CER-ETH – Center of Economic Research at ETH Zurich

On the Public Economics of Annuities with Differential Mortality

Antoine Bommier, Marie-Louise Leroux, Jean-Marie Lozachmeur

Working Paper 10/135
September 2010

Economics Working Paper Series

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

On the Public Economics of Annuities with Differential Mortality

Antoine Bommier[1], Marie-Louise Leroux[2], Jean-Marie Lozachmeur[3]

September 10, 2010

[1]ETH-Zurich (Department of Management, Technology and Economics).
[2]CORE, Université catholique de Louvain.
[3]Corresponding author. Toulouse School of Economics (CNRS-GREMAQ and IDEI). Mailing: IDEI, 21 Allée de Brienne 31000 Toulouse, France. E-mail: jean-marie.lozachmeur@univ-tlse1.fr.

Abstract

This paper studies the problem of redistribution between individuals having different mortality rates. We use a continuous time model in which there are two types of individuals characterized by different survival probability paths. Individual preferences are represented by a generalized life-cycle utility function which can exhibit temporal risk aversion. We successively compare utilitarian allocations when individuals exhibit temporal risk neutrality and temporal risk aversion. This problem is analyzed in the contexts of full information and asymmetric information on mortality rates.

JEL classification: H55, H23, I31.

Key words: Uncertain Lifetime, Redistribution, Annuities, Nonlinear Taxation.

1 Introduction

Studies on annuity and pensions usually rely on the seminal paper by Yaari (1965). This standard approach, though analytically convenient, relies on some strong assumptions on individuals preferences. In particular lifetime utilities are additively separable, which implies temporal risk neutrality. Such an assumption has major consequences. Firstly, it is found that the optimal annuity pattern is independent of the individual's mortality profile (Yaari, 1965, Levhari and Mirman, 1977, Barro and Friedman, 1977). Secondly, optimal allocation of resources between individuals with different mortality takes a very simple form. As shown for example in Sheshinski (2007), a utilitarian social planner would like to equalize instantaneous levels of consumptions between individuals with different mortality profiles. Consequently, the Social Security scheme should redistribute life-cycle income from individuals with high mortality to those with low mortality.

Accounting for temporal risk aversion is of crucial importance when considering risks that have long term consequences. The risk of death being one of them, temporal risk aversion turns out to be a key aspect of individual preferences when studying intertemporal choice under uncertain lifetime (Bommier, 2006). In particular, when temporal risk aversion is introduced, Yaari's famous result vanishes and the optimal consumption profile depends on individual's mortality. For a social planner this may be of importance for two reasons. First, the first-best objective no longer corresponds to giving the same annuity profile to all individuals, independently of their mortality. Second, since people with different mortality look for different annuity profiles, the age profile of annuity becomes an interesting policy tool that can be used to achieve some redistribution.

The present paper emphasizes the role of temporal risk aversion when designing pensions for individuals having different mortality profiles. The question of pension design has been addressed by several papers using the standard additive approach, thus assuming temporal risk neutrality. This is the case in Sheshinski (2007) and Cremer et al. (2010). Accounting for temporal risk aversion provides new perspectives. Contrarily to what is found with the standard additive approach, it is not necessary to relate heterogeneity in mortality and heterogeneity in income or wealth to have non trivial results. Thus, we decided to focus on the simple case where all agents have the same financial endowments. The problems that result from the correlation between mortality and income or wealth, which were central in these two papers, are left for further contributions.

We consider a setting with heterogenous individuals differing in their mortality profile and being on the verge of retirement. The low-type individuals are characterized by a higher mortality rate at any age. Time is continuous and we consider that the economy is characterized by an exogenous interest rate. Individuals utility may exhibit temporal risk neutrality as well as temporal risk aversion. We study the design of annuity profiles implied by a utilitarian social planner. Our main results are as follows. First, when the government can observe individuals mortality profiles, the optimum leads to a pooling allocation if individuals preferences exhibit temporal risk neutrality. However, with temporal risk aversion, low-type individuals should have a higher level of instantaneous consumption and a lower consumption growth rate. When mortality rates are private information, the strategy of the government very much depends on the characteristics of existing private markets and on the government's ability to commit. As an illustration, we consider the case where there are perfect markets for savings and annuities and the case where private savings are impossible. These are of course extreme scenarios, with reality lying in between. We also compare the situation where the government can commit to some flow of pension payments in exchange of information on mortality, with the setting where such kind of commitment is impossible. When there are perfect markets, the first best can be implemented through the appropriate design of a public pension system. This is no longer possible when markets are deficient, but we then characterize the optimal government policies. The paper is illustrated by numerical simulations based on realistic demographic data.

2 The model

We consider an economy where every individuals are endowed with the same initial wealth W_0. The population is divided into two categories. Individuals of type H are characterized by lower mortality rates than individuals of type L. Denoting $\mu^H(t)$ and $\mu^L(t)$ agents H and L hazard rate of death at age t, we thus assume that:

$$A1 : \mu^H(t) < \mu^L(t) \text{ for every } t$$

Agents of type H have therefore higher survival probability that agents of type L. We also assume that mortality rates increase with age:

$$A2 : \frac{d}{dt}\mu^i(t) > 0$$

Demographic studies indicate that this assumption is realistic when considering ages greater than 25 or 30. Since our paper deals with pensions that are typically received after retirement, such an assumption is rather unrestrictive. Assumption A_3 further states that the hazard rate of death increases more slowly for individuals of type L:

$$A3 : \frac{\mu^H(t+\varepsilon)}{\mu^H(t)} \geqslant \frac{\mu^L(t+\varepsilon)}{\mu^L(t)}$$

for any t and $t+\varepsilon$, where $\varepsilon > 0$. In other words, this assumption states that the relative difference between mortality rates is decreasing with age. Again this assumption is supported by studies on differential mortality at adult and old ages (Brown, Liebman and Pollet, 2002).

At birth, the proportion of type-i individuals is n^i, an exogenous constant. Throughout the paper, we denote $j(t)$, the exogenous return on private savings at time t, the actuarially fair return on annuity being $j(t) + \mu^H(t)$ for type-H individuals and $j(t) + \mu^L(t)$ for type-L individuals.

2.1 Individuals preferences

Yaari's standard approach consists in assuming that a life of length T with a consumption profile $\mathbf{c} = c(.)$ yields a lifetime utility:

$$U^{\text{yaari}}(\mathbf{c}, T) = \int_0^T \alpha(t)\, u(c(t))\, dt \tag{1}$$

where $\alpha(t)$ is an exogenous time discount factor. Bommier (2006) emphasized the limits of such an approach which relies on the assumption of temporal risk neutrality, a rather unappealing assumption for dealing with risks which have long term consequences, such as the risk of death. Temporal risk aversion can be introduced, without abandoning the expected utility framework, by considering utility functions of the form:

$$U(\mathbf{c}, T) = \phi\left(\int_0^T \alpha(t)\, u(c(t))\, dt\right) \tag{2}$$

where ϕ is an increasing function. With no loss of generality it can be assumed that $\phi(0) = 0$. As is known from Kihlstrom and Mirman (1974), playing with the function ϕ involves adjusting individuals risk aversion. When ϕ is concave the agent with the above utility function is simply more risk averse than the agent with Yaari's utility function (which is obtained when ϕ is linear).

A simple interpretation of the specification in (1) is that agents have a linear "lifetime felicity". Each moment of life gives them an instantaneous felicity $\alpha(t)u(c(t))$ that is additively aggregated in order to get the lifetime felicity. However, given the uncertainty about life duration (and about consumption), individuals cannot know ex-ante what will their lifetime felicity be. At most, they know the distribution of lifetime felicity. Introducing a function ϕ as in (2) enables to consider risk aversion with respect to lifetime felicity. For consumption profiles that would provide a constant flow of felicity, the function ϕ would determine individual risk aversion with respect to life duration. A linear ϕ would involve assuming risk neutrality, while a concave function ϕ would indicate a positive risk aversion. While there is no theoretical obstacle to considering risk-prone agents, we limit ourselves to the case where ϕ is concave ($\phi'' \leq 0$) and where $-\phi''/\phi'$ is a non increasing function (consistent with the idea of non increasing absolute risk aversion with respect to lifetime felicity).

Undeniably, introducing temporal risk aversion complicates the computation associated with utility maximization. This is probably one of the main reasons that led economists to focus on Yaari's specification for so many years. A major difficulty seems to appear when writing the expected utility function. Indeed, when life duration is random, the expected lifetime utility associated with a given consumption profile $\mathbf{c}$ is:

$$EU(\mathbf{c}) = \int_0^{+\infty} \mu(t) \exp\left(-\int_0^t \mu(\tau)d\tau\right) \phi\left(\int_0^t \alpha(\tau)\, u(c(\tau))\, d\tau\right) dt$$

By integration by part, this may also rewrite:

$$EU(\mathbf{c}) = \int_0^{+\infty} s(t)\alpha(t)\,u(c(t))\,\phi'\left(\int_0^t \alpha(\tau)\,u(c(\tau))\,d\tau\right)dt$$

where $s(t) = \exp\left(-\int_0^t \mu(\tau)d\tau\right)$ is the probability of being alive at age t. When ϕ is not linear, expected utility is then no longer additive, which might look like the beginning of a nightmare for economists. Bommier (2006) explains however that this difficulty can easily be avoided by making a linear approximation. This allows to avoid the pangs of endogenous discounting without losing most of the insights brought by this novel approach.

