TI 2013-091/VI
Tinbergen Institute Discussion Paper

# Flexible Pension Take-up in Social Security

*Yvonne Adema*[1,2]
*Jan Bonenkamp*[3,2]
*Lex Meijdam*[4,2]

[1] *Erasmus School of Economics, Erasmus University Rotterdam, Tinbergen Institute;*
[2] *Netspar;*
[3] *CPB Netherlands Bureau for Economic Policy Analysis, The Hague, The Netherlands;*
[4] *CentER, Tilburg University, The Netherlands.*

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031



Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579

# Flexible pension take-up in social security*

Yvonne Adema[†] Jan Bonenkamp[‡] Lex Meijdam[§]

This version: July 3, 2013

**Abstract**

This paper studies the redistribution and welfare effects of increasing the flexibility of individual pension take-up. We use an overlapping-generations model with Beveridgean pay-as-you-go pensions, where individuals differ in ability and life span. We find that introducing flexible pension take-up can induce a Pareto improvement when the initial pension scheme contains within-cohort redistribution and induces early retirement. Such a Pareto-improving reform entails the application of uniform actuarial adjustment of pension entitlements based on average life expectancy. Introducing actuarial non-neutrality that stimulates later retirement further improves such a flexibility reform.

Key words: redistribution, retirement, flexible pensions

JEL codes: H55, H23, J26

*We are grateful to Casper van Ewijk, Kees Goudswaard, Ben Heijdra and Ed Westerhout for their comments on earlier versions of this paper. We also wish to thank the participants of the pension workshop in Wuerzburg (June 5, 2012), the participants of the Netspar International Pension Workshop in Paris (June 7, 2012) and the participants of the Annual Congress of the International Institute of Public Finance in Dresden (August 16, 2012) for helpful comments and suggestions.

[†]Department of Economics, Erasmus University Rotterdam, Tinbergen Institute and Netspar, Email: adema@ese.eur.nl

[‡]CPB Netherlands Bureau for Economic Policy Analysis and Netspar, Email: j.p.m.bonenkamp@cpb.nl

[§]Department of Economics and CentER, Tilburg University, and Netspar, Email: a.c.meijdam@uvt.nl

# 1 Introduction

Since the 1970s, the effective retirement age has declined in almost all Western countries while at the same time life expectancy has increased substantially. These developments led to an increase of the average retirement period relative to the working period thereby eroding the fiscal sustainability of pension schemes. To reverse this trend, in recent years more attention has been given to pension reforms that improve labour supply incentives and encourage people to work longer. Countries like the UK and Australia, for example, introduced a flexible retirement age and increased the reward to continue working. The advantage of this type of reforms is that it not only reduces the labour market distortions caused by incentives to retire early but can also increase the sustainability of pension systems.

A potential disadvantage is, however, that these flexibility reforms are typically implemented in a *uniform* way, i.e., applied to all participants in the same way, while individuals have heterogenous characteristics (e.g., in terms of life expectancy or income level). Uniformly implemented reforms therefore probably have different welfare effects at the individual level and may affect certain types of individuals negatively.[1] 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 and Bonenkamp, 2009). For example, unfunded pension schemes, especially those of the Beveridgean type, often contain redistribution from high to low incomes. Apart from this, these pension schemes typically also redistribute from short-lived to long-lived agents because they are based on collective annuities which do not depend on individual life expectancy. This makes collective annuities subject 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, Meara et al., 2008). Pension reforms that introduce more flexibility in pension take-up will affect these redistribution effects. It is therefore important to take into account the redistribution in existing pension schemes and the fact that individuals are heterogenous when analyzing the welfare effects of pension flexibility reforms.

This paper explores the redistribution and welfare effects of the introduction of a flexible starting date for pension benefits in the context of an unfunded pension scheme with an explicit redistribution motive. That means, we consider a change from a payout scheme in which benefits start at the fixed *statutory* retirement age to a scheme where benefits start at the flexible *effective* retirement age. This flexible pension take-up is combined with actuarial adjustments of pension benefits for early or late retirement. To analyse the economic implications of this reform, we use

[1] In the Netherlands for example there was a lot of discussion whether a reform aimed at increasing the retirement age would not hurt the low-skilled too much as these people typically start working earlier and have a shorter life span and therefore prefer to retire early.

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 pay-as-you-go (PAYG) social security system is of the Beveridgean type and is characterized by life-time annuities and proportional contributions. 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. Note that, in contrast to the former, the latter type of redistribution is regressive due to the positive link between productivity and longevity. The fact that individuals are heterogenous implies that introducing pension flexibility will affect individuals differently.

Implementing pension contracts with a variable starting date for benefits, as analysed in this paper, is important for various reasons. It helps 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 but who are prevented to do that because of liquidity or borrowing constraints. Flexible pensions can also function as a hedge against all types of risks, like disability risks (Diamond and Mirrlees, 1978) or financial risks (Pestieau and Possen, 2010). This paper adds some other arguments. We will illustrate that flexible pensions can stimulate people to postpone retirement voluntary. In that case flexible pension take-up may help 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 different agents as a way to get rid of the unintended redistribution from short-lived to long-lived agents.

The main results are as follows. First, introducing a flexible pension take-up cannot be Pareto improving if the government conditions the adjustment factor of benefits on individual life expectancies. Individual actuarial adjustment eliminates the unintended redistribution from short-lived to long-lived agents. The low-skilled therefore benefit from this reform at the expense of the high-skilled. Second, introducing a flexible pension take-up can be Pareto improving if the 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. 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 contribution rate is sufficiently high, the low-skilled also gain because they receive higher pensions, enabled by the additional tax payments of the high-skilled. Third, combining uniform adjustment with actuarial non-neutrality to induce people to postpone retirement can further improve the reform, i.e., a Pareto improvement can be achieved at a lower contribution rate or for a given contribution rate the welfare effects are more positive for all individuals.

This paper is related to studies that analyse the interaction between pension schemes and retirement decisions (see e.g., Hougaard Jensen et al., 2003) 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; Hachon, 2008 and Cremer et al., 2010). In addition, our paper is also related to Fisher and Keuschnigg (2010) and Jaag et al. (2010) who investigate the labour market impact of pension reforms towards more actuarial neutrality. Most of these aforementioned studies focus on pension reforms that strengthen the link between contributions and benefits. Our study, in contrast, deals with the implementation of a flexible pension take-up.

This paper is most closely related to Cremer and Pestieau (2003). 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 income to people with low income. To obtain this outcome, both studies need that the benefit rule of the social security scheme redistributes within generations and that there is an initial retirement distortion (i.e., early retirement), the removal of which brings additional resources. However, the studies differ in the reforms they focus on. Cremer and Pestieau (2003) 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 implemented reform, the introduction of a flexible pension take-up. In our 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 paper is organized as follows. In Section 2 we introduce the benchmark model. This model contains a PAYG social security scheme with inflexible pension take-up and life-time annuities. Section 3 analyses the redistribution and welfare effects of reforms aimed at increasing the flexibility of individual pension take-up. In Section 4 we elaborate on these flexibility reforms by introducing non-neutral actuarial adjustment of benefits. Section 5 concludes the paper.

## 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 of life. 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 their working life after the statutory retirement age or to advance retirement and stop working before the statutory retirement age. So the statutory retirement age is related to the date agents receive their pension benefit, which is not necessarily equal to their effective retirement date.

### 2.1 Preferences

Preferences over first-period and second-period consumption are represented by the following utility function:

$$U(c, x) = u(c) + \pi u(x) \tag{1}$$

with $u' > 0$ and $u'' < 0$; $c$ is first-period consumption, $x$ is second-period consumption and $\pi \leq 1$ is the length of the second period. To keep the analysis as simple as possible we assume that the interest rate and the discount rate are zero.[2] Second-period consumption is defined net of the disutility of labour:

$$x = \frac{d}{\pi} - \frac{\gamma}{2}\left(\frac{z}{\pi}\right)^2 \tag{2}$$

where $d$ is total consumption of goods when old yielding a consumption stream of $d/\pi$, $z$ denotes the working period and $\gamma$ is the preference parameter for leisure. Following Casamatta et al. (2005) and Cremer and Pestieau (2003), we assume a quadratic specification for the disutility of work. This specification makes the problem more tractable, but comes with the cost that there are no income effects in labour supply. Income effects in the retirement decision are found to be small compared to substitution effects, however, 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., $z/\pi$). This implies that for a given retirement age an agent with a short life span experiences a higher disutility of work than an agent with a long life span because this short-lived agent works a relatively larger share of his remaining life time.

### 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

[2] 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. Relaxing these assumptions would not change our main results, however.

devoted to working as a high-skilled worker.

The individual-specific parameter $a$ reflects the ability of individuals to acquire high working skills. The higher is $a$, the more able is the individual, and the less time a worker needs to become high-skilled for acquiring a work skill. The parameter $a$ ranges between 0 and 1 and its cumulative distribution function is denoted by $G(\cdot)$, i.e., $G(a)$ 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 total number of individuals born in each period to be one, i.e., $G(1) = 1$.

