Redistributive effects of different pension structures when longevity varies by socioeconomic status in a general equilibrium setting*

Miguel Sanchez-Romero[1,3], Ronald D. Lee[2], and Alexia Prskawetz[1,3]

[1]Wittgenstein Centre of Demography and Global Human Capital (IIASA, VID/ÖAW, WU), Austria
[2]University of California, Berkeley (UCB), US
[3]Institute of Statistics and Mathematical Methods in Economics, Research Unit Economics, TU Wien, Austria

September 19, 2018

Abstract

Differences in life expectancy between high and low socioeconomic groups are often large and have widened in many countries in recent decades. In the US, this difference may now be as large as ten to 14 years. Such longevity gaps by socioeconomic status strongly affect the actuarial fairness and progressivity of many public pension systems, suggesting a need for policy reforms to address this issue. However, behavioral responses to longevity differences and policy change complicate the analysis of possible reforms. Here we consider how some pension reforms would perform in a general equilibrium setting when there is heterogeneity in both longevity and ability. We evaluate the redistributive effects of three Notional Defined Contribution plans and three Defined Benefit plans, calibrated on the US case, drawing on a recent National Research Council study and other information.

Keywords: Human capital, Longevity, Inequality, Life cycle, Social Security
JEL codes: E24, J10, J18, H55

*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".

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.

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 general equilibrium 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), and the parameters of our theoretical model are calibrated to match key aspects of the NASEM findings. We focus on redistributive effects of six different public pension systems: three Notional Defined Contribution (NDC) plans and three Defined Benefit (DB) plans. One version of these NDC and DB plans ignores mortality differences at retirement. 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 NASEM (2015) study of 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 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 the retirement, 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 cohort in the US, individuals would have retired between ages 61 and 64 in NDC plans, whereas individuals would have retired on average one year later in DB plans. This difference in the retirement age raises the marginal benefit of education in DB plans. Hence, the average number of years in schooling and the stock of human capital is higher in DB plans. However, we find that in 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 at retirement.

Throughout the article, we will focus on how the six public pension systems redistributive income across income groups and how individuals response to alternative pension settings. 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 assumption, data used, and parameters 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

Individuals. We assume that the mortality of each individual is completely determined by their lifetime income. We denote by $\mathcal{I} = \{1, 2, \ldots, I\}$ the set of I income levels. Let the probability of surviving to age x of an individual belonging to group $i \in \mathcal{I}$ be

$$p_i(x) = e^{-\int_0^x \mu_i(t)dt}, \tag{1}$$

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 (with $\partial\mu_i(t)/\partial t \geq 0$ for $t \geq \bar{x}$). The life expectancy at age x of an individual belonging to group i is defined as

$$\mathsf{e}_i(x) = \int_x^\omega \frac{p_i(t)}{p_i(x)} dt. \tag{2}$$

Fig. 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 by the NASEM (2015).[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 difference in life expectancy between these two income groups widens for the cohort born in 1960. In particular, the difference becomes as high as 16.2, 11.9, 9.4 years at ages 15, 50, and 65, respectively.

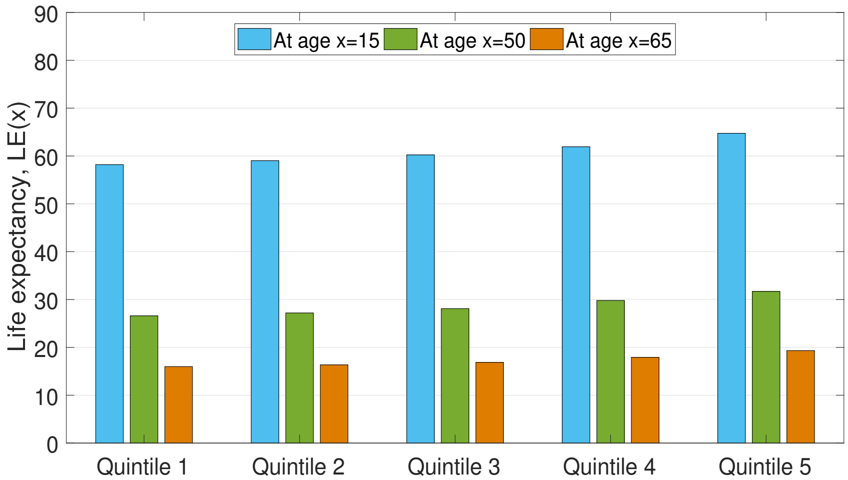
Population. To simplify the demographic analysis, we assume each longevity group grows steadily at a rate n and that the total number of births across longevity groups is the same.[2] As a consequence, the total population size at time t is

