

Redistributive effects of different pension systems when longevity varies by socioeconomic status

Miguel Sanchez-Romero^{*1,3}, Ronald D. Lee², and Alexia Prskawetz^{1,3}

¹Institute of Statistics and Mathematical Methods in Economics, Research Unit Economics, TU Wien, Austria

²University of California, Berkeley (UCB), US

³Wittgenstein Centre of Demography and Global Human Capital (Univ. Vienna, IIASA, VID/ÖAW), Austria

Abstract

We propose a general analytical framework to model the redistributive features of alternative pension systems when individuals face ex ante differences in mortality. Differences in life expectancy between high and low socioeconomic groups are often large and have widened recently in many countries. Such longevity gaps affect the actuarial fairness and progressivity of public pension systems. However, behavioral responses to longevity and policy complicate analysis of possible reforms. Here we consider how various pension systems would perform in an OLG setting with heterogeneous longevity and ability. We evaluate redistributive effects of three Notional Defined Contribution plans and three Defined Benefit plans, calibrated on the US case. Compared to a benchmark non-redistributive plan that accounts for differences in mortality, US Social Security reduces regressivity from longevity differences, but would require group-specific life tables to achieve progressivity. Moreover, without separate life tables, despite apparent accounting gains, lower income groups would suffer welfare losses and higher income groups would enjoy welfare gains through indirect effects of pension systems on labor supply.

Keywords: Human capital, Longevity, Inequality, Life cycle, Pension, Progressivity

JEL codes: E24, J10, J18, H55

*Corresponding author: miguel.sanchez@oeaw.ac.at

1 Introduction

There are large differences in mortality by socioeconomic status in rich industrial nations and in some developing countries as well, according to a growing literature which has also found that these differences have often widened in recent decades (NASEM, 2015; OECD, 2016; Waldron, 2007; Bosworth et al., 2016; Chetty et al., 2016; Rosero-Bixby and Dow, 2016; Rostron et al., 2010). While these increasing inequalities in health are themselves an urgent and critically important problem for policy, here we will focus on a different issue: These mortality differences interact with government programs, particularly those for the elderly such as public pensions, health care, and long term care. The economically advantaged groups survive for more years than those with lower income, and thereby receive more costly benefits from each of these programs. Unless tax and contribution structures, on the one hand, and benefit structures on the other, take such differences into account the result can be a net transfer of income from the poor to the rich through these programs. To the extent that programs are designed to be progressive, and intended to redistribute income from rich to poor, these mortality differences will reduce or even reverse the direction of redistribution. Effects of this sort on government programs in the United States were recently quantified and found large (National Academies of Sciences, Engineering and Medicine, 2015; henceforth NASEM, 2015). For example, the widening of the longevity gap between the top and bottom income quintile in the US raised the present value of lifetime government benefits for the top quintile relative to the bottom by \$132,000 for men and by \$157,000 for women (NASEM, 2015, :11). Consequences of this kind are surely present in other countries as well, and have begun to draw broader attention (Holzmann et al., 2020; Ayuso et al., 2017; Lee and Sanchez-Romero, 2020; OECD, 2016).

As populations age, the fiscal sustainability of government programs for the elderly has been increasingly threatened, leading to strong pressures for policy adjustments now and in the near future. Potential policy adjustments, such as raising the Normal Retirement Age for pensions or indexing each generation's benefit level to its remaining life expectancy, will have different effects on groups with different mortality, effects that will increase if the mortality differences continue to widen. The interactions of mortality differences with policy adjustments were also analyzed by the NASEM study, for selected program changes.

Analysis of these effects and interactions is far from straightforward, due both to data limitations and to the likelihood of broader behavioral responses by individuals to policies and to their own mortality risks. On the data side, assessments require calculation of taxes and benefits over an entire adult lifetime in relation to mortality differences across an entire lifetime. Since work often starts before age 20, and

because many individuals survive past 100, analysis requires longitudinal data for each generation over a span of something like 80 years, disaggregated by socioeconomic status. Such data are seldom if ever available. Empirical studies such as [NASEM \(2015\)](#) have in practice been based on a mixture of observed, simulated and projected data, reflecting many assumptions and introducing many uncertainties, and even after these efforts data on some key variables may be unavailable for parts of the lifecycle.

There are also difficult theoretical issues. Presumably these mortality differences are to some degree known to the actors, who then take them into account as they formulate their lifecycle plans for education, consumption, saving, and retirement, plans that are further complicated by individual differences in ability. Once government programs are added to the picture, all sorts of new incentives and distortions arise, with different effects for different longevity and ability groups. Analysis of the redistributive effects of government programs must also consider the way that these individual behavioral responses will affect the outcome.

In this paper we focus on public pensions rather than considering the whole range of public programs for the elderly. We develop a model for a small open economy, the population of which exhibits heterogeneous ability and mortality leading to differences in education, income, and retirement planning. Mortality differences by long term earnings are based on the analysis in [NASEM \(2015\)](#), which hereinafter we refer as *the Report*, and the parameters of our theoretical model are calibrated to match key aspects of *the Report* findings. We restrict our analysis to the steady state and develop a general analytical framework of pension systems. Within this framework we focus on redistributive effects of six different public pension systems: three Notional Defined Contribution (NDC) plans and three Defined Benefit (DB) plans. An NDC pension system is a PAYG system which is designed to mimic a standard Defined Contribution (DC) system, in the sense that participants make mandatory contributions to a notional account that earns a rate of return stipulated by the government, not determined by the stock market, and is generally chosen to equal or be close to the sustainable rate of return that a PAYG system can pay—that is, the growth rate of GDP or of the wage bill. ‘On retirement the participant uses the notional account to purchase from the government an annuity based on a similar rate of return and the remaining life expectancy, based on the life table for the participant’s generation. The NDC system provides an actuarially fair tradeoff between retirement and continuing to work, and it is expected that participants will view these contributions as an investment rather than as a tax. An NDC system should by design be approximately fiscally sustainable. In our analysis, one version of these NDC and DB plans ignores mortality differences at retirement as do all current programs. Another three versions—one NDC and two DB plans—differ in their approaches to structuring taxes/contributions or benefits so as to reduce

or avoid the program inequities arising from differences in life expectancy, or to achieve redistribution more generally. These five systems are compared to an ideal NDC plan (our benchmark) in which both contributions and retirement benefits are adjusted for the mortality of each group. For concreteness, our analysis draws in various ways on the results of *the Report* for the US case, either with a core DB system closely resembling that of the US, or a modified NDC system that shares some quantitative features of the US case. In order to focus on the role of the mortality differences, we simplify in various ways, including assuming that the systems are in long term fiscal balance.

Our main findings stem from the behavioral responses arising from the difference in retirement age between NDC and DB plans. NDC systems minimize labor market distortions by better linking contributions to pension benefits. Thus, in NDC systems earlier or later retirement ages tend to be as neutral as possible to the budget of the social security and the individual, since benefits are automatically adjusted according to the remaining years-lived in retirement. In contrast, DB systems poorly link contributions to pension benefits as life expectancy increases. In order to re-introduce actuarial fairness, DB pension systems apply penalties/rewards for early/late retirement ages. However, when these penalties/rewards are not in line with those that are actuarially fair, the pension system not only modifies retirement behavior, but it also leads to a series of other behavioral responses that affect the wealth and welfare of individuals. In particular, our estimates indicate that under the mortality regime of the 1930 birth cohort in the US, individuals in the first three income quintiles would have retired on average two years later in non-progressive DB plans, and by one and a half years later in progressive DB plans, compared to NDC plans. This difference in the retirement age raises the marginal benefit of education in non-progressive DB plans. In contrast, despite the positive difference in the retirement age between progressive DB plans and NDC plans, in progressive DB plans the marginal benefit of education does not raise because these systems impose a high effective tax on labor. We also find that in non-progressive DB systems the average increase in lifetime income is accompanied by a fall in lifetime welfare, since the increase in lifetime income comes at the expense of less leisure time during the working period and in retirement.

Throughout the article, we will focus on how the six public pension systems redistribute income across income groups and how individuals respond to alternative pension settings and mortality regimes. The paper is structured as follows. In Section 2 we detail the demographic characteristics of the population. In Section 3 we set a general accounting framework for simultaneously analyzing NDC and DB pension systems. In Section 4 we introduce a lifecycle model of labor supply in which individuals decide their education, hours worked, and the retirement age. Details about assumptions, data used, and parameter values are provided in

Section 5. The redistributive properties of each pension system by income quintile under two mortality regimes are presented in Section 6. Section 7 concludes. We provide a detailed derivation of the economic model in the Appendix.

2 Demographics

We assume that the mortality of each individual is completely determined by their lifetime income. We denote by $\mathcal{J} = \{1, 2, \dots, I\}$ the set of I income levels. Let $\mu_i(t) \geq 0$ be the mortality hazard rate at age t of an individual of group $i \in \mathcal{J}$. Table 1 shows the life expectancy at ages 15, 50, and 65 for the US male cohorts born in 1930 and 1960 based on *the Report*.¹ The data shows that the difference in life expectancy between the highest and the lowest quintiles is 6.5, 5.1, and 3.3 years at age 15, 50, and 65, respectively, for the birth cohort born in 1930. The differences in life expectancy between these two income groups widens for the cohort born in 1960 to 16.2, 11.9, 9.4 years. For details see Appendix A.

Table 1: Life expectancy at ages 15, 50, and 65 by income quintile, US males, birth cohorts of 1930 and 1960

Income level	1930 birth cohort			1960 birth cohort		
	Life expectancy at			Life expectancy at		
	age 15	age 50	age 65	age 15	age 50	age 65
Quintile 1	56.3	25.6	15.0	55.6	25.1	14.7
Quintile 2	57.1	26.2	15.3	58.5	27.3	16.0
Quintile 3	58.3	27.1	15.9	65.1	32.4	19.7
Quintile 4	60.0	28.8	16.9	70.5	36.8	23.2
Quintile 5	62.8	30.7	18.3	71.7	37.8	24.1

Notes: Authors' estimates based on data from *the Report*.

To simplify the demographic analysis, we assume each income group grows steadily at a rate n and that the total number of births across income groups is the same. These two assumptions imply that fertility is slightly higher for lower

¹See https://www.ssa.gov/oact/NOTES/as120/LifeTables_Body.html. Note that *the Report* only gives life expectancy by income quintile by age 50. To obtain survival probabilities below age 50 by income quintiles we use cohort-life tables from SSA. More specifically, for each income quintile we choose the cohort-life table from SSA that best matches the corresponding life expectancy at age 50 from *the Report*.

income groups, to overcome the lower proportion of females surviving through the reproductive ages.

3 The pension model

In order to provide comparable results across pension systems, we need a framework that allows us to compare simultaneously all pension plans. For this purpose we will use a pension point system, described below, that can reproduce both a defined benefit (DB) system and a defined contribution (DC) system (see, for instance, Börsch-Supan, 2006; OECD, 2005).

3.1 Parametric components

In order to keep the model as tractable as possible, we exclude disability benefits, survivor benefits, and widowhood benefits. The pension system only pays benefits to those workers who survive to retirement. Let us assume that workers contribute an amount $\tau y(t)$, where τ is the contribution rate and $y(t)$ is the labor income subject to payroll tax, for which workers gain pension points $\text{pp}(t)$ that entitle them to receive a pension benefit upon retirement. Suppose that workers earn ϕ pension points per unit of social contribution paid. Pension points earn a rate of return equal to an interest rate \tilde{r} plus a mortality risk premium. The risk premium arises because we exclude benefits from disability, survivorship, and widowhood. The interest rate \tilde{r} is assumed to be equal to or lower than the market interest rate, which we denote by r . Most of the interest rates, or indexes, used in pension systems fit into one of the following three cases: (i) when $\tilde{r} = 0$, past contributions are only adjusted for inflation; (ii) when $\tilde{r} = r$, past contributions are invested in the market and capitalized according to the interest rate r (i.e., funded system); and (iii) when $\tilde{r} = n + g$, where n is the population growth and g is the productivity growth rate, then past contributions are capitalized according to the growth rate of the national wage bill at the macro level, which corresponds to the intrinsic rate of return of a PAYG pension system (Samuelson, 1958).² The amount of pension points earned depends on the system. In a DC system, they are equal to the contribution paid (i.e., $\phi = 1$). In a DB system, they are equal to the yearly pension benefit accrual, which is a fraction ($\phi = \varrho/\tau$) of the contributions paid (i.e., $\phi\tau y = \varrho y$). Thus, the total number of pension points accumulate over the working life according to

$$\frac{\partial \text{pp}_i(t)}{\partial t} = (\tilde{r} + \tilde{\mu}(t))\text{pp}_i(t) + \phi\tau y_i(t) \text{ with } \text{pp}_i(0) = 0, \quad (1)$$

²Samuelson (1958) shows that the internal rate of return of a transfer system is equal to the growth rate of the contribution base of the system.

where \tilde{r} is the interest rate used by the pension system, $\tilde{\mu}(t)$ is the mortality hazard rate at age t used by the pension system, and $y_i(t)$ is the labor income. The total number of pension points in Eq. (1) receives a different name in each pension system. For instance, in a DC system, the total number of pension points before retirement is equal to the pension wealth, while in a DB system the total number of pension points at retirement is equal to the average indexed yearly earnings.

To calculate the pension benefit of a retiree, $b_i(R_i)$, the government applies a conversion factor, $f_i(R_i)$, that transforms at retirement the pension points accumulated (pp) into pension benefits

$$b_i(R_i) = f_i(R_i)\text{pp}_i(R_i). \quad (2)$$

In a DC system, the government transforms the pension wealth into an annuity using cohort-specific life tables and an effective interest rate. As a consequence, a higher life expectancy at retirement, holding the effective interest rate constant, leads to a reduction in benefits in a DC system. Thus, the conversion factor at the age of retirement, R_i , is

$$f_i(R_i) = E_i(R_i)/\tilde{A}(R_i, \tilde{r}), \quad (3)$$

where $E_i(R_i)$ is a factor that corrects for the difference in life expectancy of individuals of type $i \in \mathcal{J}$ relative to the average individual at retirement. Similar to [Ayuso et al. \(2017\)](#) we assume the correction factor is specific to the group to which the individual belongs and depends on the retirement age R_i .³ $\tilde{A}(R_i, \tilde{r})$ is the present value of a life annuity of a dollar per year, paid from age R_i onwards, calculated with an effective interest of \tilde{r} and a mortality hazard rate $\tilde{\mu}(\cdot)$.

In the DB system, the government multiplies the average indexed yearly earnings by a replacement rate, $\varphi(\text{pp})$, and then applies an adjustment factor $\beta(R_i)$ for early or late retirement to determine the pension benefit of the retiree. The replacement rate can be constant (i.e., $\varphi(\text{pp}) = \varphi$) or it can decrease as the average indexed yearly earnings increases (i.e., $\varphi'(\text{pp}) < 0$). To consider actuarial fairness we implement for the DB system the penalties/rewards for early/late retirement established in the US pension system for each birth cohort. As a result, in a DB system, the conversion factor at the age of retirement R_i is

$$f_i(R_i) = E_i(R_i)\beta(R_i)\varphi(\text{pp}_i(R_i)), \quad (4)$$

where $E_i(R_i)$ is the same correction factor introduced in Eq. (3). Thus, while the DC system takes into account the life expectancy through $\tilde{A}(R_i, \tilde{r})$, there is no such relationship in the DB system.

³We denote the actuarial present value at the exact age x , using the mortality hazard rates $\mu_i(\cdot)$, when the effective interest rate is r as $A_i(x, r) = \int_x^\omega e^{-\int_x^t r + \mu_i(j) dj} dt$. The correction factor proposed by [Ayuso et al. \(2017\)](#) is $E_i(R_i) = \tilde{A}(R_i, \tilde{r})/A_i(R_i, \tilde{r})$.

3.2 Pension wealth

Given that individuals expect to receive future benefits during retirement out of their contributions, the pension system generates a transfer wealth, which is known as the social security wealth (SSW). SSW is defined as the present value of the survival weighted stream of future benefits minus the present value of the survival weighted stream of remaining social contributions. Hence, if we compare individuals with similar stream of contributions, but different life expectancies, those individuals with high life expectancy will have a higher SSW than individuals with low life expectancy. As a consequence, individuals with high life expectancy will value their contributions more than those with low life expectancy. To explicitly account for the mortality differential effect on the value of contributing to the pension system, we compare for each individual the value of investing, from age $t < R_i$ until retirement, a dollar in the pension system to the value of investing the same dollar in the capital market

$$P_i(t) = \phi f_i(R_i) A_i(R_i, r) e^{\int_t^{R_i} (\tilde{r} + \tilde{\mu}(j) - (r + \mu_i(j))) dj}. \quad (5)$$

The term $\phi f_i(R_i) A_i(R_i, r)$ is the present value of the weighted stream of benefits from retirement until death that results from the contribution of a dollar. The exponential term accounts for the difference from age t until retirement between the rate of return of the pension system, $\tilde{r} + \tilde{\mu}$, and the rate of return of the capital market, $r + \mu_i$.

Using (5), we can rewrite the evolution of the social security wealth over the working life as follows

$$\frac{\partial \text{SSW}_i(t)}{\partial t} = \left(\tilde{r} + \tilde{\mu}(t) + \frac{1}{P_i(t)} \frac{\partial P_i(t)}{\partial t} \right) \text{SSW}_i(t) + \tau y_i(t) \quad (6)$$

with $\text{SSW}_i(\omega) = 0$. The derivation of (6) is provided in Appendix B. The term $\frac{1}{P_i(t)} \frac{\partial P_i(t)}{\partial t}$ is the evolution of the value of investing a dollar in the pension system at age $t < R_i$ to the value of investing the same dollar in the capital market, which from (5) is equal to

$$\frac{1}{P_i(t)} \frac{\partial P_i(t)}{\partial t} = (r - \tilde{r}) + (\mu_i(t) - \tilde{\mu}(t)), \quad (7)$$

and $\tau y_i(t)$ is the social contribution paid. When $\tilde{\mu}(t) = \mu_i(t)$, we have that P_i increases more rapidly with age when the market interest rate is higher than the interest used by the social security system (i.e, $r > \tilde{r}$). When $r = \tilde{r}$ and the social security system applies the same average mortality rate to all individuals, those individuals with a life expectancy below the average level (i.e., $e_i(x_0) < \tilde{e}(x_0)$)

or $\mu_i(t) > \tilde{\mu}(t) \forall t$) will begin with a low valuation of their contributions but it will increase with age, while those individuals with a life expectancy above the average level (i.e., $e_i(x_0) > \tilde{e}(x_0)$ or $\mu_i(t) < \tilde{\mu}(t) \forall t$) will begin with a higher valuation of their contributions but it will decrease with age. See Figure 1 for an illustration. Thus, the second component of (7) accounts for the redistribution of resources within the cohort from those with low life expectancy to those with high life expectancy.