A high-skilled worker provides an effective labour supply of one unit per unit of working time, while a low-skilled worker provides only $q < 1$ units of effective labour for each unit of working time. This difference in effective labour supply also applies to the second period of life. 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 by $W_y(a)$, is given by:

$$W_y(a) \equiv \begin{cases} qw & \text{for} \quad a \leq a^* \\ aw & \text{for} \quad a \geq a^* \end{cases} \tag{3}$$

where $a^*$ is the cut-off ability level to become high-skilled. It is assumed that $a^*$ is exogenous.[3] For the second period of life the maximum labour income, $W_o(a)$, equals:

$$W_o(a) \equiv \begin{cases} qw & \text{for} \quad a \leq a^* \\ w & \text{for} \quad a \geq a^* \end{cases} \tag{4}$$

### 2.3 Individual life span

Each individual lives completely the first period of life (with a length normalized to unity) but only a fraction $\pi(a) \leq 1$ of the second period. We assume 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 in line with the empirical evidence that income positively co-moves with life expectancy.[4]

Whenever necessary to parameterize the function $\pi(a)$, we will use the following specification:

$$\pi(a) = \bar{\pi}\left[1 + \lambda(a - \bar{a})\right], \quad \lambda \geq 0 \tag{5}$$

where $\bar{a} \equiv \int_0^1 a \, \mathrm{d}G$ denotes the average ability level. This simple function has the

[3] In Appendix C we work out the model with endogenous schooling like in Razin and Sadka (1999). As shown, endogenizing the skill level does not change the main results derived in the body of this paper.

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

following appealing 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, $\text{Cov}(\pi, a) = \lambda \text{Var}(a) \geq 0$. Third, consistent with empirical findings (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 gets larger if average life expectancy increases, i.e., $\pi(a = 1) - \pi(a = 0) = \lambda\bar{\pi}$; life expectancy of more able individuals increases more when the average life span rises.

### 2.4 Consumption and retirement

An individual faces the following intertemporal budget constraint:

$$c + d = (1 - \tau)W_y + (1 - \tau)zW_o + P \tag{6}$$

where $\tau$ is the social security contribution (tax) rate and $P$ denotes total pension entitlements received during old age.[5]

Maximizing life-time utility (1) over $c$, $d$ and $z$, subject to the life-time budget constraint (6) yields the following first-order conditions:

$$u'(c) = u'(x) \tag{7}$$

$$(1 - \tau)W_o = \frac{\gamma z}{\pi} \tag{8}$$

Expression (7) is the standard consumption Euler equation. Equation (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 $c$, $x$ and $z$ for the benchmark model:

$$c = \frac{1}{1 + \pi}\left[(1 - \tau)W_y + \frac{(1 - \tau)^2 W_o^2 \pi}{2\gamma} + P\right] \tag{9}$$

$$x = c \tag{10}$$

$$z = \frac{(1 - \tau)W_o \pi}{\gamma} \tag{11}$$

where $P$ denotes total pension entitlements in the benchmark model. Note that the social security tax distorts the retirement decision: the larger the contribution rate $\tau$, the earlier agents leave the labour market, i.e., the lower $z$, because it reduces the net wage (and thus the price of leisure). Notice further that our disutility specification ensures that the retirement period is proportional to longevity, i.e., $\pi - z = [1 - (1 -$

[5] It is assumed that individual abilities and life spans are not publicly observable and therefore non-uniform lump-sum transfers are not available.

$\tau)W_o/\gamma]\pi$. Hence, a longer life span is split between later retirement and a longer retirement period. Low-skilled workers retire earlier than high-skilled workers for two reasons. First, since it is assumed that $q < 1$, 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).

### 2.5 Social security

The PAYG social security scheme is of the Beveridgean type. In the benchmark model, agents receive a flat pension benefit $b$ per retirement period which starts at the statutory retirement age $h$ and lasts until the end of the individual old-age period $\pi$. Total pension entitlements $P$ are then:[6]

$$P = (\pi - h)b \tag{12}$$

The fact that the pension benefit is flat, but social security contributions $\tau$ are proportional to the wage rate implies that the pension scheme redistributes income from high-income to low-income individuals. The pension scheme also redistributes from short-lived to long-lived individuals, however, as individuals receive the flat pension benefit until their death. The positive link between ability, wages and life expectancy in our model then implies that there is also some redistribution from low incomes to high incomes, as the latter group typically has a longer life span.

A feasible social security pension scheme must satisfy the following resource constraint:[7]

$$\int_0^1 P\,\mathrm{d}G = \tau q w \int_0^{a^*} (1 + z_L)\,\mathrm{d}G + \tau w \int_{a^*}^1 (a + z_H)\,\mathrm{d}G \tag{13}$$

Using equations (5) and (12), we can rewrite this equation as:

$$b(\bar{\pi} - h) = \tau q w \int_0^{a^*} (1 + z_L)\,\mathrm{d}G + \tau w \int_{a^*}^1 (a + z_H)\,\mathrm{d}G \tag{14}$$

This condition states that the total amount of pension benefits paid out (left-hand side) equals the total amount of tax contributions received (right-hand side). 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.

As a measure for redistribution, we calculate the net benefit of participating in the pension scheme. The net benefit is the difference between the total pension benefits

[6]We impose that $\pi - h > 0$ for any $a$-individual. In other words, nobody passes away before the statutory retirement age.

[7]Throughout this paper subscript 'L' refers to low-skilled workers and subscript 'H' refers to high-skilled workers.

received and tax contributions paid:

$$NB \equiv (\pi - h)b - \tau(W_y + zW_o) \tag{15}$$

An agent is a net beneficiary if total pension benefits received exceeds contributions paid (i.e., $NB > 0$). Otherwise, the agent is a net contributor (i.e., $NB < 0$). *A priori* it is not immediately clear whether the low-skilled agents are the net beneficiaries of this Beveridgean pension system. On the one hand, low-skilled agents benefit from this pension scheme as they have a lower wage rate and generally retire earlier than high-skilled agents. On the other hand, low-skilled agents also die earlier than high-skilled agents which implies that low-skilled agents are negatively affected by the pension scheme.

Using the definition of net benefits, equation (15), the budget constraint of the pension scheme implies:

$$\int_0^{a^*} NB_L \, \mathrm{d}G + \int_{a^*}^1 NB_H \, \mathrm{d}G = 0 \tag{16}$$

The net benefits of all (young) individuals is equal to zero, reflecting the zero-sum game nature of the pension scheme.[8]

## 3 Pension flexibility reforms

As explained in the Introduction, in recent years many countries have 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 pension 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 short-lived to long-lived 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. In this section we 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, and when this person retires earlier, he receives a decrement. The imposed coincidence of pension take-up and retirement is a realistic assumption because in practice flexible pension schemes often

[8]With a positive interest rate the sum of net benefits would be negative as in that case all future generations have to pay for the windfall gain given to the old generation at the time the pension scheme was introduced.

contain legal restrictions to continue work after a person has opted for benefits.[9] We will first discuss the actuarial adjustment of benefits in general. The specific cases of individual actuarial adjustment and uniform actuarial adjustment of benefits will be discussed in Subsections 3.1 and 3.2, respectively.

**Actuarial adjustment of benefits**

Suppose the government pays benefits $p$ to an individual over his whole effective retirement period. Total pension entitlements are then equal to $P = (\pi - z)p$. Pension earnings per retirement period $p$ are given by:

$$p = m(z, \hat{\pi})b \tag{17}$$

where $b$ is the reference flat pension benefit independent of contributions and labour history. The factor $m(\cdot)$ 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 and is given by:

$$m(z, \hat{\pi}) = \frac{\bar{\pi} - h}{\hat{\pi} - z} \tag{18}$$

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* retirement period and the *individual* retirement period measured by the reference life-span parameter $\hat{\pi}$ which will be specified below. At the individual level, actuarial non-neutrality arises when $\hat{\pi}$ differs from $\pi$. The function $m(\cdot)$ is an increasing function in the individual retirement decision $z$; when an agent decides to continue to work after the statutory retirement age the pension benefit in the remaining retirement periods will be adjusted upwards.

We consider two scenarios for the life span to be used in the adjustment factor which differ with respect to the information set available to the government. In the first scenario, the government can observe individual life expectancies and uses adjustment factors based on individual life spans ($\hat{\pi} = \pi$). The government can then get rid of the adverse redistribution from short- to long-lived individuals. The implication of this is, however, that the high-skilled will be harmed by this reform while the low-skilled gain, and a Pareto improvement is not possible. In reality the government cannot observe individual life expectancies. We therefore assume in Section 3.2 that the government applies a uniform actuarial adjustment factor, based on the average life span of the population ($\hat{\pi} = \bar{\pi}$). This uniform actuarial adjustment introduces selection

[9] In countries like Portugal, Spain and France the coincidence of pension take-up and retirement is regulated by law. In the Dutch flexible second-pillar schemes the access to pension benefits is also conditional on dismissal.

effects in the retirement decision, long-lived agents have an incentive to postpone retirement, while short-lived agents have an incentive to advance retirement. We show that in this reform scenario a Pareto improvement is possible.

## 3.1 Individual actuarial adjustment of benefits

To set the scene, we take a rather extreme position in this section and assume that the government can observe individual life spans and uses this information to assess 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:

$$m = \frac{\bar{\pi} - h}{\pi - z} \tag{19}$$

$$P = (\bar{\pi} - h)b \tag{20}$$

Note from equation (19) that $m = 1$ for an agent with an average ability level ($a = \bar{a}$) who retires at the statutory retirement age $h$. For this so-called *average* individual the pension benefit per retirement period is equal to the reference benefit, i.e., $p = b$. In case this person retires later than the statutory retirement age, then $m > 1$, implying that the per-period benefit is adjusted upwards, i.e., $p > b$. On the other hand, when the person retires earlier than the statutory retirement age, we have $m < 1$ and $p < b$.

The retirement decision is actuarially neutral because the effective retirement age has no effect on the total pension entitlements $P$, i.e., $\partial P / \partial z = 0$. 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 life-time pension benefits.