The idea is to rely on what is called the assumption of a *"priceless life context"*. Basically, this assumption consists in assuming that the difference in terms of welfare between life and death is much greater than the difference between high and low levels of consumption. Under this assumption (and through an appropriate normalization of the functions ϕ and/or α), preferences can be approximated by an additive expected lifetime utility function:[1]

$$EU(\mathbf{c}) = \int_0^{+\infty} s(t)\,\alpha(t)\,\beta(t)\,u(c(t))\,dt \tag{3}$$

where

$$\beta(t) = \frac{-1}{s(t)} \int_t^{+\infty} \dot{s}(\tau)\,\phi'\left(\int_0^{\tau} \alpha(x)\,dx\right)d\tau \tag{4}$$

The main departure from Yaari comes from $\beta(t)$ which represents the time discounting factor. Note that when ϕ is linear, as in Yaari's case, $\beta(t)$ is constant and can be omitted. In the other cases, however, β is not constant and its shape depends both on the mortality risks (through the survival function $s(t)$) and on the degree of temporal risk, via the function ϕ. When ϕ is concave, β is decreasing, reflecting the fact that the combination of

[1] See Appendix A for calculations.

temporal risk aversion with mortality risks generates time discounting as explained in Bommier (2006).

All our theoretical results will rely on this linear approximation. They should therefore be considered as valid and formally proven only in the limit case where the value of life is infinite. Section 4, which presents the result of numerical estimations, helps getting an idea on how our conclusions are modified if assuming a plausible finite value of life.

We finally assume in the rest of the paper that the utility function $u(.)$ exhibits a constant intertemporal elasticity of substitution. This means that the ratio $cu''(c)/u'(c)$ is assumed to be independent of c.

2.2 Individuals types and preferences properties

Before going further, it is useful to compare both types of individuals preferences properties. We prove the following lemma in the appendix.

Lemma 1 *If individuals mortality patterns satisfy Assumptions A1-A3, then for any times* t, $\varepsilon > 0$, *we have* $\beta^H(t) \leq \beta^L(t)$ *and* $\beta^H(t+\varepsilon)/\beta^H(t) \geq \beta^L(t+\varepsilon)/\beta^L(t)$.

This lemma tells that time discounting generated by mortality profiles are such that individuals of type L value more consumption at any date. Furthermore, the time discount factor decreases at a higher rate for low-survival individuals.

Note finally that Assumptions A1 to A3 imply some monotonicity properties on individuals indifference curves. To see this write the marginal rate of substitution between consumptions at date t and $t+\varepsilon$ for any pair of consumptions $(c(t), c(t+\varepsilon))$. Differentiation of (3) gives:

$$\begin{aligned} MRS^{i}_{c(t),c(t+\varepsilon)} &= \left.\frac{dc(t)}{dc(t+\varepsilon)}\right|_{EU^i} \\ &= -\frac{s^i(t+\varepsilon)\,\alpha(t+\varepsilon)\,\beta^i(t+\varepsilon)\,u'(c(t+\varepsilon))}{s^i(t)\,\alpha(t)\,\beta^i(t)\,u'(c(t))} \end{aligned} \tag{5}$$

Assumption A2 implies $s^L(t+\varepsilon)/s^L(t) < s^H(t+\varepsilon)/s^H(t)$ and Assumption A3 implies $\beta^L(t+\varepsilon)/\beta^L(t) < \beta^H(t+\varepsilon)/\beta^H(t)$ so that $MRS^{H}_{c(t),c(t+\varepsilon)} < MRS^{L}_{c(t),c(t+\varepsilon)}$. In other words, the slope of indifference curve in the $\{c(t+\varepsilon), c(t)\}$ space is less steeper for the type-L individuals.

Another way to illustrate this point, perhaps more insightful, involves comparing what people with different mortality types would do if having the same initial endowment W_0 and having access to a perfect annuity market. Individual i's problem would then be:

$$\begin{aligned} &\max_{c(t)} EU^i(\mathbf{c}) = \int_0^{+\infty} s^i(t)\,\alpha(t)\,\beta^i(t)\,u(c(t))\,dt \\ &\text{s.to} \int_0^{+\infty} \exp\left(\int_0^t -j(\tau)+\mu^i(\tau)\,d\tau\right) c(t)\,dt \le W_0. \end{aligned}$$

The first-order condition would then write:

$$\beta^i(t)\,\alpha(t)\,u'\left(c^i(t)\right) = \lambda \exp\left(-\int_0^t j(\tau)\,d\tau\right) \tag{6}$$

where λ is the Lagrange multiplier associated with the individual's budget constraint, leading to the following result, proved in Appendix C:

Proposition 1 *If agents L and H are provided with the same wealth endowment and have access to actuarially fair annuities, then under Assumptions A1-A3, they choose consumption profiles $c^L(t)$ and $c^H(t)$ such that:*

(a) With temporal risk neutrality (i.e. when ϕ is linear): $\dot{c}^L(t)/c^L(t) = \dot{c}^H(t)/c^H(t)$ *for every* t.

(b) With temporal risk aversion (i.e. when ϕ is concave): $\dot{c}^L(t)/c^L(t) < \dot{c}^H(t)/c^H(t)$ *for every* t.

When temporal risk aversion is introduced, type-H and -L agents choose consumption paths with different growth rates. Indeed, agent H, whose mortality is low, chooses a higher rate of consumption growth (or a lower rate of consumption decline) than agent L. This reflects the relation between mortality and impatience discussed in Bommier (2008). Since agents' optimal strategies are different, we may anticipate that a social planner may be willing to provide different pension levels and different pension profiles to individuals of different types. Moreover, in the case where the type is not observable, the planner may use this heterogeneity of individuals' optimal strategies to make them reveal their type by letting them choose a pension plan among several alternatives. We address these questions below where we discuss the planner's optimal strategy, depending on whether individuals' mortality is private information or not.

3 Government's problem

For the following we assume that there is a utilitarian government whose aim is to maximize the sum of individuals' expected utility functions

$$n^H EU^H\left(\mathbf{c}^H\right) + n^L EU^L\left(\mathbf{c}^L\right)$$

and discuss what it should do depending on the information context that is considered.

3.1 Full information

Assume first that the government can perfectly observe individuals types. In this first-best problem, it maximizes social welfare under the resource constraint of the economy. Its problem is thus:

$$\begin{array}{ll} & \max_{c^H(t),c^L(t)} \; n^H EU^H\left(\mathbf{c}^H\right) + n^L EU^L\left(\mathbf{c}^L\right) \\ \text{s.to} & \int_0^{+\infty} n^H s^H(t) \exp\left(\int_0^t -j\left(\tau\right)d\tau\right) c^H(t)dt \\ + & \int_0^{+\infty} n^L s^L\left(t\right) \exp\left(\int_0^t -j\left(\tau\right)d\tau\right) c^L(t)dt \leq W_0 \end{array}$$

First-order conditions of the first-best problem are:

$$\beta^i\left(t\right)\alpha\left(t\right)u'\left(c^i\left(t\right)\right) - \lambda \exp\left(-\int_0^t j\left(\tau\right)d\tau\right) = 0 \tag{7}$$

for any $i = H, L$ and every t and λ is the Lagrange multiplier associated with the resource constraint. We prove the following proposition in appendix:

Proposition 2 *Under Assumptions A1-A2, the first-best allocation is characterized by:*

(a) With temporal risk neutrality, $c^H\left(t\right) = c^L\left(t\right)\ \forall t$.

(b) With temporal risk aversion,

(i) $c^H\left(t\right) < c^L\left(t\right)$ *for every* t.

(ii) With assumption A3, $\dot{c}^L\left(t\right)/c^L\left(t\right) < \dot{c}^H\left(t\right)/c^H\left(t\right)$ *for every* t.

Under the assumption of temporal risk neutrality, point (a) of Proposition 2 states that the optimum involves providing all individuals with the same consumption profiles. However, as stressed in point *(b)*, when individuals expected utility exhibits temporal risk aversion, the optimum is to offer a higher instantaneous consumption level for the low-survival individuals at all ages. This is because, low-type individuals have on average a lower lifetime felicity and that lifetime utility is concave in lifetime felicity. As in Proposition 1, the consumption level of type-H individuals increases (resp. decreases) at a higher (resp. lower) rate. In case (a), it is clear that there is a positive transfer of (expected) lifetime income from the low- to the high-type individuals. The level of this transfer is measured by $\left(\int_0^{+\infty} n^H s^H(t)c(t)dt - \int_0^{+\infty} n^L s^L(t)c(t)dt\right)$ where $c(t)$ is the optimal consumption profile for both types of individuals. In case (b) however, the sign of the transfer is ambiguous.

3.2 Asymmetric information

Assume now that the government is unable to tell who is of type L and who is of type H. In the following, we analyze two scenarios. In a first section, we study the case where individuals have access to a privately fair annuity market. It turns out that one can always implement the first-best optimum in this case. Then we turn to the case where individuals do not have access to savings.