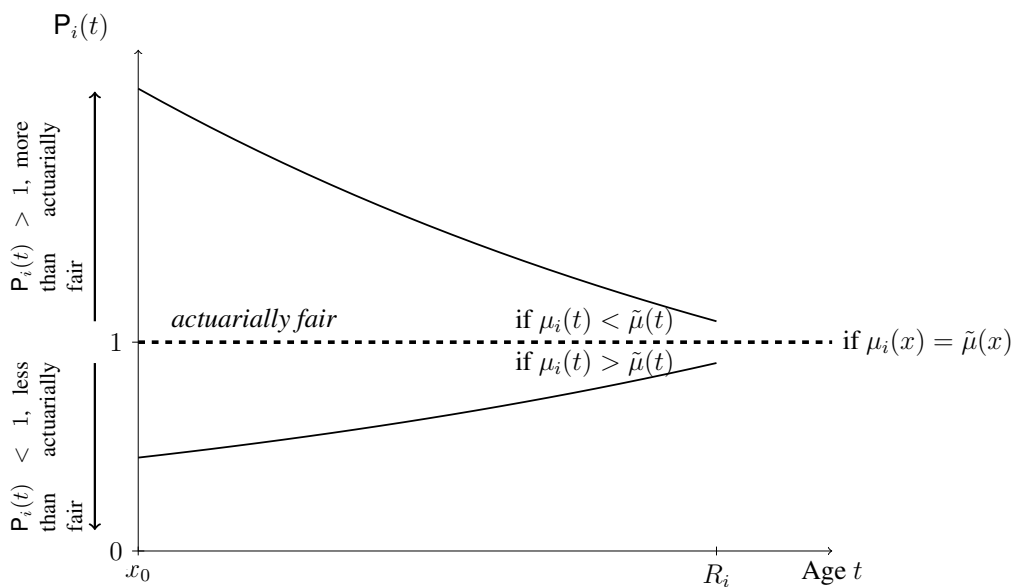


Figure 1: Stylized evolution of the relative value of a dollar contributed to the pension system at age x for an individual who plans to retire at age R_i , $P_i(t)$. Case: when $\tilde{r} = r$.

3.3 Basic components of alternative pay-as-you-go pension systems

We implement the pension model introduced in sections 3.1 and 3.2 to analyze six alternative PAYG pension systems (3 DCs and 3 DBs). To ease the comparison across pension systems and clearly show their main features, we will for the moment introduce three assumptions that are convenient for understanding how the six alternative pension systems affect eqs. (2)–(7) across individual types. First, we assume the market interest rate coincides with the internal rate of return of a

PAYG pension system (i.e., $r = \tilde{r} \equiv n + g$, so Eq. (7) depends on $\mu_i(t)$ and $\tilde{\mu}(t)$).⁴ Second, the retirement age R_i is assumed to be the same across individual types and coincides with the normal retirement age established by the pension system, which we denote by R_n . In other words, we abstract from any penalty/reward for early/late retirement (i.e., $\beta(R_n) = 1$).⁵ Third, all pension systems provide the same replacement rate for the average individual within each cohort, which implies that $\varphi = (\tau/\varrho)/\tilde{A}(R_n, \tilde{r})$. However, the first two assumptions will not hold in our simulation results, which are based on actual data for the US.

The following pension systems are implemented:

1. A standard notional defined contribution system (NDC-I) in which the government applies the same average life table for all income groups for computing the pension points and calculating the retirement benefits.
2. A notional defined contribution system (NDC-II) in which the government computes the pension points using an average life table for all income groups. However, unlike NDC-I, the government uses the income-specific life table for the calculation of the retirement benefits. This pension system mimics the one proposed by [Ayuso et al. \(2017\)](#).
3. A notional defined contribution system (NDC-III) in which the government applies the income-specific life table associated to each individual type both for the computation of the pension points and for the calculation of the retirement benefits.
4. A defined benefit system that uses all the parametric components of the US pension system, except for the replacement rate that is assumed to be constant at 0,417 (DB-I).
5. A defined benefit system with a progressive replacement rate (see Fig. 2 in [Sánchez-Romero and Prskawetz, 2017](#)). This pension system mimics the US pension system (DB-II).
6. A defined benefit system with a two-tier replacement rate (DB-III). One tier introduces a progressive replacement rate as in the US pension system, while the second tier corrects for differences in life expectancy similar to the NDC-II.

⁴Note that, for the sake of simplicity, here, we assume no difference between a funded and an unfunded pension system. Later, in Section 6 we assume that $r > \tilde{r}$.

⁵The parametric component $\beta(R_n)$ is in fact one of the ways that differential mortality affects actuarial fairness of PAYG pension systems. In Section 6 we use the actual penalty/reward function by birth cohort from the US pension system.

Table 2: Alternative PAYG pension systems and their impact on the social security wealth at age x_0 by life expectancy

		DEFINED CONTRIBUTION (DC)		
	Symbol	Avg. Life Table (LT) NDC-I	Corrected Avg. LT NDC-II	i -th LT NDC-III
Indexation	\bar{r}	$n + g$	$n + g$	$n + g$
Point factor	ϕ	1	1	1
Correction factor	E_i	1	$\bar{A}(R_n, \bar{r})/A_i(R_n, \bar{r})$	1
Replacement rate	f_i	$1/\bar{A}(R_n, \bar{r})$	$E_i(R_n)/\bar{A}(R_n, \bar{r})$	$1/A_i(R_n, \bar{r})$
Value of \$1 contributed	$P_i(x_0)$	$\begin{cases} < 1 & \text{for } e_i < \bar{e}, \\ > 1 & \text{for } e_i > \bar{e}. \end{cases}$	$\begin{cases} < 1 & \text{for } e_i < \bar{e}, \\ > 1 & \text{for } e_i > \bar{e}. \end{cases}$	1
Value of \$1 contributed	$P_i(R_n)$	$\begin{cases} < 1 & \text{for } e_i < \bar{e}, \\ > 1 & \text{for } e_i > \bar{e}. \end{cases}$	1	1
Soc. sec. wealth	$SSW_i(x_0)$	$\begin{cases} < 0 & \text{for } e_i < \bar{e}, \\ > 0 & \text{for } e_i > \bar{e}. \end{cases}$	$\begin{cases} < 0 & \text{for } e_i < \bar{e}, \\ > 0 & \text{for } e_i > \bar{e}. \end{cases}$	0

		DEFINED BENEFIT (DB)		
	Symbol	Non-Progressive DB-I	Progressive DB-II	Corrected-Progressive DB-III
Indexation	\bar{r}	$n + g$	$n + g$	$n + g$
Point factor	ϕ	ϱ/τ	ϱ/τ	ϱ/τ
Correction factor	E_i	1	1	$\bar{A}(R_n, \bar{r})/A_i(R_n, \bar{r})$
Replacement rate	f_i	$\beta(R_n)\varphi$	$\beta(R_n)\varphi(\text{pp}_i(R_n))$	$E_i(R_n)\beta(R_n)\varphi(\text{pp}_i(R_n))$
Value of \$1 contributed	$P_i(x_0)$	$\begin{cases} < 1 & \text{for } e_i < \bar{e}, \\ > 1 & \text{for } e_i > \bar{e}. \end{cases}$	≤ 1	≤ 1
Value of \$1 contributed	$P_i(R_n)$	$\begin{cases} < 1 & \text{for } e_i < \bar{e}, \\ > 1 & \text{for } e_i > \bar{e}. \end{cases}$	≤ 1	$\begin{cases} > 1 & \text{for } e_i < \bar{e}, \\ < 1 & \text{for } e_i > \bar{e}. \end{cases}$
Soc. sec. wealth	$SSW_i(x_0)$	$\begin{cases} < 0 & \text{for } e_i < \bar{e}, \\ > 0 & \text{for } e_i > \bar{e}. \end{cases}$	≤ 0	≤ 0

Notes: ‘ \bar{e} ’ denotes the life expectancy at birth of the reference population group used by the social security system, which we assume is calculated using the average survival probability of the birth cohort. ‘ e_i ’ denotes the life expectancy at birth of the individual analyzed. x_0 is the minimum working age. All the calculations are done under the following assumptions: (a) the life expectancy is positively correlated with the income level, (b) the market interest rate r is equal to $\bar{r} = n + g$, (c) the pension replacement rate φ is equal to $(\tau/\varrho)/A_s(R_n, \bar{r})$ so as to coincide with the defined contribution system, and (d) the retirement age is fixed at the normal retirement age for all population groups, which implies that $\beta(R_n) = 1$. $A_i(R_n, r)$ denotes the actuarial present value of an individual of type i at the exact age R_n when the effective interest rate is r .

Table 2 summarizes how each pension system may affect eqs. (2)–(7) across individual types. Table 2 is divided in two sections. The top section contains the information for the defined contribution systems (from NDC-I to NDC-III), while the section at the bottom provides the information for the defined benefit systems (from DB-I to DB-III). For each pension system the information is divided in two groups of individuals. Individuals with an average life expectancy below the average level

($e_i < \tilde{e}$) and individuals with an average life expectancy above the average level ($e_i > \tilde{e}$). Table 2 shows that under the presence of a heterogeneous population by life expectancy, those individuals who have *a priori* an average life expectancy below (resp. above) the average level become (a) net contributors (resp. beneficiaries) in a NDC-I, NDC-II, DB-I systems —i.e., $SSW_i(x_0) < (>)0$, they become (b) neither contributors nor beneficiaries in a NDC-III system —i.e. $SSW_i(x_0) = 0$, while (c) the sign of the social security wealth is a priori ambiguous in a DB-II and a DB-III systems. Nonetheless, if pension systems are highly progressive, then it should be expected that those individuals with low (resp. high) income, who also have a life expectancy below (resp. above) the average level, will become net beneficiaries from (resp. contributors to) the pension system.

Table 2 can also be used for understanding the impact of each pension system on the social security wealth after relaxing the first two assumptions. In particular, if we first allow $r > \tilde{r}$, the overall value of a pension point will decline in all pension systems and thus the social security wealth at age x_0 will be lower. Second, if individuals retire before (resp. after) R_n , the social security wealth value will be smaller (resp. bigger) but the signs will remain.

4 The economic model

In the previous section we presented a general accounting framework for analyzing most pension systems. Next, we explain the main features of the life cycle model implemented to construct the labor income earned, the contributions paid, the total pension points accumulated, and the pension benefits claimed by each individual type $i \in \mathcal{J}$ across the six pension systems analyzed.

4.1 The individual problem

Let us consider an individual, who belongs to quintile $i \in \mathcal{J}$, starts making decisions after finishing the compulsory educational system at age x_0 , and lives up to a maximum age ω . Assume the stock of human capital accumulates at age $t \in [x_0, S_i)$ according to a [Ben-Porath \(1967\)](#) technology

$$\frac{\partial h_i(t)}{\partial t} = \theta_i h_i(t)^\gamma - \delta h_i(t) \text{ with } h_i(x_0) = 1, \quad (8)$$

where θ_i is the learning ability of an individual belonging to group $i \in \mathcal{J}$, $\gamma \in (0, 1)$ is the returns to scale to the time devoted to education, and δ is the human capital depreciation rate. From age S_i onwards our individual earns a wage rate per hour worked w_i , which is assumed to be a function of years of schooling and years of

post-schooling experience. Let the wage rate be $w_i(S, t) = h_i(S)\bar{w}(t - S)$, where $h_i(S)$ is the stock of human capital of an individual with S years of schooling and $\bar{w}(t - S) > 0$ accounts for the returns to $t - S$ years of post-schooling experience. Assume the wage rate per hour worked increases with post-schooling experience according to the following [Mincer \(1974\)](#) equation

$$\bar{w}(t - S) = \exp(\beta_0(t - S) - \beta_1(t - S)^2) \text{ for } t \geq S, \quad (9)$$

where $\beta_0, \beta_1 > 0$ are parameters that guarantee the usual hump-shape of the wage rate per unit of human capital (see, for instance, Table 2, p. 326, in [Heckman et al., 2006](#)).

During the working period the individual supplies her/his intensive labor in exchange for the wage rate $w_i(S, t)$ and pays contributions to the pension system. Assume individuals relate their contributions to their future benefits, and so they do not see them as a tax on labor income. This assumption connects the economic model with the general pension model introduced in Section 3. Once the individual reaches the retirement age, she/he receives the corresponding pension benefits and enjoys leisure.

Budget constraint. The choice of the path of consumption c_i , hours worked ℓ_i , length of schooling S_i , and the retirement age R_i are bounded by a lifetime budget constraint.

We assume the existence of a perfect annuity market in which individuals can purchase life-insured loans, when they are in debt, and annuities in case of having positive financial wealth.⁶ Moreover, individuals start with zero assets, $a_i(x_0) = 0$, and in the terminal age ω they do not hold wealth, $a_i(\omega) = 0$. Thus, the consumption over the remaining lifespan is financed by current assets $a_i(x)$, by the present value at age x of the remaining flow of labor income, and by the social security wealth at age x , $\text{SSW}_i(x)$. The lifetime budget constraint at age x is

$$\int_x^\omega e^{-\int_x^t r + \mu_i(j) dj} c_i(t) dt = a_i(x) + \int_x^{R_i} e^{-\int_x^t r + \mu_i(j) dj} y_i(S_i, t) dt + \text{SSW}_i(x), \quad (10)$$

where r is the market interest rate, $\mu_i(t)$ is the mortality hazard rate at age t , and $y_i(S_i, t) = w_i(S_i, t)\ell_i(t)$ is the (gross) labor income earned at age t by an individual of type i after working $\ell_i(t)$ hours for a wage rate per hour worked of $w_i(S_i, t)$.

⁶In case of death the insurance keeps the debt/assets. Therefore, individuals die without assets at all ages.

Preferences. The individual chooses the consumption path c_i , the length of schooling S_i , the number of hours worked ℓ_i for the wage rate w_i , given by (8)–(9), and the retirement age R_i by maximizing the lifetime expected utility V_i at age $x \in [x_0, S)$, which is given by

$$V_i(x) = \int_x^\omega e^{-\int_x^t \rho + \mu_i(j) dj} U(c_i(t)) dt - \int_{S_i}^{R_i} e^{-\int_x^t \rho + \mu_i(j) dj} \alpha_i v(\ell_i(t)) dt - \int_x^{S_i} e^{-\int_x^t \rho + \mu_i(j) dj} \eta dt + \int_{R_i}^\omega e^{-\int_x^t \rho + \mu_i(j) dj} \varphi(t) dt. \quad (11)$$

The first two components on the right-hand side of (11) are, respectively, the lifetime utility from consumption and the lifetime disutility from work. The third term accounts for the disutility from attending school (Sánchez-Romero et al., 2016; Restuccia and Vandenbroucke, 2013; Oreopoulos, 2007) and the last term captures the utility of leisure during retirement. ρ is the subjective discount factor, $U(c)$ is an instantaneous utility function that is assumed to be twice differentiable with $U'(c) > 0$ and $U''(c) < 0$, $\alpha_i > 0$ is the weight of the disutility of the labor supplied for an individual of type i , $v(\ell)$ is the disutility of working ℓ hours (with $v'(\ell) > 0$ and $v''(\ell) < 0$), $\eta > 0$ is the marginal disutility from attending school and $\varphi(t) > 0$ (with $\varphi'(t) \geq 0$) is the marginal utility of leisure during retirement, which increases with age as the amount of retirement time is squeezed by later and later retirement.

The last three terms in (11) are key (i) for reproducing the supply of labor during the working life, which is hump-shaped, (ii) for taking into account that the return to schooling exceeds the marginal cost of education (Heckman et al., 2006), and (iii) for replicating actual retirement ages given that individuals would never retire because continuing to work would raise consumption and reduce intensive labor.

4.2 Optimal decisions

The primary objective of this section is to explain how each pension system affects the economic behavior of our heterogenous individuals with respect to mortality. We first solve the problem of maximizing lifetime utility (11) subject to the lifetime budget constraint (10) and the laws of motion (1) and (6)–(9). The definition of the Hamiltonians and the first-order conditions are reported in Appendix D.

We start by analyzing how each pension plan affects consumption and hours worked for a given length of schooling and retirement age. Next, given the optimal consumption and labor supply, we analyze the impact that each pension plan has on the length of schooling and on the age of retirement.