### Consumption and welfare effects

The retirement decisions are the same as in the benchmark social security model ($z_{ben} = z_{ind}$).[10] The aggregate budget constraint of the pension contract also does not change, implying that the pension benefit per retirement period stays the same as

[10] In the rest of this paper, subscript 'ben' refers to the *benchmark* model and subscript 'ind' to the flexible model based on *individual* actuarial adjustment of benefits. We only use subscripts if it is strictly necessary, i.e., in equations in which we compare one of the flexibility reforms with the benchmark case.

well ($b_{ben} = b_{ind}$). Only consumption changes:

$$c_{ind} = c_{ben} + \frac{(\bar{\pi} - \pi)b}{1 + \pi} \quad (21)$$

$$x_{ind} = x_{ben} + \frac{(\bar{\pi} - \pi)b}{1 + \pi} \quad (22)$$

From these equations we can immediately infer the following result:

**Proposition 1.** *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.*

The intuition for this result is that 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 the agents with long life spans (i.e., the high-skilled)[11,12].

## 3.2 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 *heterogenous*.

### Actuarial adjustment factor

With uniform adjustment, the reference life-span index is the same for each agent, $\hat{\pi} = \bar{\pi}$, so the adjustment factor and pension entitlements are:

$$m = \frac{\bar{\pi} - h}{\bar{\pi} - z} \quad (23)$$

$$P = \frac{(\pi - z)(\bar{\pi} - h)b}{\bar{\pi} - z} \quad (24)$$

[11] In theory, applying individual-specific conversion factors could result in a Pareto improvement if we would not restrict benefit adjustments to be actuarially neutral at the individual level.

[12] The government could also decide to use the skill level of individuals to determine the actuarial adjustment factor. The advantage of this approach is that skill levels are observable for the government and are in general correlated with individual life expectancies (see e.g., van Kippersluis et al., 2011). Appendix B derives this case formally and shows that skill-dependent actuarial adjustment of benefits cannot result in a Pareto improvement, just like the case where the actuarial adjustment of benefits is based on individual life expectancies.

The actuarial adjustment factor $m$ equals one for each individual who retires at the statutory retirement age, i.e, if $z = h$, so 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 (24) 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 (24):

$$\Psi(z) \equiv \frac{\partial P(z)}{\partial z} = \frac{(\pi - \bar{\pi})p}{\bar{\pi} - z} \tag{25}$$

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 life-time 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 (6), but now $P$ is defined as in equation (24). Only the first-order condition regarding retirement changes:

$$(1 - \tau)W_o + \Psi(z) = \frac{\gamma z}{\pi} \tag{26}$$

with $\Psi(z)$ given by equation (25). Consumption and retirement are then equal to:[13]

$$c_{uni} = c_{ben} + \frac{1}{1+\pi}\left[P_{uni} - P_{ben} - \frac{[\Psi(z_{uni})]^2\,\pi}{2\gamma}\right] \tag{27}$$

$$z_{uni} = z_{ben} + \frac{\Psi(z_{uni})\pi}{\gamma} \tag{28}$$

Equation (28) 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 $\Psi$, imply an additional distortion in the retirement decision. This redistribution distortion can either stimulate retirement or depress retirement, depending on the individual life span $\pi$. For individuals with

[13] Subscript 'uni' refers to *uniform* actuarial adjustment of benefits.

below-average life spans ($\pi < \bar{\pi}$), $\Psi < 0$, which implies that these people move up 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 case. The last term in equation (27) 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 (12) and (24) that total pension benefits are generally not the same in the benchmark scheme and in the flexibility reform with uniform adjustment.[14]

**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 that under certain conditions this reform can lead to a Pareto improvement.

Suppose that the reform takes place unexpectedly. First we will analyse how this reform affects utility of the current old generation. In the benchmark second-period consumption is equal to:

$$\pi x_{ben} = s_{ben} + \frac{(1-\tau)^2 W_o^2 \pi}{2\gamma} + P_{ben} \tag{29}$$

where savings are 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 (26). Using this condition, old-age consumption after the reform is:

$$\pi x_{uni} = s_{ben} + \frac{(1-\tau)^2 W_o^2 \pi}{2\gamma} + P_{uni} - \frac{[\Psi(z_{uni})]^2 \pi}{2\gamma} \tag{30}$$

The old generation is not worse off after the reform when $u(x_{uni}) - u(x_{ben}) \geq 0$, implying:

$$\pi x_{uni} - \pi x_{ben} \geq 0 \Rightarrow P_{uni} - P_{ben} - \frac{[\Psi(z_{uni})]^2 \pi}{2\gamma} \geq 0 \tag{31}$$

The current young generation and future generations are better off if $U(c_{uni}, x_{uni}) \geq U(c_{ben}, x_{ben})$ for each ability level, which implies, using equation (7), $c_{uni} \geq c_{ben}$. From

[14]This difference is not only due to the direct effect of a different adjustment factor, but also 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 flat reference pension benefit $b$.

equation (27) 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 (31) 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 set equal to the retirement age of the individual with the average ability level, i.e., $h = z(\bar{a})$.*

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, we can then derive the following result:

**Proposition 2.** *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 $\tau^*$ equal to:*

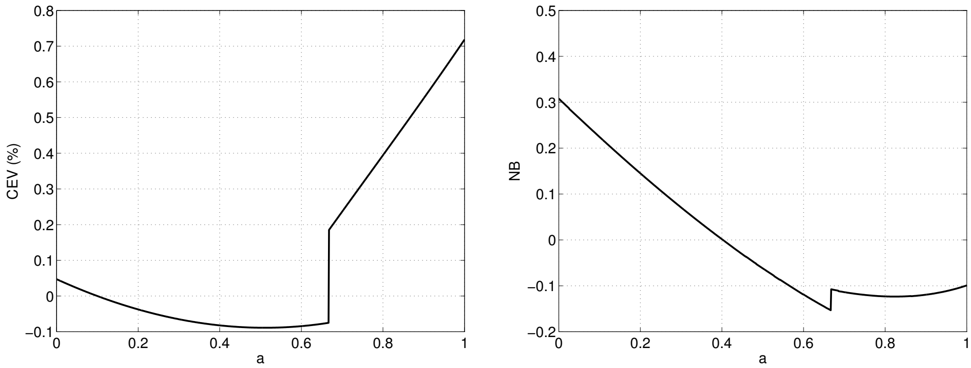
$$\tau^* = \frac{(\gamma - qw)\sqrt{\gamma - w} - (\gamma - w)\sqrt{\gamma - qw}}{w\sqrt{\gamma - qw} - qw\sqrt{\gamma - w}} \tag{32}$$

*Proof.* See Appendix A.1. □

The intuition for this result is as follows. High-skilled workers certainly gain from this reform because the adjustment factor is too high for them 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 as their actuarial adjustment factor is too low. 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.

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 the increasing fiscal burden of ageing. Moreover, if the reform is conducted properly, it will also foster redistribution from the rich to the poor. Similar to Cremer and Pestieau (2003), this 'double

Figure 1: Uniform adjustment: welfare and redistribution




(a) Welfare (b) Net benefit

*Notes*: Net benefit is expressed in absolute difference from the benchmark and certainty-equivalent variation is expressed in percent benchmark consumption.

dividend' hinges on two conditions. First, the retirement decision in the benchmark model 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.[15]

In Figure 1 we show a numerical illustration of the welfare (left graph) and redistribution effects (right graph) of a switch to a flexible scheme based on uniform actuarial adjustment. The underlying parameterization is as follows. The tax rate $\tau$ is 0.3, $w = 1$ and $\gamma = 2$. We further assume $h = 1/6$ and $\bar{\pi} = 0.7$, which implies an official retirement age of 65 and an average life span of 81 years.[16] The heterogeneity parameter $\lambda$ 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 estimates, this gives $\lambda = 1/6$. We interpret the high skill level as the highest attainable education levels in the Netherlands (i.e., higher vocational training and university) and the low skill level as the collective term of all remaining education levels. 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 fourty percent less than high-skilled agents ($q = 0.6$). Finally, we assume that ability $a$ follows a uniform distribution, i.e.,

[15] Cremer and Pestieau (2003) obtain this efficiency gain by age-dependent taxation, i.e., by giving the young a higher tax rate than the old. However, the advantage of actuarial adjustment is that it is 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).

[16] We assume that life time consists of 30 years of 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. The official retirement age is therefore $60 + 30h$ and the average life span is $60 + 30\bar{\pi}$.

$G(a) = a$, and that the utility function is logarithmic, i.e., $u(\cdot) = \ln(\cdot)$.

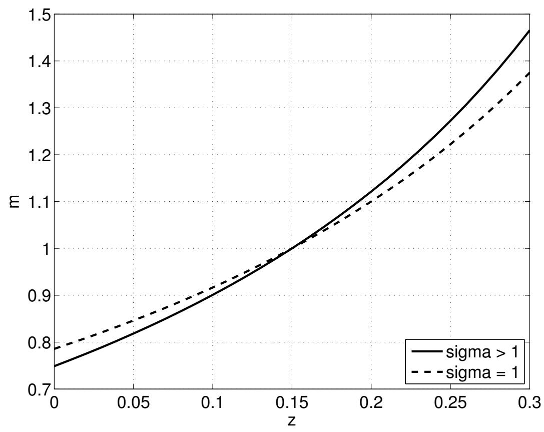
Figure 1(a) shows that the welfare effects of introducing flexible retirement with uniform adjustment are positive for all high-ability agents. These agents benefit from a lower implicit tax on continued activity due to the attractive actuarial adjustment factor and therefore choose to work longer. With these parameter settings, however, the 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 experience an increase in the net benefit from the scheme (see Figure 1(b)). To achieve a Pareto improvement, the contribution rate needs to be at least 40 percent, that is, $\tau^* = 0.4$.

