-\tau)^2 W_o^2 \pi}{2\gamma} + P_{ben} \tag{5.25}$$

with savings equal to $s = (1-\tau)W_y - c$. After the reform, the first-order condition for the retirement decision of the old generation is given by equation (5.22). Using this condition, old-age consumption after the reform is:

$$\pi x_{uni} = s_{ben} + \frac{(1-\tau)^2 W_o^2 \pi}{2\gamma} + P_{uni} - \frac{[\Psi(z)]^2 \pi}{2\gamma} \tag{5.26}$$

The old generation is not worse off after the reform if $u(x_{uni}) - u(x_{ben}) \geq 0$, implying:

$$\pi x_{uni} - \pi x_{ben} \geq 0 \;\Rightarrow\; P_{uni} - P_{ben} - \frac{[\Psi(z)]^2 \pi}{2\gamma} \geq 0 \tag{5.27}$$

The current young and future generations are better off when $U(c_{uni}, x_{uni}) \geq U(c_{ben}, x_{ben})$ for each ability level, which implies, using equation (5.7), $c_{uni} \geq c_{ben}$. From equation (5.23) we can see that the condition for young and future generations is exactly the same as that for the current old generation. This is due to the fact that there are no income effects in the retirement decision. Consequently, for a given ability level the transition generation and all future young generations retire at the same age and thus have the same amount of life-time income. Hence, when condition (5.27) is satisfied and is strictly positive for at least one a-individual, the reform is Pareto improving. To analyse the possibility of a Pareto improvement we make the following assumption:

Assumption 1 *The statutory retirement age is equal to the retirement age of the individual with the average ability level, i.e.,* $h = z(\bar{a})$.

due to the effect of the adjustment factor on the retirement decisions which, via the budget constraint of the PAYG scheme, will in general lead to a different reference pension benefit b.

This assumption implies that individuals with below-average life span have an incentive to advance retirement as from an actuarial point of view the adjustment factor of retirement postponement is too low for them. Therefore, for these people retiring *after* the statutory retirement age is not in their interest, *ceteris paribus*, as it reduces pension entitlements compared to the benchmark. For individuals with above-average life span exactly the opposite holds. These individuals have an incentive to postpone retirement because the actuarial adjustment factor is too high for them. Hence, retiring *before* the statutory retirement is not in their interest.

Suppose Assumption 1 is satisfied. Then we can derive the following result:

Proposition 5.3 *A pension reform from inflexible Beveridgean pensions towards flexible Beveridgean pensions with uniform actuarial adjustment of pension benefits is a Pareto improvement if and only if $\tau \geq \tau^*$, with τ^* (approximately) equal to:*

$$\tau^* = \frac{(\gamma - qw)\sqrt{\gamma - w} - (\gamma - w)\sqrt{\gamma - qw}}{w\sqrt{\gamma - qw} - qw\sqrt{\gamma - w}} \tag{5.28}$$

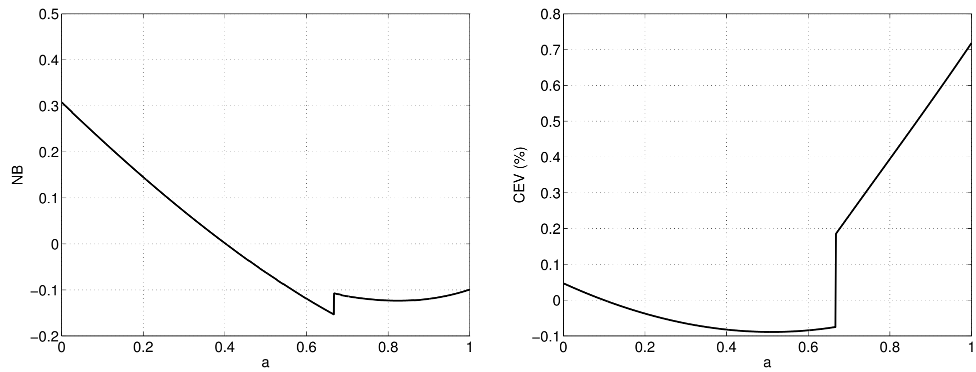
Proof See Appendix 5.A.3. ∎

The intuition for this result is as follows. High-skilled workers certainly gain from this reform because the adjustment factor is too high from an actuarial perspective. This leads to a lower implicit tax on continued activity and thus later retirement. The welfare of low-skilled workers in principle declines because they are confronted with higher implicit taxation. The only way to compensate for this loss is to give the low-skilled more social security benefits. If the contribution tax rate is sufficiently high, it is indeed possible that the continued activity of the more able generates enough resources to compensate the less able so that ultimately the welfare of all agents is higher.[21]

Uniform actuarial adjustment of benefits gives the government an instrument to reduce distortions on the extensive margin of labour supply. This is therefore one way to obtain additional resources that can be used to meet

[21] We implicitly assume here that there are no interaction effects between the intensive and extensive margin of labour supply. High-skilled people can decide to reduce the number of hours working if they know that they will retire later, which lowers the tax revenues for the government. However, one can ask if this effect is sufficiently strong in practice because there is evidence that the labour supply of younger workers is less elastic than that of older workers (see e.g., Fenge et al., 2006).

Figure 5.2: Uniform adjustment: redistribution and welfare



(a) Net benefit (b) Welfare

Notes: Net benefit is expressed in absolute difference from the benchmark and certainty-equivalent variation is expressed in percent benchmark consumption.

the increasing fiscal burden of ageing.[22] Moreover, if the reform is conducted properly, it will also foster redistribution from rich to poor. Similar to Cremer and Pestieau (2003a), this 'double dividend' hinges on two conditions. First, the retirement decision in the benchmark needs to have a downward distortion, i.e., retirement is too early, and the removal of this distortion therefore brings additional resources. Second, the pension contract needs to be redistributive from rich to poor people so that most of the cost of the reform is borne by individuals with relatively high earnings.

From a policy perspective, an attractive feature of this reform is that it allows the low-skilled people to leave the labour market earlier than with a fixed pension take-up *without* a loss in life-time pension income. If life expectancy improvements of low-skilled people (especially in terms of good health) fall short of those of high-skilled people, it will generally be more difficult for governments to implement policies to postpone retirement as the physical ability of the low-skilled to work longer is limited. Against this background, this

[22]Using a similar model, Cremer and Pestieau (2003a) show that this efficiency gain can also be obtained by age-dependent taxation, by giving the young a higher tax rate than the old. When taking a broader perspective, age-dependent taxation could also be facilitated by uniform pricing in funded pension schemes. As shown in Chapter 2, uniform pricing imposes an implicit tax on the pension accrual of the young and a subsidy on that of the old. However, the advantage of the flexibility reforms as analysed in this chapter, is that these provide an additional instrument which specifically applies to the extensive margin of labour supply. The efficiency improvement can then also arise in a more general set-up that also includes the intensive margin (see Fisher and Keuschnigg, 2010).

reform towards variable pension take-up may provide more flexibility to the low-skilled to cope with future policy measures aimed to postpone retirement.

In Figure 5.2 we show a numerical illustration of the redistribution (left graph) and welfare effects (right graph) of a switch to a flexible scheme based on uniform actuarial adjustment. The underlying parameterization is the same as used earlier (and thus includes a tax rate τ of 30%, see Section 5.3.2). The welfare effects of all high-ability agents are positive. These agents benefit from a lower implicit tax on continued activity due to the (for them) attractive actuarial adjustment factor and therefore choose to work longer. However, the resulting additional tax contributions are not sufficient to compensate all low-skilled agents for the higher implicit tax they are confronted with, although most of them will experience an increase in the net benefit from the pension scheme. To achieve a Pareto improvement, the contribution rate needs to be at least 40%, that is $\tau^* = 0.4$.[23]

5.4.3 Individual actuarial adjustment of benefits

For those policy makers who are mainly interested in improving the welfare of the less wealthy people, the application of uniform adjustment factors might not fulfil their objectives. As we have seen, this reform in principle exacerbates the unintended transfers from the low-skilled to the high-skilled unless the contribution rate is sufficiently high at the outset. There are different ways to reduce this regressive implication though. To set the scene, we first take a rather extreme position and hypothetically assume that the government can observe individual life spans. The government can then use this information in the assessment of the adjustment of benefits.

Actuarial adjustment factor

With individual adjustment, $\hat{\pi} = \pi$, the individual-specific adjustment factor m and the pension entitlements P become:

$$m = \frac{\bar{\pi} - h}{\pi - z} \tag{5.29}$$

$$P = (\bar{\pi} - h)b \tag{5.30}$$

[23]This seems a rather high number for a tax rate primarily used for old-age pensions. However, in reality redistribution from low to high incomes also occurs in other parts of the economy, like the tax or public health care system. If we also take these kinds of redistribution channels into account, a tax rate of 40% is maybe not such an unrealistic threshold value.

Note from equation (5.29) that $m = 1$ for an agent with an average ability level ($a = \bar{a}$) who retires at the statutory retirement age h. For this 'average' individual the pension benefit is equal to the reference benefit level, i.e., $p = b$. If this person retires later than the statutory retirement age, then $m > 1$, which means that he receives a benefit which is adjusted upward, i.e., $p > b$. On the other hand, if the person retires earlier than the statutory retirement age, we have $m < 1$ implying $p < b$.

With individual adjustment, the retirement decision is actuarially neutral in the sense that the effective retirement date has no effect on the total pension entitlement. To see this, note from equation (5.30):

$$\frac{\partial P}{\partial z} = 0 \tag{5.31}$$

Hence, agents cannot increase their total pension entitlements by postponing or advancing retirement. Any individual, irrespective of life span, income or skill level, receives exactly the same amount of total pension benefits over the life time.

Consumption and welfare effects

Compared to the benchmark social security model, the retirement decisions are the same. Also the aggregate budget constraint of the pension contract does not change and, consequently, also the pension benefit per retirement period stays the same. The only thing in the model that changes are the consumption decisions which can be written as:

$$c_{ind} = c_{ben} + \frac{(\bar{\pi} - \pi)b}{1 + \pi} \tag{5.32}$$

From this equation we immediately infer the following result.

Proposition 5.4 *Introducing retirement flexibility using individual actuarial adjustment of pension benefits implies that the welfare of the short-lived agents ($\pi < \bar{\pi}$) increases while the welfare of the long-lived agents ($\pi > \bar{\pi}$) decreases. This reform therefore cannot be a Pareto improvement.*

Hence, individual actuarial adjustment removes redistribution related to life-span differences. Agents with short life spans (i.e., the low-skilled) therefore benefit from this reform at the expense of those agents with long life spans (i.e., the high-skilled).[24]

[24] Theoretically, applying individual-specific conversion factors for earlier and later retirement can result in a Pareto improvement if we would not restrict benefit adjustments to be

5.4.4 Skill-dependent actuarial adjustment of benefits

A practical impediment of the previous reform is that life spans are generally not observable at the individual level. A possible solution to this information problem is to base actuarial adjustment on characteristics which are (better) observable and are at least to some extent correlated with individual life expectancies (see, e.g., Bovenberg et al., 2006). In terms of our model, this characteristic could be the skill level of agents. By choosing to become high-skilled (or to remain low-skilled), agents partly reveal information about their life span. Indeed, the life span of high-skilled agents is generally higher than that of low-skilled agents. The government can use this information by conditioning the actuarial adjustment factor on skill level. This also reduces the redistribution from low-skilled to high-skilled people.

Actuarial adjustment factor

With skill-dependent adjustment, the reference life-span measure is conditional on skill group: $\hat{\pi} = \bar{\pi}_L$ for the low-skilled group and $\hat{\pi} = \bar{\pi}_H$ for the high-skilled group. The actuarial conversion factor is:

$$m = \begin{cases} \frac{\bar{\pi}-h}{\bar{\pi}_L - z_L} & \text{if} \quad a < a^* \\ \frac{\bar{\pi}-h}{\bar{\pi}_H - z_H} & \text{if} \quad a > a^* \end{cases} \tag{5.33}$$

and pension entitlements are equal to:

$$P = \begin{cases} \frac{(\pi - z_L)(\bar{\pi}-h)b}{\bar{\pi}_L - z_L} & \text{if} \quad a < a^* \\ \frac{(\pi - z_H)(\bar{\pi}-h)b}{\bar{\pi}_H - z_H} & \text{if} \quad a > a^* \end{cases} \tag{5.34}$$

From equation (5.33) it follows that skill-dependent adjustment reduces redistribution from short-lived to long-lived ability groups, like with individual actuarial adjustment. Indeed, suppose that all agents retire at the statutory retirement date h. Then we have for the low-skilled group that $m > 1$ while for the high-skilled group $m < 1$. Hence, low-skilled agents are compensated for the fact that they have a shorter life span. However, contrary to individual adjustment, skill-dependent adjustment does not remove redistribution completely. Therefore, these transfers will still lead to distortions in the retirement decision. To see this, from equation (5.34) we have:

$$\Psi(z) \equiv \frac{\partial P(z)}{\partial z} = \begin{cases} \frac{(\pi - \bar{\pi}_L)p}{\bar{\pi}_L - z_L} & \text{if} \quad a < a^* \\ \frac{(\pi - \bar{\pi}_H)p}{\bar{\pi}_H - z_H} & \text{if} \quad a > a^* \end{cases} \tag{5.35}$$

---

actuarially neutral at the individual level.

Within skill group, agents with relatively high life spans ($\pi > \hat{\pi}$) have an incentive to delay retirement and agents with relatively low life spans ($\pi < \hat{\pi}$) still want to retire earlier. However, these selection effects are lower than with uniform adjustment because the heterogeneity in life expectancy within skill groups is obviously lower than the life-time heterogeneity in the total population.

Consumption and retirement

The expressions for consumption and retirement are similar as the corresponding expressions in case of uniform actuarial adjustment of benefits. That is,

$$c_{edu} = c_{ben} + \frac{1}{1+\pi}\left[P_{edu} - P_{ben} - \frac{[\Psi(z)]^2\pi}{2\gamma}\right] \tag{5.36}$$

$$z_{edu} = z_{ben} + \frac{\Psi(z_{edu})\pi}{\gamma} \tag{5.37}$$

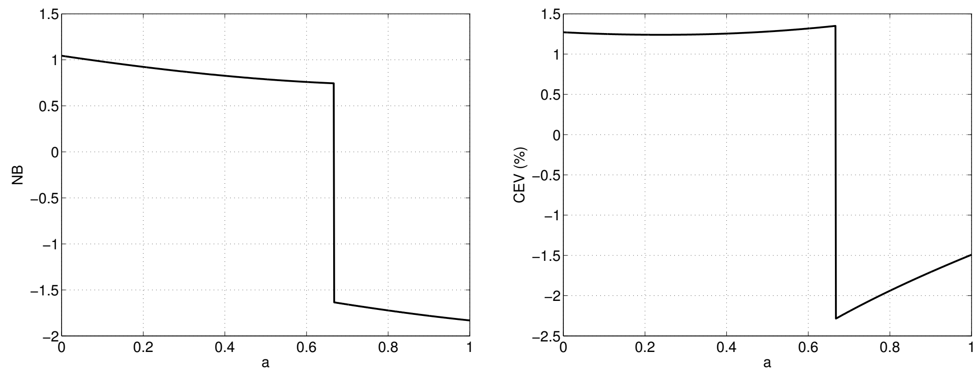
The only difference arises in the specification of the derivative Ψ because this is now based on the average life expectancy of the high-skilled or low-skilled group, see equation (5.35).

Like with uniform actuarial adjustment of benefits, retirement behaviour is again subject to two different kinds of labour-supply distortions. The first distortion is caused by the payroll tax rate τ which induces early retirement. The second distortion in the retirement decision is caused by the derivative Ψ which reflects the selection effects associated with life-span heterogeneity. Notice that the impact of this second distortion is smaller than under uniform actuarial adjustment because for the majority of the people the difference between the own life span and the skill-group average is smaller than the difference between the individual life span and the average life span of the total population.

Welfare effects

The application of skill-dependent conversion of benefits cannot result in a Pareto improvement. Since this reform removes (at least to some extent) the redistribution from short-lived to long-lived individuals, the welfare of low-skilled (*high-skilled*) agents raises (*falls*), as can also be seen in Figure 5.3. As the pension scheme is actuarially neutral on average for the high-skilled and low-skilled group, there are no selection effects that can induce efficiency improvements to compensate the losses of the high-skilled agents. In this re-

Figure 5.3: Skill-dependent adjustment: redistribution and welfare



(a) Net benefit (b) Welfare

Notes: Net benefit is expressed in absolute difference from the benchmark and certainty-equivalent variation is expressed in percent benchmark consumption.

spect, the welfare effects of skill-dependent adjustment are comparable with those of individual adjustment.

The following proposition summarizes this result.

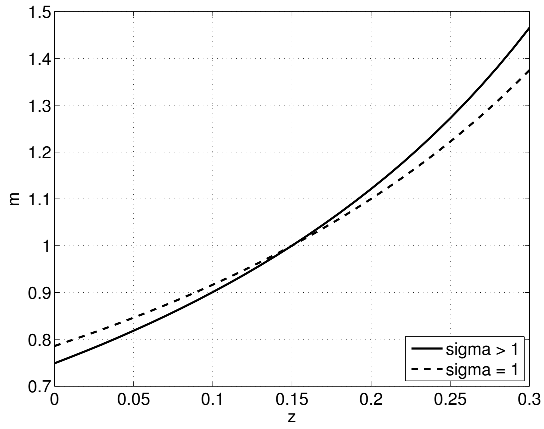
Proposition 5.5 *A pension reform from inflexible Beveridgean pensions towards flexible Beveridgean pensions with skill-dependent actuarial adjustment of pension benefits cannot be a Pareto improvement.*

Proof See Appendix 5.A.4. ∎

5.5 Introducing actuarial non-neutrality

From all flexibility reforms we have considered in the previous section, uniform adjustment is the only candidate that theoretically can improve welfare of both unskilled *and* skilled agents. A practical problem with this reform is that, to obtain this Pareto improvement, the redistributive social security scheme needs to be large (as represented by a high contribution rate). In this final section we will make the point that a Pareto improvement of a flexible pension take-up can also be achieved for lower tax critical rates. As will be shown, the key driver behind this result is the introduction of incentives to stimulate work continuation at an actuarially non-neutral way. Unfortunately, with life-time heterogeneity ($\lambda > 0$) it is not possible to show analytically that

Figure 5.4: Actuarial adjustment factor



actuarial non-neutrality provides a Pareto improvement. We therefore first abstract from life-time heterogeneity ($\lambda = 0$) and for this simplified case we prove that actuarial non-neutrality allows for a Pareto improvement if the tax rate passes a critical value. Then we numerically show that this tax critical rate is lower than for actuarial neutrality in the more general case with life-time heterogeneity.