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

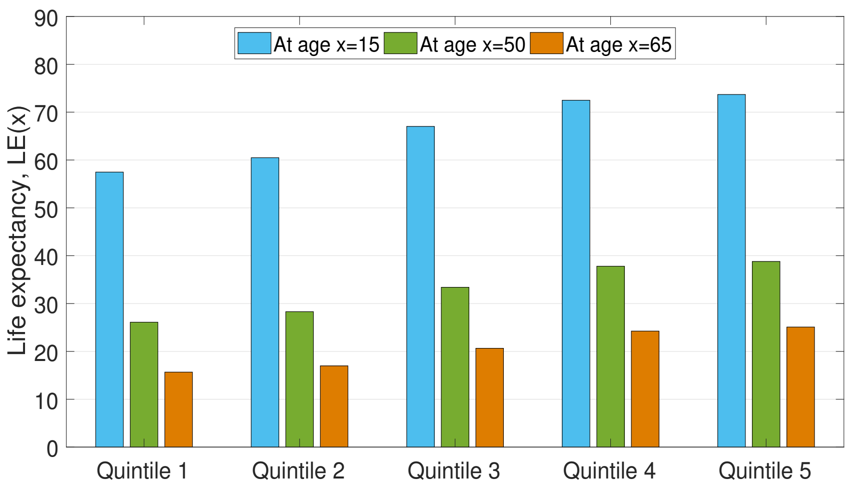
where $B(t)$ is the total number of births at time t, $p(x)$ is the average survival, which implies that the average mortality hazard rate at age x, denoted by $\mu(x)$, is $\sum_{i\in\mathcal{I}} \mu_i(x)p_i(x) / \sum_{i\in\mathcal{I}} p_i(x)$. Thus, the existence of different longevity groups implies that the average mortality hazard rate is biased with age towards the mortality hazard rate of higher income individuals.

[1] We calculate the survival probabilities associated to each birth cohort and income quintile by finding the cohort-life table from the US Social Security Administration (SSA) that matches the life expectancy at age 65 by birth cohort and income quintile reported by NASEM (2015). See https://www.ssa.gov/oact/NOTES/as120/LifeTables_Body.html.

[2] Thus, we are implicitly assuming that fertility is higher for lower income groups, to overcome the lower proportion of females surviving through the reproductive ages.



(a) Cohort 1930, US males



(b) Cohort 1960, US males

Figure 1: Life expectancy at ages 15, 50, and 65 by income quintile, US males, cohorts 1930 and 1960.

Note: Authors' estimates based on data reported in the report by the NASEM (2015).

3 The pension model

The aim of the paper is to analyze the redistributive impact that different pension plans have on life cycle decisions of heterogeneous individuals by income quintile. 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 in the pension system disability benefits, survivor benefits, and widowhood benefits. As a consequence, the pension system acts as an insurance institution that only pays benefits to those workers that survive to retirement. Let us assume that by contributing to the pension system the amount $\tau y(t)$, where τ is the contribution rate and $y(t)$ is the labor income subject to payroll tax, workers gain pension points, $\mathsf{pp}(t)$, that entitle them to receive a pension benefit upon retirement. Suppose that workers earn ϕ pension points per unit of social contribution paid. Moreover, let us assume pension points are capitalized, or indexed, according to r plus a mortality risk premium. The risk premium arises because we exclude benefits from disability, survivor, and widowhood. The capitalization factor r is assumed to be equal or lower than the market interest rate, which we denote by r. Most of the capitalization factors, or indexes, applied in pension systems fit into one of the following three cases: (i) when $\mathsf{r} = 0$, past contributions are only adjusted for inflation; (ii) when $\mathsf{r} = r$, past contributions are invested in the market and capitalized according to the interest rate r (i.e., funded system); and (iii) when $\mathsf{r} = n + g$, 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 growth rate of a PAYG pension system (Samuelson, 1958).[3] The amount of pension points earned depends on whether the system is DB or DC. In a DC system, the pension points earned each period are equal to the contribution paid (i.e., $\phi = 1$), whereas in a DB system the pension points are equal to the yearly pension benefit accrual, which is a fraction (ϱ) of the labor income earned or a fraction (ϱ/τ) of the contributions paid (i.e., $\phi = \varrho/\tau$); that is, pension