Consumption and hours worked. Given a length of schooling and retirement age, the optimization yields the following optimal dynamics of consumption and intensive labor supply (hours worked):

$$\frac{1}{c_i(t)} \frac{\partial c_i(t)}{\partial t} = \sigma_c(r - \rho), \quad (12)$$

$$\frac{1}{\ell_i(t)} \frac{\partial \ell_i(t)}{\partial t} = \sigma_l \left(\frac{\frac{\partial \bar{w}(t-S_i)}{\partial t}}{\bar{w}(t-S_i)} - (r - \rho) + \frac{\tau \bar{\mathcal{P}}_i(t)}{1 - \tau + \tau \bar{\mathcal{P}}_i(t)} \frac{\frac{\partial \bar{\mathcal{P}}_i(t)}{\partial t}}{\bar{\mathcal{P}}_i(t)} \right), \quad (13)$$

where $\sigma_c = -U'(c)/cU''(c) > 0$ is the intertemporal elasticity of substitution for consumption and $\sigma_l = v'(\ell)/\ell v''(\ell) > 0$ is the intertemporal elasticity of substitution for labor. Eq. (12) is the standard Euler equation for consumption and shows that the rate of increase of consumption is not affected by the pension system. In contrast, Eq. (13) shows that pension systems influence the evolution of the intensive labor supply through $\bar{\mathcal{P}}_i$. Indeed, $\bar{\mathcal{P}}_i$ measures the labor supply distortion implied by the pension plan. The value of $\bar{\mathcal{P}}_i$ is calculated as the marginal rate of substitution between social contributions and assets:⁷

$$\bar{\mathcal{P}}_i(t) = \frac{\partial V_i(t)/\partial (\mathbf{pp}_i(t)/\phi)}{\partial V_i(t)/\partial a_i(t)} = \mathbf{P}_i(t) (1 - \varepsilon_i). \quad (14)$$

$\bar{\mathcal{P}}_i(t)$ compares the value of an *additional* dollar invested in the pension system to the value of investing the same *additional* dollar in the capital market. Hence, $\bar{\mathcal{P}}_i$ differs from \mathbf{P}_i because, as a result of paying an additional dollar to the pension system, the pension replacement rate may fall

$$\varepsilon_i = -\frac{\mathbf{pp}_i(R_i)}{\mathbf{f}_i(R_i)} \frac{\partial \mathbf{f}_i(R_i)}{\partial \mathbf{pp}_i(R_i)} \geq 0. \quad (15)$$

We can distinguish two cases:⁸

- (i) A flat pension system in which the replacement rate is invariant to the pension points accumulated ($\varepsilon_i = 0$); i.e. $\bar{\mathcal{P}}_i = \mathbf{P}_i$. This is the case for pension systems NDC-I, NDC-II, NDC-III, and DB-I.
- (ii) A progressive pension system in which the replacement rate decreases with the number of pension points accumulated ($\varepsilon_i > 0$); i.e. $\bar{\mathcal{P}}_i < \mathbf{P}_i$. This is the case for pension systems DB-II and DB-III.

⁷A similar expression to (14) can be found in [Auerbach and Kotlikoff \(1987\)](#) and [Sánchez-Romero and Prskawetz \(2017\)](#).

⁸We exclude from our analysis any pension system that is *a priori* regressive; i.e. a pension system that *a priori* transfers resources from the poor to the rich.

From Eq. (13) it follows that an increase in \bar{P}_i over age will induce a postponement of labor supply to higher ages, whereas a decline in \bar{P}_i over age will induce an anticipation of labor supply to earlier ages. From eqs. (7) and (14) we know that an increase in $\bar{P}_i(x)$ over age occurs when $(r - \tilde{r} + \mu_i(t) - \tilde{\mu}(t)) > 0$ and it declines when $(r - \tilde{r} + \mu_i(t) - \tilde{\mu}(t)) < 0$. Therefore, pension systems incentivize individuals with $\mu_i(t) < \tilde{\mu}(t)$ to supply more labor early in the working life. Moreover, the extent to which the labor supply varies with changes in $\bar{P}_i(t)$ depend on its level. Thus, flat pension systems have a stronger effect than progressive pension systems on the path of labor supply.

Length of schooling. Given a retirement age R_i and the optimal paths of consumption and labor supply, the optimal length of schooling S^* satisfies:⁹

$$r_i^h(S^*) = r_i^w(S^*, R_i) + \frac{\eta}{U'(c_i(S^*))W_i(S^*, R_i)}. \quad (16)$$

See the derivation of the optimal length of schooling condition in Appendix D.2. The left-hand side of (16) is the return to education at the S^* th unit of schooling, $r_i^h(S_i^*) = \frac{1}{h_i(S^*)} \frac{\partial h_i(S^*)}{\partial S}$. The first term on the right-hand side of (16), $r_i^w(S^*, R_i)$, is the marginal cost of the S^* th unit of schooling expressed in terms of foregone earnings. The last term in (16) is the marginal disutility from attending schooling, η , relative to the utility given to the present value of the stream of labor earnings out of effective social contributions, $U'(c_i(S^*))W_i(S^*, R_i)$. The term

$$W_i(S^*, R_i) = \int_{S^*}^{R_i} e^{-\int_{S^*}^t r + \mu_i(j) dj} (1 - \tau_i(t)) y_i(S^*, t) dt, \quad (17)$$

and $\tau_i(t) = \tau (1 - \bar{P}_i(t))$ is the effective social contribution rate.

Assuming a logarithmic utility function in consumption and given that consumption, $c_i(S^*)$, depends on P_i —see eqs. (6) and (10)—while $W_i(S^*, R_i)$ depends on \bar{P}_i , from (16) we have that: (i) a flat pension system (i.e. $P_i = \bar{P}_i$) does not change the optimal length of schooling because the non-pecuniary cost of schooling remains constant. Thus, any change in schooling is driven by changes in the retirement age. In contrast, (ii) a progressive pension system (i.e. $P_i > \bar{P}_i$) reduces the optimal length of schooling, because this pension system increases the non-pecuniary cost of schooling.

⁹See Sánchez-Romero et al. (2016) for a detailed explanation of the influence of the nonpecuniary cost of schooling on the optimal decision making process of the individual.

Optimal retirement age. Consider now a fixed length of schooling S_i . Under this setting the optimal retirement age R^* satisfies the following condition

$$U'(c_i(R^*))y_i(S_i, R^*)(1 - \tau_i^{GW}(R^*)) = \alpha_i v(\ell_i(R^*)) + \varphi(R^*). \quad (18)$$

See the proof in Appendix D.3. Eq. (18) implies that individuals choose the retirement age that equates the marginal benefit of continue working (left-hand side) to the marginal cost of continue working (right-hand side). The right-hand side of (18) is the sum of the disutility from work and the marginal utility loss of leisure from not being retired. The term $\tau_i^{GW}(R^*)$ on the left-hand side of (18) is the effective tax/subsidy rate on additional work at older ages calculated by Gruber and Wise (1999). This rate assesses the marginal change in the social security wealth from delaying retirement, relative to the labor income the individual would have earned. Given that in Section 3 we assumed that all DB systems implement the same penalties/rewards for early/late retirement established in the US pension system, all DB systems will have similar retirement incentives. Likewise, all NDC systems will have similar incentives for early/late retirement based on the remaining life expectancy at retirement.

5 Parametrization

We calibrate our model to match data for our reference group—i.e. US males born in 1930—on the length of schooling, the retirement age, and the present value of lifetime benefits at age 50 for each income quintile.¹⁰ Data on the length of schooling and the retirement age for each income quintile is taken from the Health and Retirement Survey (HRS).¹¹ The present value of lifetime benefits (PVB) at age 50 for the cohort born in 1930 is taken from *the Report*. Table 3 summarizes the model economy parameters. We detail the assumptions and the strategy followed in our calibration process in Appendix C.

Table 4 reports the optimal length of schooling S_i and the optimal retirement age R_i for the benchmark scenario (US males born in 1930, US pension system). In Appendix E, tables 6 and 7 report the optimal length of schooling and retirement

¹⁰We calibrate our model to males born in the US because historical data on schooling, income, labor force participations, and retirement for men is more abundant than for women. In addition, the present value of expected benefits at age 50 for women is to a large extent affected by widowhood benefits. Studying women will imply modeling the probability of becoming a widow, which is out of the scope of this paper. See Haan et al. (2019) for an empirical estimation using administrative data of the internal rate of return of men and women in Germany.

¹¹Data from the HRS on length of schooling and retirement age for males born in 1930 was provided by Arda Aktas and Miguel Poblote-Cazenave.

Table 3: Model parameters

Parameter	Symbol	Value	Parameter	Symbol	Value
Demographics			Preferences		
First age at entrance	x_0	14	Subjective discount factor	ρ	0.005
Maximum age	ω	114	Utility cost of not being retired	$\varphi(t)$	$483.65(\mathbf{e}(t))^{-2.20}$
Annual population growth	n	0.005	Labor elasticity of substitution	σ_L	0.33
Minimum length of schooling	\underline{S}	10	Utility weight of labor	$\alpha(q1)$	200
Maximum length of schooling	\bar{S}	20		$\alpha(q2)$	160
Technology				$\alpha(q3)$	140
Market interest rate	r	0.030		$\alpha(q4)$	130
Labor-augmenting technological progress growth rate	g	0.015		$\alpha(q5)$	130
Social security system			Education		
Minimum retirement age	\underline{R}	NDC=55, DB=62	Returns of scale in education	γ	0.65
Maximum retirement age	\bar{R}	70	Disutility of schooling	η	3.5
Capitalization factor	\bar{r}	0.02	Mincerian eq.	β_0	0.07
Accrual rate in DB systems	ϕ	1/45		β_1	0.0011
Avg. replacement rate in DB systems	$f(\text{pp})$	0.4167	Learning ability	$\theta(q1)$	0.110
Social contribution rate				$\theta(q2)$	0.110
Cohort 1930	τ_{1930}	0.1192		$\theta(q3)$	0.110
Cohort 1960	τ_{1960}	0.1460		$\theta(q4)$	0.115
				$\theta(q5)$	0.115

ages for all pension scenarios, respectively. Life expectancy and total years-worked are calculated using the specific mortality rates for each income quintile. The last column in Table 4 shows how, in our model, individuals in higher income quintiles spend a longer period of time in retirement relative to the total number of years worked.

Table 4: In-sample performance of the model: Optimal length of schooling (S_i), retirement age (R_i), and present value of lifetime benefits (PVB) by income quintile. US males born in 1930, US pension system (DB-II)

Quintile	Schooling		Retirement		PVB at age 50 (in \$1 000s)		Life expectancy at $S_i + 6$	Years-worked	Years-retired to years-worked
	S_i		R_i				$\mathbf{e}_i(S_i + 6)$	YW_i	$\frac{\mathbf{e}_i(S_i+6) - YW_i}{YW_i}$
	Bench.	Data	Bench.	Data	Bench.	Data	Bench.	Bench.	Bench.
q1	11.30	11.20	62.70	63.18	132	126	54.31	41.91	0.30
q2	11.80	11.04	63.30	63.60	149	141	54.63	42.12	0.30
q3	12.30	12.28	64.00	63.56	170	166	55.33	42.49	0.30
q4	13.20	12.84	64.70	63.52	198	192	56.25	42.50	0.32
q5	14.20	14.55	65.20	64.23	225	226	58.03	42.76	0.36

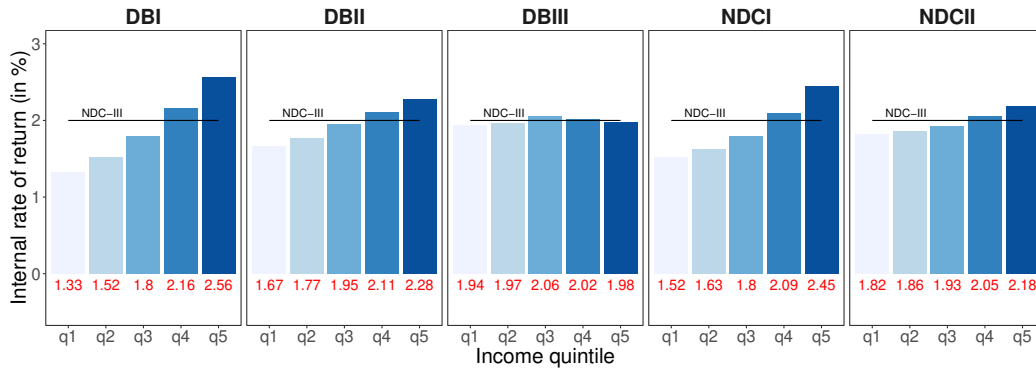
Notes: Small figures highlighted in gray are data from the HRS on length of schooling and retirement age for males born in 1930, and from the NASEM (2015) on the present value of lifetime benefits for the same cohort.

6 Redistributive effects of each pension system

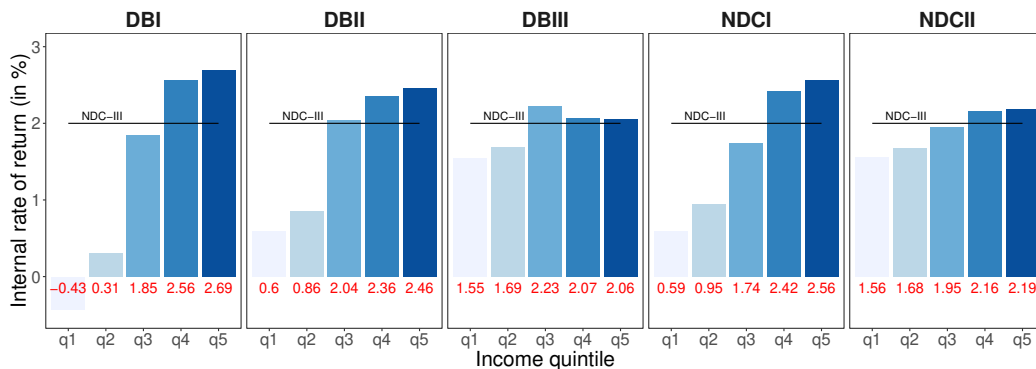
6.1 Internal rate of return

One measure to analyze the redistributive characteristics of a pension system, which is not affected by the scale of contributions (i.e., pp_i), is the internal rate of return (IRR). Thus, we report in Figure 2 the IRR values of each pension system by income quintile and birth cohort. The differences in IRR across income quintiles, shown in Fig. 2, are explained by the fact that a pension point earned by an individual with low life expectancy has a lower value than a pension point earned by an individual with higher life expectancy (see Eq. (5) and Table 2).

In a stable, mature PAYG pension system, the implicit rate of return equals the rate of growth of the population plus the rate of growth of productivity, or in this case 2.0% per year. In Fig. 2 we see that this rate of return is achieved by all income groups and mortality regimes under NDC-III —see black solid line— in which both point accumulation and the annuity rate are adjusted for the mortality of each group. This is the benchmark against which we can assess the rate of return for the groups under the other pension systems. For the NDC-I and NDC-II cases, we see that the lower income quintiles q1, q2 and q3 have $IRR < 2.0$ —see the numbers in black at the bottom of each bar— and therefore are redistributing income to the higher income q4 and q5 who have $IRR > 2.0$, and this redistribution is greater for the mortality regime of the 1960 birth cohort. The situation is more complicated for the DB systems. The actual US system, corresponding quite closely to DB-II, is explicitly designed to be redistributive from rich to poor through explicit differences in the replacement rates by income. However, we see that because of differential mortality, DB-II fails in this goal, and instead redistributes from q1 and q2 to q4 and q5 under both mortality regimes, but particularly with more unequal mortality in the 1960s. The q3 group does redistribute to others, at least slightly, under the 1930 mortality regime and becomes a net receiver under the more unequal mortality. In other words, the differential mortality completely undoes and mostly reverses the intended progressivity of the DB-II system (Sánchez-Romero and Prskawetz, 2017). Under DB-III, which both has progressive benefit levels and makes additional adjustments to benefits for differential mortality, there is a significant improvement, but the degree of progressivity is also weakened with the more unequal mortality regime.



(a) Mortality regime of 1930 cohort



(b) Mortality regime of 1960 cohort

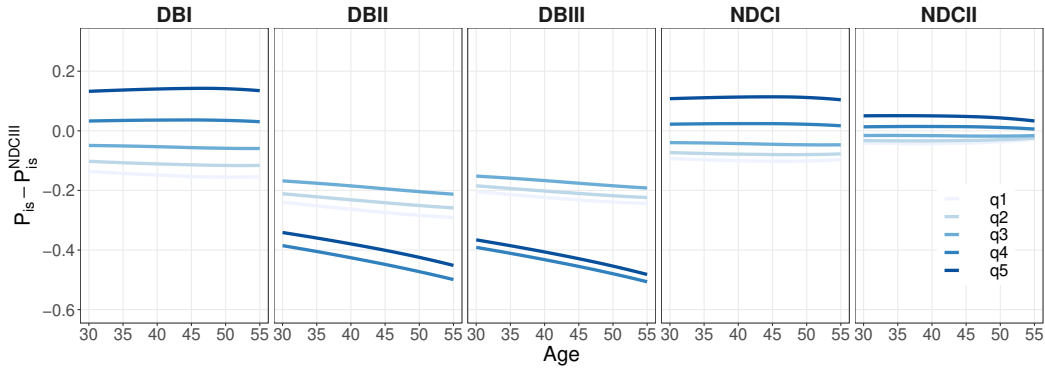
Figure 2: Internal rate of return of each pension system by income quintile (in percentage). US males, with mortality regime of birth cohorts 1930 (Panel a) and 1960 (Panel b)

Notes: Horizontal lines depict the internal rate of return of the NDC-III system. The numbers at the bottom of each column report the internal rate of return for each income quintile group.

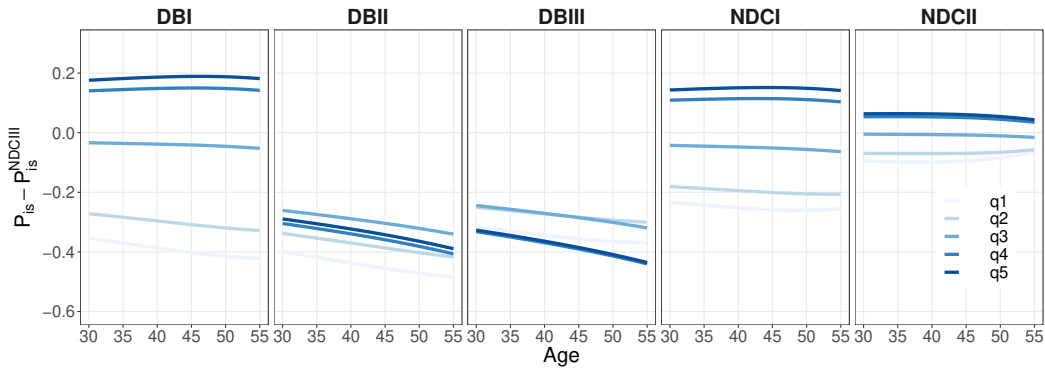
6.2 The value of investing an additional dollar in each pension system

Another method to measure the redistributive effects of a pension system is to look at the relative value of investing an additional dollar in the pension system compared to the value of investing an additional dollar in the capital market; i.e. \bar{P}_i . Hence, as shown in sections 3.2 and 4.2, \bar{P}_i can be used to study at each age and across individual types the redistributive effect of each pension system. Moreover, \bar{P}_i complements the information provided in the previous section, since the IRR measures the redistributive effects of each pension system over the whole life cy-

cle, while $\bar{\mathcal{P}}_i$ does it by age.