5.5.1 Actuarial adjustment factor

If we abstract from life-span heterogeneity, each agent, irrespective of his ability level, lives a fraction $\pi \leq 1$ of the second period. Suppose now that the actuarial adjustment factor has the following specification:

$$m(z,\pi) = \left(\frac{\pi - h}{\pi - z}\right)^{\sigma}, \quad \sigma > 1 \tag{5.38}$$

The parameter σ governs the degree of actuarial non-neutrality of the adjustment factor, as shown in Figure 5.4. If $\sigma = 1$, the adjustment of pension benefits is completely actuarially neutral with respect to the retirement decision (see Section 5.4.3). This means that pension earnings are not sensitive to the retirement date. For $\sigma > 1$, the adjustment factor is *higher* than the actuarially-neutral level if agents retire *later* than the statutory retirement age ($z > h$). On the contrary, the adjustment factor is *lower* than the actuarially-neutral level if agents retire *earlier* than the statutory retirement age ($z < h$). In other words, specification (5.38) rewards delaying retirement and discourages early retirement as long as $\sigma > 0$.

Given the specification of the conversion factor, equation (5.38), the pension entitlements P are equal to:

$$P = (\pi - h)^{\sigma}(\pi - z)^{1-\sigma} b \tag{5.39}$$

Taking the derivative of P with respect to z then gives:

$$\Psi(z) \equiv \frac{\partial P(z)}{\partial z} = (\sigma - 1)p \tag{5.40}$$

Hence, if $\sigma > 1$ then $\Psi > 0$ meaning that introducing actuarial non-neutrality will give agents an incentive to continue working as this will increase pension entitlements.

5.5.2 Consumption and retirement

The consumption decision and retirement decision are equal to:

$$c_{nan} = c_{ben} + \frac{1}{1+\pi}\left[P_{nan} - P_{ben} - \frac{[\Psi(z)]^2 \pi}{2\gamma}\right] \tag{5.41}$$

$$z_{nan} = z_{ben} + \frac{\Psi(z_{nan})\pi}{\gamma} \tag{5.42}$$

where P and Ψ are defined by equations (5.39) and (5.40), respectively. Taking the derivative of the retirement choice with respect to the neutrality parameter σ gives, evaluated at $\sigma = 1$:

$$\frac{\partial z}{\partial \sigma} = \frac{\pi p}{\gamma} > 0 \tag{5.43}$$

An increase in the neutrality parameter σ leads to later retirement. Consequently, the introduction of this type of non-neutrality in the retirement decision can undo (at least to some extent) the distortionary effect of the social security tax. This result is comparable with the situation we had before, in the flexibility reform with uniform actuarial adjustment in combination with heterogeneous life spans. In that case, however, the pension scheme is still actuarially neutral on average: high-skilled workers (with a long life span) indeed receive a subsidy on continuing work whereas low-skilled workers (with a short life span) experience a tax on delaying retirement. The current reform is different because now the pension scheme subsidizes work continuation for *all* agents, irrespective of the skill level.

5.5.3 Welfare effects

Introducing actuarial non-neutrality in the benefit calculation does not only stimulate labour supply, it also leads to a Pareto improvement if the contribution tax rate is sufficiently high.

Proposition 5.6 *Introducing actuarial non-neutrality aimed at stimulating work effort makes high-skilled workers strictly better off. In addition, the reform is Pareto improving if and only if* $\tau > \hat{\tau}$, *with:*

$$\hat{\tau} = \frac{[1 - G(a^*)] \ln\left(\frac{\pi - z_L}{\pi - z_H}\right)}{G(a^*)\frac{qw\pi}{\gamma(\pi - z_L)} + [1 - G(a^*)]\,\frac{w\pi}{\gamma(\pi - z_H)}} \tag{5.44}$$

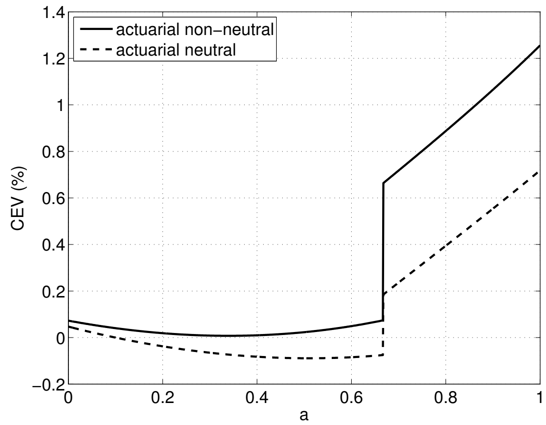
This implicit equation has a unique solution.

Proof See Appendix 5.A.5. ∎

The intuition for this result is similar as before, in the reform with uniform actuarial adjustment (see Section 5.4.2). The government can apply non-neutral actuarial conversion of benefits for late retirement as an instrument to increase the total efficiency of the economy. This subsidy reduces the existing labour-supply distortion on the extensive margin related to the contribution tax rate. With actuarial non-neutrality, however, the reward rate of retirement postponement is relatively more attractive for agents who retire later (which are the high-skilled), as can also be seen from Figure 5.4. Therefore, to ensure that the welfare of the low-skilled also improves, the contribution rate needs to be sufficiently high such that the additional tax payments of the high-skilled lead to higher pension benefits.

Figure 5.5 compares the welfare effects of uniform adjustment under actuarial neutrality (dashed line) and actuarial non-neutrality (solid line). Contrary to the analytical exposition discussed before, this figure is based on heterogeneous life spans. All parameter values are the same as those used in the previous figures. As we have concluded earlier (see Figure 5.2b), a contribution tax rate of 30% is not sufficient to ensure that uniform adjustment of benefits combined with actuarial neutrality is Pareto improving. However, if uniform adjustment will be combined with actuarial non-neutrality, the reform leads to strictly higher (and positive) welfare effects for all individuals. Hence, by introducing actuarial non-neutrality in the Beveridgean pension scheme, it is possible to achieve a Pareto improvement for more plausible contribution tax rates.

Figure 5.5: Neutral versus non-neutral actuarial adjustment



Notes: The actuarially non-neutral scenario is based on $\sigma = 1.05$.

5.6 Conclusion

In this chapter, we have studied the intragenerational redistribution and welfare effects of pension reforms that aim to improve fiscal sustainability as a response to ageing. We first considered reforms that link pension benefits or the pension entitlement age to longevity. Then we have focused on a more fundamental reform that introduces a flexible take-up of pension benefits. To analyse the redistribution and welfare effects of these reforms, we have developed a stylized two-period overlapping-generations model populated with heterogeneous agents who differ in ability and life span. The model includes a Beveridgean social security scheme with life-time annuities. Therefore, the empirically most important channels of intragenerational redistribution are taken into account: income redistribution from rich to poor and life-span redistribution from short-lived to long-lived agents.

We show that a direct link between pension benefits and longevity has positive effects on the net benefit of the low-skilled agents and negative effects on that of the high-skilled agents. For low-skilled (*high-skilled*) agents the relative decrease in the benefit is lower (*higher*) than the relative increase in the retirement period. The intragenerational redistribution effects of linking the pension entitlement age to longevity depend on the degree of life-span heterogeneity among low-skilled and high-skilled agents. Higher longevity partly induces people to retire later and therefore leads to an increase in the redistribution from high to low incomes. As long as individual differences in life span are not too high, this increased income redistribution dominates the opposite redistribution effects from short-lived to long-lived agents, implying that the unskilled agents (with short life spans and low income levels) benefit from this reform.

Our results suggest that introducing a flexible pension take-up with uniform adjustments can induce a Pareto improvement. This reform can collect additional resources without diminishing the welfare of low-skilled agents and increasing that of high-skilled agents. In that way it can also help to bear the costs of ageing in a Beveridgean pension scheme. The selection effects of uniform actuarial adjustment increase the implicit tax of the low-skilled but decrease the implicit tax of the high-skilled, who in turn decide to work longer and therefore pay more pension contributions. A necessary condition for such a Pareto improvement is that the contribution tax is sufficiently high so that the continued activity of the high-skilled generates enough tax revenues to compensate the low-skilled with higher benefits. Increasing the reward and penalty rates of later and earlier retirement in an actuarially non-neutral way can help to reduce the tax critical rate. This policy reduces the implicit tax of *all* agents (also that of the low-skilled), which means that the less-skilled agents need less compensation through the redistributive pension scheme.

In real-world pension schemes that have actuarial adjustment of pension entitlements, this adjustment is indeed independent of individual characteristics, like life expectancy or skill level. The results of this paper give a rationale for this kind of uniform flexibility reforms. In recent years, penalties and rewards for earlier or later retirement have increased in a number of countries (OECD, 2011). However, in most countries the implemented reductions of early pension benefits do still not fully correspond both to the lower amount of contributions paid by the worker and to the increase in the period over which the worker will receive pension payments (Queisser and Whitehouse, 2006). This implies that there is still room to improve the pension systems by going into the direction of complete actuarial neutrality or by moving even beyond that level, as our analysis of actuarially non-neutral adjustment suggests.

Our benchmark scheme is of the Beveridgean type and characterized by inflexible pension take-up and life-time annuities. Countries like the UK, the Netherlands and Denmark indeed follow this tradition. Other countries, like Germany, Italy and France have Bismarckian pension schemes where pension benefits are linked to former contributions. To obtain a Pareto improvement of introducing a variable starting date for pensions, we have argued that the benefit rule must satisfy two characteristics. First, it should contain an initial distortion, the removal of which brings additional resources. Second, it should have within-cohort redistribution such that most of the cost of the reform is born by the high-income people. In general, Bismarckian pension systems still contain intragenerational redistribution from short- to long-lived agents but have considerably less redistribution from rich to poor. We therefore expect that the Pareto-improving nature of a flexible pension take-up is much more difficult to realize in these types of pension schemes.

Our paper, however, provides a rationale why countries with Beveridgean pension schemes should use *uniform* rules for the adjustment of pension benefits when they introduce flexible pension take-up even though people have different skill levels and life expectancies. It is sometimes argued that it would be preferable to base the actuarial adjustment factor of pension benefits on individual life expectancy or skill level. This paper shows that even in a very simple setting the latter type of pension flexibility reform cannot be Pareto improving as some of the redistribution in the initial pension scheme (from the short- to long-lived) is removed. It is therefore important to take all types of redistribution in the initial pension scheme into account when discussing the implementation of flexible pension take-up. Applying uniform actuarial adjustment, possibly combined with non-neutral elements to increase the incentives to postpone retirement, could increase the economic efficiency of the pension system and in that way generate extra resources to cope with the costs of ageing *and* make some people better off while not hurting other people.

Appendix to Chapter 5

This appendix contains formal proofs of all propositions mentioned in this chapter (Section 5.A). It also present an extension of the basic model by endogenizing the schooling decision (Section 5.B).

5.A Proofs

5.A.1 Proof of Proposition 5.1

Proof From equation (5.14), using the definition of W_o, it follows:

$$\frac{\partial NB_L}{\partial \bar{\pi}} = (1 + \lambda a - \lambda \bar{a}) \underbrace{\left[\bar{\pi}\frac{\partial b}{\partial \bar{\pi}} + b - \frac{\tau(1-\tau)q^2 w^2}{\gamma}\right]}_{\Omega_L} - h\frac{\partial b}{\partial \bar{\pi}} \tag{5.45}$$

$$\frac{\partial NB_H}{\partial \bar{\pi}} = (1 + \lambda a - \lambda \bar{a}) \underbrace{\left[\bar{\pi}\frac{\partial b}{\partial \bar{\pi}} + b - \frac{\tau(1-\tau)w^2}{\gamma}\right]}_{\Omega_H} - h\frac{\partial b}{\partial \bar{\pi}} \tag{5.46}$$

In addition, from the budget constraint, equation (5.13), we derive:

$$\begin{aligned}\frac{\partial b}{\partial \bar{\pi}} = & -\frac{\tau q w}{(\bar{\pi}-h)^2}\int_0^{a^*}\left[1 + \frac{(1-\tau)qwh}{\gamma}\frac{\partial \pi}{\partial \bar{\pi}}\right] \mathrm{d}G \\ & -\frac{\tau w}{(\bar{\pi}-h)^2}\int_{a^*}^{1}\left[a + \frac{(1-\tau)wh}{\gamma}\frac{\partial \pi}{\partial \bar{\pi}}\right] \mathrm{d}G < 0\end{aligned} \tag{5.47}$$

Because $\pi(a) > 0$ we have $1 + \lambda(a - \bar{a}) > 0$ for each ability level a. Notice that $\Omega_L > \Omega_H$ because $q < 1$. Let $\Omega_H > 0$. Then it must be true that $\Omega_L > 0$ as well, but this contradicts the condition $\int_0^{a^*} \partial NB_L/\partial\bar{\pi}\,\mathrm{d}G + \int_{a^*}^1 \partial NB_H/\partial\bar{\pi}\,\mathrm{d}G = 0$. Hence, $\Omega_H < 0$. This result together with the assumption $\pi'(a) \geq 0$ implies that the maximum of $\partial NB_H(a)/\partial\bar{\pi}$ is smaller than the minimum value of $\partial NB_L(a)/\partial\bar{\pi}$. Therefore, $\int_0^{a^*} \partial NB_L/\partial\bar{\pi}\,\mathrm{d}G = -\int_{a^*}^1 \partial NB_H/\partial\bar{\pi}\,\mathrm{d}G > 0$. ∎

5.A.2 Proof of Proposition 5.2

Proof Taking the derivative of the net benefit with respect to $\bar{\pi}$ gives:

$$\frac{\partial NB_L}{\partial\bar{\pi}} = (1+\lambda a - \lambda\bar{a})\underbrace{\left[b - \frac{\tau(1-\tau)q^2w^2}{\gamma}\right]}_{\Delta_L} - b\frac{\partial h}{\partial\bar{\pi}} \tag{5.48}$$

$$\frac{\partial NB_H}{\partial\bar{\pi}} = (1+\lambda a - \lambda\bar{a})\underbrace{\left[b - \frac{\tau(1-\tau)w^2}{\gamma}\right]}_{\Delta_H} - b\frac{\partial h}{\partial\bar{\pi}} \tag{5.49}$$

Using the budget constraint, equation (5.13), we derive:

$$\begin{aligned} b\frac{\partial h}{\partial\bar{\pi}} = {} & \frac{\tau q w h}{\bar{\pi}(\bar{\pi}-h)}\int_0^{a^*} z_L\,\mathrm{d}G + \frac{\tau w h}{\bar{\pi}(\bar{\pi}-h)}\int_{a^*}^1 z_H\,\mathrm{d}G \\ & + \frac{\tau q w}{\bar{\pi}-h}G(a^*) + \frac{\tau w}{\bar{\pi}-h}\int_{a^*}^1 a\,\mathrm{d}G > 0 \end{aligned} \tag{5.50}$$

Note that $\Delta_L > \Delta_H$ because $q < 1$. Suppose first that $\lambda = 0$. Then it directly follows from the zero-sum requirement $\int_0^{a^*} \partial NB_L/\partial\bar{\pi}\,\mathrm{d}G + \int_{a^*}^1 \partial NB_H/\partial\bar{\pi}\,\mathrm{d}G = 0$ that $\int_0^{a^*} \partial NB_L/\partial\bar{\pi}\,\mathrm{d}G > 0$ and $\int_{a^*}^1 \partial NB_H/\partial\bar{\pi}\,\mathrm{d}G < 0$. Suppose now that $\lambda = \lambda^*$, with λ^* such that $\int_0^{a^*} \partial NB_L/\partial\bar{\pi}\,\mathrm{d}G = \int_{a^*}^1 \partial NB_H/\partial\bar{\pi}\,\mathrm{d}G = 0$. This implies:

$$\lambda^* = \frac{G(a^*)\left(b\frac{\partial h}{\partial\bar{\pi}} - \Delta_L\right)}{\Delta_L \int_0^{a^*}(a-\bar{a})\,\mathrm{d}G} > 0 \tag{5.51}$$

To show that $\lambda^* > 0$, we first derive from equation (5.13),

$$\begin{aligned} b\frac{\partial h}{\partial\bar{\pi}} - \Delta_L = {} & \frac{\tau(1-\tau)q^2w^2}{\gamma} - \frac{\tau(1-\tau)q^2w^2}{\gamma}\int_0^{a^*}(1+\lambda a - \lambda\bar{a})\,\mathrm{d}G \\ & - \frac{\tau(1-\tau)w^2}{\gamma}\int_{a^*}^1(1+\lambda a - \lambda\bar{a})\,\mathrm{d}G < 0 \end{aligned} \tag{5.52}$$

Hence, the nominator of (5.51) is negative. From equations (5.50) and (5.52) together it follows that $\Delta_L > 0$. Because $\int_0^{a^*}(a-\bar{a})\,\mathrm{d}G < 0$, the denominator is also negative, which indeed implies $\lambda^* > 0$. Since $\int_0^{a^*} \partial NB_L/\partial\bar{\pi}\,\mathrm{d}G$ and $\int_{a^*}^1 \partial NB_H/\partial\bar{\pi}\,\mathrm{d}G$ are both linear functions of λ, it turns out that λ^* is a unique solution. This completes the proof. ∎

5.A.3 Proof of Proposition 5.3

Proof We have the following condition for a Pareto improvement:

$$\begin{aligned}\Gamma &\equiv P_{uni} - P_{ben} - \frac{[\Psi(z)]^2\pi}{2\gamma} \\ &= \frac{(\pi - z)(\bar{\pi} - h)}{\bar{\pi} - z} b_{uni} - (\pi - h) b_{ben} - \frac{[\Psi(z)]^2\pi}{2\gamma} \geq 0 \end{aligned} \tag{5.53}$$

where for at least one a-individual this inequality has to hold strictly. We start from a situation in which each agent has the same life span, i.e., $\lambda = 0$. This means that $\pi(a) = \bar{\pi}$ for each a-individual and hence $\Gamma = 0$. Now we derive the following derivative at $\lambda = 0$:[25]

$$\frac{\partial \Gamma}{\partial \lambda} = \frac{z - h}{\bar{\pi} - z} \bar{\pi}(a - \bar{a}) b_{uni} + (\bar{\pi} - h)\left(\frac{\partial b_{uni}}{\partial \lambda} - \frac{\partial b_{ben}}{\partial \lambda}\right) \tag{5.54}$$

To prove that the reform is Pareto improving we have to show that $\partial\Gamma/\partial\lambda \geq 0$, and for at least one individual it should be strictly positive. Note that Assumption 1 implies that the minimum of the first term is equal to zero, i.e., for the agent with ability $a = \bar{a}$. Hence, the reform is Pareto improving if $\partial b_{uni}/\partial\lambda \geq \partial b_{ben}/\partial\lambda$.