3.2.1 With a perfect annuity market

When individuals have access to a privately fair annuity market, the government can easily decentralize the first-best allocation described in Proposition 2, by distributing uniform pensions ρ which are financed through a lump-sum tax G_0 at age 0. Indeed, pensions, which are paid contingent on survival, provide a way to redistribute between individuals without observing their type. The government has then just to pick up a level of public pension that would implement the optimal redistribution. The existence of actuarially fair annuities on the market would then allow individuals to optimally smooth consumption. Precisely, denoting W_L^* the present value of type-L individual's lifetime consumption in the first best, it is straightforward to check that ρ and G_0 such that

$$\rho = \frac{W_0 - W_L^*}{n^H \int_0^{+\infty} \left(s^H(t) - s^L(t)\right) \exp\left(-\int_0^t j(\tau)\, d\tau\right) dt} \tag{8}$$

$$G_0 = \rho \int_0^{+\infty} \left(n^H s^H(t) + n^L s^L(t)\right) \exp\left(-\int_0^t j(\tau)\, d\tau\right) dt \tag{9}$$

decentralize the first-best optimum. Whether ρ and G_0 are positive or negative depends on the sign of $W_0 - W_L^*$, that is on whether the government wishes to redistribute life-cycle income from the short-lived individuals toward the long-lived individuals or not. When $W_0 - W_L^*$ is positive, which is the case when individuals exhibit little temporal risk aversion, ρ and G_0 are positive. The first best is then implemented through a simple public pension system, which distributes constant and uniform pensions that are financed by a uniform lump-sum tax G_0 at age 0. It is worth emphasizing that, in such an idealized setting, which assumes perfect markets, the goal of the public pension system is not to help individuals to smooth their consumption - which they can do by purchasing private annuities - but to achieve the optimal redistribution between unobserved types.

For large degrees of temporal risk aversion, $W_0 - W_L^*$ may be negative. The government then wishes to redistribute life-cycle income from the long-lived toward the short-lived individuals. In such a case, ρ and G_0 are negative, and the government can implement the first-best optimum by imposing a uniform tax $-\rho$ at each age t in order to redistribute a lump-sum grant $-G_0$ at age zero.

3.2.2 No private savings

When there are no private savings the government has to directly provide individuals with consumption profiles. The strategy of the government will depend on its ability to commit for the future. Intuitively, a government who is able to commit can propose a menu of pensions which will lead the agents to reveal their types, granting them that this information will not be used against their own interest in the future. In case commitment is impossible, agents will not be willing to reveal their type, and the best the government can do involves providing all agents with the same pensions. We will study both possibilities successively.

Second best when full commitment is possible We first look at the case where the planner can offer different annuity profiles and be credible when making this offer. This means that the government is able to commit itself to paying some pensions, even though it may anticipate that given the information he will get in the future he would prefer to deviate from this plan. The government problem is then a static problem of insurance, with adverse selection, unless agents are temporal risk neutral. With temporal risk aversion, it is clear that individuals of type H would like to mimic individuals of type L in order to get a higher pension at all ages. The government has therefore to add an incentive compatibility constraint stating that type-H individuals do not get a lower utility if they reveal their true type, and its problem rewrites as follows[2]:

$$\begin{aligned}
&\max_{c^H(t),c^L(t)} \; n^H EU^H\left(\mathbf{c}^H\right)+n^L EU^L\left(\mathbf{c}^L\right)\\
\text{s.to}\quad : \;& \int_0^{+\infty} n^H s^H(t)\exp\left(\int_0^t -j(\tau)\,d\tau\right)c^H(t)dt+\\
&\int_0^{+\infty} n^L s^L(t)\exp\left(\int_0^t -j(\tau)\,d\tau\right)c^L(t)dt\leq W_0,\\
&EU^H\left(\mathbf{c}^H\right)\geq EU^H\left(\mathbf{c}^L\right).
\end{aligned}$$

Denoting γ the Lagrange multiplier associated with the incentive compatibility constraint and $\pi(t)=\left[\gamma/n^L\right]\left[s^H(t)\beta^H(t)\right]/\left[s^L(t)\beta^L(t)\right]>0$, the first-order conditions yield:

$$MRS^H_{c(t),c(t+\varepsilon)} = -\exp\left(-\int_t^{t+\varepsilon}\left(j(\tau)+\mu^H(\tau)\right)d\tau\right), \tag{10}$$

$$\begin{aligned}
MRS^L_{c(t),c(t+\varepsilon)} &= -\exp\left(-\int_t^{t+\varepsilon}\left(j(\tau)+\mu^L(\tau)\right)d\tau\right)\\
&\quad\times\left[\frac{1-\pi(t)}{1-\pi(t)\frac{\overline{MRS}^H_{c(t),c(t+\varepsilon)}}{MRS^L_{c(t),c(t+\varepsilon)}}}\right]
\end{aligned} \tag{11}$$

for any t and $t+\varepsilon$ and where $\overline{MRS}^H_{c(t),c(t+\varepsilon)}$ is the marginal rate of substitution of a type-H individual mimicking a type-L individual.

In the first best, we would simply have $MRS^H_{c(t),c(t+\varepsilon)}=-\exp\left(-\int_t^{t+\varepsilon}\left(j(\tau)+\mu^H(\tau)\right)d\tau\right)$ and $MRS^L_{c(t),c(t+\varepsilon)}=-\exp\left(-\int_t^{t+\varepsilon}\left(j(\tau)+\mu^L(\tau)\right)d\tau\right)$. Thus, what we ob-

[2] In some cases, solving such a problem involves taking zero or negative consumption for the low-type individuals after some age, which is of course meaningless. In those cases, it is necessary to consider additional constraints fixing a lower bound for consumption, which may become binding after some age. For simplicity, we shall ignore this point in this section, while we did account for it in our numerical estimations where such constraints were binding at extremely old ages.

tain in (10) for the second best is the usual result of no distortion at the top. This means that the first-best trade-off between two-period consumptions is preserved for the high-survival individual. However, the second-best optimum introduces a distortion in the trade-off between two-period consumptions for the low-type individual. As shown in Section 2.2, $MRS^H_{c(t),c(t+\varepsilon)} < MRS^L_{c(t),c(t+\varepsilon)}$ for the same pair of consumption $\{c(t), c(t+\varepsilon)\}$. Thus

$$\overline{MRS}^H_{c(t),c(t+\varepsilon)} / MRS^L_{c(t),c(t+\varepsilon)} > 1$$

and the expression in brackets of (11) is greater than one. We summarize these results in the following proposition:[3]

Proposition 3 *With Assumptions A1-A3, the second-best allocation is characterized by:*

(i) With temporal risk neutrality, the first-best solution is implementable.

(ii) With temporal risk aversion, the second-best solution is characterized by:

(a) $MRS^H_{c(t),c(t+\epsilon)} = -\left[s^H(t+\varepsilon)/s^H(t)\right]\exp\left(-\int_t^{t+\varepsilon} j(\tau)\,d\tau\right)$ *for any* t *and* $\varepsilon > 0$.

(b) $MRS^L_{c(t),c(t+\epsilon)} < -\left[s^L(t+\varepsilon)/s^L(t)\right]\exp\left(-\int_t^{t+\varepsilon} j(\tau)\,d\tau\right)$ *for any* t *and* $\varepsilon > 0$.

(c) $\left(\dot{c}^L(t)/c^L(t)\right)^{SB} < \left(\dot{c}^L(t)/c^L(t)\right)^{FB} < \left(\dot{c}^H(t)/c^H(t)\right)^{SB} = \left(\dot{c}^H(t)/c^H(t)\right)^{FB}$

where FB *and* SB *stand respectively for the first- and the second-best allocations.*

As argued above, the first-best solution is incentive compatible when individuals preferences exhibit temporal risk neutrality. With positive temporal risk aversion, point *ii(a)* states that the consumption path of individuals of type H is not distorted. This is the usual "no distortion at the top" result.

[3] See the appendix for details.

However, for type-L individuals, the marginal rate of substitution between present and future consumptions is distorted downwardly. In words, it is desirable to encourage early consumption in life as compared to the first-best trade-off. Intuitively, this property can be explained by the fact that type-L individuals have steeper indifference curves in the $\{c(t+\varepsilon), c(t)\}$ space. This implies that, starting from the first-best trade-off, a variation $dc(t+\varepsilon) < 0$ along with a variation $dc(t) = MRS^{L}_{c(t),c(t+\varepsilon)} dc(t+\varepsilon) > 0$ has no first-order effect on the utility of type-L individuals while it decreases the life-cycle utility of type-H individuals. This distortion is thus a way to relax an otherwise binding self-selection constraint. As a result, point *(c)* stresses that the variation rate of consumption of type-L individuals is lower than the one in the first best.

Pooling optimum In the scenario we developed above, the government offers a menu of pension to individuals at date zero. When picking up one particular pension profile an individual reveals his type, an information that the government would like to use in the future, providing pensions that are different from those initially scheduled. Of course, in case the government cannot commit to follow the initial offer, individuals would anticipate this and a separating equilibrium would be impossible to implement. Indeed, in the continuous time case we consider, there is no way to make all individuals reveal their type at a date t if such an information can be freely used at date $t+\varepsilon$ by the government. Any "reward" that could be consumed between time t and $t+\varepsilon$ in exchange of this information would have a negligible impact on lifetime utility when ε tends to zero and would not compensate the agents for providing information on their mortality types.