[3] Samuelson (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.

benefits built up in a year-to-year basis. Thus, the total number of pension points accumulated at the exact age $x \geq x_0$ by an individual of type $i \in \mathcal{I}$ in any pension system can be formulated as follows

$$\mathsf{pp}_{is}(x, x_0; \mathsf{r}) \equiv \mathsf{pp}_{is}(x) = \phi \int_{x_0}^{x} e^{\mathsf{r}(x-t)} \frac{p_s(t)}{p_s(x)} \tau y_i(t) dt \text{ with } \mathsf{pp}_{is}(x_0) = 0, \quad (4)$$

where x_0 is the minimum working age, r is the capitalization factor, $p_s(x)$ is the survival probability to age x used by the pension system, $y_i(t)$ is the labor income of a worker belonging to group i at age t, and $\mathcal{I}$ is the set of possible income groups. Note that the second subscript 's' denotes the survival probability used by the social security system. The total number of pension points in Eq. (4) 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 (b) of a retiree, the government applies a conversion factor, $\mathsf{f}_{is}(R_i, \mathsf{pp}_{is}(R_i))$, that transforms at age R_i the pension points accumulated (pp) into pension benefits

$$b_{is}(R_i, \mathsf{pp}_{is}(R_i)) = \mathsf{f}_{is}(R_i, \mathsf{pp}_{is}(R_i))\mathsf{pp}_{is}(R_i). \quad (5)$$

In a DC system, the government transforms the pension wealth into an annuity using cohort-specific life tables and an effective interest rate. Thus, the conversion factor at the age of retirement (R_i) is

$$\mathsf{f}_{is}(R_i, \mathsf{pp}_{is}(R_i)) = E_i(R_i)/A_s(R_i, \mathsf{r}), \quad (6)$$

where $E_i(R_i)$ is a factor that corrects for the difference in life expectancy of individuals of type $i \in \mathcal{I}$ relative to the average individual. Similar to Ayuso et al. (2017) we assume the correction factor is specific to the group that the individual belongs and it depends on the retirement age R_i. $A_s(R_i, \mathsf{r})$ is the present value of a life annuity of 1 dollar per year, paid from age R_i onwards, calculated with an effective interest of r and a survival probability $p_s(\cdot)$.[4]

In the DB system, the government multiplies the average indexed yearly earnings by a replacement rate, $\varphi(\mathsf{pp})$, 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., $\varphi(\mathsf{pp}) = \varphi$) or it can decrease as the average indexed yearly earnings increases (i.e., $\varphi'(\mathsf{pp}) < 0$).

[4] The actuarial present value of an individual of type i at the exact age t when the effective interest rate is r is given by $A_i(t, r) = \int_t^{\omega} e^{-r(x-t)} \frac{p_i(x)}{p_i(t)} dx$.

Finally, 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

$$\mathsf{f}_{is}(R_i, \mathsf{pp}_{is}(R_i)) = E_i(R_i)\beta(R_i)\varphi(\mathsf{pp}_{is}(R_i)), \tag{7}$$

where $E_i(R_i)$ is the same correction factor for the difference in life expectancy of individuals of type $i \in \mathfrak{I}$ relative to the average individual introduced in Eq. (6).