The budget constraint of the pension scheme can be written as:

$$b(\bar{\pi} - h)\Phi = X \tag{5.55}$$

with:

$$\Phi \equiv \int_0^{a^*} \left[\frac{\pi - z_L}{\bar{\pi} - z_L} - \frac{(\pi - \bar{\pi})\tau q w \pi}{\gamma(\bar{\pi} - z_L)^2}\right] \mathrm{d}G + \int_{a^*}^1 \left[\frac{\pi - z_H}{\bar{\pi} - z_H} - \frac{(\pi - \bar{\pi})\tau w \pi}{\gamma(\bar{\pi} - z_H)^2}\right] \mathrm{d}G \tag{5.56}$$

$$X \equiv \tau q w \int_0^{a^*} \left[1 + \frac{(1-\tau) q w \pi}{\gamma}\right] \mathrm{d}G + \tau w \int_{a^*}^1 \left[a + \frac{(1-\tau) w \pi}{\gamma}\right] \mathrm{d}G \tag{5.57}$$

Note that in the benchmark model $\Phi = 1$. From equation (5.55), we derive at $\lambda = 0$:

$$\frac{\partial b_{ben}}{\partial \lambda} = \frac{1}{\bar{\pi} - h}\frac{\partial X}{\partial \lambda} \tag{5.58}$$

$$\frac{\partial b_{uni}}{\partial \lambda} = \frac{1}{\bar{\pi} - h}\left(\frac{\partial X}{\partial \lambda} - X\frac{\partial \Phi}{\partial \lambda}\right) \tag{5.59}$$

[25]To avoid complex calculations that yield no additional insights, we evaluate all derivatives in this section at the initial point $\lambda = 0$.

Hence,

$$\frac{\partial b_{uni}}{\partial \lambda} - \frac{\partial b_{ben}}{\partial \lambda} = -\frac{X}{\bar{\pi} - h}\frac{\partial \Phi}{\partial \lambda} \tag{5.60}$$

From the definition of Φ above and applying Leibniz's rule, we obtain:

$$\begin{aligned}\frac{\partial \Phi}{\partial \lambda} = & \frac{\bar{\pi}}{\bar{\pi} - z_L}\left[1 - \frac{\tau q w \bar{\pi}}{\gamma(\bar{\pi} - z_L)}\right]\int_0^{a^*}(a - \bar{a})\,\mathrm{d}G \\ & + \frac{\bar{\pi}}{\bar{\pi} - z_H}\left[1 - \frac{\tau w \bar{\pi}}{\gamma(\bar{\pi} - z_H)}\right]\int_{a^*}^1(a - \bar{a})\,\mathrm{d}G\end{aligned} \tag{5.61}$$

Inserting equation (5.24) with $\lambda = 0$ in this expression, gives:

$$\frac{\partial \Phi}{\partial \lambda} = \underbrace{\frac{\gamma(\gamma - qw)}{(\gamma - qw + \tau q w)^2}}_{\Pi_L}\int_0^{a^*}(a - \bar{a})\,\mathrm{d}G + \underbrace{\frac{\gamma(\gamma - w)}{(\gamma - w + \tau w)^2}}_{\Pi_H}\int_{a^*}^1(a - \bar{a})\,\mathrm{d}G \tag{5.62}$$

Let $\tau \to 0$. Then we have that $\Pi_H > \Pi_L$ which implies that the derivative is positive and thus $\partial b_{uni}/\partial\lambda < \partial b_{ben}/\partial\lambda$ for any possible cut-off point $0 < a^* < 1$. Taking the other extreme, $\tau \to 1$, we obtain $\Pi_H < \Pi_L$ so that the derivative is negative and $\partial b_{uni}/\partial\lambda > \partial b_{ben}/\partial\lambda$ for any value $0 < a^* < 1$. The derivative is zero if and only if $\Pi_H(\tau^*) = \Pi_L(\tau^*)$ which has a unique solution $0 < \tau^* < 1$ given by equation (5.28). Hence, $\partial b_{uni}/\partial\lambda \geq \partial b_{ben}/\partial\lambda$ if and only if $\tau \geq \tau^*$. This completes the proof. ∎

5.A.4 Proof of Proposition 5.5

Proof We again have the following condition for a Pareto improvement:

$$\Gamma \equiv \frac{(\pi - z)(\bar{\pi} - h)}{\hat{\pi} - z} b_{edu} - (\pi - h) b_{ben} - \frac{[\Psi(z)]^2 \pi}{2\gamma} \geq 0 \tag{5.63}$$

where for at least one a-individual this inequality has to hold strictly. We now derive the following derivative, again evaluated at the initial position $\lambda = 0$:

$$\frac{\partial \Gamma}{\partial \lambda} = \frac{z - h}{\bar{\pi} - z}\bar{\pi}(a - \bar{a}) b_{edu} + (\bar{\pi} - h)\left(\frac{\partial b_{edu}}{\partial \lambda} - \frac{\partial b_{ben}}{\partial \lambda}\right) - \frac{\bar{\pi} - h}{\bar{\pi} - z}\frac{\partial \hat{\pi}}{\partial \lambda} b_{edu} \tag{5.64}$$

To prove that this reform cannot be a Pareto improvement, we have to show that for at least one a-individual equation (5.64) is strictly negative. Let us concentrate on a high-skilled agent with ability level $a = a^*$. Using equation (5.18), equation (5.64) then becomes:

$$\begin{aligned}\frac{\partial\Gamma(a=a^*)}{\partial\lambda} = & \underbrace{\frac{z_H-h}{\bar{\pi}-z_H}\bar{\pi}(a^*-\bar{a})b_{edu}}_{\Sigma_1} + (\bar{\pi}-h)\left(\frac{\partial b_{edu}}{\partial\lambda}-\frac{\partial b_{ben}}{\partial\lambda}\right) \\ & -\underbrace{\frac{\bar{\pi}-h}{\bar{\pi}-z_H}\frac{\int_{a^*}^1\bar{\pi}(a-\bar{a})\,\mathrm{d}G}{1-G(a^*)}b_{edu}}_{\Sigma_2>0}\end{aligned} \tag{5.65}$$

Because $\bar{\pi} > z_H(a)\ \forall a$ it follows that $|\Sigma_1| < \Sigma_2$. We therefore know that the derivative $\partial\Gamma/\partial\lambda < 0$ if $\partial b_{edu}/\partial\lambda \leq \partial b_{ben}/\partial\lambda$, which means that the flexibility reform cannot be a Pareto improving.

We again write the budget constraint in the form $b(\bar{\pi}-h)\Phi = X$, with X already defined by equation (5.57) and Φ now given by:

$$\begin{aligned}\Phi \equiv & \int_0^{a^*}\left[\frac{\pi-z_L}{\bar{\pi}_L-z_L}-\frac{(\pi-\bar{\pi}_L)\tau q w\pi}{\gamma(\bar{\pi}_L-z_L)^2}\right]\mathrm{d}G \\ & +\int_{a^*}^1\left[\frac{\pi-z_H}{\bar{\pi}_H-z_H}-\frac{(\pi-\bar{\pi}_H)\tau w\pi}{\gamma(\bar{\pi}_H-z_H)^2}\right]\mathrm{d}G\end{aligned} \tag{5.66}$$

Similar to uniform actuarial adjustment, there holds:

$$\frac{\partial b_{edu}}{\partial\lambda}-\frac{\partial b_{ben}}{\partial\lambda} = -\frac{X}{\bar{\pi}-h}\frac{\partial\Phi}{\partial\lambda} \tag{5.67}$$

From the definition of Φ, we derive at $\lambda = 0$:

$$\begin{aligned}\frac{\partial\Phi}{\partial\lambda} = & \left[\frac{1}{\bar{\pi}-z_L}-\frac{\tau q w\bar{\pi}}{\gamma(\bar{\pi}-z_L)^2}\right]\int_0^{a^*}\left(\frac{\partial\pi}{\partial\lambda}-\frac{\partial\bar{\pi}_L}{\partial\lambda}\right)\mathrm{d}G \\ & +\left[\frac{1}{\bar{\pi}-z_H}-\frac{\tau w\bar{\pi}}{\gamma(\bar{\pi}-z_H)^2}\right]\int_{a^*}^1\left(\frac{\partial\pi}{\partial\lambda}-\frac{\partial\bar{\pi}_H}{\partial\lambda}\right)\mathrm{d}G\end{aligned} \tag{5.68}$$

Notice:

$$\begin{aligned}\int_0^{a^*}\left(\frac{\partial\pi}{\partial\lambda}-\frac{\partial\bar{\pi}_L}{\partial\lambda}\right)\mathrm{d}G &= \int_0^{a^*}\left[\bar{\pi}(a-\bar{a})-\frac{\int_0^{a^*}\bar{\pi}(a-\bar{a})\,\mathrm{d}G}{G(a^*)}\right]\mathrm{d}G \\ &= \bar{\pi}\int_0^{a^*}(a-\bar{a})\,\mathrm{d}G-\bar{\pi}\int_0^{a^*}(a-\bar{a})\,\mathrm{d}G = 0\end{aligned} \tag{5.69}$$

Along the same lines, we also have:

$$\int_{a^*}^1\left(\frac{\partial\pi}{\partial\lambda}-\frac{\partial\bar{\pi}_H}{\partial\lambda}\right)\mathrm{d}G = 0 \tag{5.70}$$

Hence, $\partial\Phi/\partial\lambda = 0$ and $\partial b_{edu}/\partial\lambda = \partial b_{ben}/\partial\lambda$, which completes the proof. ∎

5.A.5 Proof of Proposition 5.6

Proof With actuarial non-neutrality, the Pareto-improving condition is:[26]

$$\Gamma \equiv (\pi - h)^{\sigma}(\pi - z)^{1-\sigma} b_{nan} - (\pi - h) b_{ben} - \frac{[\Psi(z)]^2 \pi}{2\gamma} \geq 0 \tag{5.71}$$

where for at least one a-individual this inequality should hold strictly. Suppose we start from a situation of actuarial neutrality, $\sigma = 1$, which means $\Gamma = 0$. Then we can derive the following derivative, evaluated in the initial position $\sigma = 1$:

$$\frac{\partial \Gamma}{\partial \sigma} = (\pi - h)\frac{\partial b}{\partial \sigma} + X \ln(\pi - h) - X \ln(\pi - z) \tag{5.72}$$

To prove that the reform is Pareto improving we have to show that $\partial\Gamma/\partial\sigma \geq 0$, where for at least one individual this inequality strictly holds.

Write the budget constraint of the pension scheme in the usual way:

$$b(\pi - h)\Phi = X \tag{5.73}$$

where X is already defined by equation (5.57) and with Φ equal to:

$$\begin{aligned}\Phi \equiv G(a^*) &\left[\left(\frac{\pi - z_L}{\pi - h}\right)^{1-\sigma} - \frac{(\sigma - 1)\tau q w \pi (\pi - h)^{\sigma - 1}}{\gamma(\pi - z_L)^{\sigma}}\right] \\ &+ [1 - G(a^*)]\left[\left(\frac{\pi - z_H}{\pi - h}\right)^{1-\sigma} - \frac{(\sigma - 1)\tau w \pi (\pi - h)^{\sigma - 1}}{\gamma(\pi - z_H)^{\sigma}}\right]\end{aligned} \tag{5.74}$$

From equation (5.73) it follows:

$$\frac{\partial b}{\partial \sigma} = -\frac{X}{\pi - h}\frac{\partial \Phi}{\partial \sigma} \tag{5.75}$$

Using definition (5.74), we can derive at $\sigma = 1$:

$$\begin{aligned}\frac{\partial \Phi}{\partial \sigma} = \; &\ln(\pi - h) - G(a^*)\left[\ln(\pi - z_L) + \frac{\tau q w \pi}{\gamma(\pi - z_L)}\right] \\ &- [1 - G(a^*)]\left[\ln(\pi - z_H) + \frac{\tau w \pi}{\gamma(\pi - z_H)}\right]\end{aligned} \tag{5.76}$$

Substituting equation (5.76) into equation (5.75) and inserting the resulting expression in equation (5.72) ultimately implies:

$$\begin{aligned}\frac{\partial \Gamma}{\partial \sigma} = \; &-X \ln(\pi - z) + G(a^*) X \left[\ln(\pi - z_L) + \frac{\tau q w \pi}{\gamma(\pi - z_L)}\right] \\ &+ [1 - G(a^*)]\, X \left[\ln(\pi - z_H) + \frac{\tau w \pi}{\gamma(\pi - z_H)}\right]\end{aligned} \tag{5.77}$$

[26] In this section, we abstract from life-span heterogeneity, i.e., $\lambda = 0$.

For high-skilled agents we have $z = z_H$, implying:

$$\begin{aligned}\frac{\partial \Gamma}{\partial \sigma} &= G(a^*)\frac{\tau q w \pi X}{\gamma(\pi - z_L)} + [1 - G(a^*)]\,\frac{\tau w \pi X}{\gamma(\pi - z_H)} \\ &\quad + G(a^*) X \ln\left(\frac{\pi - z_L}{\pi - z_H}\right) > 0\end{aligned} \tag{5.78}$$

Hence, high-skilled workers are strictly better off when moving from the benchmark scheme to a scheme with actuarial non-neutrality. For the low-skilled agents we have $z = z_L$, which gives:

$$\begin{aligned}\frac{\partial \Gamma}{\partial \sigma} &= G(a^*)\frac{\tau q w \pi X}{\gamma(\pi - z_L)} + [1 - G(a^*)]\,\frac{\tau w \pi X}{\gamma(\pi - z_H)} \\ &\quad - [1 - G(a^*)]\, X \ln\left(\frac{\pi - z_L}{\pi - z_H}\right)\end{aligned} \tag{5.79}$$

Suppose that $\tau \to 0$. Then $\partial\Gamma/\partial\sigma < 0$ implying that low-skilled agents are worse off after the reform. If on the other hand $\tau \to 1$, then $z_L = z_H \to 0$ so that the last term vanishes. Therefore $\partial\Gamma/\partial\sigma > 0$ which means that low-skilled also benefit from the reform. We have $\partial\Gamma/\partial\sigma = 0$ if $\tau = \hat{\tau}$, with $\hat{\tau}$ given by equation (5.44).

To prove that $\hat{\tau}$ is a unique solution, we have to show that the derivative $\partial\Gamma/\partial\sigma$ is monotonically increasing in τ at $\sigma = 1$. Rewrite equation (5.79) in $\partial\Gamma/\partial\sigma = XA$, with A equal to:

$$A \equiv G(a^*)\frac{\tau q w \pi}{\gamma(\pi - z_L)} + [1 - G(a^*)]\,\frac{\tau w \pi}{\gamma(\pi - z_H)} - [1 - G(a^*)] \ln\left(\frac{\pi - z_L}{\pi - z_H}\right)$$

Since $X > 0$ the necessary and sufficient condition for $\partial\Gamma/\partial\sigma \geq 0$ is $A \geq 0$. This implies that $\hat{\tau}$ is a unique solution if and only if A is monotonically increasing in τ. Taking the derivative of A with respect to τ ultimately gives:[27]

$$\begin{aligned}\frac{\partial A}{\partial \tau} &= G(a^*)\frac{qw(\gamma - qw)}{(\gamma - qw + \tau q w)^2} + [1 - G(a^*)]\,\frac{w(\gamma - w)}{(\gamma - w + \tau w)^2} \\ &\quad + [1 - G(a^*)]\,\frac{w\pi}{\gamma}\left(\frac{1}{\pi - z_H} - \frac{q}{\pi - z_L}\right) > 0\end{aligned} \tag{5.80}$$

This completes the proof. ∎

[27] Note that we assume $\gamma > w$ to make sure that $z < 1$ for all agents, irrespective the size of the pension scheme ($\forall\tau$).

5.B Endogenous skill level

In this appendix we show that the main welfare implications of the pension flexibility reforms still hold under endogenous schooling. We start with the derivation of the threshold ability level a^* for the benchmark model.

5.B.1 Benchmark model

An agent is indifferent in acquiring skills or not if $U_L(a) = U_H(a) \Rightarrow c_H(a^*) = c_L(a^*)$. Thus, there is a cut-off level of a, denoted a^*, which is given by:

$$a^* = q - \frac{(1-\tau)w\pi(a^*)(1-q^2)}{2\gamma} \tag{5.81}$$

Agents with ability $a < a^*$ will not invest in schooling and stay low-skilled and agents with $a > a^*$ choose to acquire extra skills and become high-skilled.

From equation (5.81), we can infer the following. First, an increase in the tax rate τ raises the fraction of low-skilled workers, i.e., $\partial a^*/\partial\tau > 0$. The tax rate induces agents to retire earlier which decreases the return period of schooling investments and, hence, reduces schooling incentives. Second, an increase in longevity raises the number of high-skilled individuals, i.e., $\partial a^*/\partial\bar{\pi} < 0$. Recall that an increase in the average life span induces agents to postpone retirement. This increases the incentive to become high-skilled because the return period of schooling investment becomes longer.

5.B.2 Uniform actuarial adjustment

We will show that uniform adjustment can either increase or decrease the schooling incentives, dependent on the fraction of low-skilled agents in the initial situation. We also show that with endogenous schooling uniform adjustment still induces a Pareto improvement if the tax rate is sufficiently high.

Cut-off ability level

From equation (5.23) we obtain:

$$a^*_{uni} = a^*_{ben} - \frac{2\gamma\Theta}{(1-\tau)w\left[2\gamma + (1-\tau)w(1-q^2)\bar{\pi}\lambda\right]} \tag{5.82}$$

with,

$$\Theta \equiv P(z_H) - P(z_L) - \frac{[\Psi(z_H)]^2\pi}{2\gamma} + \frac{[\Psi(z_L)]^2\pi}{2\gamma}$$

From equation (5.82), it follows:[28]

$$\frac{\partial a^*_{uni}}{\partial \lambda} = \frac{\partial a^*_{ben}}{\partial \lambda} - \frac{1}{(1-\tau)w}\frac{\partial \Theta}{\partial \lambda} \tag{5.83}$$

where we have used that $\Theta = 0$ if $\lambda = 0$. Using equations (5.20) and (5.21), we derive from the definition of Θ:

$$\frac{\partial \Theta}{\partial \lambda} = \frac{\bar{\pi} X(a^* - \bar{a})(z_H - z_L)}{(\bar{\pi} - z_H)(\bar{\pi} - z_L)} \tag{5.84}$$

Hence,

$$\frac{\partial a^*_{uni}}{\partial \lambda} = \frac{\partial a^*_{ben}}{\partial \lambda} - \frac{\bar{\pi} X(a^*_{ben} - \bar{a})(z_H - z_L)}{(1-\tau)w(\bar{\pi} - z_H)(\bar{\pi} - z_L)} \tag{5.85}$$

From this equation it directly follows that $\partial a^*_{uni}/\partial\lambda > \partial a^*_{ben}/\partial\lambda$ if $a^*_{ben} < \bar{a}$ and $\partial a^*_{uni}/\partial\lambda < \partial a^*_{ben}/\partial\lambda$ if $a^*_{ben} > \bar{a}$. If the marginal agent has an above-average life span, $\pi(a^*) > \bar{\pi}$, the agent has an incentive to postpone retirement. This increases the incentive to become high-skilled because later retirement raises the return period of schooling investments. When the marginal agent has a below-average life span, $\pi(a^*) < \bar{\pi}$, this person has an incentive to advance retirement which decreases the willingness to become high-skilled.

Welfare effects

To show that uniform actuarial adjustment induces a Pareto improvement, we have to distinguish between agents who switch from skill level after this reform and agents who do not switch from skill level.