In absence of commitment device the best a government can do is to provide the same pensions to all individuals. The government's problem can be expressed as:

$$\max_{c(t)} \; n^H EU^H(\mathbf{c}) + n^L EU^L(\mathbf{c})$$
$$\text{s.to} \int_0^{+\infty} \left(n^H s^H(t) + n^L s^L(t)\right) \exp\left(\int_0^t -j(\tau)\, d\tau\right) c(t)dt \leq W_0$$

The first-order condition with respect to $c(t)$ yields:

$$\bar{\beta}(t)\, \alpha(t)\, u'(c(t)) = \lambda \exp\left(\int_0^t -j(\tau)\, d\tau\right) \tag{12}$$

where λ is the Lagrange multiplier associated with the resource constraint and $\bar{\beta}(t) = \left[n^H s^H(t) \beta^H(t) + n^L s^L(t) \beta^L(t)\right] / \left[n^H s^H(t) + n^L s^L(t)\right]$. $\bar{\beta}(t)$ is a weighted sum of the $\beta^i(t)$'s with the weight given on $\beta^i(t)$ by the fraction of individuals of type i surviving at period t: $n^i s^i(t) / \sum_{j=H,L} n^j s^j(t)$. In appendix, we prove that $\dot{\beta}^L(t)/\beta^L(t) < \dot{\bar{\beta}}(t)/\bar{\beta}(t) < \dot{\beta}^H(t)/\beta^H(t)$ for every t. This implies that:

$$\left(\frac{\dot{c}^L(t)}{c^L(t)}\right)^{FB} < \frac{\dot{c}(t)}{c(t)} < \left(\frac{\dot{c}^H(t)}{c^H(t)}\right)^{FB}$$

for every t and where FB stands for the first-best allocation. In other words, the variation rate of the consumption profile in the pooling optimum lies between variation rates obtained in the first-best optimum.

4 Numerical simulations

For this illustrative part, we take a population aged above 60. This corresponds to the case in which individuals are endowed with a certain amount of capital W_0 and decide to annuitize it at the age of 60. Our types of individuals H and L have mortality rates similar to those of women and men according to the year 2000 US life table.[4] We did not choose these gender specific mortality to provide conclusion on gender issues. Actually, gender is generally well observed by the social planner and, therefore, not associated with the problem of asymmetric information that motivated the second-best approach. We only used these male and female mortality because they provide mortality rates and a differential mortality that are of a reasonable order of magnitude. It is assumed that $n^H = n^L = 0.5$.

We further assume that the subjective discount factor is such that $\alpha(x) = 1$ so that agents impatience exclusively arises from the combination of risk aversion and mortality risk, as in the "time neutral model" in Bommier (2008). We use a function $\phi(x) = [1 - \exp(-ax)]$, assuming therefore constant absolute risk aversion with respect to life duration. The parameter a is set to get plausible rates of time discounting. Precisely it has been set so that $-\dot{\beta}^L(65)/\beta^L(65) = 0.03$ per year. Finally, the utility function is given by the isoelastic utility function:

$$u(c) = 1 + \xi \frac{c^{1-\kappa}}{1-\kappa}$$

where $\kappa = 1/0.9$ assuming therefore an intertemporal elasticity of substitution of 0.9. The parameter ξ is the main determinant of the value of life. Our linear approximation which leads to equation (3) involves assuming that ξ is very small, and therefore the value of life very large. This parameter thus needs to be calibrated only for drawing Figure 1, which shows the first best when one does not rely on the additive approximation. Calibration was provided by relying on standard estimates about the Value of Statistical Life, as in Bommier (2008).

The results of the simulations are shown in Figures 1 to 4. Figure 1 shows

[4] Demographic data were taken from the Berkeley Mortality Database.

the first best when computed without making use of the linear approximation. We find that in the first best, short-lived individuals should be provided with higher pensions than long-lived individuals during most of retirement. Though we should not conclude that pensions actually redistribute resources from long-lived toward short-lived individuals. Actually, in the first best, the present value of pensions received by short-lived individuals is 5.5% smaller than the one received by the long-lived individuals because of differential mortality. The social planner is therefore willing to redistribute from short-lived individuals to long-lived individuals, though significantly less than what he would do if he were constrained to provide all individuals with the same pension. As a comparison, in the case of a pooling strategy (as in Figure 4), the present value of pensions received by short-lived individuals would be 12.3% smaller than the one received by long-lived individuals. It is worth noticing that the consumption profiles shown in Figure 1 intersect at some point (at age 84.3 precisely). This indicates that the theoretical result that $c^L(t) > c^H(t)$ which was obtained when considering an infinite value of life, does not extend to the case of a finite value of life. The conclusion about the consumption growth rates $(\dot{c}^L(t)/c^L(t) < \dot{c}^H(t)/c^H(t))$ is, however, not challenged by the result of our simulation with a finite value of life.

Figure 2 draws the pension profile we would obtain in the first best when using the linear approximation. The general pattern is comparable to that obtained in Figure 1 though we find that relying on the linear approximation leads to provide even larger pensions to short-lived individuals and smaller pensions to long-lived individuals. This is quite intuitive: the linear estimation involves assuming that the value of life is infinite and therefore overemphasizes the difference in lifetime welfare resulting from heterogeneity in life duration. For this reason, it overestimates the difference in marginal utilities between long-lived and short-lived individuals, ending up suggesting too different pension levels. The lack of precision is not completely negligible. Though, this linear approximation remains useful to provide the main intuition and a reasonable order of magnitude.

Figure 3 looks at the second best when the social planner can commit to provide some pension in exchange of the information about mortality types, and was computed relying on the linear approximation. Conform with our theoretical results the social planner provides short-lived individuals with more rapidly declining pensions than that of long-lived individuals. Long-lived individuals are indifferent between both pension profiles, while short-lived individuals are strictly happier with the more rapidly declining profile. Compared to the first-best optimum, long-lived individuals are better off and short-lived individuals are worse off. The present value of pensions received by short-lived individuals is 11% lower than that of long-lived individuals.

Lastly, in Figure 4, we look at the best pooling strategy. The government distributes the same pensions to all individuals. The social welfare is then lower than in the separating equilibrium shown in Figure 3, but one does not need to have commitment abilities to implement such a strategy. Long-lived individuals are better off in that pooling equilibrium than in the separating one, while it is the reverse for short-lived individuals. This pooling optimum is in fact the one that implements the greater redistribution from short-lived individuals to long-lived individuals and therefore the one that leads to the greater inequality in welfare.

5 Conclusion

This paper has studied the problem of redistribution between individuals differing in their survival probabilities. We have compared successively utilitarian allocations when individuals are either temporal risk neutral or temporal risk averse. In a first-best setting, we find, in the limit case where the value of life tends to infinity, that if individuals are temporal risk averse, long-lived individuals should have a lower instantaneous consumption than short-lived individuals. Conversely, with temporal risk neutrality, the pooling allocation is socially optimal.

When mortality type is a private information, the optimal government's strategy depends on the characteristics of the private markets and on the government's capacity to commit. We considered three polar cases.

If credit and insurance markets are perfect, the first best can be implemented with the use of uniform public pension system and lump-sum grant. Indeed, due to differential mortality, such a uniform pension system generates some redistribution between mortality types. Individuals can then rely on the private market for annuities to smooth their consumption.

In the absence of credit market and with asymmetric information, the first best can no longer be achieved (unless individuals exhibit temporal risk neutrality). If the government can commit to distribute in the future some given pension profiles in exchange of information about mortality, the second best is a separating equilibrium, which is obtained by offering pension profiles that vary with age differently. In absence of commitment possibilities, the government offers the same pension to everybody.

In a numerical illustration, based on realistic mortality rates, we computed the different strategies, and discussed their redistributive aspects. We found that the pooling strategy, without private savings, is actually the one which generates the largest inequalities (in terms of lifetime utilities). Alternative strategies, which would either (partially) rely on private savings or on the offer of a menu of pensions would be more progressive.

Of course, to gain in relevance the message should be refined by considering imperfect markets in order to be closer to what is observed in reality. Though, the main messages, would remain the same: (1) Differential mortality generates inequality in lifetime utilities which should be taken into account as soon as we allow for temporal risk aversion. (2) Public pension is a way to redistribute resources between individuals with different mortality patterns. (3) Redistribution should not only be achieved by playing with the level of public pension, but also by offering pensions (or menu of pensions) which vary with age. It occurs that this latter possibility is little used in practice, and we see it as an interesting avenue for future research and policy changes.

For simplicity, our paper focused on heterogeneity in mortality exclusively and did not consider wealth inequality. This was sufficient to illustrate the role of temporal risk aversion and why it suggests new policy instruments. But of course inequality in wealth is a central aspect which should be taken into account in future works. Some authors (e.g. Brunner et al., 2008 or Cremer et al., 2010) study the design of pension systems when agents differ not only in their mortality pattern but also in their income. These studies rely however on the assumption of temporal risk neutrality. With a negative correlation between mortality and wealth, these authors have shown that the design of redistributive pension systems implies a trade-off between the redistribution from high- to low-income and from short- to long-lived individuals. If one takes into account temporal risk aversion, our results show that this trade-off can be mitigated (or even eliminated) in favour of more redistribution from high- to low-income individuals.

Also, the paper is restricted to issues that are related to the annuitization of wealth at the retirement age. Allowing for an explicit account of workers is clearly a priority on our research agenda. With this additional dimension, we plan to study issues related to the formation of wealth and intergenerational redistribution but also the role of labor income taxation when individuals choose the age at which they withdraw pension benefits.

References

[1] Barro, R. J, and J. W. Friedman, 1977, “On Uncertain Lifetimes”. *The Journal of Political Economy*, 85(4):843-849.