There are good reasons to argue that in practice the tax critical rate is lower than presumed in our analysis. First, income redistribution from rich to poor runs through more channels than the pension scheme, like the tax system or public health care. Hence, when high-skilled agents are stimulated to work longer with a flexible pension take-up, the low-skilled may also be compensated through these other types of redistribution. Second, in reality the contribution tax is added to other sources of distortionary taxation. As the deadweight loss is roughly quadratic in the total tax rate, the marginal welfare improvement of introducing flexible retirement (and lowering implicit taxation) might be larger than our analysis suggests. In the next section, we show that a reduction of the tax critical rate can also be obtained by reformulating the pension reform to some extent, i.e., by setting the reward rate of retirement postponement above the actuarially-neutral level.

# 4 Introducing actuarial non-neutrality

In recent years, an increasing number of countries introduced penalties and rewards for earlier and later retirement. To stimulate work continuation, the penalty rate is typically not as high as the reward rate, i.e., the adjustment is asymmetric. In the US, for example, for each year of retirement before the statutory retirement age, the annual benefit is reduced by 6.75%. The actuarial increment for those retiring after the statutory retirement age amounts to 8%. In Japan, the difference is even larger, there the penalty rate of early retirement is 6% per year while the reward rate of later retirement is 8.4% (OECD, 2011). In this final section we therefore consider a pension flexibility reform where pension benefits are adjusted in an actuarially non-neutral way to induce people to postpone retirement. We show that under such a reform a Pareto improvement can be achieved at a lower contribution rate or that for a given contribution rate it leads to more positive welfare effects for all individuals.

Figure 2: Actuarial adjustment factor




### 4.1 Actuarial adjustment factor

To make our point as clear as possible, we abstract from life-span heterogeneity in the analytical analysis. Hence, each agent, irrespective of his ability level, lives a fraction $\pi \leq 1$ of the second period. In the simulation graphs, however, we have heterogenous life spans. The actuarial adjustment factor is specified as follows:

$$m(z, \pi) = \left(\frac{\pi - h}{\pi - z}\right)^{\sigma}, \quad \sigma > 1 \tag{33}$$

where the parameter $\sigma$ governs the degree of actuarial non-neutrality of the adjustment factor, this is also shown in Figure 2. In case $\sigma = 1$, the adjustment is completely actuarially neutral with respect to the retirement decision (see Section 3.1). 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 (33) rewards delaying retirement and discourages early retirement as long as $\sigma > 1$.

Given equation (33), the pension entitlements $P$ are equal to:

$$P = (\pi - h)^{\sigma}(\pi - z)^{1-\sigma} b \tag{34}$$

Taking the derivative of $P$ with respect to $z$ gives:

$$\Phi(z) \equiv \frac{\partial P(z)}{\partial z} = (\sigma - 1)p \tag{35}$$

Hence, if $\sigma > 1$ then $\Phi > 0$, i.e., introducing actuarial non-neutrality gives all agents an incentive to continue working as this will increase pension entitlements.

### 4.2 Consumption and retirement

The consumption decision and retirement decision are equal to:[17]

$$c_{nan} = c_{ben} + \frac{1}{1+\pi}\left[P_{nan} - P_{ben} - \frac{\left[\Psi(z_{nan})\right]^2 \pi}{2\gamma}\right] \tag{36}$$

$$z_{nan} = z_{ben} + \frac{\Psi(z_{nan})\pi}{\gamma} \tag{37}$$

where $P$ and $\Psi$ are defined by equations (34) and (35), respectively. Taking the derivative of the retirement choice with respect to the neutrality parameter $\sigma$ gives (evaluated at $\sigma = 1$):

$$\left.\frac{\partial z}{\partial \sigma}\right|_{\sigma=1} = \frac{\pi p}{\gamma} > 0 \tag{38}$$

An increase in the neutrality parameter $\sigma$ leads to later retirement. The introduction of this kind 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 in the flexibility reform with uniform actuarial adjustment and heterogeneous life spans. With uniform actuarial adjustment, however, the pension scheme is still actuarially neutral on average: high-skilled workers (with a long life span) 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 skill level.

### 4.3 Welfare effects

Introducing actuarial non-neutrality does not only stimulate labour supply, it also leads to a Pareto improvement if the tax rate is sufficiently high.

**Proposition 3.** *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:*

$$\hat{\tau} = \frac{\left[1 - G(a^*)\right] \ln\left(\frac{\pi - z_L}{\pi - z_H}\right)}{G(a^*)\frac{qw\pi}{\gamma(\pi - z_L)} + \left[1 - G(a^*)\right]\frac{w\pi}{\gamma(\pi - z_H)}} \tag{39}$$

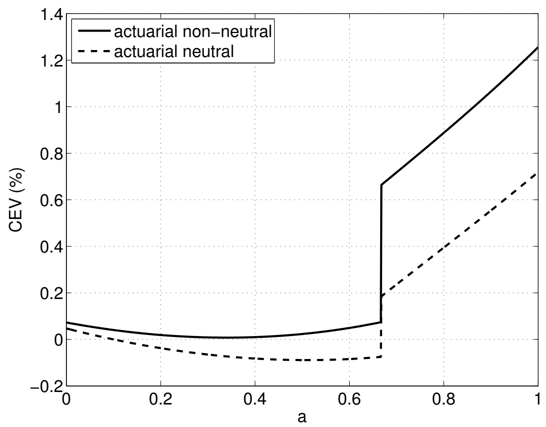
*This implicit equation has a unique solution.*

*Proof.* See Appendix A.2. □

The intuition for this result is similar as in the reform with uniform actuarial adjustment (see Section 3.2). The government can apply non-neutral actuarial conversion

[17] Subscript 'nan' refers to *not actuarially-neutral* adjustment of benefits.

Figure 3: Neutral versus non-neutral actuarial adjustment




*Notes*: The actuarially non-neutral scenario is based on $\sigma = 1.05$.

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 (i.e., the high-skilled), as can also be seen from Figure 2. Therefore, to ensure that the welfare of the low-skilled also improves, the contribution rate needs to be sufficiently high so that the additional tax payments of the high-skilled lead to higher pension benefits.

Figure 3 compares the welfare effects of a uniform adjustment under actuarial neutrality (dashed line) and actuarial non-neutrality (solid line). Contrary to the analytical exposition discussed above, this graph is based on heterogeneous life spans. All parameter values are the same as those used in the previous graphs. As we have shown before, a contribution rate of 30 percent is not sufficient to ensure that an actuarially neutral and a uniform adjustment of benefits is Pareto improving. Figure 3 shows, however, that when a uniform adjustment is combined with actuarial non-neutrality this has strictly positive welfare effects for all individuals under a contribution rate of 30 percent, i.e., the reform is Pareto improving. This implies that by introducing actuarial non-neutrality in the pension scheme it is possible to achieve a Pareto improvement for a lower contribution tax rate. The reason for this result is that an actuarial non-neutral uniform adjustment gives more incentives for the high-skilled to retire later and their labour supply in the second period will be higher than under the actuarially-neutral reform; this will generate more resources to compensate the low-skilled.

# 5 Conclusion

In this paper, we have studied the intragenerational redistribution and welfare effects of a pension reform that introduces a flexible take-up of pension benefits. To analyse the economic implications of such a pension reform, 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. In this way we take into account the empirically most important channels of intragenerational redistribution: income redistribution from rich to poor people and life-span redistribution from short-lived to long-lived agents.

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 this tax critical rate. This policy reduces the implicit tax not only of the high-skilled agents, but also of the low-skilled, implying 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 non-actuarial 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 pension benefits, 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 the rich to the 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.

Other important elements to which we have not paid attention but that might be important when analysing pension flexibility, are the role of income effects in the retirement decision or social norms. Especially in the short run, flexibility in the pension age could lead to only small changes in retirement behaviour if agents are used to retire at some socially accepted retirement age. In the long run, however, norms may change and the effects described in this paper may still apply. To what extent these kinds of issues would affect our main results, is left for future research.

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 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 the 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. In that way, this reform generates extra resources to cope with the costs of ageing *and* make some people better off while not hurting other people.

## References


Adams, P., Hurd, M., McFadden, D., Merrill, A. and Ribeiro, T. (2003), 'Healthy, wealthy, and wise? Tests for direct causal paths between health and socioeconomic status', *Journal of Econometrics* **112**, 3–56.

Bonenkamp, J. (2009), 'Measuring lifetime redistribution in Dutch occupational pensions', *De Economist* **157**, 49–77.

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 Reil-Held, A. (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.

Casamatta, G., Cremer, H. and Pestieau, P. (2005), 'Voting on pensions with endogenous retirement age', *International Tax and Public Finance* **12**, 7–28.

Cremer, H., Lozachmeur, J.-M. and Pestieau, P. (2010), 'Collective annuities and redistribution', *Journal of Public Economic Theory* **12**, 23–41.

Cremer, H. and Pestieau, P. (2003), 'The double dividend of postponing retirement', *International Tax and Public Finance* **10**, 419–434.

Diamond, P. and Mirrlees, J. (1978), 'A model of social insurance with variable retirement', *Journal of Public Economics* **10**, 295–336.

Fisher, W. and Keuschnigg, C. (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.

Hachon, C. (2008), 'Redistribution, pension systems and capital accumulation', *Financial Theory and Practice* **32**, 339–368.