Non-switching agents. The non-switching group consists of agents who are old at the time the reform is implemented and young agents who either remain low-skilled, $a < \min(a^*_{ben}, a^*_{uni})$, or high-skilled, $a > \max(a^*_{ben}, a^*_{uni})$. For this group the Pareto-improving condition is still given by equation (5.53). Hence, also with endogenous schooling the condition $\partial b_{uni}/\partial\lambda \geq \partial b_{ben}/\partial\lambda$ is a sufficient condition.

Under endogenous schooling, equation (5.60) changes into:

$$\frac{\partial b_{uni}}{\partial \lambda} - \frac{\partial b_{ben}}{\partial \lambda} = \frac{1}{\bar{\pi} - h}\left[\frac{\partial X(a^*_{uni})}{\partial \lambda} - \frac{\partial X(a^*_{ben})}{\partial \lambda} - X(a^*_{uni})\frac{\partial \Phi}{\partial \lambda}\right] \tag{5.86}$$

[28]Similar to exogenous schooling (see Sections 5.A.3 and 5.A.4), we simplify calculations by evaluating the derivatives at the initial position $\lambda = 0$ (i.e., no life-time heterogeneity).

From the definition of X, see equation (5.57), we derive at the point $\lambda = 0$:

$$\begin{aligned}\frac{\partial X}{\partial \lambda} = & -\frac{\tau(1-\tau)w^2(1-q^2)\bar{\pi}}{2\gamma}\frac{\partial a^*}{\partial \lambda} + \tau q w \int_0^{a^*} \frac{(1-\tau)qw\bar{\pi}(a-\bar{a})}{\gamma}\,\mathrm{d}G \\ & + \tau w \int_{a^*}^1 \frac{(1-\tau)w\bar{\pi}(a-\bar{a})}{\gamma}\,\mathrm{d}G\end{aligned} \tag{5.87}$$

Using equation (5.87), we obtain:

$$\begin{aligned}\frac{\partial X(a^*_{uni})}{\partial \lambda} - \frac{\partial X(a^*_{ben})}{\partial \lambda} &= \frac{\tau(1-\tau)w^2(1-q^2)\bar{\pi}}{2\gamma}\left(\frac{\partial a^*_{ben}}{\partial \lambda} - \frac{\partial a^*_{uni}}{\partial \lambda}\right) \\ &= \frac{\tau w(1-q^2)\bar{\pi}}{2\gamma}\frac{\partial \Theta}{\partial \lambda}\end{aligned} \tag{5.88}$$

where we have used equation (5.83) in going from the first line to the second line. Equation (5.86) can now rewritten in:

$$\frac{\partial b_{uni}}{\partial \lambda} - \frac{\partial b_{ben}}{\partial \lambda} = \frac{1}{\bar{\pi} - h}\left[\frac{\tau w(1-q^2)\bar{\pi}}{2\gamma}\frac{\partial \Theta}{\partial \lambda} - X\frac{\partial \Phi}{\partial \lambda}\right] \tag{5.89}$$

The derivative $\partial\Phi/\partial\lambda$ is still given by equation (5.62). That is,

$$\frac{\partial \Phi}{\partial \lambda} = \underbrace{\frac{\gamma(\gamma - qw)}{(\gamma - qw + \tau qw)^2}}_{\Pi_L}\int_0^{a^*}(a-\bar{a})\,\mathrm{d}G + \underbrace{\frac{\gamma(\gamma - w)}{(\gamma - w + \tau w)^2}}_{\Pi_H}\int_{a^*}^1(a-\bar{a})\,\mathrm{d}G \tag{5.90}$$

Let $\tau \to 0$. Then we have $\Pi_H > \Pi_L$ implying that $\partial\Phi/\partial\lambda > 0$ for any possible cut-off point $0 < a^* < 1$. From equation (5.89) then follows $\partial b_{ben}/\partial\lambda > \partial b_{uni}/\partial\lambda$. Taking the other extreme, $\tau \to 1$, we obtain $\Pi_H < \Pi_L$ so that $\partial\Phi/\partial\lambda < 0$ for any value $0 < a^* < 1$. This implies from equation (5.89) that $\partial b_{ben}/\partial\lambda < \partial b_{uni}/\partial\lambda$. As $\partial\Phi/\partial\lambda$ is continuous at $0 < \tau < 1$, there exist tax rates τ for which $\partial b_{ben}/\partial\lambda \geq \partial b_{uni}/\partial\lambda$.

Switching agents. The switching group consists of young agents who switch from *i)* either low-skilled to high-skilled, which occurs if $a^*_{ben} > \bar{a}$, or from *ii)* high-skilled to low-skilled, which occurs if $a^*_{ben} < \bar{a}$.

i) Suppose $a^*_{ben} > \bar{a}$. Then the Pareto-improving condition is given by:

$$\Gamma \equiv \underbrace{P_{uni} - P_{ben} - \frac{[\Psi(z)]^2\pi}{2\gamma}}_{Y_1} + \underbrace{(1-\tau)w(a-q) + \frac{(1-q^2)(1-\tau)^2w^2\pi}{2\gamma}}_{Y_2} \geq 0 \tag{5.91}$$

Compared to the non-switching group, we now have an additional term Y_2 which is due to the fact that switching agents are confronted with a different wage rate. Note that in the initial point $\lambda = 0$ we still have $\Gamma = 0$. Taking the derivative with respect to λ and evaluating the resulting expression at $\lambda = 0$, gives:

$$\frac{\partial \Gamma}{\partial \lambda} = \frac{\partial Y_1}{\partial \lambda} + \frac{(1-q^2)(1-\tau)^2 w^2 \bar{\pi}(a-\bar{a})}{2\gamma} \tag{5.92}$$

Let the tax rate be set such that $\partial Y_1/\partial \lambda \geq 0$. At $\lambda = 0$ we have $a = a^*_{uni} = a^*_{ben} > \bar{a}$ for the switching group. Hence, $\partial \Gamma/\partial \lambda > 0$ which means that for the switching young the Pareto-improving condition is satisfied.

ii) Suppose $a^*_{ben} < \bar{a}$. Then the Pareto-improving condition is given by:

$$\Gamma \equiv Y_1 + (1-\tau)w(q-a) + \frac{(q^2-1)(1-\tau)^2 w^2 \pi}{2\gamma} \geq 0 \tag{5.93}$$

Taking the derivative with respect to λ gives:

$$\frac{\partial \Gamma}{\partial \lambda} = \frac{\partial Y_1}{\partial \lambda} + \frac{(q^2-1)(1-\tau)^2 w^2 \bar{\pi}(a-\bar{a})}{2\gamma} \tag{5.94}$$

Suppose again that the tax rate is set such that $\partial Y_1/\partial \lambda \geq 0$. At $\lambda = 0$, we now have $a = a^*_{uni} = a^*_{ben} < \bar{a}$ for the switching group. Hence, $\partial \Gamma/\partial \lambda > 0$ which means that also for this case the Pareto-improving condition of the switching group is satisfied.

5.B.3 Skill-dependent actuarial adjustment

We show in this section that skill-dependent actuarial adjustment negatively affects the incentives to become skilled. In addition, similar to exogenous schooling, this reform cannot be a Pareto improvement.

Cut-off ability level

Skill-dependent actuarial adjustment will change schooling because it introduces an endogenous link between the schooling decision and the actuarial adjustment factor. Using equation (5.36) we can infer:

$$a^*_{edu} = a^*_{ben} - \frac{2\gamma\Theta}{(1-\tau)w\left[2\gamma + (1-\tau)w(1-q^2)\bar{\pi}\lambda\right]} \tag{5.95}$$

with,

$$\Theta \equiv P(z_H) - P(z_L) - \frac{\Psi[(z_H)]^2\pi}{2\gamma} + \frac{\Psi[(z_L)]^2\pi}{2\gamma}$$

Take the derivative of (5.95) with respect to λ, evaluated at $\lambda = 0$:

$$\frac{\partial a^*_{edu}}{\partial \lambda} = \frac{\partial a^*_{ben}}{\partial \lambda} - \frac{1}{(1-\tau)w}\frac{\partial \Theta}{\partial \lambda} \tag{5.96}$$

with,

$$\frac{\partial \Theta}{\partial \lambda} = \frac{X}{\bar{\pi} - z_H}\left[\frac{\partial \pi}{\partial \lambda} - \frac{\partial \bar{\pi}_H}{\partial \lambda}\right] - \frac{X}{\bar{\pi} - z_L}\left[\frac{\partial \pi}{\partial \lambda} - \frac{\partial \bar{\pi}_L}{\partial \lambda}\right] \tag{5.97}$$

From equations (5.5) and (5.18), we obtain for $\lambda = 0$:

$$\frac{\partial \pi(a^*)}{\partial \lambda} = \bar{\pi}(a^* - \bar{a}) \tag{5.98}$$

$$\frac{\partial \bar{\pi}_L}{\partial \lambda} = \frac{\int_0^{a^*} \bar{\pi}(a - \bar{a})\,\mathrm{d}G}{G(a^*)} \tag{5.99}$$

$$\frac{\partial \bar{\pi}_H}{\partial \lambda} = \frac{\int_{a^*}^1 \bar{\pi}(a - \bar{a})\,\mathrm{d}G}{1 - G(a^*)} \tag{5.100}$$

Substituting these expressions in equation (5.97) and rearranging, gives:

$$\frac{\partial \Theta}{\partial \lambda} = \underbrace{\frac{X}{\bar{\pi} - z_L}\frac{\int_0^{a^*} \bar{\pi}(a - a^*)\,\mathrm{d}G}{G(a^*)}}_{<0} - \underbrace{\frac{X}{\bar{\pi} - z_H}\frac{\int_{a^*}^1 \bar{\pi}(a - a^*)\,\mathrm{d}G}{1 - G(a^*)}}_{>0} < 0 \tag{5.101}$$

An increase in λ reduces Θ implying $\partial a^*_{edu}/\partial \lambda > \partial a^*_{ben}/\partial \lambda$. With educational-specific actuarial adjustment, individuals can self-select the conversion factor with their skill level. If they choose to become high-skilled this reduces *ceteris paribus* the conversion factor because this is now based on the average longevity of the high-skilled people. Individuals just at or around the margin will therefore find it less attractive to become high-skilled.

Welfare effects

Since the reform leads to fewer high-skilled agents, a high-skilled agent with ability $a = a^*_{edu}$ is also a high-skilled person in the benchmark case. Therefore, the Pareto-improving condition for this agent is still given by equation (5.63). This implies that $\partial b_{edu}/\partial \lambda \leq \partial b_{ben}/\partial \lambda$ is again a sufficient condition to reject the existence of a Pareto improvement. To show that this condition holds, notice that under endogenous schooling equation (5.67) changes into:

$$\frac{\partial b_{edu}}{\partial \lambda} - \frac{\partial b_{ben}}{\partial \lambda} = \frac{1}{\bar{\pi} - h}\left[\frac{\partial X(a^*_{edu})}{\partial \lambda} - \frac{\partial X(a^*_{ben})}{\partial \lambda} - X(a^*_{edu})\frac{\partial \Phi}{\partial \lambda}\right] \tag{5.102}$$

Using equation (5.87), we obtain in $\lambda = 0$:

$$\frac{\partial X(a^*_{edu})}{\partial \lambda} - \frac{\partial X(a^*_{ben})}{\partial \lambda} = \frac{\tau(1-\tau)w^2(1-q^2)\bar{\pi}}{2\gamma}\left(\frac{\partial a^*_{ben}}{\partial \lambda} - \frac{\partial a^*_{edu}}{\partial \lambda}\right)$$
$$= \frac{\tau w(1-q^2)\bar{\pi}}{2\gamma}\frac{\partial \Theta}{\partial \lambda} < 0 \qquad (5.103)$$

where we have used equation (5.96) in going from the first to the second line. With endogenous schooling, there still holds $\partial\Phi/\partial\lambda = 0$. Equation (5.102) then implies $\partial b_{edu}/\partial\lambda < \partial b_{ben}/\partial\lambda$.

5.B.4 Actuarially non-neutral adjustment

In this final section, we show that stimulating retirement postponement in an actuarially non-neutral way improves the incentives to become skilled. We also show that such a reform can lead to a Pareto improvement.

Cut-off ability level

The cut-off point is determined by the condition $U_H(a^*) = U_L(a^*) \Rightarrow c_H(a^*) = c_L(a^*)$. From equation (5.41) we can infer:

$$a^*_{nan} = a^*_{ben} - \frac{\Theta}{(1-\tau)w} \qquad (5.104)$$

with Θ again defined as,

$$\Theta \equiv P(z_H) - P(z_L) - \frac{[\Psi(z_H)]^2\pi}{2\gamma} + \frac{[\Psi(z_L)]^2\pi}{2\gamma}$$

Take the derivative of equation (5.104) with respect to σ, evaluated at the initial point $\sigma = 1$:[29]

$$\frac{\partial a^*_{nan}}{\partial \sigma} = \frac{\partial a^*_{ben}}{\partial \sigma} - \frac{1}{(1-\tau)w}\frac{\partial \Theta}{\partial \sigma} \qquad (5.105)$$

where we have used that $\Theta = 0$ if $\sigma = 1$. Using equations (5.39) and (5.40), we derive from the definition of Θ:

$$\frac{\partial \Theta}{\partial \sigma} = X \ln\left(\frac{\pi - z_L}{\pi - z_H}\right) > 0 \qquad (5.106)$$

Hence, $\partial a^*_{nan}/\partial\sigma < \partial a^*_{ben}/\partial\sigma$.

[29] Similar to exogenous schooling (see Section 5.A.5), we abstract from life-span heterogeneity and evaluate derivatives at the initial position $\sigma = 1$ (i.e., actuarial neutrality).

Welfare effects

Non-switching agents. The non-switching group consists of agents who are old at the time the reform is implemented and of young agents who either remain low-skilled, $a < \min(a^*_{ben}, a^*_{uni})$, or high-skilled, $a > \max(a^*_{ben}, a^*_{uni})$. For this group the Pareto-improving condition is still given by equation (5.71).

With endogenous schooling, equation (5.75) changes into:

$$\frac{\partial b}{\partial \sigma} = \frac{1}{\pi - h}\left(\frac{\partial X}{\partial \sigma} - X\frac{\partial \Phi}{\partial \sigma}\right) \tag{5.107}$$

Using equation (5.57), we have at $\sigma = 1$:

$$\begin{aligned}\frac{\partial X}{\partial \sigma} &= -\frac{\tau(1-\tau)w^2(1-q^2)\pi}{2\gamma}\frac{\partial a^*}{\partial \sigma} \\ &= \frac{\tau w(1-q^2)\pi X}{2\gamma}\ln\left(\frac{\pi - z_L}{\pi - z_H}\right) > 0\end{aligned} \tag{5.108}$$

where we have used equation (5.106) in going from the first to the second line. Note that this derivative is positive because $z_H > z_L$. The derivative $\partial\Phi/\partial\sigma$ is still given by equation (5.76). Substituting this equation together with equation (5.108) into equation (5.107) and inserting the resulting expression in equation (5.72) ultimately implies:

$$\begin{aligned}\frac{\partial \Gamma}{\partial \sigma} = {}& G(a^*)X\left[\ln(\pi - z_L) + \frac{\tau q w \pi}{\gamma(\pi - z_L)}\right] + \frac{\tau w(1-q^2)\pi X}{2\gamma}\ln\left(\frac{\pi - z_L}{\pi - z_H}\right) \\ &+ \left[1 - G(a^*)\right] X\left[\ln(\pi - z_H) + \frac{\tau w \pi}{\gamma(\pi - z_H)}\right] - X\ln(\pi - z)\end{aligned} \tag{5.109}$$

For high-skilled agents we have $z = z_H$, implying:

$$\begin{aligned}\frac{\partial \Gamma}{\partial \sigma} = {}& G(a^*)\frac{\tau q w \pi X}{\gamma(\pi - z_L)} + \left[1 - G(a^*)\right]\frac{\tau w \pi X}{\gamma(\pi - z_H)} \\ &+ X\left[G(a^*) + \frac{\tau w(1-q^2)\pi}{2\gamma}\right]\ln\left(\frac{\pi - z_L}{\pi - z_H}\right) > 0\end{aligned} \tag{5.110}$$

Hence, high-skilled workers are strictly better off when moving from the benchmark scheme to a scheme with actuarial non-neutrality. For low-skilled people with $z = z_L$ the condition becomes

$$\begin{aligned}\frac{\partial \Gamma}{\partial \sigma} = {}& G(a^*)\frac{\tau q w \pi X}{\gamma(\pi - z_L)} + \left[1 - G(a^*)\right]\frac{\tau w \pi X}{\gamma(\pi - z_H)} \\ &+ X\left[\frac{\tau w(1-q^2)\pi}{2\gamma} - 1 + G(a^*)\right]\ln\left(\frac{\pi - z_L}{\pi - z_H}\right)\end{aligned} \tag{5.111}$$

Suppose that $\tau \to 0$. Then $\partial\Gamma/\partial\sigma < 0$ implying that low-skilled agents are worse off after the reform. If on the other hand $\tau \to 1$, then $z_L = z_H \to 0$ so that the last term vanishes. Therefore $\partial\Gamma/\partial\sigma > 0$ which means that low-skilled also benefit from the reform. As $\partial\Gamma/\partial\sigma$ is continuous at $0 < \tau < 1$, there exist tax rates τ such that $\partial\Gamma/\partial\sigma > 0$.

Switching agents. With actuarially non-neutral adjustment the switching group only consists of young agents who choose to become high-skilled, $a^*_{nan} < a < a^*_{ben}$. The Pareto-improving condition then equals:

$$\Gamma \equiv P_{nan} - P_{ben} - \frac{[\Psi(z)]^2\pi}{2\gamma} + (1-\tau)w(a-q) + \frac{(1-q^2)(1-\tau)^2w^2\pi}{2\gamma} \quad (5.112)$$

Notice that the derivative of Γ with respect to σ is exactly the same as for the non-switching group and given by equation (5.110).

CHAPTER 6

SUMMARY, POLICY IMPLICATIONS AND RESEARCH AGENDA

In many modern countries, population ageing jeopardizes the sustainability and shock resistance of social security systems and pension funds. For pension funds, sustainability is further reduced by the recent global financial crisis and the Euro crisis. At the same time, the social environment in which pension institutions are operating has changed considerably last decades. Labour mobility has increased as well as the heterogeneity of labour markets, with a growing importance of temporary contracts and part-time employment (especially among women). Against this demographic, economic and social background, many countries have reformed or are planning to reform their pension systems. Although countries differ a lot in the kind of reforms they (want to) undertake, the underlying motives are the same: improving the financial sustainability and shock resistance of the pension system, encouraging labour supply and, accommodating more individual choice (OECD, 2007).