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. Depending on the individual characteristics and the pension scheme, the social security wealth might not only change savings, but it might also affect consumption, the supply of labor, and even the accumulation of human capital (Sánchez-Romero and Prskawetz, 2017).

Assuming that an individual will retire at age R_i, we define the social security wealth at age $x \leq R_i$ of an individual of type i, denoted by $\mathsf{SSW}_{is}(x, R_i)$, as

$$\begin{aligned}\mathsf{SSW}_{is}(x, R_i) \equiv \mathsf{SSW}_{is}(x) = e^{-r(R_i - x)}\frac{p_i(R_i)}{p_i(x)} b_{is}(R_i, \mathsf{pp}_{is}(R_i))A_i(R_i, r) - \\ - \int_x^{R_i} e^{-r(t-x)}\frac{p_i(t)}{p_i(x)}\tau y_i(t)dt. \end{aligned} \tag{8}$$

The first component of (8) is the present value, survival weighted, at age x of the future benefits during retirement, while the second component of (8) is the present value, survival weighted, of the remaining pension contributions to pay from age x until retirement.

Eq. (8) is standard in pension economics and finance literature and helps to study how the pension system affects the decisions on saving and retirement (Feldstein, 1974; Gruber and Wise, 1999). However, (8) does not provide information about how the social security will influence on the intensive labor supply. For this reason, we propose rewriting (8) in terms of implicit taxes on labor income. In particular, using (4)–(5) the social security wealth can also be written as follows:

$$\mathsf{SSW}_{is}(x) = \mathsf{P}_{is}(x)\frac{\mathsf{pp}_{is}(x)}{\phi} - \int_x^{R_i} e^{-r(t-x)}\frac{p_i(t)}{p_i(x)}\mathsf{t}_{is}(t)y_i(t)dt, \tag{9}$$

The derivation of (9) from (8) is provided in Appendix A. The first component of (9) is the monetary value given by an individual of type i to the social security contributions paid until age x, while the second component of (9) is the present value, survival weighted, at age x of all the implicit taxes/subsidies on labor income that individuals will expect to pay/receive until retirement.

Eq. (9) contains two new terms that are next explained in detail. First, the term $\mathsf{P}_{is}(x, S_i, R_i)$ is the **value of one dollar contributed to the pension system** at age x for an individual of type i, who would retire at age R_i and has been contributing since age S_i

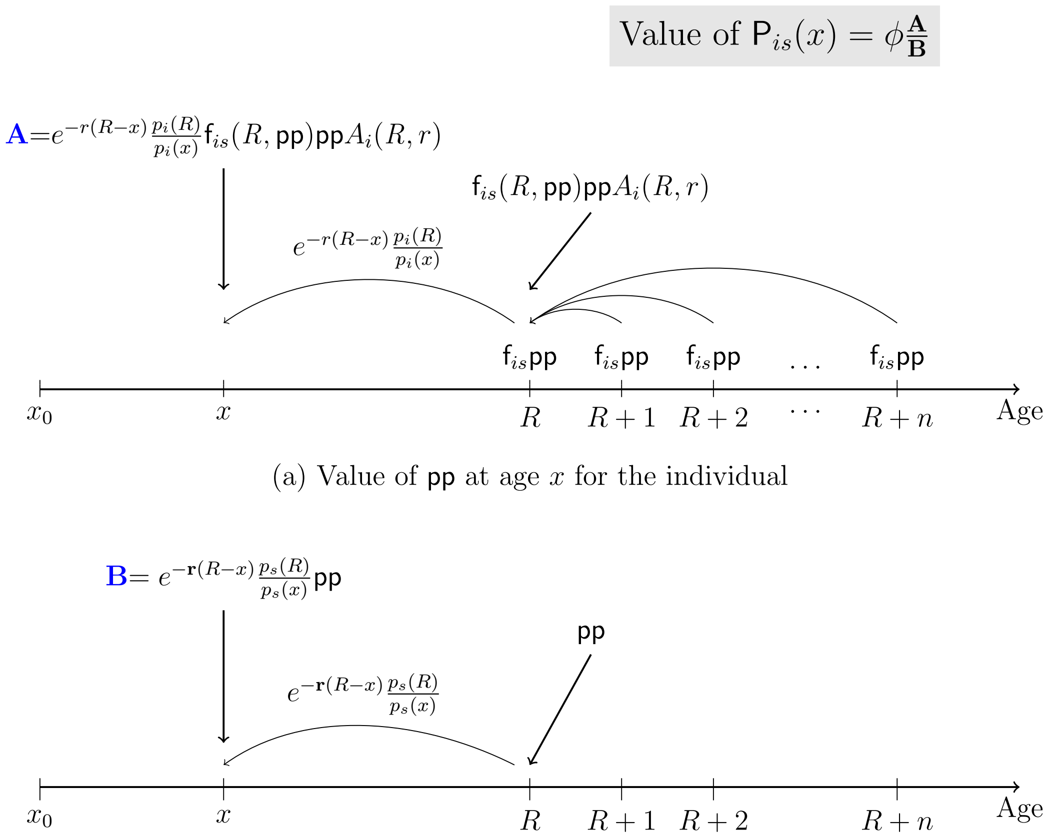
$$\mathsf{P}_{is}(x, S_i, R_i) \equiv \mathsf{P}_{is}(x) = \phi e^{(\mathsf{r}-r)(R_i-x)} \frac{p_i(R_i)}{p_i(x)} \frac{p_s(x)}{p_s(R_i)} \mathsf{f}_{is}(R_i, \mathsf{pp}_{is}(R_i)) A_i(R_i; r) > 0. \tag{10}$$