Hougaard Jensen, S., Lau, M. and Poutvaara, P. (2003), 'Efficiency and equity aspects of alternative social security rules', *FinanzArchiv* **60**, 325–358.

Jaag, C., Keuschnigg, C. and Keuschnigg, M. (2010), 'Pension reform, retirement, and life-cycle unemployment', *International Tax and Public Finance* **17**, 556–585.

Krueger, A. and Pischke, S. (1992), 'The effects of social security on labor supply: a cohort analysis of the notch generation', *Journal of Labor Economics* **10**, 412–437.

Mackenbach, J., Bos, V., Andersen, O., Cardano, M., Costa, G., Harding, S., Reid, A., Hemström, O., Valkonen, T. and Kunst, A. (2003), 'Widening socioeconomic inequalities in mortality in six Western European countries', *International Journal of Epidemiology* **32**, 830–837.

Meara, E. R., Richards, S. and Cutler, D. M. (2008), 'The gap gets bigger: changes in mortality and life expectancy, by education, 1981-2000', *Health Affairs* **27**, 350–360.

OECD (2011), *Pensions at a glance: retirement-income systems in OECD and G20 countries*, Paris: OECD Publishing.

Pappas, G., Queen, S., Hadden, W. and Fisher, G. (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.

Pestieau, P. and Possen, U. (2010), 'Retirement as a hedge'. Mimeo.

Queisser, M. and Whitehouse, E. (2006), 'Neutral or fair? Actuarial concepts and pension-system design'. Mimeo.

Razin, A. and Sadka, E. (1999), 'Migration and pension with international capital mobility', *Journal of Public Economics* **74**, 141–150.

van Kippersluis, H., O'Donnell, O. and van Doorslaer, E. (2011), 'Long run returns to education: Does schooling lead to an extended old age?', *Journal of Human Resources* **46**(4), 695–721.

# Technical appendix

This appendix contains formal proofs of all propositions mentioned in this paper (Appendix A). It also derives the effects of a pension flexibility reform where the actuarial adjustment of benefits is based on the skill level of individuals (Appendix B). Finally, we present an extension of the basic model by endogenizing the schooling decision (Appendix C).

## A Proofs

### A.1 Proof of Proposition 2

*Proof.* We have the following condition for a Pareto improvement:

$$\begin{aligned}\Gamma &\equiv P_{uni} - P_{ben} - \frac{[\Psi(z)]^2\,\pi}{2\gamma} \\ &= \frac{(\pi - z)(\bar{\pi} - h)}{\bar{\pi} - z} b_{uni} - (\pi - h) b_{ben} - \frac{[\Psi(z)]^2\,\pi}{2\gamma} \geq 0 \qquad \text{(A.1)}\end{aligned}$$

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 $\lambda = 0$:[18]

$$\frac{\partial \Gamma}{\partial \lambda} = \frac{z - h}{\bar{\pi} - z} \bar{\pi}(a - \bar{a}) b_{uni} + (\bar{\pi} - h)\left(\frac{\partial b_{uni}}{\partial \lambda} - \frac{\partial b_{ben}}{\partial \lambda}\right) \qquad \text{(A.2)}$$

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:

$$b(\bar{\pi} - h)\Phi = X \qquad \text{(A.3)}$$

with:

$$\Phi \equiv \int_0^{a^*} \left[\frac{\pi - z_L}{\bar{\pi} - z_L} - \frac{(\pi - \bar{\pi})\tau q w \pi}{\gamma(\bar{\pi} - z_L)^2}\right] \mathrm{d}G + \int_{a^*}^{1} \left[\frac{\pi - z_H}{\bar{\pi} - z_H} - \frac{(\pi - \bar{\pi})\tau w \pi}{\gamma(\bar{\pi} - z_H)^2}\right] \mathrm{d}G \qquad \text{(A.4)}$$

$$X \equiv \tau q w \int_0^{a^*} \left[1 + \frac{(1-\tau) q w \pi}{\gamma}\right] \mathrm{d}G + \tau w \int_{a^*}^{1} \left[a + \frac{(1-\tau) w \pi}{\gamma}\right] \mathrm{d}G \qquad \text{(A.5)}$$

[18] To avoid complex calculations that yield no additional insights, we evaluate all derivatives in this section at the initial point $\lambda = 0$.

Note that in the benchmark model $\Phi = 1$. From equation (A.3), we derive at $\lambda = 0$:

$$\frac{\partial b_{ben}}{\partial \lambda} = \frac{1}{\bar{\pi} - h} \frac{\partial X}{\partial \lambda} \tag{A.6}$$

$$\frac{\partial b_{uni}}{\partial \lambda} = \frac{1}{\bar{\pi} - h} \left( \frac{\partial X}{\partial \lambda} - X \frac{\partial \Phi}{\partial \lambda} \right) \tag{A.7}$$

Hence,

$$\frac{\partial b_{uni}}{\partial \lambda} - \frac{\partial b_{ben}}{\partial \lambda} = -\frac{X}{\bar{\pi} - h} \frac{\partial \Phi}{\partial \lambda} \tag{A.8}$$

From the definition of $\Phi$ above and applying Leibniz rule, we obtain:

$$\begin{aligned} \frac{\partial \Phi}{\partial \lambda} = & \frac{\bar{\pi}}{\bar{\pi} - z_L} \left[ 1 - \frac{\tau q w \bar{\pi}}{\gamma(\bar{\pi} - z_L)} \right] \int_0^{a^*} (a - \bar{a})\, \mathrm{d}G \\ & + \frac{\bar{\pi}}{\bar{\pi} - z_H} \left[ 1 - \frac{\tau w \bar{\pi}}{\gamma(\bar{\pi} - z_H)} \right] \int_{a^*}^1 (a - \bar{a})\, \mathrm{d}G \end{aligned} \tag{A.9}$$

Inserting equation (28) with $\lambda = 0$ in this expression, gives:

$$\frac{\partial \Phi}{\partial \lambda} = \underbrace{\frac{\gamma(\gamma - qw)}{(\gamma - qw + \tau q w)^2}}_{\Pi_L} \int_0^{a^*} (a - \bar{a})\, \mathrm{d}G + \underbrace{\frac{\gamma(\gamma - w)}{(\gamma - w + \tau w)^2}}_{\Pi_H} \int_{a^*}^1 (a - \bar{a})\, \mathrm{d}G \tag{A.10}$$

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 (32). Hence, $\partial b_{uni}/\partial \lambda \geq \partial b_{ben}/\partial \lambda$ if and only if $\tau \geq \tau^*$. This completes the proof. □

## A.2 Proof of Proposition 3

*Proof.* With actuarial non-neutrality, the Pareto-improving condition is[19]:

$$\Gamma \equiv (\pi - h)^{\sigma} (\pi - z)^{1-\sigma} b_{nan} - (\pi - h) b_{ben} - \frac{[\Psi(z)]^2 \pi}{2\gamma} \geq 0 \tag{A.11}$$

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 derive the following derivative, evaluated in the initial position $\sigma = 1$:

$$\frac{\partial \Gamma}{\partial \sigma} = (\pi - h) \frac{\partial b}{\partial \sigma} + (\pi - h) b \ln(\pi - h) - (\pi - h) b \ln(\pi - z) \tag{A.12}$$

[19]In this section, we abstract from life-span heterogeneity, i.e., $\lambda = 0$.

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:

$$b(\pi - h)\Phi = X \tag{A.13}$$

where $X$ is already defined by equation (A.5) and with $\Phi$ equal to:

$$\begin{aligned}\Phi \equiv G(a^*) &\left[\left(\frac{\pi - z_L}{\pi - h}\right)^{1-\sigma} - \frac{\tau q w \pi(\sigma - 1)(\pi - h)^{\sigma - 1}}{\gamma(\pi - z_L)^{\sigma}}\right] \\ &+ [1 - G(a^*)]\left[\left(\frac{\pi - z_H}{\pi - h}\right)^{1-\sigma} - \frac{\tau w \pi(\sigma - 1)(\pi - h)^{\sigma - 1}}{\gamma(\pi - z_H)^{\sigma}}\right]\end{aligned} \tag{A.14}$$

Note that $\Phi = 1$ if $\sigma = 1$, implying that equation (A.12) can be written as:

$$\frac{\partial\Gamma}{\partial\sigma} = (\pi - h)\frac{\partial b}{\partial\sigma} + X\ln(\pi - h) - X\ln(\pi - z) \tag{A.15}$$

From equation (A.13) it follows:

$$\frac{\partial b}{\partial\sigma} = -\frac{X}{\pi - h}\frac{\partial\Phi}{\partial\sigma} \tag{A.16}$$

Using definition (A.14), we can derive at $\sigma = 1$:

$$\begin{aligned}\frac{\partial\Phi}{\partial\sigma} = &\ \ln(\pi - h) - G(a^*)\left[\ln(\pi - z_L) + \frac{\tau q w \pi}{\gamma(\pi - z_L)}\right] \\ &- [1 - G(a^*)]\left[\ln(\pi - z_H) + \frac{\tau w \pi}{\gamma(\pi - z_H)}\right]\end{aligned} \tag{A.17}$$

Substituting equation (A.17) into equation (A.16) and inserting the resulting expression in equation (A.12) ultimately implies:

$$\begin{aligned}\frac{\partial\Gamma}{\partial\sigma} = &\ -X\ln(\pi - z) + G(a^*)X\left[\ln(\pi - z_L) + \frac{\tau q w \pi}{\gamma(\pi - z_L)}\right] \\ &+ [1 - G(a^*)]\,X\left[\ln(\pi - z_H) + \frac{\tau w \pi}{\gamma(\pi - z_H)}\right]\end{aligned} \tag{A.18}$$

For high-skilled agents we have $z = z_H$, implying:

$$\begin{aligned}\frac{\partial\Gamma}{\partial\sigma} = G(a^*)&\frac{\tau q w \pi X}{\gamma(\pi - z_L)} + [1 - G(a^*)]\,\frac{\tau w \pi X}{\gamma(\pi - z_H)} \\ &+ G(a^*)X\ln\left(\frac{\pi - z_L}{\pi - z_H}\right) > 0\end{aligned} \tag{A.19}$$

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:

$$\begin{aligned} \frac{\partial \Gamma}{\partial \sigma} = G(a^*) \frac{\tau q w \pi X}{\gamma(\pi - z_L)} + [1 - G(a^*)] \, \frac{\tau w \pi X}{\gamma(\pi - z_H)} \\ - [1 - G(a^*)] \, X \ln \left( \frac{\pi - z_L}{\pi - z_H} \right) \end{aligned} \tag{A.20}$$

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 $z_L = z_H = 0$ 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 (39).