(a) Mortality regime of 1930 cohort



(b) Mortality regime of 1960 cohort

Figure 3: Relative value of investing an additional dollar in the pension system, $\bar{\mathcal{P}}_i(x)$, from age 30 to 55 by income quintile relative to NDC-III. US males, with mortality regime of birth cohorts 1930 (Panel a) and 1960 (Panel b)

To understand the redistributive properties by age, we compare the value of $\bar{\mathcal{P}}_i$ for each pension system to our benchmark (NDC-III). We report in Figure 3 the difference in the evolution of $\bar{\mathcal{P}}_i$ for each pension plan and that in the NDC-III by income quintile and birth cohort. Vertical axes in Fig. 3 reflect whether the contribution to the system represents a subsidy from other income groups (positive values) or transfer to other income groups (negative values) and whether the contribution or subsidy changes with age. We see that NDC-I, NDC-II, and DB-I (pension plans with a flat replacement rate) redistribute income from poor (q1, q2, and q3) to rich income groups (q4 and q5). The situation is reversed in the progressive pension systems DB-II and DB-III, in which we see that the higher income quintiles q4

Table 5: Marginal and average replacement rates at the normal retirement age

Case	Marginal replacement rate $f_i(R_n, pp_i(R_n))(1 - \varepsilon_i)$	Replacement rate $f_i(R_n, pp_i(R_n))$
DB-II DB-III	$\begin{cases} 0.90 & \text{for } pp_i \leq \bar{y}/6, \\ 0.32 & \text{for } \bar{y}/6 < pp_i < \bar{y}, \\ 0.15 & \text{for } \bar{y} < pp_i \leq 2\bar{y}, \\ 0.00 & \text{for } 2\bar{y} < pp_i, \end{cases}$	$\begin{cases} 0.90 & \text{for } pp_i \leq \bar{y}/6, \\ 0.32 + \frac{0.58}{6} \frac{\bar{y}}{pp_i} & \text{for } \bar{y}/6 < pp_i < \bar{y}, \\ 0.15 + \frac{1.60}{6} \frac{\bar{y}}{pp_i} & \text{for } \bar{y} < pp_i \leq 2\bar{y}, \\ \frac{3.40}{6} \frac{\bar{y}}{pp_i} & \text{for } 2\bar{y} < pp_i, \end{cases}$
DB-I NDC-I NDC-II NDC-III	0.417	0.417

Notes: The term \bar{y} denotes the average labor income of the economy.

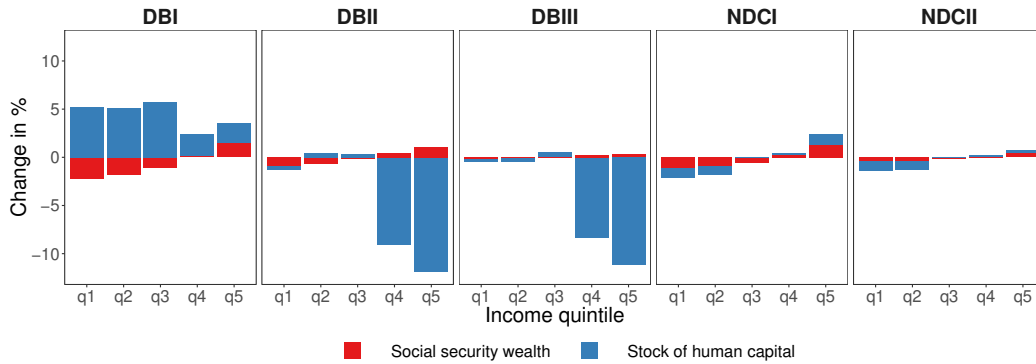
and q5 have a lower value of \bar{P}_i than lower income quintiles q1 and q2 under the mortality regime of the 1930 cohort. The fact that $\bar{P}_i - \bar{P}_i^{NDC-III}$ is negative for all income quintiles, in DB-II and DB-III systems, is due to the fact that the marginal replacement rate of DB-II and DB-III systems compared to the marginal replacement rate of the NDC-III system —see first column in Table 5— is 23% (=1-0.32/0.417) lower, for individuals with pension points between one-sixth and one average labor income, and 64% (=1-0.15/0.417) lower, for individuals with pension points between one and two times the average labor income. Comparing the results between the two mortality regimes (cf. panels (a) and (b)), we see further redistribution of income from poor individuals to rich individuals under NDC-I, NDC-II and DB-I plans and similar implicit taxes between poor and rich individuals in the more unequal mortality regime.

6.3 Wealth

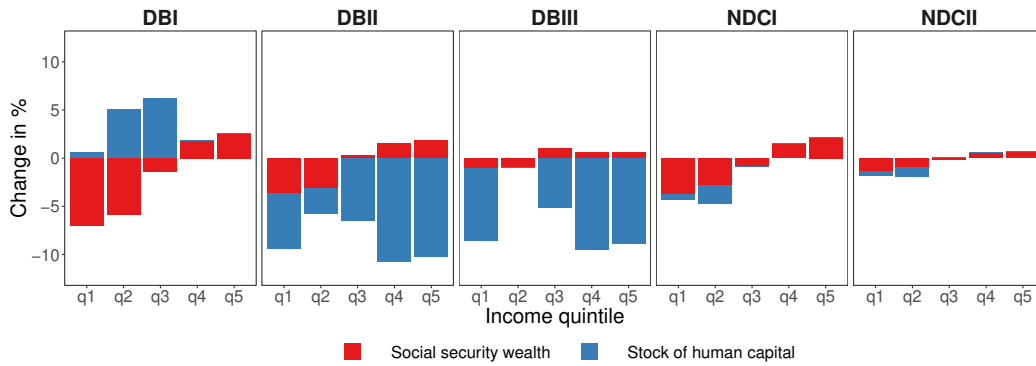
The fact that pension systems redistribute income across income groups leads individuals to respond in order to cope with the increase/loss of wealth. The lifetime wealth measure (LW) gives the most comprehensive assessment of the effects of the different pension designs on economic wellbeing, because it includes the general behavioral responses. These responses are reflected in the LW —see Eq. (10)— through changes in the social security wealth at age x_0 , $SSW(x_0)$, and through changes in the stock of human capital (HK) at age x_0 valued as the present value of

expected lifetime earnings

$$HK_i(x_0) = \int_{S_i}^{R_i} e^{-\int_{x_0}^t r + \mu_i(j) dj} y_i(S_i, t) dt. \quad (19)$$



(a) Mortality regime of 1930 cohort



(b) Mortality regime of 1960 cohort

Figure 4: Effects of each pension system and mortality regime on human capital and Social Security wealth by income quintile (measured in percentage change with respect to the results in the NDC-III system). US males, mortality regimes of birth cohorts 1930 (Panel a) and 1960 (Panel b).

Notes: Bars are plotted in ‘stacked’ format. When bars have opposite signs, lifetime wealth is the difference between both bars. When bars have similar signs, lifetime wealth equals the height of the two bars.

To analyze the changes in LW across pension systems by income quintiles, we again use the NDC-III system as a benchmark against which we can assess these changes —see Table 10 in Appendix E. Figure 4 shows the percentage change in

SSW (red bars) and HK (blue bars) by income quintile between the alternative pension systems and the NDC-III system. The negative values of the SSW for lower quintiles as opposed to positive, or rather negligible, values of SSW for higher income quintiles shows the redistribution of income from low to high income groups across all pension systems (see Fig. 4). At the same time, behavioral reactions induce changes in the length of schooling and retirement, which modify the stock of human capital (HK) (see tables 6–7 in Appendix E).¹² Looking at each pension plan, we can see that the DB-I system leads to an increase in HK for all income quintiles under the mortality regime of the 1930 cohort. Under the mortality regime of the cohort born in 1960, the DB-I system benefits mostly the higher income quintiles due to the decrease in SSW for low income quintiles. In all cases the increase is driven by the postponement in the retirement age, the additional years of schooling, and by the higher wage rate due to further investments in human capital compared to the NDC-III system. Instead, the actual US pension system —i.e., DB-II— reduces HK of q4–q5 under the mortality regime of the cohort born in 1930 due to the significant drop in the relative value of investing an additional dollar in the pension system for these two income groups (see Fig. 3). Under the mortality regime of the cohort born in 1960 the overall effect of the DB-II plan on HK is negative. This is because individuals respond to the decline in the relative value of investing an additional dollar in the pension system and in the returns to education by reducing the length of schooling.

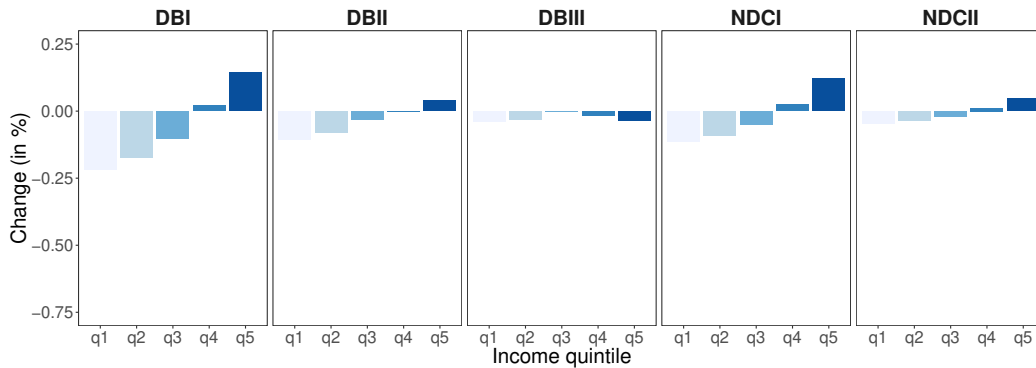
The impact of the DB-III system on HK is similar to that in DB-II but now the progressivity of the pension system is re-introduced. It is striking that in DB plans the indirect effects on HK arising from incentives for school, work, and retirement, usually are far larger than in NDC plans. This is because DB plans, with the retirement incentives designed in the US pension system, produce an incentive to retire at later ages, which does not necessarily coincide with incentives of the NDC plans. As a consequence, the increase in the retirement age leads to an increase in the length of schooling in the DB-I plan. However, in DB-II and DB-III plans, the decrease in the value of investing an additional dollar in the pension system reduces the returns of education, leading to a lower length of schooling relative to that in the NDC-III plan.

Under the NDC-I and NDC-II plans, lower income quintiles q1–q3 experience a reduction in LW, whereas the LW of higher income quintiles increases. Under both mortality regimes, the difference in LW under the NDC-I and NDC-II across income quintiles is explained by the change in SSW compared to the NDC-III plan.

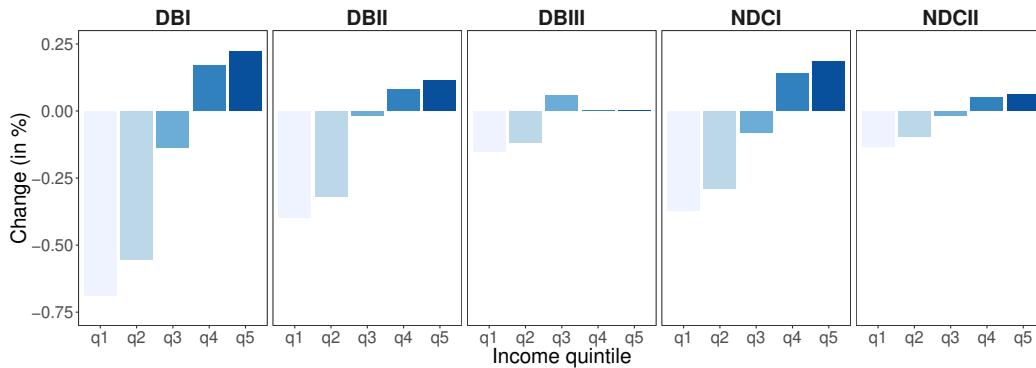
¹²Following the suggestion of anonymous reviewers, we provide in figs. 6–7 in Appendix E an alternative calculation of the changes in human capital by comparing the lifetime wealth calculation with and without behavioral reactions in which we keep the schooling, the labor supply and the retirement calibrated for the DB-II system.

6.4 Welfare

We have shown in the previous section that differences in lifetime wealth come with changes in schooling and leisure time, through age at retirement. Given that our lifetime utility (see Eq. 11) includes disutility of schooling, labor, and the utility from retirement, we can assess the impact of the different pension plans on lifetime welfare by income quintile. We will not attempt to assess how well the pension system solves the underlying problem that some people cannot provide for their own retirement. Instead, we see how lifetime utility under each pension program would differ from lifetime utility under the neutral NDC-III program structure (NDC with separate life tables) for each income quintile. These differences are shown in Figure 5.



(a) Mortality regime of 1930 cohort



(b) Mortality regime of 1960 cohort

Figure 5: Impact of each pension system on welfare by income quintile (relative to the NDC-III system). US males, with mortality regime of birth cohorts 1930 (Panel a) and 1960 (Panel b).

In Figure 5 we show the impact on lifetime welfare of each pension system by income quintile relative to the NDC-III plan. By comparing outcomes to those for the NDC-III system for each income quintile, we isolate the impact on welfare of each pension plan relative to a non redistributive pension system. The first important result can be seen by comparing in Fig. 5 the differences between the NDC plans and DB plans. Recall that all DB plans are implemented with the penalties/rewards for early/late retirement established in the US pension system, which give individuals an incentive to retire at a later age than NDC plans. In our particular case, under the mortality regime of the 1930 cohort, individuals retire between ages 61 and 66 in NDC plans, whereas individuals retire at older ages under the DB systems—see Table 7 in Appendix E. This difference in the retirement age between NDC and DB plans accounts for the strong behavioral response, its impact on the stock of human capital (see Fig. 4), and the welfare loss through the heavy cost in leisure.

In NDC plans we do not observe significant differences in the length of schooling. Thus, the sign of the change in welfare across income quintiles is explained by the impact of the pension system on SSW and hence on consumption. Moreover, given that the NDC-II is closer than the NDC-I to the NDC-III, the NDC-II plan creates smaller welfare differences across income quintiles than the NDC-I.

Comparing the results across DB plans is slightly more complicated. First, we know that the DB-I plan gives individuals higher incentives to retire later—increasing their marginal benefit of education—and to stay longer in schooling. This explains the increase in the stock of human capital (see Fig. 4) under the mortality regime of the 1930 cohort. However, the increase in human capital comes at the expense of facing a higher disutility from schooling and a loss in leisure. Only those in the highest income quintile are better off due to the strong redistribution of resources from short-lived and poor individuals to long-lived and rich individuals (see Fig. 2). Unlike the DB-I plan, the US pension system (DB-II plan) introduces a high implicit tax on work to all income quintiles. As a consequence, individuals retire in the DB-II earlier than in DB-I, though still later than in the NDC plans due to the penalties on early retirement. Moreover, given that the DB-II system produces a high implicit tax on labor, decreasing the marginal benefit of education, the length of schooling is shortened. The combination of these three behavioral reactions explains the reduction in human capital (see Fig. 4) and the increase in welfare to all income quintiles relative to DB-I, except for q5 that now transfers resources to short-lived and poor individuals. DB-III corrects for differences in life expectancy, leaving the short-lived and poor individuals better off, compared to the DB-II, and worsens the situation for long-lived and rich individuals.

7 Conclusion

Public pension systems are intended to provide a stable source of post retirement income, given that individuals have well-known difficulties saving for retirement. Some pension systems are also designed to redistribute income from individuals with higher lifetime incomes to those with lower. Almost all public pension systems are PAYG, delivering an average rate of return to participants equal to the rate of growth of the economy, which is typically lower than the market rate of interest, and consequently participants may view their contributions at least partially as a tax on labor. Pension systems modify labor supply incentives in two important ways. First, the perceived tax on labor may lead participants to work less than otherwise. Second, in DB systems the benefit structure has often created incentives for early retirement, and built in progressivity may provide further disincentives for labor. NDC systems were developed to avoid the early retirement incentive effects by mimicking funded DC programs, but they cannot avoid the “tax on labor” disincentive so long as they are PAYG.

It is increasingly realized that socioeconomic differences in longevity add a regressive element on the benefit side of pensions, so long as systems use a one-size-fits-all life table to establish actuarial tradeoffs and set the benefit rate and normal retirement age. Researchers and policy makers are seeking policy options to offset this regressive effect. However, it is important to keep in mind that policy adjustments will have both direct and indirect effects on systems and their progressivity through the behavioral responses of socioeconomic groups, effects which can be evaluated using actuarial calculations. The policy adjustments will also alter the decisions and behavior of individuals in different socioeconomic groups, because incentives for getting education, for hours of work, and for retirement age, will all be affected.

Here we have assessed the direct and indirect effects of a variety of policy adjustments to DB and NDC pension programs, which are assumed to have existed over the long term, operating in environments of more or less mortality heterogeneity (reflecting the mortality regimes of the 1930 vs 1960 birth cohorts). We have not yet attempted to investigate transitions from one program to another, although that is the situation that policy makers must face. Our analysis is based on an OLG model for a small open economy in which wages and interest rates are set by international markets, while individuals make optimizing choices for education, labor effort, age at retirement, and consumption trajectories. We have a number of important findings.

1. We replicate, in our simulations, the regressive effect of socioeconomic differences in mortality, and the large increase in regressivity moving from the

mortality regimes of the 1930 and 1960 birth cohorts, for single life table systems, whether DB or NDC.

2. Taking an actuarial approach (not general equilibrium) we find that the progressivity in benefits built into the US Social Security system greatly reduces the regressivity that mortality variation imparts, under either mortality regime. However, only when each group has its own appropriate life table is that regressivity overcome, resulting in a slightly progressive system as measured by the IRR (internal rate of return). Apparently achieving progressivity in lifetime benefits would require more than the current progressivity in annual benefits in combination with life tables for each group.
3. If we also take into account the behavioral responses to policy adjustments, these have sizable indirect effects. One way to assess these is through their impact on lifetime wealth. Under all NDC versions, the indirect effects of policy adjustments on wealth are relatively small and regressive. For the DB systems, the indirect effects are stronger. The indirect effects can be positive with non-progressive pension plans when the difference in life expectancy across income quintiles is small. However, this same system induces highly regressive effects under an environment with more mortality heterogeneity. The indirect effects of adding progressive benefits are strongly negative at higher incomes, and these change but little when group-specific life tables are added. In general these indirect effects are quite similar in the two mortality regimes in combination with life tables for each group.
4. But variations in lifetime wealth may mask offsetting variations in leisure, and the most complete assessment of policy effects emerges from comparisons of lifetime utility across pension programs. Relative to the neutral NDC-III lifetime utility, we find welfare losses for lower incomes and small welfare gains for higher incomes. This pattern arises from the utility cost of harder and longer work, which comes at a heavy cost in leisure. A reduction in welfare losses for the lower incomes and welfare gains for richer incomes can be achieved through the progressivity of benefits. However, only when benefits are adjusted using different life tables for each group do we observe slight welfare gains for the lower income groups relative to NDC-III.
5. Pension systems that introduce separate life tables for each group achieve the best welfare outcomes for the bottom three quintiles, in both NDC systems and DB systems. This finding suggests that policy makers should seriously consider pension policies of this sort.