[2] Berkeley Mortality Database, http://www.demog.berkeley.edu/~bmd/states.html.

[3] Bommier, A., 2006. "Uncertain lifetime and intertemporal choice: risk aversion as a rationale for time discounting". International Economic Review 47(4), 1223-1246.

[4] Bommier, A., 2008, "Rational Impatience". Working Paper university of Toulouse.

[5] Brown, J. R., Liebman J. B., and J. Pollet, 2002, "Estimating Life Tables That Reflect Socioeconomic Differences in Mortality", in *The Distributional Aspects of Social Security and Social Security Reform*, M. Feldstein and J. B. Liebman, eds, The University of Chicago Press.

[6] Brunner, J. K and S. Pech, 2008, "Optimum taxation of life annuities", *Social Choice and Welfare,* 30 , 285–303.

[7] Cremer, H., J.M Lozachmeur and P. Pestieau, 2010. "Collective annuities and redistribution", *Journal of Public Economic Theory*, 12 (1), 23–41.

[8] Kihlstrom, R.E. and L.J. Mirman, 1974, "Risk Aversion with many Commodities". *Journal of Economic Theory* 8:361-88.

[9] Levhari, D. and L. J. Mirman, 1977, "Savings and Consumption with an Uncertain Horizon". *The Journal of Political Economy*, 85(2):265-281.

[10] Sheshinski, E., 2007. "Utilitarian pricing of annuities", Chapter 13 in "The economic theory of annuities". Princeton University Press.

[11] Yaari, M., 1965. "Uncertain lifetime, life insurance and the theory of the consumer". The Review of Economic Studies, 32, 2,137-150.

Appendix

A Expected lifetime utility

To be able to get back to a simple additive specification, we make the assumption of a "priceless life context". As defined in Bommier (2006), it corresponds to a situation where $u(c(t)) = 1 + \lambda\omega(c(t))$ where λ is a (small) scalar and $\omega(.)$ is bounded. The lifetime expected utility function can then be rewritten as:

$$\begin{aligned} EU(\mathbf{c}) &= \int_0^{+\infty} s(t)\alpha(t)\,\phi'\left(\int_0^t \alpha(x)\left(1+\lambda\omega(c(x))\right)dx\right)dt \\ &+\lambda\int_0^{+\infty} s(t)\alpha(t)\,\omega(c(t))\,\phi'\left(\int_0^t \alpha(x)\left(1+\lambda\omega(c(x))\right)dx\right)dt. \end{aligned}$$

We assume that $\lambda \to 0$ which means that the individual would agree to give up most of his consumption to live longer. Taking the Taylor expansion of the function ϕ' and keeping only the terms of order zero and one in λ, this yields:

$$\begin{aligned} EU(\mathbf{c}) &= \int_0^{+\infty} s(t)\alpha(t)\,\phi'\left(\int_0^t \alpha(x)\,dx\right)dt \\ &+\lambda\int_0^{+\infty} s(t)\alpha(t)\,\omega(c(t))\,\phi'\left(\int_0^t \alpha(x)\,dx\right)dt \\ &+\lambda\int_0^{+\infty} s(t)\alpha(t)\left(\int_0^t \alpha(x)\,\omega(c(x))\,dx\right)\phi''\left(\int_0^t \alpha(x)\,dx\right)dt. \end{aligned}$$

Denoting the constant $\Psi = \int_0^{+\infty} s(t)\alpha(t)\,\phi'\left(\int_0^t \alpha(x)\,dx\right)dt$ and switching the order of integration of the third term, the expected utility function can be approximated by:

$$
\begin{aligned}
EU(\mathbf{c}) &\approx \Psi + \lambda \int_0^{+\infty} s(t)\alpha(t)\,\omega(c(t))\,\phi'\left(\int_0^t \alpha(x)\,dx\right) dt \\
&\quad + \lambda \int_0^{+\infty} \alpha(t)\,\omega(c(t)) \left(\int_t^{++\infty} s(x)\,\alpha(x)\,\phi"\left(\int_0^x \alpha(\tau)\,d\tau\right) dx\right) dt \\
&\approx \Psi + \lambda \int_0^{+\infty} \alpha(t)\,\omega(c(t)) \begin{bmatrix} s(t)\phi'\left(\int_0^t \alpha(x)\,dx\right) + \\ \int_t^{+\infty} s(x)\,\alpha(x)\,\phi"\left(\int_0^x \alpha(\tau)\,d\tau\right) dx \end{bmatrix} dt.
\end{aligned}
$$

Integrating by part the term in brackets yields:

$$
EU(\mathbf{c}) \approx \Psi + \lambda \int_0^{+\infty} s(t)\alpha(t)\,\omega(c(t))\,\beta(t)\,dt
$$

where $\beta(t) = \int_t^{+\infty} -(\dot{s}(\tau)/s(t))\,\phi'\left(\int_0^\tau \alpha(x)\,dx\right) d\tau$. Using $\omega(c(t)) = u(c(t)) - 1/\lambda$ and forgetting the constant, the expected lifetime utility can thus be approximated by the following additive utility function:

$$
EU(\mathbf{c}) = \int_0^{+\infty} s(t)\,\alpha(t)\,\beta(t)\,u(c(t))\,dt.
$$

Finally, denoting respectively $\dot{\beta}(t)$ and $\dot{s}(t)$ the derivatives of $\beta(t)$ and $s(t)$ with respect to t yields:

$$
\begin{aligned}
\dot{\beta}(t) &= \frac{\dot{s}(t)}{s(t)^2} \int_t^{+\infty} \dot{s}(\tau)\,\phi'\left(\int_0^\tau \alpha(x)\,dx\right) d\tau + \frac{\dot{s}(t)}{s(t)}\phi'\left(\int_0^t \alpha(x)\,dx\right) \\
&= -\frac{\dot{s}(t)}{s(t)} \int_t^{+\infty} \frac{s(\tau)}{s(t)}\phi"\left(\int_0^\tau \alpha(x)\,dx\right) d\tau \qquad (13)
\end{aligned}
$$

where we made use of integration by-part.

B Proof of Lemma 1

Let's denote $\Delta(t) = -\dot{\beta}^L(t)/\beta^L(t) - \left(-\dot{\beta}^H(t)/\beta^H(t)\right)$. We want to show that this term is positive for any t. Using the definition of $\mu(t)$, one has:

$$\Delta(t) = \frac{\mu^L(t)\int_t^{+\infty}\frac{s^L(\tau)}{s^L(t)}\left(-\phi"\left(\int_0^\tau \alpha(x)\,dx\right)\right)d\tau}{\int_t^{+\infty}\mu^L(\tau)\frac{s^L(\tau)}{s^L(t)}\phi'\left(\int_0^\tau \alpha(x)\,dx\right)d\tau} - \frac{\mu^H(t)\int_t^{+\infty}\frac{s^H(\tau)}{s^H(t)}\left(-\phi"\left(\int_0^\tau \alpha(x)\,dx\right)\right)d\tau}{\int_t^{+\infty}\mu^H(\tau)\frac{s^H(\tau)}{s^H(t)}\phi'\left(\int_0^\tau \alpha(x)\,dx\right)d\tau}.$$

Using Assumption A3 and $\phi" < 0$, we thus have the following inequality:

$$\Delta(t) \geq \mu^L(t)\left[\frac{\int_t^{+\infty}\frac{s^L(\tau)}{s^L(t)}\left(-\phi"\left(\int_0^\tau \alpha(x)\,dx\right)\right)d\tau}{\int_t^{+\infty}\mu^L(\tau)\frac{s^L(\tau)}{s^L(t)}\phi'\left(\int_0^\tau \alpha(x)\,dx\right)d\tau} - \frac{\int_t^{+\infty}\frac{s^H(\tau)}{s^H(t)}\left(-\phi"\left(\int_0^\tau \alpha(x)\,dx\right)\right)d\tau}{\int_t^{+\infty}\mu^L(\tau)\frac{s^H(\tau)}{s^H(t)}\phi'\left(\int_0^\tau \alpha(x)\,dx\right)d\tau}\right]. \tag{14}$$

Using the following notations:

$$\begin{aligned}
g(\tau) &= \mu^L(\tau)\frac{s^L(\tau)}{s^L(t)}\phi'\left(\int_0^\tau \alpha(x)\,dx\right), \\
k(\tau) &= -\frac{\phi"\left(\int_0^\tau \alpha(x)\,dx\right)}{\phi'\left(\int_0^\tau \alpha(x)\,dx\right)}\frac{1}{\mu^L(\tau)}, \\
h(\tau) &= \frac{s^H(\tau)/s^H(t)}{s^L(\tau)/s^L(t)},
\end{aligned}$$

the inequality (14) can be rewritten as:

$$\Delta(t) \geq \mu^L(t)\left[\frac{\int_t^{+\infty} g(\tau)\,k(\tau)\,d\tau}{\int_t^{+\infty} g(\tau)\,d\tau} - \frac{\int_t^{+\infty} g(\tau)\,k(\tau)\,h(\tau)\,d\tau}{\int_t^{+\infty} g(\tau)\,h(\tau)\,d\tau}\right]$$

where the functions $g(.)$, $k(.)$ and $h(.)$ are non negative. Rearranging the terms in brackets yields:

$$\Delta(t) \geq \mu^L(t) \left[\frac{\int_t^{+\infty} g(\tau) k(\tau) d\tau \int_t^{+\infty} g(\tau) h(\tau) d\tau - \int_t^{+\infty} g(\tau) k(\tau) h(\tau) d\tau \int_t^{+\infty} g(\tau) d\tau}{\int_t^{+\infty} g(\tau) d\tau \int_t^{+\infty} g(\tau) h(\tau) d\tau} \right]$$

where the denominator in brackets is positive. Define the function $f(x) = \int_t^x g(\tau) k(\tau) d\tau \int_t^x g(\tau) h(\tau) d\tau - \int_t^x g(\tau) k(\tau) h(\tau) d\tau \int_t^x g(\tau) d\tau$. By Assumption A2, h is non decreasing, k is non increasing since $-\phi"/\phi'$ is non increasing and $\mu^L(\tau)$ is increasing. This implies that $f(x)$ is non decreasing with x and therefore non negative for any $x \geqslant t$. Then $\Delta(t)$ is positive for any t which proves the result.