This thesis studies the implications of pension reforms for two main characteristics of collective pension schemes, i.e., redistribution and risk sharing. We make a clear distinction between these two concepts: *redistribution* are transfers across or within generations which are independent of a certain shock occurring; *risk sharing* are transfers which are conditional on the occurrence of a shock. Our analysis recognizes that redistribution and risk sharing are no exogenous processes but a rather complex interplay of individual characteristics (like e.g., age or life expectancy), the design of the pension scheme and economic decisions. The economic decisions we focus on are consumption decisions, portfolio allocation, labour-supply choices and, in particular, the *retirement* decision. The analysis is performed in the light of three important world-wide pension reforms, which are also on the Dutch policy agenda, namely the switch from collective defined-benefit (DB) pension schemes to individual defined-contribution (DC) pension schemes, a general increase in the pension entitlement age and, finally, the introduction of (more) flexibility in this entitlement age.

Redistribution, especially from rich to poor, is often viewed as one of the main objectives of first-pillar pension systems. The same holds for consumption smoothing and risk sharing, functions which are typically performed by second-pillar pensions. While redistribution and risk sharing certainly have advantages in terms of preventing old-age poverty and completing incomplete markets, they may also come along with a welfare cost. Pension contributions are usually levied proportional to labour income. Then the transfers (arising from redistribution or risk sharing) break down the link between contributions and benefits and therefore distort the labour-leisure decision. Hence, collective pension schemes always face a trade-off between on the one hand providing redistribution and/or risk sharing and on the other hand minimizing labour-supply distortions. The aim of this thesis is to get a better understanding of how this trade-off will be affected by pension reforms that alter the risk properties of pension benefits or increase the level and flexibility of the pension entitlement age.

The analysis in this thesis boils down to four conceptual chapters. Chapter 2 takes-off by quantifying the degree of redistribution across generations (intergenerational redistribution) and within generations (intragenerational redistribution) in Dutch occupational pension schemes. Chapter 3 takes a normative point of view and calculates the optimal amount of intergenerational risk sharing in case of a funded pension scheme with defined benefits. This chapter should be placed in light of the world-wide trend towards individual DC contracts in which intergenerational risk-sharing possibilities are no longer possible. Chapter 4 and Chapter 5 focus on the interaction between retirement flexibility on the one hand and, respectively, risk sharing and redistribution on the other hand. These two chapters refer to the observed pension reforms towards increasing the statutory retirement age and accommodating more individual choice in the retirement age.

In the remainder of this final chapter, we provide a summary of each of these conceptual chapters (Section 6.1), discuss their main policy implications (Section 6.2) and, finally, sketch some directions for further research (Section 6.3).

6.1 Summary

When analysing the economic implications of redistribution, a first step is to get insight in the level and direction of the transfers. Surprisingly, a clear picture of this is missing. Moreover, little attention is given to intragenerational transfers which may also be important. The few studies that quantify redistribution focus on pay-as-you-go (PAYG) pensions (Börsch-Supan and Reil-Held, 2001 and Ter Rele, 2007) while redistribution might also be important for collective funded schemes. The aim of **Chapter 2** is to fill these gaps and to measure redistribution in Dutch funded occupational pensions for various socioeconomic groups. To control for age-specific effects, we measure redistribution on a life-time basis. Redistribution is defined in terms of net benefit, i.e., as the difference between the expected present values of pension benefits and contributions.

Redistribution in collective pension schemes originates from applying *uniform* policy rules to *heterogeneous* participants. In the Netherlands, occupational earnings-related pensions in the second pillar are funded via a uniform contribution and accrual rate (denoted uniform pricing), determined as a fraction of the wage earned. This uniform contribution and accrual system drives a wedge between the market price of the annuity contract and the actual contributions charged. The market value depends on individual characteristics, like age and gender, which, by definition, does not hold true for uniform pricing. Differences between the market price of a pension scheme and the costs imply redistribution between groups of participants.

To quantify redistribution, we distinguish between inter- and intragenerational life-time redistribution. We use the level of educational attainment, gender and age to classify the pension fund population. Hence, intragenerational redistribution takes place between males and females on the one hand and between individuals with different levels of education on the other hand. Life-time intergenerational redistribution relates to an implicit tax imposed on future generations to service the gains given away to the generations living at the time uniform pricing was introduced.

We find that the Dutch occupational pension schemes contain sizable transfers from males to females and from low-educated employees to high-educated employees. On a life-time basis, the impact of intergenerational redistribution seems to be modest. If we account for the expected convergence of life expectancies between males and females in the coming decades, the cross-gender redistribution will reduce to some extent. Our analysis reveals that differences in life expectancy are far more important for the observed intra-generational redistribution effects than differences in income profile or the development of labour force participation.

In **Chapter 3** we turn from redistribution to risk sharing and we switch from a positive to a normative analysis. The purpose of this chapter is to analyse the added value of intergenerational risk sharing in collective funded DB schemes in face of the growing popularity of individual DC schemes. *A priori* it is not immediately clear that a move from collective funded DB pensions to more individual funded DC pensions improves social welfare. On the one hand, funded DB pension schemes allow for welfare-improving intergenerational risk sharing. As demonstrated in e.g., Diamond (1977) and Gordon and Varian (1988) in competitive financial markets currently living generations are not able to share risks with those who are not born yet. Mandatory participation in a collective funded DB pension scheme can (at least partly) solve this market incompleteness. The main feature of such a pension scheme is that it smoothes shocks over and beyond the life time of a single generation by disconnecting individual contributions and benefits. On the other hand, collective funded DB pensions often involve distortions on labour markets, an aspect that certainly decreases welfare. In most pension plans contributions are related to labour income. A disconnection between individual contributions and benefits then implies that the contribution rate contains an implicit tax (or subsidy) which distorts the labour-supply decision.

The existing literature dealing with intergenerational risk sharing in funded pension schemes report substantial welfare gains (see, e.g., Teulings and De Vries, 2006; Gollier, 2008 and Cui et al., 2011). However, these studies may overstate the welfare gains from intergenerational risk sharing for two reasons. First, labour supply is assumed to be exogenous so that distortions associated with implicit pension taxes are not taken into account. Second, most of these studies only focus on capital market risks and assume risk-free human capital. It is well known, however, that human capital is risky. Moreover, the returns to stocks and human capital are positively correlated in the long run which may decrease the preference of young people to absorb equity risk through their pension deal (Benzoni et al., 2007).

This chapter aims to accommodate these shortcomings and allows for wage risk and endogenous labour supply to calculate the merits of intergenerational risk sharing. We develop a model that represents a small open economy populated with two overlapping generations and a pension fund. The economy is subject to two potentially correlated macroeconomic risk factors, capital market risk and wage risk. The two overlapping generations cannot trade risks because the young are not able to participate in the capital market before shocks occur. The young generation decides upon the amount of private saving, labour supply and the portfolio allocation in order to maximize expected life-time utility. The old generation is retired. The pension fund provides risk-free benefits and raises state-contingent contributions proportional to individual labour income. Hence, the young generation bears the full mismatch risk between the safe benefits provided to the old and the accumulated pension assets. Labour is assumed to be perfectly immobile so that agents are not able to avoid implicit taxes by moving abroad. Taking into account the behavioural response of the consumer to its actions, the pension fund optimally chooses the portfolio allocation in order to maximize social welfare.

We analytically show for Cobb-Douglas utility that the introduction of a collective funded scheme with defined benefits involves an *ex ante* Pareto improvement, *even* if pension contributions are distortionary and *even* if the returns on equity and human capital are positively correlated. Using numerical simulations, we show that this result also holds for more general preferences and alternative model settings. The driving force behind this result is the optimal investment policy of the pension fund. With this instrument, the fund is able to control both the diversification gains from risk sharing and the size of the labour-supply distortions. As *ultimum remedium* the pension fund can always replicate the private economy in which intergenerational risk sharing and labour-supply distortions are absent by investing all contributions in the risk-free asset. This investment strategy rules out the possibility that labour-supply distortions exceed the welfare gain from insurance.

Notwithstanding this result, we show that endogenous labour supply decreases the demand for risky assets if contributions are distortionary. This result contrasts with existing studies on the interaction between labour supply and portfolio choice (see e.g. Bodie et al., 1992; Choi and Shim, 2006 and Farhi and Panageas, 2007). These studies show that labour-supply flexibility offers insurance against adverse shocks which justifies more risky asset portfolios. The idea is that income effects in labour-supply behaviour cause a negative correlation between asset returns and labour income allowing individuals to take more risk. This chapter, however, shows that income-related intergenerational transfers also introduce substitution effects. These substitution effects work in the opposite direction and generate a positive correlation between labour income and asset returns. Hence, labour supply is subject to pro-cyclical pressure which reduces the risk-bearing capacity of consumers.

In the two last conceptual chapters of this thesis, we build a bridge between, on the one side, risk sharing (Chapter 4) and redistribution (Chapter 5) and, on the other side, the move towards flexible pension schemes. In **Chapter 4**, we raise the question how this trend from inflexible to flexible pension contracts will affect consumption and portfolio decisions during working life. As stressed in the literature, the important advantage of retirement flexibility is that it provides insurance against all types of risks, like disability risk (Diamond and Mirrlees, 1978) or stock market risk (Pestieau and Possen, 2010). The general idea of flexible retirement is that it gives individuals the ability to adjust working life to their own preferences and to avoid abrupt changes in life-time consumption. Viewed in this way, retirement flexibility serves as a hedge against adverse investment outcomes which allows for more risk taking in pension assets (see e.g. Bodie et al., 1992). Similar to Chapter 3, the basic mechanism behind this result is the negative correlation between asset returns and labour income due to wealth effects in the retirement decision. Our analysis reveals that factors like the type of risk, the willingness of consumers to substitute consumption for leisure, and general equilibrium effects have an important impact on the insurance provided by retirement flexibility. Different positions about these factors change existing views from the literature.

Compared to the existing literature, we add three important elements to the analysis on portfolio choice and retirement. First, we complement the partial equilibrium approach with a general equilibrium one. A general equilibrium perspective seems the most natural road to take because the move to flexible pensions clearly is an international phenomenon. With general equilibrium, we explicitly recognize that consumption and labour-supply decisions affect factor prices which, in turn, influence the insurance effect of retirement flexibility. Second, we distinguish between productivity and depreciation risk and these risk factors are directly linked to production. This distinction is important because both risk factors constitute a rather different effect on income and substitution effects in labour supply. Third, we allow for more general preferences which are characterized by a constant elasticity of substitution function of consumption and leisure. This specification allows the elasticity of substitution between labour and leisure to take any positive number.

To analyse the interaction between portfolio choice, consumption and retirement decisions, we use a two-period overlapping-generations (OLG) model of a closed economy. The model includes government debt and incorporates endogenous retirement. In our framework, the young working generation decides upon its consumption and portfolio allocation. Agents can either invest in risk-free government bonds or in risky firm stocks. Retirement is endogenous and we compare two different retirement settings: under *flexible* retirement, the old generation can freely postpone or advance retirement in the second period after a realization of shocks; under *fixed* retirement, this generation has to make this decision already before shocks are revealed. Once set, this decision cannot be subsequently changed when new information becomes available.

Our analysis shows that the positive relation between retirement flexibility and more risk taking is weakened and under some conditions even turned around if not only depreciation shocks but also productivity shocks are considered. Depreciation shocks mainly affect the return on capital and through the income effect these shocks contribute to the traditional view that retirement flexibility increases risk-taking behaviour. Productivity shocks, in contrast, do not only affect capital returns but also influence wages. Consequently, productivity shocks also induce substitution effects in labour supply which work in the opposite direction. These substitution effects generate a positive correlation between asset returns and labour income, thereby reducing the risk-bearing capacity of consumers.

In addition, confining the analysis to Cobb-Douglas utility, as most of the existing studies do, ignores the essential role of the elasticity of substitution between consumption and leisure in studying retirement flexibility. This elasticity of substitution governs the relative strength of income and substitution effects in labour supply and, hence, determines the insurance provided by retirement flexibility. Our analysis clearly shows that flexible retirement amplifies consumption volatility if substitution effects are important, a notion also put forward by Basak (1999).

Finally, we find that general equilibrium effects play an important role in the interaction between portfolio choice and retirement. Ignoring these effects by sticking to a partial equilibrium framework can either overstate or understate the hedging effect of retirement flexibility, dependent on the willingness of consumers to substitute between consumption and leisure. If the elasticity of substitution is high, agents choose to supply less labour after a negative productivity shock. In general equilibrium, this labour-supply response exacerbates the direct fall in the return on capital due to the productivity contraction. Compared to partial equilibrium, this higher sensitivity of the capital return for productivity risk results in lower portfolio shares invested in equity. Of course, for low elasticities of substitution just the opposite holds: then the insurance effect is more effective in general than in partial equilibrium, leading to higher equity shares.

In the last conceptual chapter of this thesis, **Chapter 5**, we again turn to redistribution. Having focused on the level and direction of redistribution effects in collective pension schemes in Chapter 2, this final chapter concentrates on the economic implications of redistribution, in particular with respect to the retirement decision. The existing literature dealing with the economic implications of redistribution mainly focuses on pension reforms that strengthen the link between contributions and benefits, i.e., a reform from Beveridgean to Bismarckian pension systems (see e.g. Hougaard Jensen et al., 2003; Lau and Poutvaara, 2006; Cremer et al., 2010; Fisher and Keuschnigg, 2010). The analysis of this chapter is conducted in face of actual pension reforms that aim to reduce the generosity of pension entitlements (by e.g. cutting benefits or increasing the pension entitlement age) and to improve work incentives (by allowing for more flexibility in individual pension take-up). The goal of this chapter is to shed light on the question how these type of pension reforms will affect intragenerational redistribution and welfare of heterogeneous agents.

To investigate this, we use a two-period overlapping-generations model populated with heterogeneous agents who differ in age, ability and life span. It is assumed that the life span of an individual is positively linked to his productivity. As such, our model is consistent with empirical evidence that finds a strong positive association between longevity and socioeconomic status, either measured in terms of income or education level (see e.g., Adams et al., 2003). The benchmark PAYG social security system is of the Beveridgean type and is characterized by life-time annuities combined with inflexible pension take-up. In this way, the pension scheme includes two types of intragenerational redistribution, from high to low incomes and from short-lived to long-lived agents.

To introduce flexible pension take-up, policy makers have to determine how pension benefits are adjusted if people postpone or advance retirement. We analyse three scenarios which differ in the information set available to the government. In the first scenario, we assume that individual abilities and life spans are not publicly observable. The government then determines the actuarial adjustment of benefits on the basis of the average longevity of the total population. However, this uniform adjustment might be accompanied with selection effects because from an actuarial point of view the reward rate of retirement postponement is too low for people with short life spans (the low-skilled) and too high for people with long life spans (the high-skilled). Therefore, we also analyse two alternative reforms which intend to reduce this regressive impact. As a rather extreme alternative, we first assume that individual life spans are observable and the government uses this information to determine the actuarial adjustment factor. Then we take a more realistic perspective and impose that the government can only observe education levels which partly reveal information about longevity.

We show that automatically linking the pension benefit level to longevity benefits the low-skilled and harms the high-skilled. For low-skilled (*high-skilled*) agents the relative decrease in the benefit is lower (*higher*) than the relative increase in the payout period. The intragenerational redistribution effects related to a direct link between the pension entitlement age and longevity depend on the degree of life-span heterogeneity among low-skilled and high-skilled agents. The net benefit of agents is subject to two opposite forces. On the one hand, the increase in life span is in absolute terms higher for high-skilled agents than for low-skilled agents as long there as there is life-span heterogeneity. This implies that the payout period of the high-skilled increases more than that of the low-skilled. On the other hand, like discussed before, an increase in longevity induces people to retire later which increases the redistribution effects from high-income to low-income earners. If heterogeneity in life spans is low (*high*), this last (*first*) effect is dominating meaning that high-skilled (*low-skilled*) agents suffer from the reform.

More importantly, we find that introducing a flexible pension take-up can be a Pareto-improving reform if actuarial adjustment of benefits occurs in a uniform way (i.e., based on the average life expectancy). Uniform benefit adjustment leads to selection effects in the retirement decision which may reduce initial tax distortions. Indeed, for the high-skilled individuals the uniform reward rate for later retirement is too high from an actuarial point of view, which reduces their implicit tax and stimulates them to continue working. This shows that uniform actuarial adjustment is an instrument for the government to effectively vary taxes over the life cycle. If the payroll tax is sufficiently high, the low-skilled also gain because the continued activity of the high-skilled generates more tax revenues and thus higher pension benefits. If the government would use non-uniform benefit adjustment instead, for example by conditioning the adjustment factor on individual or skill-group life expectancies, flexible pension take-up cannot be Pareto improving. With such a more actuarially-neutral approach, selection effects are less important and therefore also the opportunities to reduce existing distortions. Non-uniform actuarial adjustment then only eliminates the unintended transfers from short-lived to long-lived individuals and, hence, make the pension scheme less regressive (because life span is positively related with earnings).

Finally, combining uniform adjustment with actuarial non-neutrality aimed at stimulating labour supply can further improve the flexibility reform. Indeed, it enables a Pareto improvement with even lower contribution rates. The intuition is that actuarial non-neutrality can further reduce the downward distortion of the contribution tax rate, the removal of which brings additional resources.

6.2 Policy implications

What are the main policy implications of the insights obtained in this thesis? At this point, we should urge some caution because most results have been derived using stylized two-period models. That being said, it is possible to sketch some lines of thought which could be relevant for policy makers and which could pave the way for further research.

This thesis has indicated that the application of uniform pricing leads to large unintended transfers from males to females and from low-educated to high-educated people (Chapter 2). For the European legislator uniform pricing has always been considered a necessary element of solidarity to justify the lack of competition caused by the mandatory participation for companies and individuals. From a social point of view, however, this solidarity could be questioned because it results in transfers from groups of people with a high mortality risk (i.e., usually the low-educated) to groups of people with a low mortality risk (i.e., usually the high-educated). This unintended outcome should stimulate policy makers to reduce socioeconomic inequalities in mortality, for example, by improving labour conditions of physically demanding jobs or by encouraging low-educated people to stop smoking.

Apart from this more fundamental solution, there are also conceivable options to change the pension scheme in such a way that the unintended form of solidarity reduces. For example, pension institutions could choose to differentiate more to homogeneous groups of participants in the sharing of mortality risk. This option would imply the use of group-specific instruments and therefore sets high standards with regard to the information of the participants available to the pension fund. A probably less complicated alternative is to introduce a progressive contribution system or to cut-off the accrual of pension entitlements above a certain income threshold. It is empirically well known that there exists a positive relation between life expectancy and income. Hence, by introducing progressive contributions or by limiting pension accrual up to certain income thresholds, the pension fund can reduce the implicit subsidy (*tax*) imposed to long-lived (*short-lived*) people.

Apart from redistribution issues, this thesis also provides valuable insights on the issue of intergenerational risk sharing through pension institutions. In many developed countries we observe a move away from collective funded or unfunded pension schemes with defined benefits towards the establishment of individual funded schemes with defined contributions. In the face of population ageing, this development makes sense because in a defined-benefit context a decreasing group of young workers should bear the risks of an increasing group of elderly. However, the question is whether a fully individualized defined-contribution scheme should really be the end point of this development. The insights obtained in this thesis might be relevant for policy makers that have to deal with this question. In fact, our results emphasize that individual pension schemes that do not share risks among generations may not be optimal (Chapter 3). Collective funded pension schemes with well-structured intergenerational risk sharing are preferable from a welfare point of view, even if the losses from labour-supply distortions are taken into account as well as correlated shocks at the capital market and the labour market.