Besides the above mentioned findings, in this paper we also propose a general pension framework for simultaneously analyzing existing pension systems. In this general framework we exploit the relative value of investing an additional dollar in the pension system, which can be used for studying the redistributive properties of each pension system as well as the behavioral response on education, hours worked, retirement, and consumption caused by each pension system. This general pension framework can be used for addressing some currently debated pension reforms that hinge on life expectancy heterogeneity (Breyer et al., 2010; NASEM, 2015). For instance, we show that pension plans in which benefits are based on different life tables (Ayuso et al., 2016, 2017; OECD, 2017; Holzmann et al., 2020) still provide a higher return to individuals with higher lifetime income than to those with lower. Thus, to restore an equal treatment of the pension system to all income groups requires additional measures.

Acknowledgements

This project has received funding from the Austrian National Bank (OeNB) under Grant no. 17647. Ronald Lee's research was supported by the grant NIA 5R24-AG045055. This project has also received funding from the European Union's Seventh Framework Program for research, technological development and demonstration under grant agreement no. 613247: "Ageing Europe: An application of National Transfer Accounts (NTA) for explaining and projecting trends in public finances". Alan Auerbach, Friedrich Breyer, Martin Kandler, Michael Kuhn, Klaus Prettnner, and participants at the Ausschusses Bevölkerungsökonomik 2018, The Economics of Ageing and Inequality workshop at Hohenheim, EPC 2018, PAA 2019, and the workshop on longevity, heterogeneity and pension design at UCL for providing useful comments. The content is solely the responsibility of the authors and does not necessarily represent the views of any of the above.

References

- Arrazola, M., de Hevia, J., 2004. More on the estimation of the human capital depreciation rate. *Applied Economic Letters* 11 (3), 145–148.
- Auerbach, A. J., Kotlikoff, L. J., 1987. *Dynamic fiscal policy*. Cambridge University Press, Cambridge.
- Ayuso, M., Bravo, J., Holzmann, R., 2016. On the heterogeneity in longevity among

- socioeconomic groups: Scope, trends, and implications for earnings-related pension schemes. *IZA Discussion Paper* No. 10060.
- Ayuso, M., Bravo, J., Holzmann, R., 2017. Addressing Longevity Heterogeneity in Pension Scheme Design. *Journal of Finance and Economics* 6 (1), 1–24.
- Ben-Porath, Y., 1967. The production of human capital and the life cycle of earnings. *The Journal of Political Economy* 75 (4), 352–365.
- Bosworth, B., Burtless, G., Zhang, K., 2015. Sources of increasing differential mortality among the aged by socioeconomic status. CRR WP 2015-10.
- Bosworth, B., Burtless, G., Zhang, K., 2016. Later retirement, inequality in old age, and the growing gap in longevity between rich and poor. (The Brookings Institution)
- Börsch-Supan, A., 2006. What are NDC systems? What do they bring to reform strategies?. *Pension reform: Issues and prospects for non-financial defined contribution (NDC) schemes*, 35-55.
- Breyer, F., Hupfeld, S., 2010. On the fairness of early-retirement provisions. *German Economic Review* 11 (1), 60–77.
- Burtless, G., 2013. The impact of population aging and delayed retirement on workforce productivity. CRR WP 2013-11.
- Cervellati, M., Sunde, U., 2013. Life expectancy, schooling, and lifetime labor supply: Theory and evidence revisited. *Econometrica* 81 (5), 2055–2086.
- Chetty, R., 2006. A new method of estimating risk aversion. *American Economic Review*, 96 (5), 1821–1834.
- Chetty, R., Stepner, M., Abraham, S., Lin, S., Scuderi, B., Turner, N., Bergeron, A., Cutler, D., 2016. The association between income and life expectancy in the United States, 2001–2014. *The Journal of the American Medical Association*, 315 (14), 1750–1766.
- Committee on the Long-Run Macroeconomic Effects of the Aging U.S. Population Phase II, 2015. *The growing gap in life expectancy by income: Implications for federal programs and policy responses*. Washington, D. C.: The National Academies Press.
- Gruber, J., Wise, D., 1999. Social security programs and retirement around the World. *Research in Labor Economics*, 18, 1–40.

- Haan, P., Kemptner, D., Lüthen, H., 2019. The rising longevity gap by lifetime earnings—Distributional implications for the pension system. *The Journal of the Economics of Ageing*, 100199.
- Heckman, J. J., Lochner, L. J., Todd, P. E., 2006. Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond. In: E. Hanushek and F. Welch, Editor(s), *Handbook of the Economics of Education* 1, 307–458.
- Heijdra, B. J., Reijnders, L. S.-M., 2016. Human capital accumulation and the macroeconomy in an ageing society. *De Economist*, 164, 297–334.
- Holzmann, R., Alonso-García, J., Labit-Hardy, H., Villegas, A. M., 2020. NDC schemes and heterogeneity in longevity: Proposals for re-design. In: R. Holzmann, E. Palmer, R. Palacios, and S. Sacchi (Eds), *Progress and Challenges of Nonfinancial Defined Pension Schemes. Vol 1: Addressing Marginalization, Polarization, and the Labor Market*. Washington, DC: The World Bank.
- Holzmann, R., Palmer, E., Palacios, R., Sacchi, S. (eds), 2020. *Progress and Challenges of Nonfinancial Defined Pension Schemes*. Washington, D.C.: The World Bank.
- Lee, R. D., Sánchez-Romero, M., 2020. Overview of the relationship of heterogeneity in life expectancy to pension outcomes and lifetime income. In: R. Holzmann, E. Palmer, R. Palacios, and S. Sacchi (Eds), *Progress and Challenges of Nonfinancial Defined Pension Schemes. Vol 1: Addressing Marginalization, Polarization, and the Labor Market*. Washington, D.C.: The World Bank.
- Mincer, J., 1974. *Schooling, Experience and Earnings*. New York: Columbia University Press.
- OECD, 2005. *Pensions at a glance. Public policies across OECD countries*. Chapter 2. OECD Publishing, Paris, 27–35.
- OECD, 2016. *Fragmentation of retirement markets due to differences in life expectancy*. In OECD Business and Finance Outlook 2016, Chapter 6. OECD Publishing, Paris, 177–205.
- OECD, 2017. *Preventing ageing unequally*. OECD Publishing, Paris.
- Oreopoulos, P., 2007. Do dropouts drop too soon? Wealth, health and happiness from compulsory schooling. *Journal of Public Economics* 91, 2213–2229.
- Oreopoulos, P., Salvanes, K. G., 2011. Priceless: The nonpecuniary benefits of schooling. *Journal of Economic Perspectives*, 25 (1), 159–84.

- Restuccia, D., Vandenbroucke, G., 2013. A century of human capital and hours. *Economic Inquiry* 51 (3), 1849–1866.
- Rosero-Bixby, L., Dow, W. H., 2016. Exploring why Costa Rica outperforms the United States in life expectancy: A tale of two inequality gradients. *PNAS*, 113 (5), 1130–1137.
- Rostron, B. L., Boies, J. L., Arias, E., 2010. Education reporting and classification on death certificates in the United States. *Vital and Health Statistics Series 2*, 151, 1–21.
- Samuelson, P., 1958. An exact consumption-loan model of interest with or without the social contrivance of money. *The Journal of Political Economy*, 6 (December): 467–482.
- Sánchez-Romero, M., d’Albis, H., Prskawetz, A., 2016. Education, lifetime labor supply, and longevity improvements. *Journal of Economics Dynamics and Control* 73, 118–141.
- Sánchez-Romero, M., Prskawetz, A., 2017. Redistributive effects of the US pension system among individuals with different life expectancy. *The Journal of the Economics of Ageing* 10, 51–74.
- Sullivan, D., 2013. Trends In Labor Force Participation. Federal Reserve Bank of Chicago.
- Tomiyama, K., 1985. A two-stage optimal control problems and optimality conditions. *Journal of Economic Dynamics and Control* 9, 317–337.
- Waldron, H., 2007. Trends in mortality differentials and life expectancy for male Social Security-covered workers, by socioeconomic status. *Social Security Bulletin*, 67 (3), 1–28.

“For Online Publication”

Appendix: Redistributive effects of different pension systems when longevity varies by socioeconomic status

by Miguel Sanchez-Romero, Ronald D. Lee, Alexia Prskawetz

A Demographics

Individuals. Let the probability of surviving from birth to age x of an individual belonging to income group $i \in \mathcal{I} = \{1, 2, \dots, I\}$ be

$$p_i(x) = e^{-\int_0^x \mu_i(t) dt}, \quad (\text{A.1})$$

with $p_i(0) = 1$, $p_i(\omega) = 0$, $\omega \in (0, \infty)$ denotes the maximum age, and $\mu_i(t) \geq 0$ is the mortality hazard rate at age t of an individual of group i . The life expectancy at age x of an individual belonging to income group i is defined as

$$e_i(x) = \int_x^\omega \frac{p_i(t)}{p_i(x)} dt. \quad (\text{A.2})$$

Population. Given that all income groups grow steadily at the same rate n , the total population size at time t is

$$P(t) = B(t) \int_0^\omega e^{-nx} p(x) dx, \text{ with } p(x) = \frac{1}{I} \sum_{i \in \mathcal{I}} p_i(x), \quad (\text{A.3})$$

where $B(t)$ is the total number of births at time t , $p(x)$ is the average survival, which implies that the average mortality hazard rate at age x , denoted by $\mu(x)$, is $\sum_{i \in \mathcal{I}} \mu_i(x) p_i(x) / \sum_{i \in \mathcal{I}} p_i(x)$. The existence of a positive relationship between the income group and the life expectancy —see Table 1— implies that the average mortality hazard rate is biased with age towards the mortality hazard rate of higher income individuals.

B Derivation of Eq. (6)

Assuming that an individual will retire at age R_i , we define the social security wealth at age $S_i < t \leq R_i$ of an individual of type i , denoted by $\text{SSW}_i(t)$, as

$$\text{SSW}_i(t) = b_i(R_i) \int_{R_i}^\omega e^{-\int_t^x r + \mu_i(j) dj} dx - \int_t^{R_i} e^{-\int_t^x r + \mu_i(j) dj} \tau y_i(x) dx. \quad (\text{B.1})$$

The first component of (B.1) is the present value at age t of the survival weighted stream of future benefits during retirement, while the second component of (B.1) is the present value at age t of the survival weighted remaining pension contributions to pay until retirement.

Eq. (6) can easily be derived from (1)–(4) by elementary algebraic manipulations of (B.1). Integrating Eq. (1) from S_i until retirement R_i gives

$$\text{pp}_i(R_i) = \int_{S_i}^{R_i} e^{\int_x^{R_i} \tilde{r} + \tilde{\mu}(j) dj} \phi \tau y_i(x) dx. \quad (\text{B.2})$$

Splitting in two the integral (B.2) gives

$$\begin{aligned} \text{pp}_i(R_i) &= \int_{S_i}^t e^{\int_x^{R_i} \tilde{r} + \tilde{\mu}(j) dj} \phi \tau y_i(x) dx + \int_t^{R_i} e^{\int_x^{R_i} \tilde{r} + \tilde{\mu}(j) dj} \phi \tau y_i(x) dx = \\ &= e^{\int_t^{R_i} \tilde{r} + \tilde{\mu}(j) dj} \int_{S_i}^t e^{\int_x^t \tilde{r} + \tilde{\mu}(j) dj} \phi \tau y_i(x) dx + \int_t^{R_i} e^{\int_x^{R_i} \tilde{r} + \tilde{\mu}(j) dj} \phi \tau y_i(x) dx. \end{aligned} \quad (\text{B.3})$$

Substituting (2) in (B.1), using the definition of $A_i(R_i, r)$, and rearranging terms, gives

$$\text{SSW}_i(t) = e^{-\int_t^{R_i} r + \mu_i(j) dj} \mathbf{f}_i(R_i) A_i(R_i, r) \text{pp}_i(R_i) - \int_t^{R_i} e^{-\int_t^x r + \mu_i(j) dj} \tau y_i(x) dx. \quad (\text{B.4})$$

By substituting (B.3) in (B.4) and after some manipulations we have

$$\begin{aligned} \text{SSW}_i(t) &= \mathbf{f}_i(R_i) A_i(R_i, r) e^{\int_t^{R_i} \tilde{r} + \tilde{\mu}(j) - (r + \mu_i(j)) dj} \int_{S_i}^t e^{\int_x^t \tilde{r} + \tilde{\mu}(j) dj} \phi \tau y_i(x) dx \\ &\quad + e^{-\int_t^{R_i} r + \mu_i(j) dj} \phi \mathbf{f}_i(R_i) A_i(R_i, r) \int_t^{R_i} e^{\int_x^{R_i} \tilde{r} + \tilde{\mu}(j) dj} \phi \tau y_i(x) dx \\ &\quad - \int_t^{R_i} e^{-\int_t^x r + \mu_i(j) dj} \tau y_i(x) dx. \end{aligned} \quad (\text{B.5})$$

By multiplying and dividing by ϕ the first term on the right-hand side of (B.5) we obtain that the first term of (B.5) is $\mathbf{P}_i(t)$ times the total value of all the contributions paid until age t , $\text{pp}_i(t)/\phi$. Moreover, we know from (5) that $\mathbf{P}_i(t) e^{-\int_t^{R_i} \tilde{r} + \tilde{\mu}(j) dj} = \phi \mathbf{f}_i(R_i) A_i(R_i, r) e^{-\int_t^{R_i} r + \mu_i(j) dj}$. Thus,

$$\begin{aligned} \text{SSW}_i(t) &= \mathbf{P}_i(t) \frac{\text{pp}_i(t)}{\phi} + \mathbf{P}_i(t) e^{-\int_t^{R_i} \tilde{r} + \tilde{\mu}(j) dj} \int_t^{R_i} e^{\int_x^{R_i} \tilde{r} + \tilde{\mu}(j) dj} \tau y_i(x) dx \\ &\quad - \int_t^{R_i} e^{-\int_t^x r + \mu_i(j) dj} \tau y_i(x) dx. \end{aligned} \quad (\text{B.6})$$

Taking into account that $\frac{P_i(t)}{P_i(x)} = e^{\int_t^x \tilde{r} + \tilde{\mu}(j) - (r + \mu_i(j)) dj}$, the social security wealth at time t in (B.6) can also be written as

$$\text{SSW}_i(t) = P_i(t) \frac{pp_i(t)}{\phi} - \int_t^{R_i} e^{-\int_t^x r + \mu_i(j) dj} \tau (1 - P_i(x)) y_i(x) dx. \quad (\text{B.7})$$

The first component of (B.7) is the monetary value given to the stream of contributions until age t , while the second component of (B.7) is the present value, survival weighted, at age t of the stream of future implicit taxes/subsidies on labor income.

By differentiating (B.7) with respect to t we have

$$\begin{aligned} \frac{\partial \text{SSW}_i(t)}{\partial t} &= \left(\tilde{r} + \tilde{\mu}(t) + \frac{1}{P_i(t)} \frac{\partial P_i(t)}{\partial t} \right) P_i(t) \frac{pp_i(t)}{\phi} \\ &\quad - (r + \mu_i(t)) \int_t^{R_i} e^{-\int_t^x r + \mu_i(j) dj} \tau (1 - P(x)) y_i(x) dx + \tau y_i(t). \end{aligned} \quad (\text{B.8})$$

Finally, using (7) in the second term of (B.8) and the definition of SSW in (B.7) we obtain Eq. (6).

C Parametrization

We impose the following set of assumptions with respect to the economic variables. First, we assume a risk-free market discount factor (r) of 3%. This market interest rate coincides with that assumed in the report by the National Academies of Science (NASEM, 2015), that we use to parametrize this model. Second, the population is assumed to grow at an annual constant rate (n) of 0.5% and the growth rate of labor productivity (g) is set at 1.5% per year, corresponding closely to the US case. Third, the annual capitalization factor of the unfunded pension system (\tilde{r}) is set at 2% ($=n + g$), which is lower than the market discount factor. Therefore, since a return of 1% ($=3\% - 2\%$) is lost annually, contributions to the pension system are implicitly considered by individuals as a tax on labor income. From (7) we know that this assumption implies that the marginal value of a dollar contributed to the pension system is an increasing function with respect to age and, as a consequence, all pension systems will provide an incentive to supply more labor early in the working life and to reduce labor before retirement. Moreover, since all pension systems have a similar increase in $\bar{\mathcal{P}}_i$, pension systems will have a similar direct impact on the length of schooling. Fourth, unless otherwise indicated, we assume that the social security system uses the average survival probability to calculate the pension benefits; i.e., $\tilde{p}(x) := p(x) = \sum_{i \in \mathcal{J}} p_i(x) / I$.