As Figure 2 shows the term P_{is} is the result of multiplying the value of a pension point per unit of social contribution, ϕ, by the ratio of two present values. First, the present value, survival weighted, at age x of the retirement benefits which would be claimed from age R by an individual of type i from her/his pension points pp —see the top panel of Figure 2. This present value is also known as (gross) pension wealth at age $x \in (0, R)$. Second, the present value, survival weighted, at age x of the same pension points pp at age R calculated from the point of view of the social security system—see the bottom panel of Figure 2.

Eq. (10) implies that when the individual and the social security system value the pension points pp equally, then the value of one dollar contributed to the pension system (P_{is}) is unity. However, when the value of pp for an individual of type i is greater (resp. lower) than that by the social security system, then P_{is} is greater (resp. lower) than unity. A value of P_{is} greater than one implies that an individual of type i receives a greater return from her/his contribution to the pension system than investing the same dollar in the capital market, and vice versa. As a consequence, P_{is} provides information about the rate of return of a pension system, regardless of the pension points accumulated.

Using (9) and (10), we can calculate the value of one dollar contributed to the pension system and how it evolves with age for an individual of type i. Specifically, the value of P_{is} at the age of retirement, R_i, for a worker of type i is

$$\mathsf{P}_{is}(R_i) = \frac{\mathsf{SSW}_{is}(R_i)}{\mathsf{pp}_{is}(R_i)/\phi} = \phi \mathsf{f}_{is}(R_i, \mathsf{pp}_{is}(R_i)) A_i(R_i; r). \tag{11}$$



(a) Value of pp at age x for the individual

(b) Value of pp at age x for the pension system

Figure 2: The value of one dollar contributed to the pension system at age x for an individual who plans to retire at age R, $\mathsf{P}_{is}(x)$

In an actuarially fair pension system —i.e, $\phi\mathsf{f}_{is}(R_i, \mathsf{pp}_{is}(R_i)) = 1/A_i(R_i; r)$— the value of P_{is} is one. Also, from (11) we can see that in a pension system that guarantees the same social security wealth SSW to all individuals who have accumulated the same amount of pension points pp, the value of P_{is} at age R is the same for all individual types $i \in \mathcal{I}$. In addition, differentiating (10) with respect to age x gives the evolution over the lifecycle of P_{is}

$$\frac{1}{\mathsf{P}_{is}(x)}\frac{\partial \mathsf{P}_{is}(x)}{\partial x} = (r - \mathsf{r}) + (\mu_i(x) - \mu_s(x)). \tag{12}$$