C Proof of Proposition 1

The proof is similar to the one provided for Proposition 7 in Bommier (2008). Differentiating (6) with respect to t yields:

$$\begin{aligned} & \dot{\alpha}(t) \beta^i(t) u'\left(c^i(t)\right) + \alpha(t) \left[\dot{c}^i(t) u''\left(c^i(t)\right) \beta^i(t) + u'\left(c^i(t)\right) \dot{\beta}^i(t)\right] \\ = & -\lambda \exp\left(-\int_0^t j(\tau) d\tau\right) j(t) \end{aligned}$$

which, after some manipulation, gives:

$$\frac{\dot{c}^i(t)}{c^i(t)} = -\frac{u'(c^i(t))}{u''(c^i(t)) c^i(t)} \left[j(t) + \frac{\dot{\alpha}(t)}{\alpha(t)} + \frac{\dot{\beta}^i(t)}{\beta^i(t)} \right]$$

where $u'(c^i(t)) / u''(c^i(t)) c^i(t)$ is a constant by assumption.

(a) With temporal risk neutrality, $\beta^i(t)$ is equal to a constant so that $-\dot{c}^i(t)/c^i(t) =$

$j(t) + \dot{\alpha}(t) / \alpha(t)$ for $i = H, L$.

(b) With temporal risk aversion, Lemma 1 implies $\dot{c}^L(t) / c^L(t) < \dot{c}^H(t) / c^H(t)$.

D Proof of Proposition 2

Differentiation of (7) with respect to t yields:

$$\dot{\alpha}^i(t) \left[u'\left(c^i(t)\right) \beta^i(t)\right] + \alpha(t) \left[\begin{array}{c} \dot{c}^i(t)\, u''\left(c^i(t)\right) \beta^i(t) \\ + u'\left(c^i(t)\right) \dot{\beta}^i(t) \end{array}\right] = -\lambda \exp\left(-\int_0^t j(\tau)\, d\tau\right) j(t)$$

which after some manipulation gives:

$$\frac{\dot{c}^i(t)}{c^i(t)} = -\frac{u'\left(c^i(t)\right)}{u''\left(c^i(t)\right) c^i(t)} \left[j(t) + \frac{\dot{\alpha}(t)}{\alpha(t)} + \frac{\dot{\beta}^i(t)}{\beta^i(t)}\right] \tag{15}$$

where $u'\left(c^i(t)\right) / u''\left(c^i(t)\right) c^i(t)$ is a constant by assumption.

(a) With temporal risk neutrality, $\beta^i(t)$ is equal to a constant so that (7) implies $c^i(t) = c(t)$ for every t and $i = H, L$.

(b) With temporal risk aversion, (7) implies $c^H(t) < c^L(t)$. Equation (15) and Lemma 1 imply $\dot{c}^L(t) / c^L(t) < \dot{c}^H(t) / c^H(t)$.

E Second-best optimum

E.1 Proof of point (ii) of Proposition 3

First-order conditions of the second-best problem are:

$$\frac{\partial EU^H}{\partial c^H(t)}\left(1+\frac{\gamma}{n^H}\right)-\lambda s^H(t)\exp\left(-\int_0^t j(\tau)\,d\tau\right)=0, \tag{16}$$

$$\frac{\partial EU^L}{\partial c^L(t)}-\lambda s^L(t)\exp\left(-\int_0^t j(\tau)\,d\tau\right)-\frac{\gamma}{n^L}\frac{\partial EU^H}{\partial c^L(t)}=0. \tag{17}$$

(a) Straightforward rearrangement of (16) taken at time t and $t+\epsilon$ gives the marginal rate of substitution between two-period consumptions (10) for the high-type individual which proves point (a).

(b) Denoting $EU^j_{c(t)}$ the expected marginal utility of consumption at date t and evaluating (17) at time t and $t+\epsilon$, we get:

$$\frac{EU^L_{c(t+\varepsilon)}}{EU^L_{c(t)}}\left[1-\frac{\gamma}{n^L}\frac{\overline{EU}^H_{c(t+\varepsilon)}}{EU^L_{c(t)}}\frac{EU^L_{c(t)}}{EU^L_{c(t+\varepsilon)}}\right]=$$
$$\frac{s^L(t+\varepsilon)}{s^L(t)}\exp\left(-\int_0^t j(\tau)\,d\tau\right)\times\left[1-\frac{\gamma}{n^L}\frac{\overline{EU}^H_{c(t)}}{EU^L_{c(t)}}\right]$$

where $\overline{EU}^H_{c(t)}$ is the expected marginal utility of a type-H individual mimicking a type-L individual. Multiplying the second term in brackets of the LHS by $\overline{EU}^H_{c(t)}/\overline{EU}^H_{c(t)}$ yields:

$$\frac{EU^L_{c(t+\varepsilon)}}{EU^L_{c(t)}}\left[1-\frac{\gamma}{n^L}\frac{\overline{EU}^H_{c(t+\varepsilon)}}{EU^L_{c(t)}}\frac{EU^L_{c(t)}}{EU^L_{c(t+\varepsilon)}}\frac{\overline{EU}^H_{c(t)}}{\overline{EU}^H_{c(t)}}\right]=$$
$$\frac{s^L(t+\varepsilon)}{s^L(t)}\exp\left(-\int_0^t j\left(\tau\right)d\tau\right)\times\left[1-\frac{\gamma}{n^L}\frac{\overline{EU}^H_{c(t)}}{EU^L_{c(t)}}\right].$$

This can be rewritten as:

$$MRS^L_{c(t),c(t+\varepsilon)}\left[1-\frac{\gamma}{n^L}\frac{\overline{MRS}^H_{c(t),c(t+\varepsilon)}}{MRS^L_{c(t),c(t+\varepsilon)}}\frac{s^H\left(t\right)\beta^H\left(t\right)}{s^L\left(t\right)\beta^L\left(t\right)}\right]=$$
$$\frac{s^L(t+\varepsilon)}{s^L(t)}\exp\left(-\int_0^t j\left(\tau\right)d\tau\right)\times\left[1-\frac{\gamma}{n^L}\frac{s^H\left(t\right)\beta^H\left(t\right)}{s^L\left(t\right)\beta^L\left(t\right)}\right]$$

which yields (11).

(c) The first-order conditions (16) and (17) can be rewritten as:

$$\alpha\left(t\right)u'\left(c^H\left(t\right)\right)\beta^H\left(t\right)\left(1+\frac{\gamma}{n^H}\right)=\lambda\exp\left(-\int_0^t j\left(\tau\right)d\tau\right), \tag{18}$$
$$\alpha\left(t\right)u'\left(c^L\left(t\right)\right)\left[\beta^L\left(t\right)-\frac{\gamma}{n^L}\frac{s^H\left(t\right)}{s^L\left(t\right)}\beta^H\left(t\right)\right]=\lambda\exp\left(-\int_0^t j\left(\tau\right)d\tau\right). \tag{19}$$

Differentiating (18) with respect to time yields $\dot{c}^H\left(t\right)/c^H\left(t\right)$ to be the same as in (15) whereas differentiation of (19) yields:

$$\begin{aligned}
\frac{\dot{c}^L(t)}{c^L(t)} &= -\frac{u'\left(c^L(t)\right)}{c^L(t)\,u''\left(c^L(t)\right)} \times \\
&\left(j(t) + \frac{\dot{\alpha}(t)}{\alpha(t)} + \frac{\dot{\beta}^L(t) - \gamma\left(\frac{s^H(t)}{s^L(t)}\dot{\beta}^H(t) + \left(\frac{\dot{s}^H(t)}{s^L(t)} - \frac{\dot{s}^L(t)}{s^L(t)}\frac{s^H(t)}{s^L(t)}\right)\beta^H(t)\right)}{\beta^L(t) - \gamma\frac{s^H(t)}{s^L(t)}\beta^H(t)}\right) \\
= \; & -\frac{u'\left(c^L(t)\right)}{c^L(t)\,u''\left(c^L(t)\right)} \times \\
&\left(j(t) + \frac{\dot{\alpha}(t)}{\alpha(t)}. + \frac{\dot{\beta}^L(t)}{\beta^L(t)}\left[\frac{1 - \gamma\left(\frac{s^H(t)}{s^L(t)}\frac{\dot{\beta}^H(t)}{\dot{\beta}^L(t)} + \left(\frac{\dot{s}^H(t)}{s^L(t)} - \frac{\dot{s}^L(t)}{s^L(t)}\frac{s^H(t)}{s^L(t)}\right)\frac{\beta^H(t)}{\dot{\beta}^L(t)}\right)}{1 - \gamma\frac{s^H(t)}{s^L(t)}\frac{\beta^H(t)}{\beta^L(t)}}\right]\right)
\end{aligned} \tag{20}$$