To prove that $\hat{\tau}$ is a unique solution, we have to show that the derivative $\partial \Gamma / \partial \sigma$ is monotonically increasing in $\tau$ at $\sigma = 1$. Rewrite equation (A.20) 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 $\tau$. Taking the derivative of $A$ with respect to $\tau$ gives, after some algebraic manipulations[20]:

$$\begin{aligned} \frac{\partial A}{\partial \tau} = G(a^*) \frac{q w (\gamma - q w)}{(\gamma - q w + \tau q w)^2} + [1 - G(a^*)] \, \frac{w(\gamma - w)}{(\gamma - w + \tau w)^2} \\ + [1 - G(a^*)] \, \frac{w \pi}{\gamma} \left( \frac{1}{\pi - z_H} - \frac{q}{\pi - z_L} \right) > 0 \end{aligned} \tag{A.21}$$

This completes the proof. □

# B Skill-dependent actuarial adjustment

A practical impediment of the reform where the actuarial adjustment of benefits is based on individual life spans 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. In terms of our model, this characteristic could be the skill level of agents. The fact that agents are high- or low-skilled reveals 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 in-

[20]Note that we assume $\gamma > w$ to make sure that $z < 1$ for all agents irrespective the size of the pension scheme (so $\forall \tau$).

formation by conditioning the actuarial adjustment factor on skill level. This also reduces the redistribution from low-skilled to high-skilled people.

### Actuarial adjustment factor

The adjustment factor is made conditional on skill level with the following life span indicator:

$$\hat{\pi} = \begin{cases} \bar{\pi}_L \equiv \int_0^{a^*} \frac{\pi}{G(a^*)}\,\mathrm{d}G & \text{if} \quad a < a^* \\ \bar{\pi}_H \equiv \int_{a^*}^1 \frac{\pi}{1-G(a^*)}\,\mathrm{d}G & \text{if} \quad a > a^* \end{cases} \tag{A.22}$$

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 adjustment factor is:

$$m = \begin{cases} \frac{\bar{\pi}-h}{\bar{\pi}_L - z_L} & \text{if} \quad a < a^* \\ \frac{\bar{\pi}-h}{\bar{\pi}_H - z_H} & \text{if} \quad a > a^* \end{cases} \tag{A.23}$$

and pension entitlements are equal to:

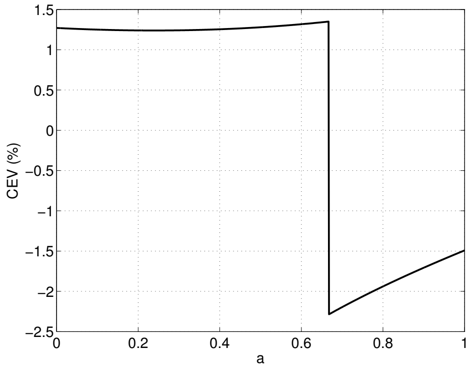
$$P = \begin{cases} \frac{(\pi - z_L)(\bar{\pi}-h)b}{\bar{\pi}_L - z_L} & \text{if} \quad a < a^* \\ \frac{(\pi - z_H)(\bar{\pi}-h)b}{\bar{\pi}_H - z_H} & \text{if} \quad a > a^* \end{cases} \tag{A.24}$$

From equation (A.23) it follows that skill-dependent adjustment reduces redistribution from short-lived to long-lived ability groups, like with individual actuarial adjustment. Indeed, when all agents retire at the statutory retirement date $h$, it holds that $m > 1$ for the low-skilled group while for the high-skilled group $m < 1$. Low-skilled agents are compensated for the fact that they have a shorter life span. Contrary to individual adjustment, however, skill-dependent adjustment does not remove life-span redistribution completely. Therefore, these transfers will lead to distortions in the retirement decision. To see this, from equation (A.24) we have:

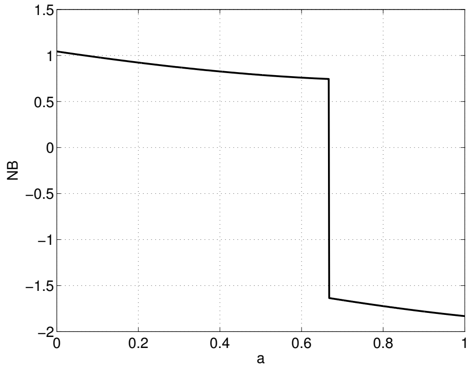
$$\Psi(z) \equiv \frac{\partial P(z)}{\partial z} = \begin{cases} \frac{(\pi - \bar{\pi}_L)p}{\bar{\pi}_L - z_L} & \text{if} \quad a < a^* \\ \frac{(\pi - \bar{\pi}_H)p}{\bar{\pi}_H - z_H} & \text{if} \quad a > a^* \end{cases} \tag{A.25}$$

Within a skill group, agents with relatively long life spans ($\pi > \hat{\pi}$) still have an incentive to delay retirement and agents with relatively short life spans ($\pi < \hat{\pi}$) still prefer to retire early. These selection effects are smaller, however, 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.

Figure 4: Skill-dependent adjustment: welfare and redistribution




(a) Welfare




(b) Net benefit

*Notes*: Net benefit is expressed in absolute difference from the benchmark and certainty-equivalent variation is expressed in percent benchmark consumption.

### Consumption and retirement

The expressions for consumption and retirement are given by:[21]

$$c_{edu} = c_{ben} + \frac{1}{1+\pi}\left[P_{edu} - P_{ben} - \frac{[\Psi(z_{edu})]^2\,\pi}{2\gamma}\right] \tag{A.26}$$

$$z_{edu} = z_{ben} + \frac{\Psi(z_{edu})\pi}{\gamma} \tag{A.27}$$

Retirement behaviour is subject to two different labour supply distortions. The first distortion is caused by the contribution rate $\tau$ which induces early retirement. The second distortion is caused by the derivative $\Psi$ which reflects the selection effects associated with life-span heterogeneity.

### Welfare effects

Applying skill-dependent adjustment 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 agents rises while the welfare of the high-skilled falls, as can also be seen in Figure 4. 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 respect, the welfare effects of skill-dependent adjustment are comparable with those of individual adjustment.

The following proposition summarizes this result:

[21]Subscript 'edu' refers to *educational* actuarial adjustment of benefits.

**Proposition 4.** *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.* The condition for a Pareto improvement is:

$$\begin{aligned}\Gamma &\equiv P_{edu} - P_{ben} - \frac{[\Psi(z)]^2 \pi}{2\gamma} \\ &= \frac{(\pi - z)(\bar{\pi} - h)}{\hat{\pi} - z} b_{edu} - (\pi - h) b_{ben} - \frac{[\Psi(z)]^2 \pi}{2\gamma} \geq 0 \end{aligned} \tag{A.28}$$

where for at least one $a$-individual this inequality has to hold strictly. We now derive the following derivative, evaluated at the initial position $\lambda = 0$:

$$\frac{\partial \Gamma}{\partial \lambda} = \frac{z - h}{\bar{\pi} - z} \bar{\pi}(a - \bar{a}) b_{edu} + (\bar{\pi} - h) \left( \frac{\partial b_{edu}}{\partial \lambda} - \frac{\partial b_{ben}}{\partial \lambda} \right) - \frac{\bar{\pi} - h}{\bar{\pi} - z} \frac{\partial \hat{\pi}}{\partial \lambda} b_{edu} \tag{A.29}$$

To prove that this reform cannot be a Pareto improvement, we have to show that for at least one $a$-individual equation (A.29) is strictly negative. Let us concentrate on a high-skilled agent with ability level $a = a^*$. Using equation (A.22), equation (A.29) then becomes:

$$\begin{aligned}\frac{\partial \Gamma(a = a^*)}{\partial \lambda} &= \underbrace{\frac{z_H - h}{\bar{\pi} - z_H} \bar{\pi}(a^* - \bar{a}) b_{edu}}_{\Sigma_1} + (\bar{\pi} - h) \left( \frac{\partial b_{edu}}{\partial \lambda} - \frac{\partial b_{ben}}{\partial \lambda} \right) \\ &\quad - \underbrace{\frac{\bar{\pi} - h}{\bar{\pi} - z_H} \frac{\int_{a^*}^{1} \bar{\pi}(a - \bar{a})\, \mathrm{d}G}{1 - G(a^*)} b_{edu}}_{\Sigma_2 > 0} \end{aligned} \tag{A.30}$$

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.