Fifth, we assume for the NDC systems a minimum retirement age of 55 and a maximum retirement age of 70 for all $i \in \mathcal{J}$. For the DB systems, we restrict the minimum retirement age to 62 and the maximum to 70, similar to the US pension system. Sixth, the social security budget is balanced¹³

$$\sum_{i \in \mathcal{J}} \int_{S_i}^{R_i} e^{-nt} p_i(t) \tau w_i(S_i, t) \ell_i(t) dt = \sum_{i \in \mathcal{J}} e^{-nR_i} \int_{R_i}^{\omega} e^{-(n+g)(t-R_i)} p_i(t) b_i(R_i) dt. \quad (\text{C.1})$$

Using Eq. (C.1) we adjust in the DB systems the social contribution rate in order to support all pension benefits claimed by the surviving retirees, while in the NDC systems we adjust the overall pension replacement rate, or generosity of the pension system. For the sake of comparison across the alternative pension systems, we use for all the NDC systems the social contribution rate obtained for the US pension system (DB-II). In particular, we obtain that the necessary social contribution rate to balance the US pension system with our assumed population structure is 11.92%, under the hypothetical assumption that the population faces the survival probabilities of the cohort born in 1930. While τ must be set at 14.60% in the case of using the survival probabilities of the cohort born in 1960. In addition, we assume all individuals have similar preferences, except for the disutility of labor (α_i) that is specific to the income quintile of the individual. This assumption reflects the fact that individuals in different quintiles have different health and labor market trajectories. The instantaneous utility of consumption is assumed to be logarithmic ($\sigma_c = 1$), as found empirically by Chetty (2006), the intertemporal elasticity of substitution on labor (σ_ℓ) is set at 0.33, so that workers supply on average thirty five percent of their available time for labor (excluding sleep time), and the subjective discount factor (ρ) is set at 0.005. As a result, the cross-sectional consumption profile increases with age by one percent, similar to the consumption pattern reported in the NTA accounts for the US in 2003 (see www.ntaccounts.org). The marginal utility of leisure during retirement $\varphi(\cdot)$ is assumed to be constant across income groups and birth cohorts and monotonically increasing with age. Thus, we consider the marginal utility of leisure to be a function of the average life expectancy of the 1930 birth cohort; i.e $\varphi(t) = \varphi_0 (\mathbf{e}(t))^{-\varphi_1}$ with $\varphi_0, \varphi_1 > 0$. To match the wage rate per unit of human capital for the cohort born in 1930, we take the parameters of the Mincerian equation reported in Table 2 in Heckman et al. (2006). Nevertheless, the parameter β_0 is adjusted in order to take into account the effect of productivity growth. As in Cervellati and Sunde (2013) we fix the returns to scale in education (γ) at 0.65. Following Sánchez-Romero et al. (2016) we set the disutility of schooling

¹³Note in Eq. (C.1) that similar to the US pension system, we assume that pension benefits are held constant (in real terms) after retirement and thus they do not increase with productivity.

(η) at 3.5 for all income groups, which corresponds to the most likely value of η for an average return to schooling between 11 and 12 percent given the life expectancy of US males born in 1930. The learning ability (θ_i) for each income quintile group is calibrated to replicate the length of schooling and the retirement age from the Health and Retirement Survey (HRS).¹⁴ Finally, the weight of disability cost (α_i) is used to replicate the present value of lifetime benefits (PVB) at age 50 reported in [NASEM \(2015\)](#) for the cohort born in 1930. See [Table 9](#) in [Appendix E](#).

D Economic problem

We solve the problem of maximizing the lifetime utility (11) subject to the constraints (1)–(10) using the Hamiltonian before age S_i , during the working period (S_i, R_i), and after the retirement age R_i , or periods 1, 2, and 3, respectively ([Tomiyama, 1985](#)). We denote with the letter \mathcal{H}^j the Hamiltonian associated to period $j = \{1, 2, 3\}$.

Period 1. Given a length of schooling S_i and retirement age R_i , the Hamiltonian of an individual type $i \in \mathcal{J}$ before working ($t \leq S_i$) is defined as

$$\begin{aligned} \mathcal{H}^1 = & e^{-\int_{x_0}^t \rho + \mu_i(j) dj} [U(c_i(t)) - \eta] + \lambda_a^1(t) [(r + \mu_i(t))a_i(t) - c_i(t)] + \\ & + \lambda_h[\theta_i h_i(t)^\gamma - \delta h_i(t)] \end{aligned} \quad (\text{D.1})$$

where $\lambda_a^1(t)$ and $\lambda_h(t)$ are the costate or adjoint variables associated to the dynamics of each state variable $\{a_i(t), h_i(t)\}$ for period 1. The first-order conditions (FOCs) for an interior consumption is:

$$\frac{\partial \mathcal{H}^1}{\partial c_i} = e^{-\int_{x_0}^t \rho + \mu_i(j) dj} U'(c_i(t)) - \lambda_a^1(t) = 0. \quad (\text{D.2})$$

The dynamic laws of the costate or adjoint variables are:

$$\frac{\partial \lambda_a^1(t)}{\partial t} = -\lambda_a^1(t)(r + \mu_i(t)), \quad (\text{D.3})$$

$$\frac{\partial \lambda_h^1(t)}{\partial t} = -\lambda_h^1(t)(\gamma \theta_i h_i(t)^{\gamma-1} - \delta), \quad (\text{D.4})$$

¹⁴Data from the HRS on length of schooling and retirement age for males born in 1930 was provided by Arda Aktas and Miguel Poblete-Cazenave.

Period 2. Given a length of schooling S_i and a retirement age R_i , the Hamiltonian of an individual type $i \in \mathcal{J}$ during the working period ($S_i < t < R_i$) is defined as

$$\begin{aligned} \mathcal{H}^2 = & e^{-\int_{x_0}^t \rho + \mu_i(j) dj} [U(c_i(t)) - \alpha_i v(\ell_i(t))] + \\ & + \lambda_a^2(t) [(r + \mu_i(t))a_i(t) + (1 - \tau)h_i(S_i)\bar{w}(t - S_i)\ell_i(t) - c_i(t)] + \\ & + \lambda_p^2(t) [(\tilde{r} + \tilde{\mu}(t))\mathbf{pp}_i(t) + \phi\tau h_i(S_i)\bar{w}(t - S_i)\ell_i(t)], \end{aligned} \quad (\text{D.5})$$

where $\lambda_a^2(t)$, $\lambda_h^2(t)$, and $\lambda_p^2(t)$ are the costate or adjoint variables associated to the dynamics of each state variable $\{a_i(t), h_i(S_i), \mathbf{pp}_i(t)\}$ for period 2. The first-order conditions (FOCs) for an interior consumption and hours worked are:

$$\frac{\partial \mathcal{H}^2}{\partial c_i} = e^{-\int_{x_0}^t \rho + \mu_i(j) dj} U'(c_i(t)) - \lambda_a^2(t) = 0, \quad (\text{D.6})$$

$$\begin{aligned} \frac{\partial \mathcal{H}^2}{\partial \ell_i} = & -e^{-\int_{x_0}^t \rho + \mu_i(j) dj} \alpha_i v'(\ell_i(t)) + \lambda_a^2(t)(1 - \tau)h_i(S_i)\bar{w}(t - S_i) + \\ & + \lambda_p^2(t)\phi\tau h_i(S_i)\bar{w}(t - S_i) = 0. \end{aligned} \quad (\text{D.7})$$

The dynamic laws of the costate or adjoint variables are:

$$\frac{\partial \lambda_a^2(t)}{\partial t} = -\lambda_a^2(t)(r + \mu_i(t)), \quad (\text{D.8})$$

$$\frac{\partial \lambda_h^2(t)}{\partial t} = -\lambda_a^2(t)(1 - \tau)\bar{w}(t - S_i)\ell_i(t) - \lambda_p^2(t)\phi\tau\bar{w}(t - S_i)\ell_i(t), \quad (\text{D.9})$$

$$\frac{\partial \lambda_p^2(t)}{\partial t} = -\lambda_p^2(t)(\tilde{r} + \tilde{\mu}(t)). \quad (\text{D.10})$$

Period 3. Given a length of schooling S_i and a retirement age R_i , the Hamiltonian of an individual type $i \in \mathcal{J}$ during retirement ($t \geq R_i$) is defined as

$$\begin{aligned} \mathcal{H}^3 = & e^{-\int_{x_0}^t \rho + \mu_i(j) dj} [U(c_i(t)) + \varphi(t)] \\ & + \lambda_a^3(t) [(r + \mu_i(t))a_i(t) + \mathbf{f}_i(R_i)\mathbf{pp}_i(R_i) - c_i(t)], \end{aligned} \quad (\text{D.11})$$

where $\lambda_a^3(t)$ and $\lambda_p^3(t)$ are the costate or adjoint variables associated to the dynamics of the state variables $\{a_i(t), \mathbf{pp}_i(t)\}$ for period 3. The first-order conditions (FOCs) for an interior consumption and hours worked are:

$$\frac{\partial \mathcal{H}^3}{\partial c_i} = e^{-\int_{x_0}^t \rho + \mu_i(j) dj} U'(c_i(t)) - \lambda_a^3(t) = 0. \quad (\text{D.12})$$

Using the definition (15), the dynamic laws of the costate or adjoint variables are:

$$\frac{\partial \lambda_a^3(t)}{\partial t} = -\lambda_a^3(t)(r + \mu_i(t)), \quad (\text{D.13})$$

$$\frac{\partial \lambda_p^3(t)}{\partial t} = -\lambda_a^3(t)\mathbf{f}_i(R_i)(1 - \varepsilon_i). \quad (\text{D.14})$$

Moreover, the following matching conditions hold at the switching ages S_i and R_i for the costate or adjoint variables

$$\lambda_a(S_i) := \lambda_a^1(S_i) = \lambda_a^2(S_i), \quad (\text{D.15a})$$

$$\lambda_h(S_i) := \lambda_h^1(S_i) = \lambda_h^2(S_i), \quad (\text{D.15b})$$

$$\lambda_a(R_i) := \lambda_a^2(R_i) = \lambda_a^3(R_i), \quad (\text{D.15c})$$

$$\lambda_p(R_i) := \lambda_p^2(R_i) = \lambda_p^3(R_i), \quad (\text{D.15d})$$

and for the Hamiltonians

$$\mathcal{H}(S_i) := \mathcal{H}^1(S_i) = \mathcal{H}^2(S_i), \quad (\text{D.16a})$$

$$\mathcal{H}(R_i) := \mathcal{H}^2(R_i) = \mathcal{H}^3(R_i). \quad (\text{D.16b})$$

Taking into account the above matching conditions, let us define the marginal rate of substitution of assets for social contributions $\bar{\mathcal{P}}(t) = \phi \lambda_p(t) / \lambda_a(t)$ for periods $\{2, 3\}$. Differentiating $\bar{\mathcal{P}}(t)$ with respect to time t , and using the dynamics of the adjoint variables, gives

$$\frac{\partial \bar{\mathcal{P}}(t)}{\partial t} = \begin{cases} \bar{\mathcal{P}}(t)(r - \tilde{r} + \mu_i(t) - \tilde{\mu}(t)) & \text{for } S_i < t < R_i, \\ \bar{\mathcal{P}}(t)(r + \mu_i(t)) - \phi \mathbf{f}_i(R_i)(1 - \varepsilon_i) & \text{for } t \geq R_i. \end{cases} \quad (\text{D.17})$$

Solving (D.17) and using the fact that $\bar{\mathcal{P}}(\omega) = 0$, the marginal rate of substitution of assets for a dollar of social contribution is:

$$\bar{\mathcal{P}}(t) = \phi \mathbf{f}_i(R_i)(1 - \varepsilon_i) A_i(R_i, r) e^{\int_t^{R_i} \tilde{r} + \tilde{\mu}(j) - (r + \mu_i(j)) dj}, \quad (\text{D.18})$$

which is equivalent to $\bar{\mathcal{P}}(t) = \mathbf{P}(t)(1 - \varepsilon_i)$.

D.1 Optimal paths of consumption (c) and hours worked (ℓ)

From the FOCs, the optimal consumption and hours worked at age x satisfy

$$e^{-\int_{x_0}^t \rho + \mu_i(j) dj} U'(c_i(t)) = \lambda_a(t), \quad (\text{D.19})$$

$$e^{-\int_{x_0}^t \rho + \mu_i(j) dj} \alpha_i v'(\ell_i(t)) = \lambda_a(t) h_i(S_i) \bar{w}(t - S_i)(1 - \tau + \tau \bar{\mathcal{P}}(t)). \quad (\text{D.20})$$

Taking logarithms in both sides of the equation and differentiating with respect to t gives

$$-(\rho + \mu_i(t)) + \frac{c_i(t)U''(c_i(t))}{U'(c_i(t))} \frac{1}{c_i(t)} \frac{\partial c_i(t)}{\partial t} = \frac{1}{\lambda_a(t)} \frac{\partial \lambda_a(t)}{\partial t}, \quad (\text{D.21})$$

$$-(\rho + \mu_i(t)) + \frac{\ell_i(t)v''(\ell_i(t))}{v'(\ell_i(t))} \frac{1}{\ell_i(t)} \frac{\partial \ell_i(t)}{\partial t} = \frac{1}{\lambda_a(t)} \frac{\partial \lambda_a(t)}{\partial t} + \frac{\bar{w}'(t - S_i)}{\bar{w}(t - S_i)} + \frac{\tau \bar{\mathcal{P}}'(t)}{1 - \tau + \tau \bar{\mathcal{P}}(t)}. \quad (\text{D.22})$$

Using the envelope condition on assets held and rearranging terms, we obtain (12) and (13), respectively.

D.2 Optimal length of schooling (S_i)

Given an optimal retirement age R_i we first differentiate the expected utility $V_i(x_0)$ w.r.t. S and making it equal to zero

$$\begin{aligned} \int_{x_0}^{\omega} e^{-\int_{x_0}^t \rho + \mu_i(j) dj} U'(c_i(t)) \frac{\partial c_i(t)}{\partial S} dt - \int_S^{R_i} e^{-\int_{x_0}^t \rho + \mu_i(j) dj} \alpha_i v'(\ell_i(t)) \frac{\partial \ell_i(t)}{\partial S} dt \\ = e^{-\int_{x_0}^S \rho + \mu_i(j) dj} (\eta - \alpha_i v(\ell_i(S))). \end{aligned} \quad (\text{D.23})$$

Substituting the FOCs in the previous equation gives

$$\begin{aligned} \int_{x_0}^{\omega} \lambda_a(t) \frac{\partial c_i(t)}{\partial S} dt - \int_S^{R_i} \lambda_a(t) (1 - \tau + \tau \bar{\mathcal{P}}(t)) w_i(S_i, t) \frac{\partial \ell_i(t)}{\partial S} dt \\ = e^{-\int_{x_0}^S \rho + \mu_i(j) dj} (\eta - \alpha_i v(\ell_i(S))). \end{aligned} \quad (\text{D.24})$$

Solving the envelope condition on assets gives $\lambda_a(t) = \lambda_a(x_0) e^{-\int_{x_0}^t r + \mu_i(j) dj}$. Substituting this last result in (D.24) and dividing by $\lambda_a(x_0)$ gives

$$\begin{aligned} \int_{x_0}^{\omega} e^{-\int_{x_0}^t r + \mu_i(j) dj} \frac{\partial c_i(t)}{\partial S} dt - \int_S^{R_i} e^{-\int_{x_0}^t r + \mu_i(j) dj} (1 - \tau + \tau \bar{\mathcal{P}}(t)) w_i(S_i, t) \frac{\partial \ell_i(t)}{\partial S} dt \\ = e^{-\int_{x_0}^S r + \mu_i(j) dj} \frac{\eta - \alpha_i v(\ell(S))}{\lambda_a(x_0)}. \end{aligned} \quad (\text{D.25})$$

Second, we differentiate the budget constraint (10) at age x_0 w.r.t. S

$$\begin{aligned} \int_{x_0}^{\omega} e^{-\int_{x_0}^t r + \mu_i(j) dj} \frac{\partial c_i(t)}{\partial S} dt = \int_S^{R_i} e^{-\int_{x_0}^t r + \mu_i(j) dj} w_i(S, t) \frac{\partial \ell_i(t)}{\partial S} dt \\ + \int_S^{R_i} e^{-\int_{x_0}^t r + \mu_i(j) dj} \frac{\partial w_i(S, t)}{\partial S} \ell_i(t) dt \\ - e^{-\int_{x_0}^S r + \mu_i(j) dj} w_i(S, S) \ell_i(S) + \frac{\partial \text{SSW}_i(x_0)}{\partial S}. \end{aligned} \quad (\text{D.26})$$

From (B.7) we have

$$\text{SSW}_i(x_0) = - \int_S^{R_i} e^{-\int_{x_0}^t r+\mu_i(j) dj} \tau(1 - \mathbf{P}(t)) w_i(S, t) \ell_i(t) dt. \quad (\text{D.27})$$

First, differentiating (D.27) with respect to S gives

$$\begin{aligned} \frac{\partial \text{SSW}_i(x_0)}{\partial S} &= - \int_S^{R_i} e^{-\int_{x_0}^t r+\mu_i(j) dj} \tau(1 - \mathbf{P}(t)) w_i(S, t) \frac{\partial \ell_i(t)}{\partial S} dt \\ &\quad - \int_S^{R_i} e^{-\int_{x_0}^t r+\mu_i(j) dj} \tau(1 - \mathbf{P}(t)) \frac{\partial w_i(S, t)}{\partial S} \ell_i(t) dt \\ &\quad + e^{-\int_{x_0}^S r+\mu_i(j) dj} \tau(1 - \mathbf{P}(S)) w_i(S, S) \ell_i(S) \\ &\quad + \int_S^{R_i} e^{-\int_{x_0}^t r+\mu_i(j) dj} \tau \frac{\partial \mathbf{P}(t)}{\partial S} w_i(S, t) \ell_i(t) dt. \end{aligned} \quad (\text{D.28})$$

Second, differentiating (5) with respect to S gives

$$\frac{\partial \mathbf{P}_i(t)}{\partial S} = \phi \frac{\partial \mathbf{f}_i(R_i)}{\partial \text{pp}_i(R_i)} \frac{\partial \text{pp}_i(R_i)}{\partial S} A_i(R_i, r) e^{\int_t^{R_i} \tilde{r} + \tilde{\mu}(j) - (r + \mu_i(j)) dj}. \quad (\text{D.29})$$

Using (5) and (15) in (D.29) gives

$$\frac{\partial \mathbf{P}_i(t)}{\partial S} = - \frac{1}{\text{pp}_i(R_i)} \frac{\partial \text{pp}_i(R_i)}{\partial S} \varepsilon_i \mathbf{P}_i(t). \quad (\text{D.30})$$