Note that by Assumption A2 and Lemma 1, one has:

$$\frac{\dot{s}^H(t)}{s^H(t)} + \frac{\dot{\beta}^H(t)}{\beta^H(t)} > \frac{\dot{\beta}^L(t)}{\beta^L(t)} + \frac{\dot{s}^L(t)}{s^L(t)}$$

so that

$$\left(\frac{s^H(t)}{s^L(t)}\frac{\dot{\beta}^H(t)}{\dot{\beta}^L(t)} + \left(\frac{\dot{s}^H(t)}{s^L(t)} - \frac{\dot{s}^L(t)}{s^L(t)}\frac{s^H(t)}{s^L(t)}\right)\frac{\beta^H(t)}{\dot{\beta}^L(t)}\right) > \frac{s^H(t)}{s^L(t)}\frac{\beta^H(t)}{\beta^L(t)}$$

which implies

$$\frac{1 - \gamma\left(\frac{s^H(t)}{s^L(t)}\frac{\dot{\beta}^H(t)}{\dot{\beta}^L(t)} + \left(\frac{\dot{s}^H(t)}{s^L(t)} - \frac{\dot{s}^L(t)}{s^L(t)}\frac{s^H(t)}{s^L(t)}\right)\frac{\beta^H(t)}{\dot{\beta}^L(t)}\right)}{1 - \gamma\frac{s^H(t)}{s^L(t)}\frac{\beta^H(t)}{\beta^L(t)}} > 1.$$

Thus, comparing (20) with (15) yields $\left(\dot{c}^L(t)/c^L(t)\right)^{SB} < \left(\dot{c}^L(t)/c^L(t)\right)^{FB}$.

E.2 The pooling optimum

Differentiating (12) with respect to t yields

$$\frac{\dot{c}(t)}{c(t)} = -\frac{u'(c(t))}{u''(c(t))\,c(t)}\left[j(t) + \frac{\overset{\bullet}{\bar{\beta}}(t)}{\bar{\beta}(t)} + \frac{\dot{\alpha}(t)}{\alpha(t)}\right] \tag{21}$$

where

$$\frac{\overset{\bullet}{\bar{\beta}}(t)}{\bar{\beta}(t)} = \frac{\begin{array}{l}\Big\{\Big[n^H\left(\dot{s}^H(t)\,\beta^H(t) + s^H(t)\,\dot{\beta}^H(t)\right) \\ +n^L\left(\dot{s}^L(t)\,\beta^L(t) + s^L(t)\,\dot{\beta}^L(t)\right)\Big]\left[n^H s^H(t) + n^L s^L(t)\right] \\ -\left[n^H s^H(t)\,\beta^H(t) + n^L s^L(t)\,\beta^L(t)\right]\left[n^H \dot{s}^H(t) + n^L \dot{s}^L(t)\right]\Big\}\end{array}}{\left[n^H s^H(t) + n^L s^L(t)\right]\left[n^H s^H(t)\,\beta^H(t) + n^L s^L(t)\,\beta^L(t)\right]}$$

Developing the numerator in the above expression and rearranging terms yields:

$$\frac{\overset{\bullet}{\bar{\beta}}(t)}{\bar{\beta}(t)} = \frac{\begin{array}{l}\Big\{\left(n^H s^H(t)\,\dot{\beta}^H(t) + n^L s^L(t)\,\dot{\beta}^L(t)\right)\left(n^H s^H(t) + n^L s^L(t)\right) \\ +n^H n^L\left(\beta^H(t) - \beta^L(t)\right)\left(\dot{s}^H(t)\,s^L(t) - \dot{s}^L(t)\,s^H(t)\right)\Big\}\end{array}}{\left[n^H s^H(t) + n^L s^L(t)\right]\left[n^H s^H(t)\,\beta^H(t) + n^L s^L(t)\,\beta^L(t)\right]}$$

We want to compare $\overset{\bullet}{\bar{\beta}}(t)/\bar{\beta}(t)$ with $\dot{\beta}^H(t)/\beta^H(t)$ and $\dot{\beta}^L(t)/\beta^L(t)$.

Let first denote Λ, the difference between $\overset{\bullet}{\bar{\beta}}(t)/\bar{\beta}(t)$ and $\dot{\beta}^H(t)/\beta^H(t)$.

It has the following expression

$$\Lambda = \frac{\begin{array}{c}\left\{n^L s^L(t)\left(n^H s^H(t)+n^L s^L(t)\right)\left(\dot{\beta}^L(t)-\beta^L(t)\frac{\dot{\beta}^H(t)}{\beta^H(t)}\right)\right.\\ \left.+n^H n^L\left(\beta^H(t)-\beta^L(t)\right)\left(\dot{s}^H(t)s^L(t)-\dot{s}^L(t)s^H(t)\right)\right\}\end{array}}{\left[n^H s^H(t)+n^L s^L(t)\right]\left[n^H s^H(t)\beta^H(t)+n^L s^L(t)\beta^L(t)\right]}$$

which is always negative by Assumption A1 and Lemma 1. This implies $\overset{\bullet}{\bar{\beta}}(t)/\bar{\beta}(t) < \dot{\beta}^H(t)/\beta^H(t)$.

Let us now denote Υ the difference between $\overset{\bullet}{\bar{\beta}}(t)/\bar{\beta}(t)$ and $\dot{\beta}^L(t)/\beta^L(t)$. It yields:

$$\Upsilon = \frac{\begin{array}{r}n^H s^H(t)\left\{\left(n^H s^H(t)+n^L s^L(t)\right)\left(\dot{\beta}^H(t)-\beta^H(t)\frac{\dot{\beta}^L(t)}{\beta^L(t)}\right)\right.\\ \left.+n^L s^L(t)\left(\beta^H(t)-\beta^L(t)\right)\left(\frac{\dot{s}^H(t)}{s^H(t)}-\frac{\dot{s}^L(t)}{s^L(t)}\right)\right\}\end{array}}{\left[n^H s^H(t)+n^L s^L(t)\right]\left[n^H s^H(t)\beta^H(t)+n^L s^L(t)\beta^L(t)\right]}$$

where the first part in the numerator is positive while the second one is negative. Equivalently,

$$\begin{aligned}\Upsilon = \chi\Big\{&\mu^H(t)\left(\frac{\dot{\beta}^H(t)}{\mu^H(t)}\left(n^H s^H(t)+n^L s^L(t)\right)+n^L s^L(t)\left[\beta^L(t)-\beta^H(t)\right]\right)\\ &-\mu^L(t)\left(\frac{\beta^H(t)}{\beta^L(t)}\frac{\dot{\beta}^L(t)}{\mu^L(t)}\left(n^H s^H(t)+n^L s^L(t)\right)+n^L s^L(t)\left[\beta^L(t)-\beta^H(t)\right]\right)\Big\}\end{aligned}$$

with $\chi = n^H s^H(t)/\left[\left(n^H s^H(t)+n^L s^L(t)\right)\left(n^H s^H(t)\beta^H(t)+n^L s^L(t)\beta^L(t)\right)\right] > 0$. First note that, using equation (13), $\dot{\beta}^i(t)$ can be rewritten as

$$\dot{\beta}^i(t) = \mu^i(t)\left(\beta^i(t)-\phi'\left(\int_0^t \alpha(x)\,dx\right)\right)$$

where $\beta^i(t) - \phi'\left(\int_0^t \alpha(x)\,dx\right) < 0$. Using this expression and rearranging terms, this yields:

$$\frac{\dot{\beta}^H(t)}{\mu^H(t)}\left(n^H s^H(t) + n^L s^L(t)\right) + n^L s^L(t)\left[\beta^L(t) - \beta^H(t)\right] =$$
$$n^H s^H(t)\left(\beta^H(t) - \phi'\left(\int_0^t \alpha(x)\,dx\right)\right) + n^L s^L(t)\left(\beta^L(t) - \phi'\left(\int_0^t \alpha(x)\,dx\right)\right)$$

which is always negative. Using Assumption A1 and Lemma 1, one also has

$$\begin{aligned} & \frac{\beta^H(t)}{\beta^L(t)}\frac{\dot{\beta}^L(t)}{\mu^L(t)}\left(n^H s^H(t) + n^L s^L(t)\right) + n^L s^L(t)\left[\beta^L(t) - \beta^H(t)\right] \\ < \; & \frac{\dot{\beta}^H(t)}{\mu^H(t)}\left(n^H s^H(t) + n^L s^L(t)\right) + n^L s^L(t)\left[\beta^L(t) - \beta^H(t)\right] < 0 \end{aligned}$$

so that $\Upsilon > 0$ and $\dot{\beta}^L(t)/\beta^L(t) < \dot{\bar{\beta}}(t)/\bar{\beta}(t)$.

Using expressions (15) and $\dot{\beta}^L(t)/\beta^L(t) < \dot{\bar{\beta}}/\bar{\beta}(t) < \dot{\beta}^H(t)/\beta^H(t)$ yields:

$$\frac{\dot{c}^L(t)}{c^L(t)}^{FB} < \frac{\dot{c}(t)}{c(t)} < \frac{\dot{c}^H(t)}{c^H(t)}^{FB}$$

which proves the result.