The budget constraint of the pension scheme can be written as

$$b(\bar{\pi} - h)\Phi = X \tag{A.31}$$

with $X$ already defined by equation (A.5) and with $\Phi$ equal to:

$$\begin{aligned}\Phi \equiv &\int_0^{a^*} \left[ \frac{\pi - z_L}{\bar{\pi}_L - z_L} - \frac{(\pi - \bar{\pi}_L)\tau q w \pi}{\gamma(\bar{\pi}_L - z_L)^2} \right] \mathrm{d}G \\ &+ \int_{a^*}^{1} \left[ \frac{\pi - z_H}{\bar{\pi}_H - z_H} - \frac{(\pi - \bar{\pi}_H)\tau w \pi}{\gamma(\bar{\pi}_H - z_H)^2} \right] \mathrm{d}G \end{aligned} \tag{A.32}$$

From equation (A.31), we derive:

$$\frac{\partial b_{ben}}{\partial \lambda} = \frac{1}{\bar{\pi} - h} \frac{\partial X}{\partial \lambda} \tag{A.33}$$

$$\frac{\partial b_{edu}}{\partial \lambda} = \frac{1}{\bar{\pi} - h} \left( \frac{\partial X}{\partial \lambda} - X \frac{\partial \Phi}{\partial \lambda} \right) \tag{A.34}$$

Hence,

$$\left. \frac{\partial b_{edu}}{\partial \lambda} - \frac{\partial b_{ben}}{\partial \lambda} \right|_{\lambda=0} = -\frac{X}{\bar{\pi} - h} \frac{\partial \Phi}{\partial \lambda} \tag{A.35}$$

From the definition of $\Phi$ and applying Leibniz rule, we derive at $\lambda = 0$:

$$\begin{aligned} \frac{\partial \Phi}{\partial \lambda} = & \left[ \frac{1}{\bar{\pi} - z_L} - \frac{\tau q w \bar{\pi}}{\gamma(\bar{\pi} - z_L)^2} \right] \int_0^{a^*} \left( \frac{\partial \pi}{\partial \lambda} - \frac{\partial \bar{\pi}_L}{\partial \lambda} \right) \mathrm{d}G \\ & + \left[ \frac{1}{\bar{\pi} - z_H} - \frac{\tau w \bar{\pi}}{\gamma(\bar{\pi} - z_H)^2} \right] \int_{a^*}^1 \left( \frac{\partial \pi}{\partial \lambda} - \frac{\partial \bar{\pi}_H}{\partial \lambda} \right) \mathrm{d}G \end{aligned} \tag{A.36}$$

Notice:

$$\begin{aligned} \int_0^{a^*} \left( \frac{\partial \pi}{\partial \lambda} - \frac{\partial \bar{\pi}_L}{\partial \lambda} \right) \mathrm{d}G &= \int_0^{a^*} \left[ \bar{\pi}(a - \bar{a}) - \frac{\int_0^{a^*} \bar{\pi}(a - \bar{a}) \, \mathrm{d}G}{G(a^*)} \right] \mathrm{d}G \\ &= \bar{\pi} \int_0^{a^*} (a - \bar{a}) \, \mathrm{d}G - \bar{\pi} \int_0^{a^*} (a - \bar{a}) \, \mathrm{d}G = 0 \end{aligned} \tag{A.37}$$

Along the same lines, we also have:

$$\int_{a^*}^1 \left( \frac{\partial \pi}{\partial \lambda} - \frac{\partial \bar{\pi}_H}{\partial \lambda} \right) \mathrm{d}G = 0 \tag{A.38}$$

Hence, we have $\partial \Phi / \partial \lambda = 0$. Therefore, $\partial b_{edu} / \partial \lambda = \partial b_{ben} / \partial \lambda$, which completes the proof. $\square$

# C 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.

## C.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:

$$a^*_{ben} = q - \frac{(1 - \tau) w \pi(a^*_{ben})(1 - q^2)}{2\gamma} \tag{A.39}$$

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 (A.39), we can infer the following. First, an increase in the tax rate $\tau$ 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.

### C.2 Uniform actuarial adjustment

We will show that uniform adjustment can either increase or decrease the schooling incentives, depending 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 (27) we obtain:

$$a^*_{uni} = a^*_{ben} - \frac{2\gamma\Theta}{(1-\tau)w\left[2\gamma + (1-\tau)w(1-q^2)\bar{\pi}\lambda\right]} \tag{A.40}$$

with,

$$\Theta \equiv P(z_H) - P(z_L) - \frac{\left[\Psi(z_H)\right]^2\pi}{2\gamma} + \frac{\left[\Psi(z_L)\right]^2\pi}{2\gamma}$$

From equation (A.40), it follows[22]:

$$\frac{\partial a^*_{uni}}{\partial\lambda} = \frac{\partial a^*_{ben}}{\partial\lambda} - \frac{1}{(1-\tau)w}\frac{\partial\Theta}{\partial\lambda} \tag{A.41}$$

where we have used that $\Theta = 0$ if $\lambda = 0$. Using equations (24) and (25), we derive from the definition of $\Theta$:

$$\frac{\partial\Theta}{\partial\lambda} = \frac{\bar{\pi}X(a^* - \bar{a})(z_H - z_L)}{(\bar{\pi} - z_H)(\bar{\pi} - z_L)} \tag{A.42}$$

Hence,

$$\frac{\partial a^*_{uni}}{\partial\lambda} = \frac{\partial a^*_{ben}}{\partial\lambda} - \frac{\bar{\pi}X(a^*_{ben} - \bar{a})(z_H - z_L)}{(1-\tau)w(\bar{\pi} - z_H)(\bar{\pi} - z_L)} \tag{A.43}$$

[22]Similar to exogenous schooling (see Appendices A.1 and B), we simplify calculations by evaluating the derivatives at the initial position $\lambda = 0$ (i.e., no life-time heterogeneity).

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, $\pi(a^*) > \bar{\pi}$, this 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 change their skill level after this reform and agents who do not change their 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 (A.1). 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 (A.8) changes into:

$$\frac{\partial b_{uni}}{\partial\lambda} - \frac{\partial b_{ben}}{\partial\lambda}\bigg|_{\lambda=0} = \frac{1}{\bar{\pi}-h}\left[\frac{\partial X(a^*_{uni})}{\partial\lambda} - \frac{\partial X(a^*_{ben})}{\partial\lambda} - X(a^*_{uni})\frac{\partial\Phi}{\partial\lambda}\right] \tag{A.44}$$

From the definition of $X$, see equation (A.5), we derive at the point $\lambda = 0$:

$$\begin{aligned}\frac{\partial X}{\partial\lambda} = &-\frac{\tau(1-\tau)w^2(1-q^2)\bar{\pi}}{2\gamma}\frac{\partial a^*}{\partial\lambda} + \tau q w\int_0^{a^*}\frac{(1-\tau)qw\bar{\pi}(a-\bar{a})}{\gamma}\,\mathrm{d}G \\ &+ \tau w\int_{a^*}^1\frac{(1-\tau)w\bar{\pi}(a-\bar{a})}{\gamma}\,\mathrm{d}G\end{aligned} \tag{A.45}$$

Using equation (A.45), we obtain:

$$\begin{aligned}\frac{\partial X(a^*_{uni})}{\partial\lambda} - \frac{\partial X(a^*_{ben})}{\partial\lambda} &= \frac{\tau(1-\tau)w^2(1-q^2)\bar{\pi}}{2\gamma}\left(\frac{\partial a^*_{ben}}{\partial\lambda} - \frac{\partial a^*_{uni}}{\partial\lambda}\right) \\ &= \frac{\tau w(1-q^2)\bar{\pi}}{2\gamma}\frac{\partial\Theta}{\partial\lambda}\end{aligned} \tag{A.46}$$

where we have used equation (A.41) in going from the first line to the second line. Equation (A.44) can now rewritten in:

$$\frac{\partial b_{uni}}{\partial\lambda} - \frac{\partial b_{ben}}{\partial\lambda} = \frac{1}{\bar{\pi}-h}\left[\frac{\tau w(1-q^2)\bar{\pi}}{2\gamma}\frac{\partial\Theta}{\partial\lambda} - X\frac{\partial\Phi}{\partial\lambda}\right] \tag{A.47}$$

The derivative $\partial\Phi/\partial\lambda$ is still equal to equation (A.10). That is,

$$\frac{\partial\Phi}{\partial\lambda} = \underbrace{\frac{\gamma(\gamma - qw)}{(\gamma - qw + \tau qw)^2}}_{\Pi_L} \int_0^{a^*} (a - \bar{a})\, \mathrm{d}G + \underbrace{\frac{\gamma(\gamma - w)}{(\gamma - w + \tau w)^2}}_{\Pi_H} \int_{a^*}^{1} (a - \bar{a})\, \mathrm{d}G \tag{A.48}$$

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 (A.47) 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 (A.47) 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 $\tau$ 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:

$$\Gamma \equiv \underbrace{P_{uni} - P_{ben} - \frac{[\Psi(z)]^2\pi}{2\gamma}}_{\Upsilon_1} + \underbrace{(1-\tau)w(a-q) + \frac{(1-q^2)(1-\tau)^2 w^2 \pi}{2\gamma}}_{\Upsilon_2} \geq 0 \tag{A.49}$$

Compared to the non-switching group, we now have an additional term $\Upsilon_2$ 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 $\lambda$ and evaluating the resulting expression at $\lambda = 0$, gives:

$$\frac{\partial\Gamma}{\partial\lambda} = \frac{\partial\Upsilon_1}{\partial\lambda} + \frac{(1-q^2)(1-\tau)^2 w^2 \bar{\pi}(a-\bar{a})}{2\gamma} \tag{A.50}$$

Let the tax rate be set such that $\partial\Upsilon_1/\partial\lambda \geq 0$. 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:

$$\Gamma \equiv \Upsilon_1 + (1-\tau)w(q-a) + \frac{(q^2-1)(1-\tau)^2 w^2 \pi}{2\gamma} \geq 0 \tag{A.51}$$

Taking the derivative with respect to $\lambda$ gives:

$$\frac{\partial\Gamma}{\partial\lambda} = \frac{\partial\Upsilon_1}{\partial\lambda} + \frac{(q^2-1)(1-\tau)^2 w^2 \bar{\pi}(a-\bar{a})}{2\gamma} \tag{A.52}$$

Suppose again that the tax rate is set such that $\partial\Upsilon_1/\partial\lambda \geq 0$. 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.