Substituting (D.30) on the third term on the right-hand side of Eq. (D.28), and using the fact that $\mathbf{P}_i(t) = \mathbf{P}_i(R_i) e^{\int_t^{R_i} \tilde{r} + \tilde{\mu}(j) - (r + \mu_i(j)) dj}$ gives, after canceling terms,

$$\frac{\partial \mathbf{P}_i(t)}{\partial S} = - \frac{\partial \text{pp}_i(R_i)}{\partial S} \frac{\varepsilon_i \mathbf{P}_i(R_i)}{\phi} e^{-\int_{x_0}^{R_i} r + \mu_i(j) dj}. \quad (\text{D.31})$$

Differentiating the total pension points at age R_i with respect to S gives

$$\begin{aligned} \frac{\partial \text{pp}_i(R_i)}{\partial S} &= \int_S^{R_i} e^{\int_t^{R_i} \tilde{r} + \tilde{\mu}(j) dj} \phi \tau w_i(S, t) \frac{\partial \ell_i(t)}{\partial S} dt \\ &\quad + \int_S^{R_i} e^{\int_t^{R_i} \tilde{r} + \tilde{\mu}(j) dj} \phi \tau \frac{\partial w_i(S, t)}{\partial S} \ell_i(t) dt \\ &\quad - e^{\int_S^{R_i} \tilde{r} + \tilde{\mu}(j) dj} \phi \tau w_i(S, S) \ell_i(S). \end{aligned} \quad (\text{D.32})$$

Plugging (D.32) in (D.31), and using the fact that $\mathbf{P}_i(R_i) = \mathbf{P}_i(t)e^{-\int_t^{R_i} \tilde{r} + \bar{\mu}(j) - (r + \mu_i(j))dj}$, gives

$$\begin{aligned} \frac{\partial \mathbf{P}_i(t)}{\partial S} &= -\frac{\partial \mathbf{pp}_i(R_i)}{\partial S} \frac{\mathbf{P}_i(R_i)}{\phi} \varepsilon_i e^{-\int_{x_0}^{R_i} r + \mu_i(j) dj} \\ &= -\int_S^{R_i} e^{-\int_{x_0}^t r + \mu_i(j) dj} \tau \varepsilon_i \mathbf{P}_i(t) w_i(S, t) \frac{\partial \ell_i(t)}{\partial S} dt \\ &\quad - \int_S^{R_i} e^{-\int_{x_0}^t r + \mu_i(j) dj} \tau \varepsilon_i \mathbf{P}_i(t) \frac{\partial w_i(S, t)}{\partial S} \ell_i(t) dt \\ &\quad + e^{-\int_{x_0}^S r + \mu_i(j) dj} \tau \varepsilon_i \mathbf{P}_i(S) w_i(S, S) \ell_i(S). \end{aligned} \quad (\text{D.33})$$

Substituting (D.33) in the last term in (D.28) gives

$$\begin{aligned} \frac{\partial \text{SSW}_i(x_0)}{\partial S} &= -\int_{x_0}^{R_i} e^{-\int_{x_0}^t r + \mu_i(j) dj} \tau (1 - \bar{\mathcal{P}}_i(t)) w_i(S, t) \frac{\partial \ell_i(t)}{\partial S} dt \\ &\quad - \int_{x_0}^{R_i} e^{-\int_{x_0}^t r + \mu_i(j) dj} \tau (1 - \bar{\mathcal{P}}_i(t)) \frac{\partial w_i(S, t)}{\partial S} \ell_i(t) dt \\ &\quad + e^{-\int_{x_0}^S r + \mu_i(j) dj} \tau (1 - \bar{\mathcal{P}}_i(S)) w_i(S, S) \ell_i(S). \end{aligned} \quad (\text{D.34})$$

By plugging (D.34) into (D.26) we have

$$\begin{aligned} &\int_{x_0}^{\omega} e^{-\int_{x_0}^t r + \mu_i(j) dj} \frac{\partial c_i(t)}{\partial S} dt - \int_S^{R_i} e^{-\int_{x_0}^t r + \mu_i(j) dj} (1 - \tau + \tau \bar{\mathcal{P}}_i(t)) w_i(S, t) \frac{\partial \ell_i(t)}{\partial S} dt \\ &= \int_S^{R_i} e^{-\int_{x_0}^t r + \mu_i(j) dj} (1 - \tau + \tau \bar{\mathcal{P}}_i(t)) \frac{\partial w_i(S, t)}{\partial S} \ell_i(t) dt \\ &\quad - e^{-\int_{x_0}^S r + \mu_i(j) dj} (1 - \tau + \tau \bar{\mathcal{P}}_i(t)) w_i(S, S) \ell_i(S). \end{aligned}$$

Using the fact that the left-hand side of (D.25) and (D.31) are equal, then

$$\begin{aligned} &e^{-\int_{x_0}^S r + \mu_i(j) dj} \frac{\eta - \alpha_i v(\ell_i(S))}{\lambda_a(x_0)} \\ &= \int_S^{R_i} e^{-\int_{x_0}^t r + \mu_i(j) dj} (1 - \tau + \tau \bar{\mathcal{P}}_i(t)) \frac{\partial w_i(S, t)}{\partial S} \ell_i(t) dt \\ &\quad - e^{-\int_{x_0}^S r + \mu_i(j) dj} (1 - \tau + \tau \bar{\mathcal{P}}_i(t)) w_i(S, S) \ell_i(S). \end{aligned} \quad (\text{D.35})$$

Now, differentiating $w_i(S, t)$ w.r.t. S gives

$$\begin{aligned} \frac{\partial w_i(S, t)}{\partial S} &= \frac{\partial \bar{w}(t - S)}{\partial S} h_i(S) + \bar{w}(t - S) \frac{\partial h_i(S)}{\partial S} \\ &= -\frac{1}{\bar{w}(t - S)} \frac{\partial \bar{w}(t - S)}{\partial t} w_i(S, t) + \frac{1}{h_i(S)} \frac{\partial h_i(S)}{\partial S} w_i(t, S). \end{aligned} \quad (\text{D.36})$$

Third, we use (D.36) in (D.35)

$$\begin{aligned}
& e^{-\int_{x_0}^S \rho + \mu_i(j) dj} \frac{\eta - \alpha_i v(\ell_i(S))}{\lambda_a(x_0)} \\
&= \frac{\partial h_i(S)}{\partial S} \int_S^{R_i} e^{-\int_{x_0}^t r + \mu_i(j) dj} (1 - \tau + \tau \bar{\mathcal{P}}_i(t)) w_i(S, t) \ell_i(t) dt \\
&- \int_S^{R_i} e^{-\int_{x_0}^t r + \mu_i(j) dj} \frac{\partial \bar{w}(t-S)}{\partial t} (1 - \tau + \tau \bar{\mathcal{P}}_i(t)) w_i(S, t) \ell_i(t) dt \\
&- e^{-\int_{x_0}^S r + \mu_i(j) dj} (1 - \tau + \tau \bar{\mathcal{P}}_i(t)) w_i(S, S) \ell_i(S). \quad (\text{D.37})
\end{aligned}$$

Therefore, after rearranging terms, the optimal length of schooling satisfies the following condition

$$\begin{aligned}
& \frac{\partial h_i(S)}{\partial S} \int_S^{R_i} e^{-\int_{x_0}^t r + \mu_i(j) dj} (1 - \tau + \tau \bar{\mathcal{P}}_i(t)) w_i(S, t) \ell_i(t) dt \\
&= \int_S^{R_i} e^{-\int_{x_0}^t r + \mu_i(j) dj} \frac{\partial \bar{w}(t-S)}{\partial t} (1 - \tau + \tau \bar{\mathcal{P}}_i(t)) w_i(S, t) \ell_i(t) dt \\
&+ e^{-\int_{x_0}^S r + \mu_i(j) dj} (1 - \tau + \tau \bar{\mathcal{P}}_i(t)) w_i(S, S) \ell_i(S) \\
&+ e^{-\int_{x_0}^S r + \mu_i(j) dj} \frac{\eta - \alpha_i v(\ell_i(S))}{\lambda_a(x_0)}. \quad (\text{D.38})
\end{aligned}$$

Defining human capital at age S net of effective labor income taxes as

$$W_i(S, R_i) = \int_S^{R_i} e^{-\int_S^t r + \mu_i(j) dj} (1 - \tau + \tau \bar{\mathcal{P}}_i(t)) w_i(S, t) \ell_i(t) dt \quad (\text{D.39})$$

and dividing both sides of (D.38) by $W_i(S, R_i)$, multiplying by $e^{\int_{x_0}^S r + \mu_i(j) dj}$, we obtain the optimal length of schooling condition

$$r_i^h(S) = \bar{r}_i(S, R_i) + \frac{\eta - \alpha_i v(\ell_i(S))}{U'(c_i(S)) W_i(S, R_i)}. \quad (\text{D.40})$$

$r_i^h(S)$ is the return to education at age S for an individual of type i

$$r_i^h(S) = \frac{1}{h_i(S)} \frac{\partial h_i(S)}{\partial S}, \quad (\text{D.41})$$

$\bar{r}_i(S, R)$ is the rate of return lost from not working at age S or the marginal cost of the S th unit of schooling for an individual of type i

$$\bar{r}_i(S, R_i) = \int_S^{R_i} \frac{\partial \bar{w}(t-S)}{\partial t} \psi_i(t) dt + \psi_i(S), \quad (\text{D.42})$$

where

$$\psi_i(t) = \frac{e^{-\int_{x_0}^t r + \mu_i(j) dj} (1 - \tau + \tau \bar{\mathcal{P}}_i(t)) w_i(S, t) \ell_i(t)}{\int_S^{R_i} e^{-\int_{x_0}^u r + \mu_i(j) dj} (1 - \tau + \tau \bar{\mathcal{P}}_i(u)) w_i(S, u) \ell_i(u) du}. \quad (\text{D.43})$$

From (D.42) and (D.43) we have $\int_S^{R_i} \psi_i(t) dt = 1$ and $\lim_{S \rightarrow R_i} \bar{r}_i(S, R_i) = 1$.

D.3 Optimal retirement age (R_i)

Similar to the previous subsection we start assuming that the optimal length of schooling S_i is given. Then, we differentiate the expected utility $V_i(x_0)$ w.r.t. the optimal retirement age R and equate the result to the derivative of the lifetime budget constrain w.r.t. to the optimal retirement age.

Proof. Given an optimal length of schooling S_i we first differentiate the expected utility $V_i(x_0)$ w.r.t. R and making it equal to zero

$$\begin{aligned} \int_{x_0}^{\omega} e^{-\int_{x_0}^t \rho + \mu_i(j) dj} U'(c_i(t)) \frac{\partial c_i(t)}{\partial R} dt - \alpha_i \int_{S_i}^R e^{-\int_{x_0}^t \rho + \mu_i(j) dj} v'(\ell_i(t)) \frac{\partial \ell_i(t)}{\partial R} dt \\ = e^{-\int_{x_0}^R \rho + \mu_i(j) dj} (\alpha_i v(\ell_i(R)) + \varphi(R)). \end{aligned} \quad (\text{D.44})$$

Substituting the FOCs in the previous equation gives

$$\begin{aligned} \int_{x_0}^{\omega} \lambda_a(t) \frac{\partial c_i(t)}{\partial R} dt - \int_{S_i}^R \lambda_a(t) (1 - \tau + \tau \bar{\mathcal{P}}_i(t)) w_i(S_i, t) \frac{\partial \ell_i(t)}{\partial R} dt \\ = e^{-\int_{x_0}^R \rho + \mu_i(j) dj} (\alpha_i v(\ell_i(R)) + \varphi(R)). \end{aligned} \quad (\text{D.45})$$

Using the envelope conditions on assets, which gives $\lambda_a(t) = \lambda_a(x_0) e^{-\int_{x_0}^t r + \mu_i(j) dj}$, dividing both sides of the equation by $\lambda_a(x_0)$ and rearranging terms gives

$$\begin{aligned} \int_{x_0}^{\omega} e^{-\int_{x_0}^t r + \mu_i(j) dj} \frac{\partial c_i(t)}{\partial R} dt - \int_{S_i}^R e^{-\int_{x_0}^t r + \mu_i(j) dj} (1 - \tau + \tau \bar{\mathcal{P}}_i(t)) w_i(S_i, t) \frac{\partial \ell_i(t)}{\partial R} dt \\ = e^{-\int_{x_0}^R \rho + \mu_i(j) dj} \frac{\alpha_i v(\ell_i(R)) + \varphi(R)}{\lambda_a(x_0)}. \end{aligned} \quad (\text{D.46})$$

Second, we differentiate the budget constraint (10) at age x_0 w.r.t. R

$$\begin{aligned} \int_{x_0}^{\omega} e^{-\int_{x_0}^t r + \mu_i(j) dj} \frac{\partial c_i(t)}{\partial R} dt = \int_{S_i}^R e^{-\int_{x_0}^t r + \mu_i(j) dj} w_i(S_i, t) \frac{\partial \ell_i(t)}{\partial R} dt \\ + e^{-\int_{x_0}^R r + \mu_i(j) dj} w_i(S_i, R) \ell_i(R) + \frac{\partial \text{SSW}_i(x_0)}{\partial R}. \end{aligned} \quad (\text{D.47})$$

Differentiating (B.7) at age x_0 with respect to R gives

$$\begin{aligned} \frac{\partial \text{SSW}_i(x_0)}{\partial R} &= - \int_{S_i}^R e^{-\int_{x_0}^t r + \mu_i(j) dj} \tau(1 - \mathbf{P}_i(t)) w_i(S_i, t) \frac{\partial \ell_i(t)}{\partial R} dt \\ &\quad - e^{-\int_{x_0}^R r + \mu_i(j) dj} \tau(1 - \mathbf{P}_i(R)) w_i(S_i, R) \ell_i(R) \\ &\quad + \int_{S_i}^R e^{-\int_{x_0}^t r + \mu_i(j) dj} \tau \frac{\partial \mathbf{P}_i(t)}{\partial R} w_i(S_i, t) \ell_i(t) dt. \end{aligned} \quad (\text{D.48})$$

Now, by differentiating (5) with respect to R we have

$$\begin{aligned} \frac{\partial \mathbf{P}_i(t)}{\partial R} &= \mathbf{P}_i(t) \left(\frac{1}{\mathbf{f}_i(R)} \frac{\partial \mathbf{f}_i(R)}{\partial R} + \frac{1}{A_i(R, r)} \frac{\partial A_i(R, r)}{\partial R} + \tilde{r} + \tilde{\mu}(R) - (r + \mu_i(R)) \right) \\ &= \mathbf{P}_i(t) \left(\frac{1}{\mathbf{f}_i(R)} \frac{\partial \mathbf{f}_i(R)}{\partial R} - \frac{1}{A_i(R, r)} + \tilde{r} + \tilde{\mu}(R) \right). \end{aligned} \quad (\text{D.49})$$

Given that $\mathbf{f}_i(R)$ depends on both R and \mathbf{pp} , the total derivative of $\mathbf{f}_i(R)$ with respect to R is

$$\frac{\partial \mathbf{f}_i(R)}{\partial R} = \frac{\partial \mathbf{f}_i(R)}{\partial R} \Big|_{\mathbf{pp}} + \frac{\partial \mathbf{f}_i(R)}{\partial \mathbf{pp}_i(R)} \frac{\partial \mathbf{pp}_i(R)}{\partial R}. \quad (\text{D.50})$$

Thus, from (15) and (D.50), Eq. (D.49) can be rewritten as follow

$$\frac{\partial \mathbf{P}_i(t)}{\partial R} = \mathbf{P}_i(t) \left(\frac{1}{\mathbf{f}_i(R)} \frac{\partial \mathbf{f}_i(R)}{\partial R} \Big|_{\mathbf{pp}} - \frac{1}{A_i(R, r)} + \tilde{r} + \tilde{\mu}(R) \right) - \mathbf{P}_i(t) \varepsilon_i \frac{1}{\mathbf{pp}_i(R)} \frac{\partial \mathbf{pp}_i(R)}{\partial R}. \quad (\text{D.51})$$

Substituting (D.51) on (D.48) we get

$$\begin{aligned} \frac{\partial \text{SSW}_i(x_0)}{\partial R} &= - \int_{S_i}^R e^{-\int_{x_0}^t r + \mu_i(j) dj} \tau(1 - \mathbf{P}_i(t)) w_i(S_i, t) \frac{\partial \ell_i(t)}{\partial R} dt \\ &\quad - e^{-\int_{x_0}^R r + \mu_i(j) dj} \tau(1 - \mathbf{P}_i(R)) w_i(S_i, R) \ell_i(R) \\ &\quad + \left(\frac{1}{\mathbf{f}_i(R)} \frac{\partial \mathbf{f}_i(R)}{\partial R} \Big|_{\mathbf{pp}} - \frac{1}{A_i(R, r)} + \tilde{r} + \tilde{\mu}(R) \right) \int_{S_i}^R e^{-\int_{x_0}^t r + \mu_i(j) dj} \tau \mathbf{P}_i(t) w_i(S_i, t) \ell_i(t) dt \\ &\quad - \frac{1}{\mathbf{pp}_i(R)} \frac{\partial \mathbf{pp}_i(R)}{\partial R} \int_{S_i}^R e^{-\int_{x_0}^t r + \mu_i(j) dj} \tau \mathbf{P}_i(t) \varepsilon_i w_i(S_i, t) \ell_i(t) dt. \end{aligned} \quad (\text{D.52})$$

Using the relation $\mathbf{P}_i(t) = \mathbf{P}_i(R)e^{\int_t^R \tilde{r} + \tilde{\mu}(j) - (r + \mu_i(j))dj}$ in the last two terms of (D.52), and rearranging terms, we have

$$\begin{aligned} \frac{\partial \text{SSW}_i(x_0)}{\partial R} &= - \int_{S_i}^R e^{-\int_{x_0}^t r + \mu_i(j) dj} \tau (1 - \mathbf{P}_i(t)) w_i(S_i, t) \frac{\partial \ell_i(t)}{\partial R} dt \\ &\quad - e^{-\int_{x_0}^R r + \mu_i(j) dj} \tau (1 - \mathbf{P}_i(R)) w_i(S_i, R) \ell_i(R) \\ &\quad + e^{-\int_{x_0}^R r + \mu_i(j) dj} \left(\frac{1}{\mathbf{f}_i(R)} \frac{\partial \mathbf{f}_i(R)}{\partial R} \Big|_{\text{pp}} - \frac{1}{A_i(R, r)} + \tilde{r} + \tilde{\mu}(R) \right) \mathbf{P}_i(R) \frac{\text{pp}_i(R)}{\phi} \\ &\quad - e^{-\int_{x_0}^R r + \mu_i(j) dj} \frac{\partial \text{pp}(R)}{\partial R} \frac{\mathbf{P}_i(R) \varepsilon_i}{\phi}. \quad (\text{D.53}) \end{aligned}$$