The development towards individual flexible pension schemes also raises new policy issues. The notion that the investment policy in these schemes should be based on individual preferences for retirement will become increasingly important. This thesis shows that risky investment strategies are not always in the interest of individuals, even if they have the flexibility to choose their own retirement date (Chapter 4). This is in particular the case if shocks to pension wealth and wages are positively correlated or if consumers view leisure and consumption as close substitutes. Empirical evidence indeed suggests that the stock market and human capital are positively correlated (Baxter and Jermann, 1997; Benzoni et al., 2007) and that substitution effects are more important than income effects in retirement behaviour (Gruber and Wise, 1999; Coile and Gruber, 2001 and French, 2005). These results are relevant for private or public pension institutions, like corporate pension funds, trust funds or life-insurance companies, to which individuals have dedicated or will dedicate their saving and investment decisions.

Our research emphasizes that retirement flexibility is not only relevant for funded pensions, but also for unfunded pensions. This thesis shows that the introduction of a variable date of pension take-up in first-pillar Beveridgean pension schemes may have two positive effects: it is likely to restore the financial balance of the system (in the face of the growing burden of ageing), and it may foster redistribution from rich to poor (Chapter 5). To obtain this double advantage, an important condition is that actuarial adjustment of pension entitlements (when retirement is postponed or advanced) should be done in a uniform way. In many real-world pension schemes, actuarial adjustment is indeed implemented independent of individual characteristics, like life expectancy or skill level. The results of this chapter might give a rationale for this kind of flexibility reforms, based on uniform actuarial adjustment.

Despite all kinds of measures undertaken to stimulate work, in most countries the penalty rates of early retirement are still below the actuarially-neutral level (Queisser and Whitehouse, 2006). This implies that there is still room to improve pension schemes by going into the direction of complete actuarial neutrality. In fact, our analysis would suggest increasing the reward rates of retirement postponement (or the penalty rates of early retirement) to levels even beyond the actuarially-neutral one. Stimulating work continuation in an actuarially non-neutral way could partly undo the existing distortions on the retirement decisions related to the proportional contribution tax and could therefore further improve a reform towards flexible pension take-up.

6.3 Research agenda

Our study suggests that there are large transfers from low-educated groups to high-educated groups within funded second-pillar pension schemes based on uniform pricing (Chapter 2). Other studies, in contrast, have pointed out a reverse redistribution from high-educational groups to low-educational groups within unfunded first-pillar pension schemes (see e.g., Ter Rele, 2007). Recent research for the Netherlands suggests that this redistribution from high-educated to low-educated persons in the first pillar dominates the reverse redistribution in the second pillar (Bonenkamp and Ter Rele, 2013). Preferably, such an integrated analysis also incorporates other parts of the public sector, like the tax system, health care and education.

Considering empirical challenges, it would be interesting to study the role of inter- and intragenerational redistribution on labour-supply incentives. Empirical evidence shows that the labour-supply elasticity of women is generally higher than that of men (Evers et al., 2008). This indicates that the subsidy women generally receive from men in collective funded pension schemes positively affects labour supply. However, there is still a lack of empirical knowledge to what extent labour-supply elasticities depend on age or skill level. The few studies which are available find a positive relation between the labour-supply elasticity and age (see French, 2005 and Fenge et al., 2006). This suggests that the transfers from young to old generations in pension schemes with uniform pricing may stimulate total labour supply.

Regarding the welfare gains of intergenerational risk sharing provided by collective funded pension schemes, our analysis focuses on the alleviation of the constraint that prevents generations to trade in financial markets before they are born (Chapter 3). Considering theoretical challenges, further research is needed to evaluate the role of other sources of market incompleteness. In practice, pension benefits are often linked to current wages enabling older people to acquire a claim on the human capital of the young. In this way, the pension scheme alleviates another market inefficiency related to the non-tradability of human capital. Perhaps this market incompleteness is even more important than the inability to trade risks before agents are born.

Our analysis focuses on the *optimal* degree of risk sharing provided by pension funds. In future research, it would be interesting to study the *actual* degree of risk sharing provided by pension funds and how this compares with the optimal amount. How much do pension funds in practice contribute to risk sharing between current and future generations? And what is the government contribution to intergenerational risk sharing? So far, there is very little empirical evidence on these issues. At a more fundamental level, an important question is which institution, governments or pension funds, is better equipped to share risks between generations. The advantage of the government is that it has the power of taxation to commit future generations. For pension funds it is more difficult to enforce commitment as people can in principle move to another fund. On the other hand, pension schemes carried out by the government are probably more sensitive to all kinds of political risks.

With respect to the interaction between retirement flexibility and portfolio allocation (Chapter 4), first and foremost we need more empirical knowledge. Our analysis raises a number of interesting empirical questions. First, will retirement (or labour) flexibility induce a greater risk taking in an individual's asset portfolio? Answering this question requires an identification of a measure of flexibility. A possibility is to compare job categories with a fixed amount of hours with job categories that offer opportunities for working extra hours. Concerning retirement flexibility, a possibility is to focus on evidence from actual policy reforms that move from inflexible to flexible pension contracts. A second relevant empirical question is whether a higher riskiness of individual's human capital indeed leads to less risk taking, as our analysis suggests. One way to answer this question is by investigating whether there is a negative association between the correlation between wages and asset returns on the one hand and risk-taking investment behaviour on the other hand.

From a theoretical perspective, a direction of further research is to analyse the influence of alternative social security schemes on retirement and investment decisions. Retirement flexibility and social security schemes have in common that they both can protect retirees against adverse financial shocks. Since pension systems (notably unfunded ones) serve as an additional asset which implicit return may be imperfectly correlated with asset returns, they will certainly affect individual portfolio and retirement decisions. Moreover, social security schemes may also affect these economic decisions through their impact on the implicit marginal tax wedge or through their provision of intergenerational risk sharing.

As regards to redistributive pension schemes, more empirical research is needed on how intergenerational and intragenerational transfers influence actual economic decisions (Chapter 5). In the context of our analysis of flexible pension take-up in Beveridgean pensions, the question is to what extent any difference in retirement behaviour of individuals with rather different socioeconomic characteristics can be attributed to these redistribution effects? Our analysis presumes that individuals not only know how their relative socioeconomic position interferes with the pension scheme in which they are participating, they also (implicitly or explicitly) use this information in their economic decisions. However, there is empirical evidence that calls the rationality of people on these issues into question (see e.g., Van Rooij, 2008).

Finally, as a theoretical contribution, the relevance of a flexible first-pillar pension scheme model can also be studied in a political economy model with endogenous retirement and heterogeneous agents. It is interesting to analyse the political support for a switch to flexible pension take-up and to characterize the equilibrium (majority voting) size of the pension scheme for a given bias in the benefit formula.

BIBLIOGRAPHY

Aarssen, K. and B. Kuipers (2007), "Everyone gains, but some more than others," in O. Steenbeek and S. Van Der Lecq (eds.), *Costs and benefits of collective pension systems*, Springer, Berlin, pp. 137–156.

Adams, P., M. Hurd, D. McFadden, A. Merrill, and T. Ribeiro (2003), "Healthy, wealthy, and wise? Tests for direct causal paths between health and socioeconomic status," *Journal of Econometrics*, **112**, 3–56.

Adema, Y. (2008), *The international spillover effects of ageing and pensions*, PhD dissertation, Chapter 4, Tilburg University.

Andersen, T. (2005), "Social security and longevity," Mimeo.

Asch, B., S. Haider, and J. Zissimopoulos (2005), "Financial incentives and retirement: evidence from federal civil service workers," *Journal of Public Economics*, **89**, 427–440.

Attanasio, O. and G. Weber (1995), "Is consumption growth consistent with intertemporal optimization? Evidence for the consumer expenditure survey," *Journal of Political Economy*, **103**, 1121–1157.

Ball, L. and N. Mankiw (2007), "Intergenerational risk sharing in the spirit of Arrow, Debreu, and Rawls, with applications to social security design," *Journal of Political Economy*, **115**, 523–547.

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

Basak, S. (1999), "On the fluctuations in consumption and market returns in the presence of labor and human capital," *Journal of Economic Dynamics and Control*, **23**, 1029–1064.

Baxter, M. and U. Jermann (1997), "The international diversification puzzle is worse than you think," *American Economic Review*, **87**, 170–180.

Beetsma, R. and A. Bovenberg (2009), "Pensions and intergenerational risk-sharing in general equilibrium," *Economica*, **76**, 364–386.

Beetsma, R., W. Romp, and S. Vos (2013), "Intergenerational risk sharing, pensions and endogenous labour supply in general equilibrium," *Scandinavian Journal of Economics*, **115**, 141–154.

Benzoni, L., P. Collin-Dufresne, and R. Goldstein (2007), "Portfolio choice over the life-cycle when the stocks and labor markets are cointegrated," *Journal of Finance*, **62**, 2123–2167.

Blanchard, O. and S. Fisher (1989), *Lectures on Macroeconomics*, The MIT Press, Cambridge.

Bloemen, H. (2011), "The effect of private wealth on the retirement rate: an empirical analysis," *Economica*, **78**, 637–655.

Blundell, R. and T. MaCurdy (1999), "Labor supply: a review of alternative approaches," in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, North Holland, Amsterdam, pp. 1559–1695.

Bodie, Z., R. Merton, and W. Samuelson (1992), "Labor supply flexibility and portfolio choice in a life cycle model," *Journal of Economic Dynamics and Control*, **16**, 427–449.

Boeijen, D., C. Jansen, C. Kortleve, and J. Tamerus (2007), "Intergenerational solidarity in the uniform contribution," in O. Steenbeek and S. Van Der Lecq (eds.), *Costs and benefits of collective pension systems*, Springer, Berlin, pp. 119–136.

Bohn, H. (2009), "Intergenerational risk sharing and fiscal policy," *Journal of Monetary Economics*, **56**, 805–816.

Bonenkamp, J. (2009), "Measuring lifetime redistribution in Dutch occupational pensions," *De Economist*, **157**, 49–77.

Bonenkamp, J. and H. Ter Rele (2013), "Herverdeling door pensioenregelingen: een integrale analyse van de AOW en de aanvullende pensioenen," *TPEdigitaal*, **7**, 51–65 (in Dutch).

Borck, R. (2007), "On the choice of public pensions when income and life expectancy are correlated," *Journal of Public Economic Theory*, **9**, 711–725.

Börsch-Supan, A. and A. Reil-Held (2001), "How much is transfer and how much is insurance in a pay-as-you-go system? The German case," *Scandinavian Journal of Economics*, **103**, 505–524.

Bovenberg, A., R. Koijen, T. Nijman, and C. Teulings (2007), "Saving and investing over the life cycle and the role of collective pension funds," *De Economist*, **155**, 347–415.

Bovenberg, A., J. Mackenbach, and R. Mehlkopf (2006), "Een eerlijk en vergrijzingbestendig ouderdomspensioen," *Economische Statistische Berichten*, **91**, 648–651 (in Dutch).

Bovenberg, A. and R. Mehlkopf (2013), "De risico-opslag in de nieuwe reële pensioencontracten," Mimeo.

Bovenberg, A. and H. Uhlig (2008), "Pension systems and the allocation of macroeconomic risk," in L. Reichlin and K. West (eds.), *NBER international seminar on macroeconomics 2006*, University of Chicago Press, Chicago, pp. 241–344.

Bovenberg, A. and C. Van Ewijk (2012), "The future of multi-pillar pension systems," in A. Bovenberg, C. Van Ewijk, and E. Westerhout (eds.), *The future of multi-pillar pensions*, Cambridge University Press, Cambridge, pp. 373–418.

Campbell, J. (1994), "Inspecting the mechanism: an analytical approach to the stochastic growth model," *Journal of Monetary Economics*, **33**, 463–506.

Campbell, J. and L. Viceira (2002), *Strategic asset allocation; portfolio choice for long-term investors*, Oxford University Press, New York.

Campbell, J. and L. Viceira (2005), "The term structure of the risk-return tradeoff," *Financial Analysts Journal*, **61**, 34–44.

Casamatta, G., H. Cremer, and P. Pestieau (2000), "The political economy of social security," *Scandinavian Journal of Economics*, **102**, 503–522.

Casamatta, G., H. Cremer, and P. Pestieau (2005), "Voting on pensions with endogenous retirement age," *International Tax and Public Finance*, **12**, 7–28.

Choi, K. and G. Shim (2006), "Disutility, optimal retirement, and portfolio selection," *Mathematical Finance*, **16**, 443–467.

Choi, K., G. Shim, and Y. Shin (2008), "Optimal portfolio, consumption-leisure and retirement choice problem with CES utility," *Mathematical Finance*, **18**, 445–472.

Coile, C. and J. Gruber (2001), "Social security incentives for retirement," in D. Wise (ed.), *Themes in the economics of aging*, University of Chicago Press, Chicago, pp. 311–341.

Collard, F. and M. Juillard (2001), "Accuracy of stochastic perturbations methods: the case of asset pricing models," *Journal of Economic Dynamics and Control*, **25**, 979–999.

Conde-Ruiz, J. and P. Profeta (2007), "The redistributive design of social security systems," *The Economic Journal*, **117**, 686–712.

Constantinides, G., J. Donaldson, and R. Mehra (2002), "Junior can't borrow: a new perspective on the equity premium puzzle," *Quarterly Journal of Economics*, **117**, 269–296.

Cörvers, F., R. Euwals, and A. de Grip (2011), *Labour market flexibility in the Netherlands*, De Swart, The Hague.

Cremer, H., J.-M. Lozachmeur, and P. Pestieau (2010), "Collective annuities and redistribution," *Journal of Public Economic Theory*, **12**, 23–41.

Cremer, H. and P. Pestieau (2003a), "The double dividend of postponing retirement," *International Tax and Public Finance*, **10**, 419–434.

Cremer, H. and P. Pestieau (2003b), "Social insurance competition between Bismarck and Beveridge," *Journal of Urban Economics*, **54**, 181–196.

Cubeddu, L. (2000), "Intragenerational redistribution in unfunded pension systems," *IMF Staff Papers*, **47**, 90–115.

Cui, J., F. De Jong, and E. Ponds (2011), "Intergenerational risk sharing within funded pension schemes," *Journal of Pension Economics and Finance*, **10**, 1–29.

Deboosere, P. and S. Gadeyne (2002), "Levensverwachting en sterftekansen naar geslacht, gewest en onderwijs-niveau in België, 1991-1996," *Bevolking en gezin*, **31**, 47–74 (in Dutch).

Demange, G. (2002), "On optimality in intergenerational risk sharing," *Economic Theory*, **20**, 1–27.

Diamond, P. (1965), "National debt in a neoclassical growth model," *American Economic Review*, **55**, 1126–1150.

Diamond, P. (1977), "A framework for social security analysis," *Journal of Public Economics*, **8**, 275–298.

Diamond, P. and J. Mirrlees (1978), "A model of social insurance with variable retirement," *Journal of Public Economics*, **10**, 295–336.

DNB (2009), Statistisch Bulletin (december), De Nederlandsche Bank, Amsterdam (in Dutch).

Enders, W. and H. Lapan (1982), "Social security taxation and intergenerational risk sharing," *International Economic Review*, **23**, 647–658.

Epstein, L. and S. Zin (1991), "Substitution, risk aversion, and the temporal behavior of consumption and asset returns: an empirical analysis," *Journal of Political Economy*, **99**, 263–286.

Euwals, R. and D. Van Vuuren (2005), "Arbeidsaanbod tot 2050: een beleidsneutraal scenario," CPB Memorandum 134, Centraal Planbureau, The Hague (in Dutch).

Evers, M., R. De Mooij, and D. Van Vuuren (2008), "The wage elasticity of labour supply: a synthesis of empirical estimates," *De Economist*, **156**, 25–43.

Farhi, E. and S. Panageas (2007), "Saving and investing for early retirement: a theoretical analysis," *Journal of Financial Economics*, **83**, 87–121.

Fehr, H. and C. Habermann (2008), "Risk sharing and efficiency implications of progressive pension arrangements," *Scandinavian Journal of Economics*, **110**, 419–443.

Feldstein, M. (1985), "The optimal level of social security benefits," *Quarterly Journal of Economics*, **100**, 303–320.

Fenge, R., S. Uebelmesser, and M. Werding (2006), "On the optimal timing of implicit social security taxes over the life cycle," *FinanzArchiv*, **62**, 68–107.

Fisher, W. and C. Keuschnigg (2010), "Pension reform and labor market incentives," *Journal of Population Economics*, **23**, 769–803.

French, E. (2005), "The effects of health, wealth, and wages on labour supply and retirement behaviour," *Review of Economic Studies*, **72**, 395–427.

French, E. and J. Jones (2012), "Public pensions and labor supply over the life cycle," *International Tax and Public Finance*, **19**, 268–287.

Galasso, V. and P. Profeta (2002), "The political economy of social security: a survey," *European Journal of Political Economy*, **18**, 1–29.

Gollier, C. (2008), "Intergenerational risk-sharing and risk-taking of a pension fund," *Journal of Public Economics*, **92**, 1463–1485.

Gordon, R. and H. Varian (1988), "Intergenerational risk sharing," *Journal of Public Economics*, **37**, 185–202.

Gorski, M., T. Krieger, and T. Lange (2007), "Pensions, education and life expectancy," Mimeo.

Gottardi, P. and F. Kubler (2011), "Social security and risk sharing," *Journal of Economic Theory*, **146**, 1078–1106.

Gruber, J. and D. Wise (1999), *Social security and retirement around the world*, University of Chicago Press, Chicago.

Hachon, C. (2008), "Redistribution, pension systems and capital accumulation," *Financial Theory and Practice*, **32**, 339–368.

Hansen, L. and K. Singleton (1983), "Stochastic consumption, risk aversion, and the temporal behavior of asset returns," *Journal of Political Economy*, **91**, 249–265.

Hári, N., R. Koijen, and T. Nijman (2006), "The determinants of the money's worth of participation in collective pension schemes," Mimeo.

Hoevenaars, R. and E. Ponds (2008), "Valuation of intergenerational transfers in funded collective pension schemes," *Insurance: Mathematics and Economics*, **42**, 578–593.

Hougaard Jensen, S. and O. Jørgensen (2008), "Uncertain demographics, longevity adjustment of the retirement age and intergenerational risk-sharing," in J. Alho, S. Hougaard Jensen, and J. Lassila (eds.), *Uncertain demographics and fiscal sustainability*, Cambridge University Press, Cambridge, pp. 239–257.

Hougaard Jensen, S., M. Lau, and P. Poutvaara (2003), "Efficiency and equity aspects of alternative social security rules," *FinanzArchiv*, **60**, 325–358.