Figure 1: First Best

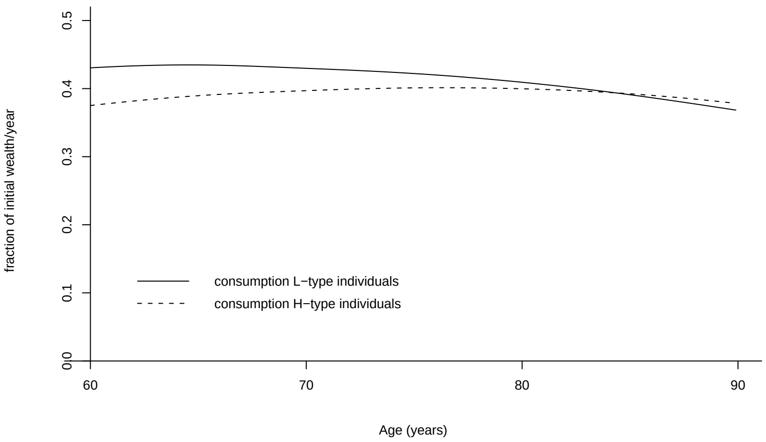


Figure 2: First Best − Linear Approximation

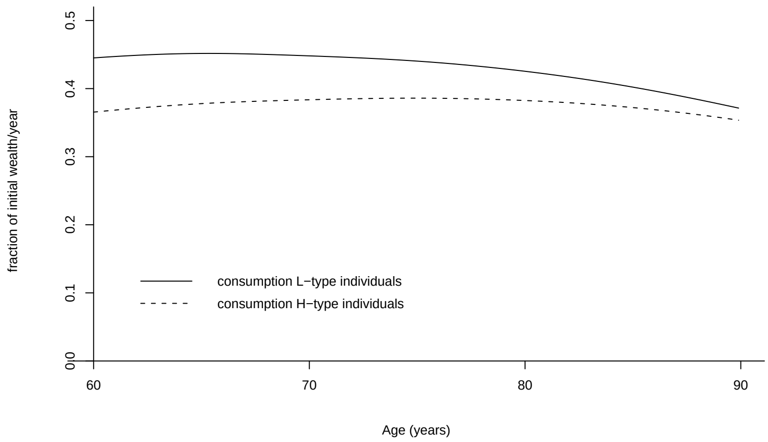


Figure 3: Second Best

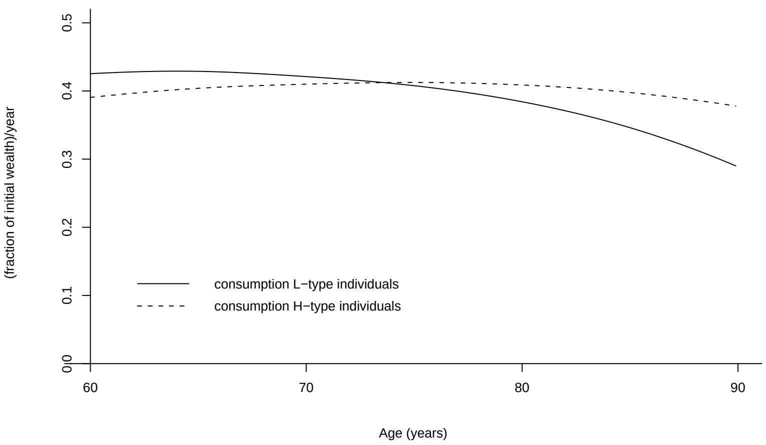
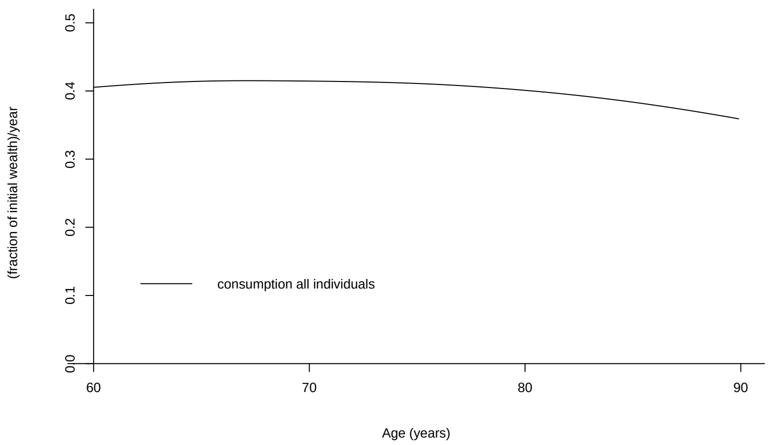


Figure 4: Pooling Optimum



Working Papers of the Center of Economic Research at ETH Zurich

(PDF-files of the Working Papers can be downloaded at www.cer.ethz.ch/research).

10/134 A. Bommier, A. Chassagnon and F. Le Grand
Comparative Risk Aversion: A Formal Approach with Applications to Saving Behaviors

10/133 A. Bommier and B. Villeneuve
Risk Aversion and the Value of Risk to Life

10/132 L. Bretschger and S. Valente
Endogenous Growth, Asymmetric Trade and Resource Taxation

10/131 H. Gersbach and N. Surulescu
Default Risk in Stochastic Volatility Models

10/130 F. Schwark
Economics of Endogenous Technical Change in CGE Models - The Role of Gains from Specialization

10/129 L. Bretschger, R. Ramer and F. Schwark
Long-Run Effects of Post-Kyoto Policies: Applying a Fully Dynamic CGE model with Heterogeneous Capital

10/128 M. T. Schneider, C. Traeger and R. Winkler
Trading Off Generations: Infinitely-Lived Agent Versus OLG

10/127 V. Kappel
The Effects of Financial Development on Income Inequality and Poverty

10/126 M. T. Schneider
The Larger the Better? The Role of Interest-Group Size in Legislative Lobbying

10/125 A. Ziegler
Individual Characteristics and Stated Preferences for Alternative Energy Sources and Propulsion Technologies in Vehicles: A Discrete Choice Analysis

10/124 P. F. Peretto and S. Valente
Resource Wealth, Innovation and Growth in the Global Economy

09/123 H. Gersbach and M. T. Schneider
Tax Contracts and Elections

09/122 V. Hahn
Why the Publication of Socially Harmful Information May Be Socially Desirable

09/121 A. Ziegler
Is it Beneficial to be Included in a Sustainability Stock Index? A Panel Data Study for European Firms

09/120 K. Pittel and L. Bretschger
The Implications of Heterogeneous Resource Intensities on Technical Change and Growth

09/119 E. J. Balistreri, R. H. Hillberry and T. F. Rutherford
Trade and Welfare: Does Industrial Organization Matter?

09/118 H. Gersbach, G. Sorger and C. Amon
Hierarchical Growth: Basic and Applied Research

09/117 C. N. Brunnschweiler
Finance for Renewable Energy: An Empirical Analysis of Developing and Transition Economies

09/116 S. Valente
Optimal Policy and Non-Scale Growth with R&D Externalities

09/115 T. Fahrenberger
Short-term Deviations from Simple Majority Voting

09/114 M. Müller
Vote-Share Contracts and Learning-by-Doing

09/113 C. Palmer, M. Ohndorf and I. A. MacKenzie
Life's a Breach! Ensuring 'Permanence' in Forest Carbon Sinks under Incomplete Contract Enforcement

09/112 N. Hanley and I. A. MacKenzie
The Effects of Rent Seeking over Tradable Pollution Permits

09/111 I. A. MacKenzie
Controlling Externalities in the Presence of Rent Seeking

09/110 H. Gersbach and H. Haller
Club Theory and Household Formation

09/109 H. Gersbach, V. Hahn and S. Imhof
Constitutional Design: Separation of Financing and Project Decision

09/108 C. N. Brunnschweiler
Oil and Growth in Transition Countries

09/107 H. Gersbach and V. Hahn
Banking-on-the-Average Rules

09/106 K. Pittel and D.T.G. Rübbelke
Decision Processes of a Suicide Bomber – Integrating Economics and Psychology

08/105 A. Ziegler, T. Busch and V.H. Hoffmann
Corporate Responses to Climate Change and Financial Performance: The Impact of Climate Policy

09/104 S. Valente
Endogenous Growth, Backstop Technology Adoption and Optimal Jumps

09/103 K. Pittel and D. Rübbelke
Characteristics of Terrorism

09/102 J. Daubanes
Taxation of Oil Products and GDP Dynamics of Oil-rich Countries

09/101 S. Valente
Accumulation Regimes in Dynastic Economies with Resource Dependence and Habit Formation

08/100 A. Ziegler
Disentangling Specific Subsets of Innovations: A Micro-Econometric Analysis of their Determinants

08/99 M. Bambi and A. Saïdi
Increasing Returns to Scale and Welfare: Ranking the Multiple Deterministic Equilibria

08/98 M. Bambi
Unifying time-to-build theory

08/97 H. Gersbach and R. Winkler
International Emission Permit Markets with Refunding

08/96 K. Pittel and L. Bretschger
Sectoral Heterogeneity, Resource Depletion, and Directed Technical Change: Theory and Policy

08/95 M. D. König, S. Battiston, M. Napoletano and F. Schweitzer
The Efficiency and Evolution of R&D Networks

08/94 H. Gersbach and F. Mühe
Vote-Buying and Growth

08/93 H. Gersbach
Banking with Contingent Contracts, Macroeconomic Risks, and Banking Crises