Now, from (B.2) we differentiate the total pension points at age R with respect to R , i.e. $\frac{\partial \text{pp}(R)}{\partial R}$, which gives

$$\begin{aligned} \frac{\partial \text{pp}_i(R_i)}{\partial S} &= \int_{S_i}^R e^{\int_t^R \tilde{r} + \tilde{\mu}(j) dj} \phi \tau w_i(S, t) \frac{\partial \ell_i(t)}{\partial R} dt \\ &\quad + (\tilde{r} + \tilde{\mu}(R)) \text{pp}_i(R) + \phi \tau w_i(S_i, R) \ell_i(R). \quad (\text{D.54}) \end{aligned}$$

Plugging (D.54) in (D.53) and rearranging terms gives

$$\begin{aligned} \frac{\partial \text{SSW}_i(x_0)}{\partial R} &= - \int_{S_i}^R e^{-\int_{x_0}^t r + \mu_i(j) dj} \tau (1 - \mathbf{P}_i(t)(1 - \varepsilon_i)) w_i(S_i, t) \frac{\partial \ell_i(t)}{\partial R} dt \\ &\quad - e^{-\int_{x_0}^R r + \mu_i(j) dj} \tau (1 - \mathbf{P}_i(R)(1 - \varepsilon_i)) w_i(S_i, R) \ell_i(R) \\ &\quad + e^{-\int_{x_0}^R r + \mu_i(j) dj} \left(\frac{1}{\mathbf{f}_i(R)} \frac{\partial \mathbf{f}_i(R)}{\partial R} \Big|_{\text{pp}} - \frac{1}{A_i(R, r)} + (\tilde{r} + \tilde{\mu}(R))(1 - \varepsilon_i) \right) \mathbf{P}_i(R) \frac{\text{pp}_i(R)}{\phi}. \quad (\text{D.55}) \end{aligned}$$

Next, plugging (D.55) in (D.47) gives

$$\begin{aligned} &\int_{x_0}^{\omega} e^{-\int_{x_0}^t r + \mu_i(j) dj} \frac{\partial c_i(t)}{\partial R} dt - \int_{S_i}^R e^{-\int_{x_0}^t r + \mu_i(j) dj} (1 - \tau + \tau \bar{\mathbf{P}}_i(t)) w_i(S_i, t) \frac{\partial \ell_i(t)}{\partial R} dt \\ &= e^{-\int_{x_0}^R r + \mu_i(j) dj} (1 - \tau + \tau \bar{\mathbf{P}}_i(t)) w_i(S_i, R) \ell_i(R) \\ &\quad + e^{-\int_{x_0}^R r + \mu_i(j) dj} \left(\frac{1}{\mathbf{f}_i(R)} \frac{\partial \mathbf{f}_i(R)}{\partial R} \Big|_{\text{pp}} - \frac{1}{A_i(R, r)} + (\tilde{r} + \tilde{\mu}(R))(1 - \varepsilon_i) \right) \mathbf{P}_i(R) \frac{\text{pp}_i(R)}{\phi}. \quad (\text{D.56}) \end{aligned}$$

Using the fact that the left-hand side of (D.3) and (D.56) are equal, we obtain

$$\begin{aligned}
& e^{-\int_{x_0}^R \rho + \mu_i(j) dj} \frac{\alpha_i v(\ell_i(R)) + \varphi(R)}{\lambda_a(x_0)} = e^{-\int_{x_0}^R r + \mu_i(j) dj} (1 - \tau + \tau \bar{\mathcal{P}}_i(t)) w_i(S_i, R) \ell_i(R) \\
& + e^{-\int_{x_0}^R r + \mu_i(j) dj} \left(\frac{1}{\mathbf{f}_i(R)} \frac{\partial \mathbf{f}_i(R)}{\partial R} \Big|_{\text{pp}} - \frac{1}{A_i(R, r)} + (\tilde{r} + \tilde{\mu}(R))(1 - \varepsilon_i) \right) \mathbf{P}_i(R) \frac{\text{pp}_i(R)}{\phi}.
\end{aligned} \tag{D.57}$$

Multiplying both sides of (D.57) by $e^{\int_{x_0}^R r + \mu_i(j) dj}$ gives

$$\begin{aligned}
& e^{(r-\rho)(R-x_0)} \frac{\alpha_i v(\ell_i(R)) + \varphi(R)}{\lambda_a(x_0)} = (1 - \tau + \tau \bar{\mathcal{P}}_i(t)) w_i(S_i, R) \ell_i(R) \\
& + \left(\frac{1}{\mathbf{f}_i(R)} \frac{\partial \mathbf{f}_i(R)}{\partial R} \Big|_{\text{pp}} - \frac{1}{A_i(R, r)} + (\tilde{r} + \tilde{\mu}(R))(1 - \varepsilon_i) \right) \mathbf{P}_i(R) \frac{\text{pp}_i(R)}{\phi}.
\end{aligned} \tag{D.58}$$

Defining the implicit tax rate on retirement $\tau_i^{GW}(R)$ as

$$\begin{aligned}
& \tau_i^{GW}(R) = \tau(1 - \bar{\mathcal{P}}_i(t)) \\
& - \left(\frac{1}{\mathbf{f}_i(R)} \frac{\partial \mathbf{f}_i(R)}{\partial R} \Big|_{\text{pp}} - \frac{1}{A_i(R, r)} + (\tilde{r} + \tilde{\mu}(R))(1 - \varepsilon_i) \right) \frac{\mathbf{P}_i(R) \frac{\text{pp}_i(R)}{\phi}}{w_i(S_i, R) \ell_i(R)},
\end{aligned} \tag{D.59}$$

where the first term is the implicit tax paid by working one additional period and the second term is the relative change in the social security wealth caused by postponing the retirement age.

Then, using (D.59) and the fact that $U'(c_i(R)) = e^{-(r-\rho)(R-x_0)} \lambda_a(x_0)$ we obtain that the optimal retirement age condition satisfies

$$\alpha_i v(\ell_i(R)) + \varphi(R) = U'(c_i(R)) w_i(S_i, R) \ell_i(R) (1 - \tau_i^{GW}(R)), \tag{D.60}$$

which coincides with (18). ■

E Additional simulated data

Length of schooling S_i .

Table 6: Optimal length of schooling by income quintile (S_i), US male birth cohorts 1930 and 1960

	Defined Contribution (NDC)			Defined Benefit		
	Avg. LT	Corrected Avg. LT	i -th LT	Non- progressive	Progressive	Progressive Corrected
	NDC-I	NDC-II	NDC-III	DB-I	DB-II	DB-III
Cohort 1930						
Quintile 1	11.4	11.4	11.5	11.9	11.3	11.3
Quintile 2	11.8	11.8	11.9	12.3	11.8	11.7
Quintile 3	12.4	12.4	12.4	12.9	12.3	12.3
Quintile 4	14.4	14.4	14.4	14.6	13.2	13.3
Quintile 5	15.8	15.7	15.7	15.9	14.2	14.3
Cohort 1960						
Quintile 1	11.3	11.3	11.3	11.3	10.6	10.4
Quintile 2	12.4	12.5	12.6	13.0	12.2	12.4
Quintile 3	15.7	15.7	15.7	16.3	14.7	14.9
Quintile 4	18.8	18.8	18.8	18.8	17.2	17.4
Quintile 5	19.0	19.0	19.0	19.0	17.5	17.7

Retirement ages R_i .

Table 7: Optimal retirement age by income quintile (R_i), US male birth cohorts 1930 and 1960

	Defined Contribution (NDC)			Defined Benefit		
	Avg. LT	Corrected Avg. LT	i -th LT	Non- progressive	Progressive	Progressive Corrected
	NDC-I	NDC-II	NDC-III	DB-I	DB-II	DB-III
Cohort 1930						
Quintile 1	60.9	60.9	61.1	63.2	62.7	62.8
Quintile 2	61.5	61.5	61.7	63.8	63.3	63.3
Quintile 3	62.5	62.5	62.5	64.6	64.0	64.1
Quintile 4	64.6	64.6	64.5	65.3	64.7	64.7
Quintile 5	66.8	66.6	66.4	67.1	65.2	65.2
Cohort 1960						
Quintile 1	60.1	60.2	60.6	62.0 [†]	62.0 [†]	62.0 [†]
Quintile 2	62.2	62.4	62.7	64.8	63.4	64.7
Quintile 3	67.3	67.5	67.4	69.8	68.4	68.6
Quintile 4	70.0	70.0	70.0	70.0	70.0	70.0
Quintile 5	70.0	70.0	70.0	70.0	70.0	70.0

Notes: Symbol '†' represents cases in which individuals exit the labor market before they start claiming retirement benefits. In particular, individuals subject to the mortality regime of the 1960 cohort retire on average at age 61.1 in the DB-I system, 60.5 in the DB-II system and at age 60.4 in the DB-III system.

Hours worked ℓ_i .

Table 8: Average optimal hours worked over the working life by income quintile (S_i), US male birth cohorts 1930 and 1960

	Defined Contribution (NDC)			Defined Benefit		
	Avg. LT	Corrected Avg. LT	i -th LT	Non- progressive	Progressive	Progressive Corrected
	NDC-I	NDC-II	NDC-III	DB-I	DB-II	DB-III
Cohort 1930						
Quintile 1	0.253	0.253	0.253	0.252	0.250	0.250
Quintile 2	0.268	0.268	0.268	0.267	0.265	0.265
Quintile 3	0.278	0.278	0.278	0.277	0.275	0.275
Quintile 4	0.286	0.286	0.286	0.285	0.280	0.280
Quintile 5	0.287	0.287	0.287	0.287	0.282	0.283
Cohort 1960						
Quintile 1	0.252	0.252	0.252	0.252	0.250	0.249
Quintile 2	0.269	0.269	0.269	0.268	0.266	0.266
Quintile 3	0.283	0.283	0.283	0.282	0.278	0.278
Quintile 4	0.293	0.293	0.293	0.293	0.286	0.287
Quintile 5	0.294	0.294	0.294	0.294	0.287	0.288

Present value of lifetime benefits.

Table 9: Present value of lifetime benefits at age 50 by income quintile and pension system, US males, birth cohorts 1930 and 1960 (in \$1 000s)

	Defined Contribution (NDC)			Defined Benefit (DB)		
	Avg. LT	Corrected Avg. LT	i -th LT	Non- progressive	Progressive	Progressive Corrected
	NDC-I	NDC-II	NDC-III	DB-I	DB-II	DB-III
Cohort 1930						
Quintile 1	127.02	138.55	147.60	133.83	132.05	142.90
Quintile 2	143.21	153.11	161.27	154.97	149.70	157.25
Quintile 3	163.27	169.63	173.41	182.28	170.34	175.56
Quintile 4	217.69	214.81	211.62	240.67	197.60	193.42
Quintile 5	270.78	247.35	234.01	300.49	224.57	205.72
Cohort 1960						
Quintile 1	117.55	155.42	177.58	95.91	107.58	136.57
Quintile 2	150.46	188.10	208.66	146.66	143.91	182.55
Quintile 3	253.62	270.85	274.81	307.32	258.47	270.30
Quintile 4	408.52	376.59	358.61	474.74	356.91	320.30
Quintile 5	431.70	385.33	363.10	501.68	375.31	326.15

Lifetime wealth.

Table 10: Distribution of lifetime wealth (LW) at age 14 by income quintile and mortality regime in the NDC-III system (in \$1 000s)

	Cohort 1930			Cohort 1960		
	SSW	HK	LW	SSW	HK	LW
	I	II	III=I+II	I	II	III=I+II
Quintile 1	-15.61	503.23	487.62	-18.51	489.98	471.48
Quintile 2	-17.24	554.66	537.42	-22.67	594.77	572.10
Quintile 3	-18.84	604.38	585.53	-32.24	826.47	794.22
Quintile 4	-22.96	740.85	717.89	-43.29	1 108.48	1 065.20
Quintile 5	-26.31	842.44	816.14	-44.25	1 129.04	1 084.79

Notes: SSW stands for social security wealth, HK denotes the stock of human capital, and LW is the lifetime wealth.

Table 10 reports the lifetime wealth (detrended by productivity) by income quintile relative to that obtained for the income group q3 under the mortality regime 1930. We can see in Tab. 10 that the higher income quintiles q3–q5 experience an average increase over twenty percent in their lifetime wealth with the more unequal mortality regime, q2 experiences an increase of seven percent in the lifetime wealth, and q1 has five percent less wealth. Moreover, since the NDC-III system provides the same internal rate of return across income groups, we have that the ratio between the social security wealth and the stock of human capital is the same across income group.

Behavioral effects

Figure 6 shows the difference in lifetime wealth between a model with and without behavioral reactions by pension system and mortality regime. Figure 6 shows that those in the highest income quintiles experience a significant increase in their stock of human capital (HK) when they react to demographic and economic changes, while those in the lowest income quintiles experience almost no change in their stock of human capital.

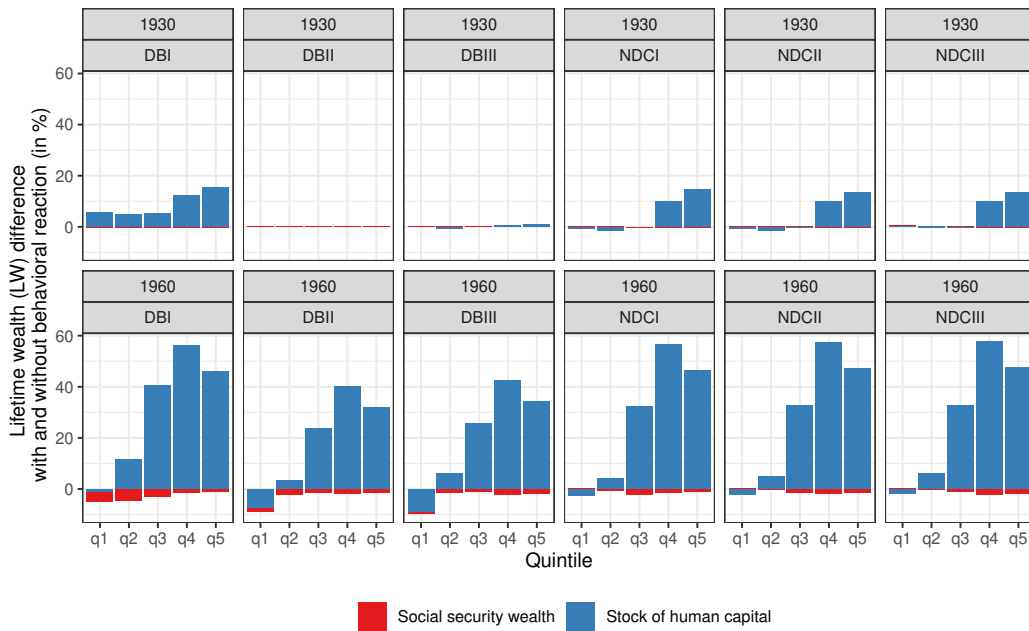


Figure 6: Impact of behavior on lifetime wealth by income quintile and pension system. US males, mortality regimes of birth cohorts 1930 (top panels) and 1960 (bottom panels).

Figure 7 shows the welfare difference of a model with behavioral reactions relative to that without behavioral reactions for each pension system and mortality regime. This figure clearly shows that when individuals of different quintiles can optimally react to changes in the economic environment, those with higher income and longer life expectancy experience welfare gains, while those with lower income and shorter life expectancy either have small welfare gains in NDC systems or have welfare losses in DB systems. The lower increase in welfare for those in quintile 5 relative to those in quintile 4 is caused by the upper bound in the retirement age.

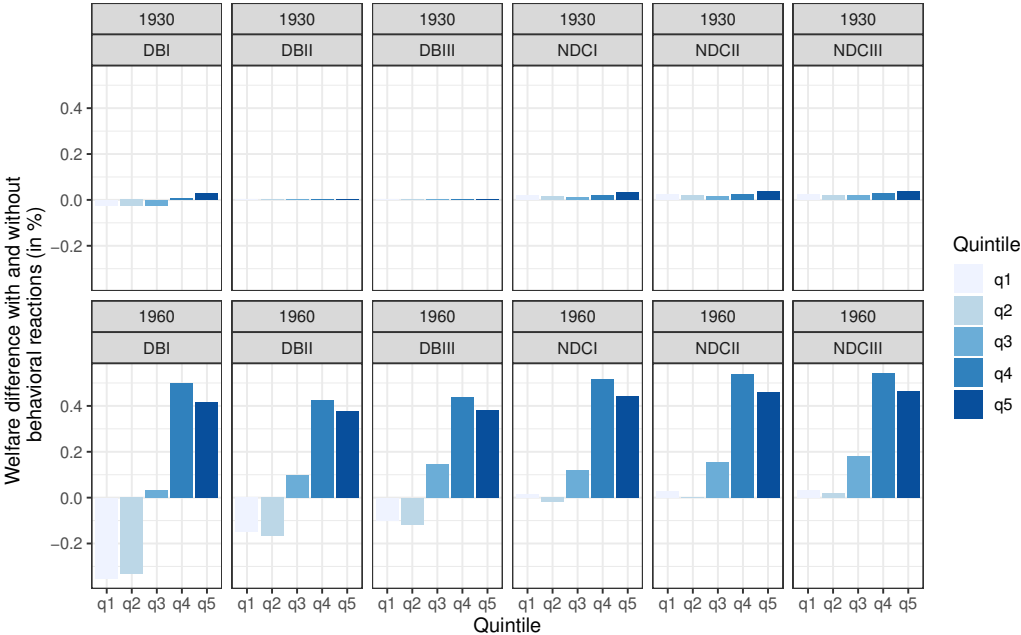


Figure 7: Impact of behavior on welfare by income quintile and pension system. US males, mortality regimes of birth cohorts 1930 (top panels) and 1960 (bottom panels).