Hoyert, D., E. Arias, B. Smith, S. Murphy, and K. Kochanek (2001), "Deaths: final data for 1999," *National Vital Statistics Reports*, **49**, 1–116.

Jaag, C., C. Keuschnigg, and M. Keuschnigg (2010), "Pension reform, retirement, and life-cycle unemployment," *International Tax and Public Finance*, **17**, 556–585.

Juillard, M. and O. Kamenik (2005), "Solving SDGE models: approximation about the stochastic steady state," *Computing in Economics and Finance 2005*, **106**, Society for Computational Economics.

King, R., C. Plosser, and S. Rebelo (2002), "Production, growth and business cycles: technical appendix," *Computational Economics*, **20**, 87–116.

Knoef, M. (2006), "An APC analysis of female labour force participation in the Netherlands," Mimeo.

Koethenburger, M., P. Poutvaara, and P. Profeta (2008), "Why are more redistributive social security systems smaller? A median voter approach," *Oxford Economic Papers*, **60**, 275–292.

Kolmar, M. (2007), "Beveridge versus Bismarck public-pension systems in integrated markets," *Regional Science and Urban Economics*, **37**, 649–669.

Krueger, A. and S. Pischke (1992), "The effects of social security on labor supply: a cohort analysis of the Notch generation," *Journal of Labor Economics*, **10**, 412–437.

Krueger, D. and F. Kubler (2006), "Pareto-improving social security reform when financial markets are incomplete!?" *American Economic Review*, **96**, 737–755.

Kuné, J. (2005), "Billijkheid en doelmatigheid in het systeem van de (aanvullende) pensioenvoorziening," *Financiële en monetaire studies*, **23**, (in Dutch).

Lachance, M.-E. (2004), "Retirement income insurance: a do-it-yourself approach," Mimeo.

Lau, M. and P. Poutvaara (2006), "Social security incentives and human capital investment," *Finnish Economic Papers*, **19**, 16–24.

Lutjens, E. (2007), "Doorsneepremie: solidariteit of discriminatie?" *Tijdschrift voor pensioenvraagstukken*, **3**, 84–89 (in Dutch).

Mackenbach, J., V. Bos, O. Andersen, M. Cardano, G. Costa, S. Harding, A. Reid, O. Hemström, T. Valkonen, and A. Kunst (2003), "Widening socioeconomic inequalities in mortality in six Western European countries," *International Journal of Epidemiology*, **32**, 830–837.

Matsen, E. and Ø. Thøgersen (2004), "Designing social security - a portfolio choice approach," *European Economic Review*, **48**, 883–904.

Meara, E., S. Richards, and D. Cutler (2008), "The gap gets bigger: changes in mortality and life expectancy, by education, 1981-2000," *Health Affairs*, **27**, 350–360.

Mehlkopf, R. (2011), *Risk sharing with the unborn*, PhD dissertation, Tilburg University.

Merton, R. (1969), "Lifetime portfolio selection under uncertainty: the continuous time case," *Review of Economics and Statistics*, **51**, 247–257.

Natali, D. (2008), *Pensions in Europe, European pensions*, PIE-Peter Lang, Brussels.

Nelissen, J., H. Verbon, D. Van Kampen, and P. Vermaseren (2011), "Herverdelingsaspecten van het pensioenakkoord," *Economische Statistische Berichten*, **96**, 198–201 (in Dutch).

Nishiyama, S. and K. Smetters (2007), "Does social security privatization produce efficiency gains?" *Quarterly Journal of Economics*, **122**, 1677–1719.

OECD (2006), *Live longer, work longer*, OECD Publishing, Paris.

OECD (2007), *Pensions at a glance: public policies across OECD countries*, OECD Publishing, Paris.

OECD (2009), *Pensions at a glance: retirement-income systems in OECD countries*, OECD Publishing, Paris.

OECD (2011), *Pensions at a glance: retirement-income systems in OECD and G20 countries*, OECD Publishing, Paris.

Olovsson, C. (2010), "Quantifying the risk-sharing welfare gains of social security," *Journal of Monetary Economics*, **57**, 364–375.

Pappas, G., S. Queen, W. Hadden, and G. Fisher (1993), "The increasing disparity in mortality rates between socioeconomic groups in the United States: 1960 and 1986," *New England Journal of Medicine*, **329**, 103–109.

Persson, M. (1983), "The distribution of abilities and the progressive income tax," *Journal of Public Economics*, **22**, 73–88.

Pestieau, P. and U. Possen (2010), "Retirement as a hedge," Mimeo.

Queisser, M. and E. Whitehouse (2006), "Neutral or fair? Actuarial concepts and pension-system design," Mimeo.

Razin, A. and E. Sadka (1999), "Migration and pension with international capital mobility," *Journal of Public Economics*, **74**, 141–150.

Samuelson, P. (1958), "An exact consumption loan model of interest with or without the social contrivance of money," *Journal of Political Economy*, **66**, 467–482.

Samuelson, P. (1969), "Lifetime portfolio selection by dynamic stochastic programming," *Review of Economic and Statistics*, **51**, 239–246.

Samuelson, P. (1970), "The fundamental approximation theorem of portfolio analysis in terms of means, variances, and higher moments," *Review of Economic Studies*, **37**, 537–542.

Sánchez-Marcos, V. and A. Sánchez-Martín (2006), "Can social security be welfare improving when there is demographic uncertainty?" *Journal of Economic Dynamics and Control*, **30**, 1615–1646.

Sandmo, A. (1970), "The effect of uncertainty on savings decisions," *Review of Economic Studies*, **3**, 353–360.

Schmitt-Grohé, S. and M. Uribe (2004), "Solving dynamic general equilibrium models using a second-order approximation to the policy function," *Journal of Economic Dynamics and Control*, **28**, 755–775.

Shiller, R. (1999), "Social security and institutions for intergenerational, intragenerational and international risk sharing," *Carnegie-Rochester Conference Series on Public Policy*, **50**, 165–204.

Sinn, H.-W. (2000), "Why a funded pension system is useful and why it is not useful," *International Tax and Public Finance*, **7**, 389–410.

Sommacal, A. (2006), "Pension systems and intragenerational redistribution when labor supply is endogenous," *Oxford Economic Papers*, **58**, 379–406.

Ter Rele, H. (2007), "Measuring the lifetime redistribution achieved by Dutch taxation, cash transfer and non-cash benefits programs," *Review of Income and Wealth*, **53**, 335–362.

Teulings, C. and C. De Vries (2006), "Generational accounting, solidarity and pension losses," *De Economist*, **154**, 63–83.

Towers Watson (2011), *Global pension asset study 2011*.

Uhlig, H. (1999), "A toolkit for analysing nonlinear dynamic stochastic models easily," in R. Marimom and A. Scott (eds.), *Computational methods for the study of dynamic economies*, Oxford University Press, Oxford, pp. 30–61.

Van Ewijk, C. (2005), "Reform of occupational pensions in the Netherlands," *De Economist*, **153**, 331–347.

Van Herten, L., K. Oudshoorn, R. Perenboom, Y. Mulder, N. Hoeymans, and D. Deeg (2002), "Gezonde levensverwachting naar sociaal-economische status," TNO rapport 2002.170, Leiden.

Van Rooij, M. (2008), *Financial literacy, retirement provisions, and household portfolio behavior: four empirical contributions*, PhD dissertation, Utrecht University.

Van Vuuren, D. (2011), "Flexible retirement," Mimeo.

Varian, H. (1980), "Redistributive taxation as social insurance," *Journal of Public Economics*, **14**, 49–68.

Viceira, L. (2001), "Optimal portfolio choice for long-horizon investors with nontradable labor income," *Journal of Finance*, **56**, 433–470.

World Bank (1994), *Averting the old age crisis: policies to protect the old and promote growth*, World Bank, Washington D.C.

SAMENVATTING (SUMMARY IN DUTCH)

Nederlandse pensioenfondsen staan voor grote uitdagingen. Als gevolg van nog steeds dalende sterftekansen, worden mensen gemiddeld genomen steeds ouder. De levensverwachting bij geboorte bedraagt momenteel 79,2 jaar voor mannen en 82,9 jaar vrouwen. Ter vergelijking: in 2000 was dat nog 75,5 jaar voor mannen tegenover 80,6 jaar voor vrouwen. Dat mensen steeds langer leven is ontegenzeggelijk een groot goed. Voor de financiële houdbaarheid van pensioenfondsen is het echter minder goed nieuws omdat uitkeringen over een langere periode betaald moeten worden. Daar komt bij dat door de economische crisis van de afgelopen jaren pensioenfondsen er financieel sowieso niet florissant voor staan. Afgezien van het oplossen van deze financiële problemen, zullen fondsen ook een antwoord moeten bieden op ontwikkelingen van meer sociaaleconomische aard, zoals de toegenomen arbeidsmobiliteit, de individualisering van de samenleving en, daaruit voortvloeiend, een mogelijk afnemend draagvlak voor solidariteit.

In het licht van deze demografische, financiële en sociaaleconomische ontwikkelingen, zijn de afgelopen jaren in veel landen plannen ontwikkeld of reeds uitgevoerd om het pensioenstelsel te hervormen. De invulling van die hervormingen verschilt vaak sterk van land tot land. Sommige landen hebben enkele grote hervormingen doorgevoerd of aangekondigd, zoals een algemene stijging van de pensioenleeftijd of een koppeling van de pensioenuitkering aan de levensverwachting. Andere landen kiezen voor een reeks kleinere aanpassingen, zoals het afschaffen van regelingen voor vroegpensioen of een verhoging van de pensioenleeftijd alleen voor vrouwen, die bij elkaar opgeteld ook een groot effect hebben op de toekomstig pensioenen. Hoe sterk de concrete invulling ook kan verschillen, de onderliggende motieven van de hervormingen zijn hetzelfde: het verbeteren van de houdbaarheid en schokbestendigheid van het pensioenstelsel, het stimuleren om langer door te werken en het introduceren van (meer) individuele keuzevrijheid.

Hervormen van collectieve pensioenstelsels gaat vrijwel nooit zonder slag of stoot. Toen de Nederlandse regering in 2005 aankondigde om de fiscale voordelen van prepensioenregelingen af te schaffen, kwamen vakbonden massaal in het geweer. De vakbonden zagen deze plannen als inbreuk op de pensioenrechten van oudere deelnemers waarvoor jarenlang premie was betaald. Deze sociale bewogenheid bij collectieve pensioenhervormingen is niet vreemd als we een blik werpen op de kale cijfers: in de meeste landen bedraagt het aandeel van collectieve uitkeringen (eerste en tweede pijler samen) in het totale pensioeninkomen meer dan 80%. In Nederland ligt dit percentage zelfs op ruim 90%. Toch is dit niet het hele verhaal. Door de combinatie van verplichte deelname en de toepassing van vaak *uniforme* regelingen op *heterogene* deelnemers, is het onomkoombaar dat collectieve pensioenregelingen premiegelden herverdelen, zowel binnen als tussen generaties. Een hervorming kan daardoor voor een bepaalde groep beter uitpakken dan voor een andere.

Dergelijke belangentegenstellingen zouden kunnen verhullen dat collectieve pensioenen ook elementen in zich herbergen die in principe voor alle belanghebbenden voordelig zijn. Een belangrijk voorbeeld daarvan is intergenerationele risicodeling: door verplichte deelname zijn collectieve pensioenstelsels in staat om risico's over meerdere generaties uit te smeren. Door schokken over een grotere groep te spreiden, kan de individuele bijdrage lager zijn hetgeen de welvaart van alle verzekerden kan verhogen.

Herverdeling en risicodeling vormen twee wezenskenmerken van collectieve pensioenstelsels. Dit proefschrift analyseert de effecten ervan op het economisch gedrag van individuen, zowel op gebied van consumptie en sparen als ook (en vooral) op gebied van arbeidsaanbod en pensionering. Hoewel dat in de praktijk vaak niet eenduidig te bepalen is, houden we een strikte scheiding aan tussen herverdeling en risicodeling. Onder *risicodeling* verstaan we een overdracht tussen deelnemers die plaatsvindt conditioneel op een bepaalde schok in het pensioenstelsel. Deze vorm van overdracht vormt het basisprincipe van verzekeren. Het is te vergelijken met een brandverzekering op een woonhuis die alleen tot uitbetaling komt indien er daadwerkelijk brandschade is ontstaan. *Herverdeling*, echter, betreft een overdracht tussen deelnemers die plaatsvindt onafhankelijk van het optreden van een schok. Dit is een vorm van onconditionele solidariteit die weinig tot niets met verzekeren te maken heeft. Een voorbeeld van herverdeling zijn de premie-overdrachten van de hoge naar de lage inkomens in een pensioenstelsel waarin de uitkering voor iedereen hetzelfde is (zoals de AOW in Nederland).

De analyse van de effecten van risicodeling en herverdeling in collectieve pensioenen moet geplaatst worden in het licht van drie actuele pensioenhervormingen die ook voor de Nederlandse beleidsdiscussie van belang zijn: de overstap van collectieve *defined-benefit* (DB) regelingen (uitkeringen zeker, premies onzeker) op individuele *defined-contribution* (DC) regelingen (uitkeringen onzeker, premies zeker), het verhogen van de pensioenleeftijd en het introduceren van flexibele pensioenregelingen waarin deelnemers zelf kunnen beslissen op welke leeftijd de uitkering ingaat. Afhankelijk van de pensioenhervorming in kwestie, heeft de analyse betrekking op een omslagstelsel (waarin pensioenen worden betaald uit de lopende premies) of op een kapitaaldekkingsstelsel (waarin pensioenen worden betaald uit eigen gespaarde premies). Ten aanzien van kapitaaldekkingsstelsels beoogt dit proefschrift ten eerste inzicht te verschaffen in de omvang en richting van de herverdelingseffecten (hoofdstuk 2). Ten tweede, in het licht van de afnemende populariteit van DB regelingen, speelt de vraag hoeveel waarde gehecht moet worden aan intergenerationele risicodeling die met deze ontwikkeling immers verloren dreigt te gaan (hoofdstuk 3). Ten derde gaan wij in op de vraag in hoeverre flexibele pensioencontracten individuen een verzekering kunnen bieden voor onverwachte schokken op de kapitaal- of arbeidsmarkt (hoofdstuk 4). Wat betreft omslagstelsels, analyseren we in hoeverre een directe koppeling tussen de pensioenleeftijd en een toename in de levensverwachting de herverdeling beïnvloedt tussen laag- en hoogopgeleiden. Ook relevant in dit verband is de vraag of het flexibiliseren van de pensioenleeftijd kan leiden tot een pensioenregeling die voor alle deelnemers per saldo beter uitpakt (hoofdstuk 5).

Inzicht in de herverdeling tussen groepen is van groot belang omdat zowel de omvang als de richting ervan van invloed zijn op het maatschappelijk draagvlak voor het pensioenstelsel. **Hoofdstuk 2** kwantificeert de herverdeling in het Nederlandse stelsel van aanvullende pensioenen. Herverdeling wordt gemeten aan de hand van het netto profijt dat een deelnemer ontleent aan deelname, waarbij netto profijt is gedefinieerd als het verschil tussen de contante waarde van alle toekomstige uitkeringen en pensioenpremies. Een positief (*negatief*) netto profijt betekent dus dat een deelnemer meer (*minder*) ontvangt dan deze persoon heeft ingelegd. Door netto profijt op levensloop basis te definiëren, wordt gecorrigeerd voor leeftijdsafhankelijke overdrachten die gedurende de carrière zowel positief als negatief kunnen uitvallen.

In het Nederlandse stelsel van aanvullende pensioenen bouwt elke deelnemer hetzelfde percentage van het pensioengevend salaris op als pensioenaanspraak (doorsnee-opbouw) en betaalt daarvoor procentueel gezien dezelfde premie (doorsneepremie). Deze zogenoemde doorsneesystematiek drijft een wig tussen de kostprijs die voor iedere deelnemer hetzelfde is en de actuariële waarde van de pensioenaanspraak die afhankelijk is van persoonlijke karakteristieken, zoals leeftijd, geslacht of levensverwachting. Zo is de waarde van een opgebouwd pensioenrecht voor een deelnemer met een hoge levensverwachting hoger dan voor een deelnemer met een lage levensverwachting, simpelweg omdat gezonde mensen naar verwachting langer van hun pensioen kunnen genieten. Omdat beide personen dezelfde premie betalen (als percentage van het inkomen), vindt er herverdeling plaats van de deelnemer met de lage naar de deelnemer met de hoge levensverwachting.

Om de omvang van de herverdeling over de levensloop te bepalen, maken we onderscheid tussen herverdeling tussen generaties (intergenerationele herverdeling) en herverdeling binnen generaties (intragenerationele herverdeling). *Intragenerationele* herverdeling vloeit voort uit verschillen in persoonlijke karakteristieken die voor de financiering van de pensioenen relevant zijn (zoals levensverwachting). Deze herverdeling wordt gesplitst in overdrachten tussen mannen en vrouwen en in overdrachten tussen deelnemers met een verschillend opleidingsniveau. *Intergenerationele* herverdeling heeft betrekking op een impliciete belasting die wordt geheven op alle toekomstige generaties ter compensatie van te weinig betaalde premies in het verleden. Een kapitaaldekkingsstelsel gebaseerd op de doorsneesystematiek bevat immers een omslagelement: de generaties die relatief oud waren op het moment dat het stelsel is ingevoerd, hebben volledig geprofiteerd van de voor hen gunstige doorsneepremie terwijl zij in hun jonge jaren geen extra afdrachten hebben gedaan. De rekening van deze subsidie is, onbewust, neergelegd bij de toekomstige generaties in de vorm van een hogere pensioenpremie.

Uit dit hoofdstuk komt naar voren dat de Nederlandse aanvullende pensioenen over de hele levensloop bezien significante intragenerationele herverdeling bevatten. Bij een volledige pensioenopbouw is de omvang van de intergenerationele herverdeling echter beperkt. Intragenerationele overdrachten vinden plaats enerzijds van mannen naar vrouwen en anderzijds van laag- naar hoogopgeleiden. Het netto profijt van een laagopgeleide man is negatief en gelijk aan 3,3% van zijn levensinkomen. Een hoogopgeleide man heeft nog steeds een negatief netto profijt maar dat is met 0,9% al duidelijk hoger. Voor vrouwen is het netto profijt positief, variërend van 0,1% voor laagopgeleiden tot 2,4% voor hoogopgeleiden. De belangrijkste oorzaak van deze verschillen is het feit dat vrouwen een gemiddeld hogere levensverwachting hebben dan mannen; hetzelfde geldt voor hoog- ten opzichte van laagopgeleiden.

In **hoofdstuk 3** draaien we ons vizier van herverdeling naar risicodeling. We stellen ons de vraag wat de meerwaarde is van intergenerationele risicodeling in collectieve kapitaalgedekte DB pensioenstelsels. Deze vraag is actueel in het licht van de groeiende populariteit van individuele DC contracten waarin geen (of minder) plaats is voor risicodeling. Op voorhand is geenszins duidelijk dat een overstap op individuele DC regelingen tot meer welvaart leidt. Met behulp van fluctuerende collectieve buffers en pensioenpremies zijn DB contracten immers in staat om overschotten en tekorten over meerdere generaties uit te smeren, ook over toekomstige generaties. Een afzonderlijke generatie draagt zodoende een kleiner deel van het risico wat in principe welvaartsverhogend is. In individuele DC regelingen is deze risicodeling niet mogelijk en draait elke generatie volledig zelf voor de risico's op.