When $\mu_s(x) = \mu_i(x)$, we have that P_{is} increases more rapidly with age when the market interest rate is higher than the capitalization factor of the social security system (i.e, $r > \mathsf{r}$). When $r = \mathsf{r}$, if the social security system applies the same average mortality rate to all individual types, those individuals with a life expectancy below the average level (i.e., $\mathsf{e}_i(x_0) < \mathsf{e}_s(x_0)$ or $\mu_i(x) > \mu_s(x)\,\forall x$) 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., $\mathsf{e}_i(x_0) > \mathsf{e}_s(x_0)$ or $\mu_i(x) < \mu_s(x)\,\forall x$) will begin with a higher valuation of their contributions but it will decrease with age. See Figure 3 for an illustration. Thus, this second component accounts for the redistribution of resources within the cohort from those with low life expectancy to those with high life expectancy.

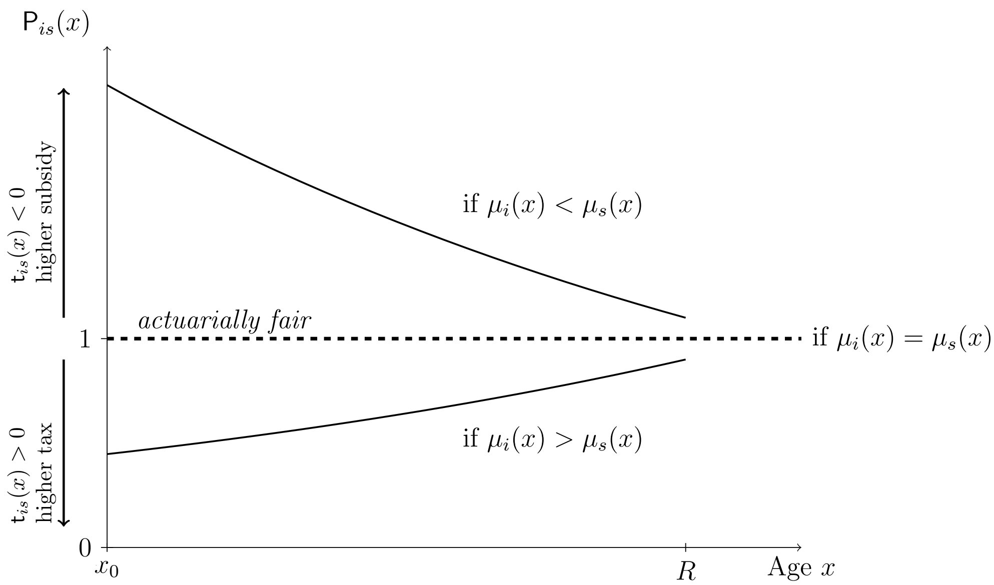


Figure 3: Stylized evolution of the monetary price of one dollar contributed to the pension system at age x for an individual who plans to retire at age R, $\mathsf{P}_{is}(x)$. Case: when $\mathsf{r} = r$.

The second term in (9) includes $\mathsf{t}_{is}(x, S_i, R_i)$; i.e., the **implicit tax/subsidy rate on labor income** faced at age x by an individual of type i, who retires at age R_i and has been contributing to the system since age S_i

$$\mathsf{t}_{is}(x, S_i, R_i) \equiv \mathsf{t}_{is}(x) = \tau\,(1 - \mathsf{P}_{is}(x))\,. \tag{13}$$

From (13) we can see that t_{is} might be either positive (tax) or negative (subsidy) according to the value of P_{is}

$$\mathsf{t}_{is}(x)\begin{cases} > 0 & \text{if } \mathsf{P}_{is}(x) < 1, \\ < 0 & \text{if } \mathsf{P}_{is}(x) > 1. \end{cases} \tag{14}$$

This relationship is shown graphically in the vertical axis in Fig. 3. From (14) we have three important relationships. First, a pension system is actuarially fair if it satisfies that $\mathsf{P}_{is}(x) = 1$ for all $x \in (S_i, R_i)$. Second, P_{is} provides information about the implicit tax/subsidy on labor income, which can be used for comparing alternative pension systems.[5] Third, differentiating (13) with respect to age gives

$$\frac{\partial}{\partial x}(1 - \mathsf{t}_{is}(x)) = \tau \frac{\partial \mathsf{P}_{is}(x)}{\partial x}. \tag{15}$$

Thus, from (15) we can see that the evolution of one dollar net of implicit tax/subsidies is the proportional to the evolution of P_{is}.