## C.3 Actuarially non-neutral adjustment

In this section, we show that stimulating retirement postponement in 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 (36) we can infer:

$$a^*_{nan} = a^*_{ben} - \frac{\Theta}{(1-\tau)w} \tag{A.53}$$

with $\Theta$ 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 (A.53) with respect to $\sigma$, evaluated at the initial point $\sigma = 1$:[23]

$$\frac{\partial a^*_{nan}}{\partial\sigma} = \frac{\partial a^*_{ben}}{\partial\sigma} - \frac{1}{(1-\tau)w}\frac{\partial\Theta}{\partial\sigma} \tag{A.54}$$

where we have used that $\Theta = 0$ if $\sigma = 1$. Using equations (34) and (35), we derive from the definition of $\Theta$:

$$\frac{\partial\Theta}{\partial\sigma} = X \ln\left(\frac{\pi - z_L}{\pi - z_H}\right) > 0 \tag{A.55}$$

Hence, $\partial a^*_{nan}/\partial\sigma < \partial a^*_{ben}/\partial\sigma$.

### 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 (A.11).

[23] Similar to exogenous schooling (see Section A.2), we abstract from life-span heterogeneity and evaluate derivatives at the initial position $\sigma = 1$ (i.e., actuarial neutrality).

With endogenous schooling, equation (A.16) changes into:

$$\frac{\partial b}{\partial \sigma} = \frac{1}{\pi - h}\left(\frac{\partial X}{\partial \sigma} - X\frac{\partial \Phi}{\partial \sigma}\right) \tag{A.56}$$

Using equation (A.5), we have at $\sigma = 1$:

$$\begin{aligned}\frac{\partial X}{\partial \sigma} &= -\frac{\tau(1-\tau)w^2(1-q^2)\pi}{2\gamma}\frac{\partial a^*}{\partial \sigma} \\ &= \frac{\tau w(1-q^2)\pi X}{2\gamma}\ln\left(\frac{\pi - z_L}{\pi - z_H}\right) > 0\end{aligned} \tag{A.57}$$

where we have used equation (A.55) when 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 (A.17). Substituting this equation together with equation (A.57) into equation (A.56) and inserting the resulting expression in equation (A.12) ultimately implies:

$$\begin{aligned}\frac{\partial \Gamma}{\partial \sigma} = G(a^*)X\left[\ln(\pi - z_L) + \frac{\tau q w \pi}{\gamma(\pi - z_L)}\right] + \frac{\tau w(1-q^2)\pi X}{2\gamma}\ln\left(\frac{\pi - z_L}{\pi - z_H}\right) \\ + \left[1 - G(a^*)\right] X \left[\ln(\pi - z_H) + \frac{\tau w \pi}{\gamma(\pi - z_H)}\right] - X\ln(\pi - z)\end{aligned} \tag{A.58}$$

For high-skilled agents we have $z = z_H$, implying:

$$\begin{aligned}\frac{\partial \Gamma}{\partial \sigma} &= G(a^*)\frac{\tau q w \pi X}{\gamma(\pi - z_L)} + \left[1 - G(a^*)\right]\frac{\tau w \pi X}{\gamma(\pi - z_H)} \\ &\quad + X\left[G(a^*) + \frac{\tau w(1-q^2)\pi}{2\gamma}\right]\ln\left(\frac{\pi - z_L}{\pi - z_H}\right) > 0\end{aligned} \tag{A.59}$$

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

$$\begin{aligned}\frac{\partial \Gamma}{\partial \sigma} &= G(a^*)\frac{\tau q w \pi X}{\gamma(\pi - z_L)} + \left[1 - G(a^*)\right]\frac{\tau w \pi X}{\gamma(\pi - z_H)} \\ &\quad + X\left[\frac{\tau w(1-q^2)\pi}{2\gamma} - 1 + G(a^*)\right]\ln\left(\frac{\pi - z_L}{\pi - z_H}\right)\end{aligned} \tag{A.60}$$

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 $\tau$ 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:

$$\Gamma \equiv P_{nan} - P_{ben} - \frac{[\Psi(z)]^2 \pi}{2\gamma} + (1-\tau)w(a-q) + \frac{(1-q^2)(1-\tau)^2 w^2 \pi}{2\gamma} \tag{A.61}$$

Notice that the derivative of $\Gamma$ with respect to $\sigma$ is exactly the same as for the non-switching group and given by equation (A.59).

## C.4 Skill-dependent actuarial adjustment

In this final section, we show in this section that skill-dependent actuarial adjustment negatively affects the incentives to become skilled. In addition, analogue to exogenous schooling, this reform cannot be a Pareto improvement.

### Cut-off ability level

Skill-dependent actuarial adjustment will changes schooling because it introduces an endogenous link between the schooling decision and the actuarial adjustment factor. Using equation (A.26) we can infer:

$$a^*_{edu} = a^*_{ben} - \frac{2\gamma\Theta}{(1-\tau)w\left[2\gamma + (1-\tau)w(1-q^2)\bar{\pi}\lambda\right]} \tag{A.62}$$

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 (A.62) with respect to $\lambda$, evaluated at $\lambda = 0$:

$$\frac{\partial a^*_{edu}}{\partial \lambda} = \frac{\partial a^*_{ben}}{\partial \lambda} - \frac{1}{(1-\tau)w}\frac{\partial \Theta}{\partial \lambda} \tag{A.63}$$

From the definition of $\Theta$ it follows:

$$\frac{\partial \Theta}{\partial \lambda} = \frac{X}{\bar{\pi} - z_H}\left[\frac{\partial \pi}{\partial \lambda} - \frac{\partial \bar{\pi}_H}{\partial \lambda}\right] - \frac{X}{\bar{\pi} - z_L}\left[\frac{\partial \pi}{\partial \lambda} - \frac{\partial \bar{\pi}_L}{\partial \lambda}\right] \tag{A.64}$$

From equations (5) and (A.22), we obtain for $\lambda = 0$:

$$\frac{\partial \pi(a^*)}{\partial \lambda} = \bar{\pi}(a^* - \bar{a}) \tag{A.65}$$

$$\frac{\partial \bar{\pi}_L}{\partial \lambda} = \frac{\int_0^{a^*} \bar{\pi}(a - \bar{a})\, \mathrm{d}G}{G(a^*)} \tag{A.66}$$

$$\frac{\partial \bar{\pi}_H}{\partial \lambda} = \frac{\int_{a^*}^1 \bar{\pi}(a - \bar{a})\, \mathrm{d}G}{1 - G(a^*)} \tag{A.67}$$

Substituting these expressions in equation (A.64) and rearranging, gives:

$$\left.\frac{\partial \Theta}{\partial \lambda}\right|_{\lambda=0} = \underbrace{\frac{X}{\bar{\pi} - z_L} \frac{\int_0^{a^*} \bar{\pi}(a - a^*)\, \mathrm{d}G}{G(a^*)}}_{<0} - \underbrace{\frac{X}{\bar{\pi} - z_H} \frac{\int_{a^*}^1 \bar{\pi}(a - a^*)\, \mathrm{d}G}{1 - G(a^*)}}_{>0} < 0 \tag{A.68}$$

An increase in $\lambda$ reduces $\Theta$ and using equation (A.63) this implies that $\partial a^*_{edu}/\partial \lambda > \partial a^*_{ben}/\partial \lambda$. With educational-specific actuarial adjustment, individuals can self-select the actuarial adjustment 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 (A.28). 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 (A.35) changes into:

$$\frac{\partial b_{edu}}{\partial \lambda} - \frac{\partial b_{ben}}{\partial \lambda} = \frac{1}{\bar{\pi} - h} \left[ \frac{\partial X(a^*_{edu})}{\partial \lambda} - \frac{\partial X(a^*_{ben})}{\partial \lambda} - X(a^*_{edu}) \frac{\partial \Phi}{\partial \lambda} \right] \tag{A.69}$$

Using equations (A.45) and (A.63), we obtain in $\lambda = 0$:

$$\begin{aligned} \frac{\partial X(a^*_{edu})}{\partial \lambda} - \frac{\partial X(a^*_{ben})}{\partial \lambda} &= \frac{\tau(1-\tau)w^2(1-q^2)\bar{\pi}}{2\gamma} \left( \frac{\partial a^*_{ben}}{\partial \lambda} - \frac{\partial a^*_{edu}}{\partial \lambda} \right) \\ &= \frac{\tau w(1-q^2)\bar{\pi}}{2\gamma} \frac{\partial \Theta}{\partial \lambda} < 0 \end{aligned} \tag{A.70}$$

With endogenous schooling, there still holds $\partial \Phi/\partial \lambda = 0$. Equation (A.69) then implies $\partial b_{edu}/\partial \lambda < \partial b_{ben}/\partial \lambda$.