Tegenover dit voordeel van DB contracten staan echter ook nadelen. Ten eerste worden in de meeste contracten premies geheven naar rato van het arbeidsinkomen. Dit betekent dat zodra een generatie door middel van een herstelpremie wordt gevraagd mee te delen in de tekorten van andere generaties, de prikkel om te werken afneemt. Herstelpremies uit hoofde van risicodeling verstoren op deze manier de arbeidsaanbodbeslissing met welvaartsverliezen tot gevolg. Ten tweede worden premiebetalers dikwijls geconfronteerd met gecorreleerde risico's op de arbeidsmarkt. Het is aanneembaar dat een grote klap op de financiële markten, zoals de Grote Recessie in 2008, uiteindelijk ook negatieve effecten heeft op de reële economie en de arbeidsmarkt in bijzonder. Er is inderdaad empirisch bewijs dat op de lange termijn kapitaalmarkten en arbeidsmarkten positief samenhangen. Dit maakt intergenerationele risicodeling via het pensioenfonds minder aantrekkelijk omdat deelnemers dubbel worden getroffen na een negatieve schok, zowel via de kapitaalmarkt (hogere herstelpremies) als via de arbeidsmarkt (lagere lonen).

Om een zuiver beeld te krijgen van de waarde van intergenerationele risicodeling gebruiken we een gestileerd model dat zowel de voordelen ervan (diversificatiewinsten) als de nadelen (arbeidsmarktverstoringen en gecorreleerde risico's) meeneemt. Op basis van dit model kunnen we analytisch bewijzen, voor een specifieke maar veel gehanteerde nutsfunctie (waarmee de preferenties van huishoudens worden beschreven), dat de introductie van een collectief kapitaalgedekt DB stelsel in verwachting voor iedere generatie welvaartsverhogend is, zelfs als rekening wordt gehouden met de rol van arbeidsmarktverstoringen en gecorreleerde risico's op de aandelen- en kapitaalmarkt. Dit resultaat blijft staan als we overgaan op numerieke simulaties met algemenere modelveronderstellingen en alternatieve parameterwaarden.

Uit de analyse komt tevens naar voren dat arbeidsmarktverstoringen en gecorreleerde risico's een significant negatief effect kunnen hebben op de welvaartswinst van intergenerationele risicodeling. Desalniettemin overheersen de positieve diversificatiewinsten. De drijvende kracht achter dit resultaat is het uitgangspunt dat het pensioenfonds het belang van de deelnemer bovenaan stelt: het kiest zijn beleggingsmix zodanig dat het verwachte nut van deelname maximaal is. Met het percentage van de premies belegd in risicovolle vermogenstitels, heeft het pensioenfonds een uniek instrument in handen om zowel de voor- als nadelen van risicodeling te sturen. Als *ultimum remedium* kan het pensioenfonds altijd terugvallen op een risicovrije beleggingsstrategie, bijvoorbeeld wanneer arbeidsmarktverstoringen groot zijn of als aandelen- en loonrisico's vrijwel perfect positief gecorreleerd zijn. Er vindt dan nauwelijks tot geen intergenerationele risicodeling meer plaats waardoor de welvaartswinst van het pensioenfonds uiteindelijk naar nul tendeert.

Hoofdstuk 2 en hoofdstuk 3 zijn primair gericht op de rol van, respectievelijk, herverdeling en risicodeling in collectieve kapitaalgedekte pensioenen. In de laatste twee inhoudelijke hoofdstukken van dit proefschrift slaan we een brug tussen risicodeling (hoofdstuk 4) en herverdeling (hoofdstuk 5) aan de ene kant en de uittreedbeslissing aan de andere kant. In **hoofdstuk 4** gaan we in op de vraag in hoeverre flexibiliteit in de uittreedbeslissing een belangrijk instrument is om onverwachte schokken op te vangen. Deze rol van de uittreedbeslissing als verzekering krijgt steeds meer aandacht in het licht van recente pensioenhervormingen die aansturen op langer doorwerken en meer keuzevrijheid in het tijdstip van pensionering. De meeste pensioenregelingen kennen tegenwoordig een flexibele uittreedleeftijd met in meer of mindere mate actuariële herrekening van de jaarlijkse uitkering bij eerder of later stoppen dan de wettelijke pensioenleeftijd.

In de bestaande economische literatuur wordt inderdaad vaak gesteld dat flexibele uittreding een goede verzekering biedt tegen risico's, zoals arbeidsongeschiktheidsrisico's, het risico op echtscheiding of kapitaalmarktrisico's. Deze rol van flexibele uittreding als 'uitlaatklep' aan het einde van de loopbaan stelt mensen bovendien in staat om gedurende het actieve leven meer beleggingsrisico te nemen. Het idee hierachter is dat een negatieve (*positieve*) vermogensschok de marginale waarde van werken vergroot (*verkleint*) waardoor mensen besluiten langer (*korter*) door te werken. Dit inkomenseffect in de uittreedbeslissing fungeert dus als een financiële buffer en leidt tot een negatieve samenhang tussen arbeids- en kapitaalinkomen. Flexibele uittreding vergroot op deze manier het risicodragend vermogen van individuen die daardoor extra kunnen profiteren van de risicopremie op aandelen.

Het doel van dit hoofdstuk is om deze bestaande inzichten (opnieuw) tegen het licht gehouden, maar dan in een algemenere modelstructuur. Zo onderzoeken we het belang van het type risicofactor waaraan individuen blootgesteld staan: net als in hoofdstuk 3 onderscheiden we zowel financiële schokken als productiviteitsschokken. Ook bezien we in hoeverre bestaande inzichten gevoelig zijn voor het type economie dat als uitgangspunt wordt genomen: we vergelijken een partieel evenwichtsmodel (factorprijzen exogeen) met een algemeen evenwichtsmodel (factorprijzen endogeen). Tot slot onderzoeken we de rol van de nutsfunctie.

Dit hoofdstuk laat zien dat de positieve samenhang tussen flexibele uittreding en risicovol beleggen minder wordt, of zelfs negatief, als niet alleen financiële risico's maar ook productiviteitssrisico's worden meegenomen. Een productiviteitsschok beïnvloedt in principe zowel de beloning van kapitaal als die van arbeid. Als de productiviteit bijvoorbeeld onverwacht daalt, leidt dit niet alleen tot lagere aandelenrendementen maar ook tot lagere lonen. Dat betekent dat de prikkel om langer door te werken kleiner wordt en mensen wellicht eerder met pensioen gaan, waardoor er een positieve samenhang ontstaat tussen arbeids- en kapitaalinkomen. Dit substitutie-effect in de uittreedbeslissing werkt dus precies de andere kant op als het eerder beschreven inkomenseffect: het zorgt ervoor dat flexibele uittreding een minder goede bescherming biedt tegen schokken met als gevolg dat de bereidheid om risicovol te beleggen daalt.

Tevens laten we zien dat de samenhang tussen flexibele uittreding en risicovol beleggen sterk afhangt van de vraag of individuen consumptie van goederen en vrije tijd als substituten, dan wel als complementaire goederen beschouwen. In het eerste geval zal men sneller besluiten de arbeidsmarkt te verlaten als lonen (de prijs van vrije tijd) laag zijn ten gevolge van een negatieve productiviteitsschok. Deze grotere gevoeligheid van de uittreedbeslissing voor veranderingen in de loonvoet vermindert de bescherming die flexibele uittreding biedt tegen onverwachte schokken en verlaagt daardoor de bereidheid risicovol te beleggen. Tot slot laat dit hoofdstuk zien dat algemene evenwichtseffecten de verzekering van flexibele uittreding tegen schokken zowel kunnen vergroten als verkleinen, afhankelijk van de mate van substitutie tussen consumptie en vrije tijd.

In het laatste conceptuele hoofdstuk, **hoofdstuk 5**, keren we terug bij het startpunt, herverdeling. Waar het in hoofdstuk 2 vooral te doen was om de omvang en richting van de herverdeling, gaat het in dit hoofdstuk om de economische gedragseffecten ervan, met name ten aanzien van de uittreedbeslissing. De volgende twee vragen staan centraal: wat voor effect heeft een koppeling van de wettelijke pensioenleeftijd aan de levensverwachting op de herverdeling tussen hoog- en laagopgeleiden? En, hoe pakt het flexibel maken van de ingangsdatum van de pensioenuitkering uit voor de welvaart van de diverse opleidingsgroepen? Beide vragen spelen op dit moment ook een rol in de Nederlandse beleidsdiscussie als het gaat om het hervormen van de AOW. Om die reden heeft dit hoofdstuk, in tegenstelling tot de drie voorgaande hoofdstukken, betrekking op een omslagstelsel.

Dit hoofdstuk maakt op gestileerde wijze inzichtelijk dat herverdelingseffecten niet exogeen zijn maar een samenspel van menselijk gedrag, persoonlijke karakteristieken (zoals levensverwachting) en de vormgeving van het pensioenstelsel. Individuen verschillen in ons model in levensverwachting en opleidingsniveau, waarbij een positief verband tussen beide is verondersteld. Een hoogopgeleide verdient doorgaans meer loon dan een laagopgeleide en zodoende levert het model ook een positieve samenhang op tussen inkomen en levensverwachting. We zien deze relatie in werkelijkheid terug. Het veronderstelde pensioenstelsel is vergelijkbaar met de AOW: premies worden proportioneel geheven over inkomen en iedereen ontvangt dezelfde uitkering die op een vaste leeftijd ingaat en wordt uitgekeerd zolang de persoon leeft. Dit stelsel leidt dus tot herverdeling van hoge naar lage inkomens (inkomensherverdeling) maar ook van kort- naar langlevenden (langlevenherverdeling). Omdat levensverwachting en inkomen positief samenhangen, maakt langlevenherverdeling het pensioenstelsel regressiever.

Hoe pakt een verhoging van de pensioenleeftijd als reactie op de gestegen levensverwachting uit voor de verschillende opleidingsgroepen? Uit de analyse komt op deze vraag geen eenduidig antwoord. De omvang en richting van de herverdelingseffecten zijn sterk afhankelijk van de mate waarin de levensverwachting tussen individuen verschilt. Net als in hoofdstuk 2 wordt herverdeling gemeten in termen van netto profijt, dat wil zeggen, het verschil in contante waarde tussen uitkeringen en premies. Er is empirisch bewijs dat de recente stijging van de levensverwaching (in absolute zin) sterker is geweest voor hoog- dan voor laagopgeleiden. Een uniforme aanpassing van de pensioenleeftijd betekent dan dat de verwachte pensioenduur voor laagopgeleiden minder sterk is toegenomen dan die van hoogopgeleiden: dit vergroot dus de (perverse) langlevenherverdeling. Daar staat tegenover dat het aannemelijk is dat de extra levensduur van hoogopgeleiden deels zal worden benut om langer te blijven werken: dit vergroot de inkomensherverdeling. Indien de heterogeniteit in levensverwachting niet te groot is, blijkt dat het tweede effect domineert en laagopgeleiden er per saldo op vooruit gaan als de pensioenleeftijd uniform wordt verhoogd. Op basis van een indicatieve berekening lijkt dit voor Nederland inderdaad het geval te zijn.

Welke opleidingsgroepen hebben baat bij het flexibiliseren van de ingangsdatum van de pensioenuitkering? Het antwoord op deze vraag hangt af van de wijze waarop de overheid de actuariële omrekening bij eerder of later stoppen dan de wettelijke pensioenleeftijd vormgeeft. Als dat gebeurt op basis van de individuele levensverwachting of eventueel een gemiddelde levensverwachting per opleidingsgroep, dan gaan zonder meer de laagopgeleiden er op vooruit ten koste van de hoogopgeleiden. De toename van de jaarlijkse uitkering bij uitstel van pensioen zal immers voor laagopgeleiden hoger zijn, als compensatie voor het feit dat zij naar verwachting korter leven. Voor overheden met een sterke voorkeur voor inkomensnivellering via de eerste pijler kan individueel-specifieke omrekening dus een effectief instrument zijn om de progressiviteit van het stelsel te vergroten.

In werkelijk is het zeer de vraag of een overheid zoveel informatie heeft dat de omrekenfactor gedifferentieerd kan worden naar groepen laat staan individuen. Het is in dit opzicht veelzeggend dat in de praktijk actuariële omrekening uitsluitend uniform gebeurt. Verrassend genoeg laat dit hoofdstuk zien dat uniforme omrekening per saldo voor *alle* opleidingsgroepen een voordeel kan opleveren. Uniforme omrekening impliceert dat voor hoogopgeleiden de beloning van langer doorwerken, in de vorm van een ophoging van hun jaarlijkse pensioenuitkering, te hoog is. Hoogopgeleiden gaan er dus sowieso in welvaart op vooruit: zij worden geconfronteerd met een lagere impliciete belasting wat hen stimuleert langer door te werken. Voor laagopgeleiden geldt juist dat uniforme omrekening de impliciete belasting vergroot, hetgeen een negatief effect heeft op hun welvaart. Daar staat tegenover dat het later uittreden van de hoogopgeleiden meer belastingsinkomsten oplevert en dus hogere pensioenuitkeringen. Dit verhoogt de welvaart van de laagopgeleiden.

Wij laten zien dat, indien de belastingvoet hoog genoeg is, het tweede effect domineert zodat ook de welvaart van de laagopgeleiden stijgt.

Tot slot laten we in dit hoofdstuk zien dat een uniforme flexibilisering van de pensioenleeftijd nog verder verbeterd kan worden door langer doorwerken extra te stimuleren (of eerder stoppen juist extra te ontmoedigen). Extra stimuleren betekent in dit verband méér dan actuariële neutraliteit zou voorschrijven. Een subsidie op langer doorwerken compenseert de bestaande verstoring van de arbeidsaanbodbeslissing als gevolg van proportionele belastingheffing. Voor alle opleidingsgroepen daalt dan de impliciete belasting op doorwerken waardoor de welvaart nog verder toeneemt.

Wat zijn de belangrijkste **beleidsimplicaties** van de bevindingen uit dit proefschrift? Op dit punt is enige voorzichtigheid geboden. De resultaten zijn verkregen op basis van veelal gestileerde modellen die zich moeilijk één-op-één laten vertalen naar de praktijk. Het is echter wel mogelijk enkele denkrichtingen aan te reiken die voor beleidsbepalers relevant zijn en tevens een aansporing zijn voor nader onderzoek.

Uit dit proefschrift blijkt dat het gebruik van de doorsneesystematiek in het Nederlandse stelsel van aanvullende pensioenen met significante intragenerationele overdrachten gepaard gaat (hoofdstuk 2). Vanuit juridisch oogpunt is de doorsneesystematiek altijd gezien als een noodzakelijke vorm van solidariteit om het gebrek aan concurrentie ten gevolge van de verplichte deelname te rechtvaardigen. Vanuit sociaal oogpunt kunnen echter vraagtekens gezet worden bij de wenselijkheid van deze solidariteit omdat deze grotendeels resulteert in overdrachten van groepen met een hoog sterfterisico (veelal laagopgeleiden) naar groepen met een laag sterfterisico (veelal hoogopgeleiden).

Er zijn verschillende manieren denkbaar waarop deze perverse herverdeling verminderd kan worden: pensioenfondsen zouden meer kunnen differentiëren naar homogene groepen in het verevenen van het langlevenrisico. Dit stelt echter hoge eisen aan de informatie die pensioenfondsen ter beschikking hebben. Minder gecompliceerde alternatieven zijn het invoeren van progressieve premies of het aftoppen van de pensioenopbouw boven een bepaalde inkomensgrens. Door gebruik te maken van de positieve samenhang tussen levensverwachting en inkomen zou een fonds langs deze lijnen de subsidie aan langlevenden kunnen beperken. Hoe dan ook kunnen onze bevindingen voor beleidsmakers een aansporing zijn om beleid te ontwikkelen gericht op het verkleinen van sociaaleconomische ongelijkheden in sterftekansen.

Naast herverdeling, biedt dit proefschrift ook ten aanzien van het andere wezenskenmerk van collectieve pensioenen, intergenerationele risicodeling, beleidsrelevante inzichten (hoofdstuk 3). In veel landen zie je een beweging van collectieve DB contracten naar meer individuele DC contracten. In het licht van de vergrijzing is dit een logische ontwikkeling aangezien de risico's van een steeds groter wordende groep ouderen moet worden beschermd door een steeds kleiner wordende groep jongeren. De vraag is wel of een volledig individueel DC contract zonder intergenerationele risicodeling dan het eindstation moet zijn. Ons onderzoek draagt bij aan het inzicht dat het behoud van enige vorm van intergenerationele risicodeling welvaartsverhogend kan zijn, zelfs als rekening wordt gehouden met de negatieve effecten ervan in termen van arbeidsmarktverstoringen en gecorreleerde risico's op aandelen- en kapitaalmarkt.

De ontwikkeling naar meer op individuele leest geschoeide contracten zal beleidsbepalers ook voor nieuwe uitdagingen stellen. Dit proefschrift wijst op het toenemende belang dat het beleggingsbeleid van pensioeninstellingen (zoals pensioenfondsen en verzekeraars) in overeenstemming is met de voorkeuren van de deelnemers, zeker in het licht van een ontwikkeling van min of meer vaste naar flexibele uittreedleeftijden (hoofdstuk 4). Het idee dat deelnemers er belang bij hebben om pensioengelden risicovoller te beleggen wanneer zij nog de mogelijkheid hebben om met hun uittreedbeslissing flexibel in te spelen op onverwachte gebeurtenissen, gaat niet altijd op. Als de kans reëel is dat deelnemers in de toekomst geconfronteerd worden met gecorreleerde risico's op de aandelen- en kapitaalmarkt, zoals tijdens de Grote Recessie, of wanneer hun uittreedbeslissing zeer gevoelig is voor loonschokken, dan zullen zij eerder kiezen voor een defensievere beleggingsstrategie.

Tot slot onderschrijft ons onderzoek het belang van het flexibiliseren van de pensioenleeftijd in de eerste pensioenpijler (hoofdstuk 5). In potentie kan de overheid daarmee twee vliegen in één klap slaan: de maatregel kan het verstorende effect van belastingen verkleinen waardoor middelen worden vrijgespeeld om bijvoorbeeld de kosten van de vergrijzing op te vangen. Daarnaast vergroot het de herverdeling van hoge naar lage inkomens waardoor de inkomensongelijkheid afneemt. In de meeste landen zijn pensioencontracten met een flexibele opname nog steeds niet volledig actuarieel neutraal in de zin dat ze een prikkel bevatten om eerder te stoppen. Onze analyse zou eerder pleiten voor verdergaande actuariële neutraliteit van deze contracten, eventueel zelfs met een subsidie op langer doorwerken.