Finally, note from (9) that we can write the social security wealth at age x_0 (the age at start making decisions) in terms of the lifetime implicit taxes/subsidies paid

$$\mathsf{SSW}_{is}(x_0) = -\int_{x_0}^{R_i} e^{-r(t-x_0)} \frac{p_i(t)}{p_i(x_0)} \mathsf{t}_{is}(t) y_i(t) dt. \tag{16}$$

From (16) we can see that the social security wealth is negative when contributions are seen as an implicit tax on labor income —i.e. $\mathsf{t}_{is}(t) > 0$ for all $t \in (0, R_i)$—, while the social security wealth becomes positive when contributions are implicitly seen as a subsidy on labor income —i.e. $\mathsf{t}_{is}(t) < 0$ for all $t \in (0, R_i)$.

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

In this paper we analyze six alternative PAYG pension systems (3 DCs and 3 DBs). To easy the comparison across pension systems and clearly show their main features, we introduce three assumptions that are convenient to understand how the six alternative pension systems affect eqs. (5)–(16) across

[5] A similar metric for analyzing alternative pension points is the internal rate of return. However, unlike the internal rate of return, the monetary value of a pension point also gives information about the wealth of the individual if we multiply P by the total number of pension points accumulated (pp).

individual types. First, we assume the market interest rate coincides with the internal rate of return of a PAYG pension system (i.e., $r = \mathsf{r} \equiv n+g$).[6] 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., $\beta(R_n) = 1$).[7] Third, all pension systems provide the same replacement rate for the average individual within each cohort, which implies that $\varphi = (\tau/\varrho)/A_s(R_n, \mathsf{r})$. 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 that 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.

[6]Note 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 $r > \mathsf{r}$.

[7]The 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 1: Alternative PAYG pension systems and their impact on the social security wealth at age x_0 by life expectancy

		Defined Contribution (DC)		
		Avg. Life Table (LT)	Corrected Avg. LT	i-th LT
	Symbol	NDC-I	NDC-II	NDC-III
Indexation	r	$n+g$	$n+g$	$n+g$
Point factor	ϕ	1	1	1
Correction factor	E_i	1	$A_s(R_n, \mathsf{r})/A_i(R_n, \mathsf{r})$	1
Replacement rate	f_{is}	$1/A_s(R_n, \mathsf{r})$	$E_i(R_n)/A_s(R_n, \mathsf{r})$	$1/A_i(R_n, \mathsf{r})$
Value of \$1 contributed	$\mathsf{P}_{is}(x_0)$	$\begin{cases} <1 & \text{for } \mathsf{e}_i < \mathsf{e}_s, \\ >1 & \text{for } \mathsf{e}_i > \mathsf{e}_s. \end{cases}$	$\begin{cases} <1 & \text{for } \mathsf{e}_i < \mathsf{e}_s, \\ >1 & \text{for } \mathsf{e}_i > \mathsf{e}_s. \end{cases}$	1
Value of \$1 contributed	$\mathsf{P}_{is}(R_n)$	$\begin{cases} <1 & \text{for } \mathsf{e}_i < \mathsf{e}_s, \\ >1 & \text{for } \mathsf{e}_i > \mathsf{e}_s. \end{cases}$	1	1
Implicit tax/subsidy	t_{is}	$\begin{cases} >0 & \text{for } \mathsf{e}_i < \mathsf{e}_s, \\ <0 & \text{for } \mathsf{e}_i > \mathsf{e}_s. \end{cases}$	$\begin{cases} \geq 0 & \text{for } \mathsf{e}_i < \mathsf{e}_s, \\ \leq 0 & \text{for } \mathsf{e}_i > \mathsf{e}_s. \end{cases}$	0
Soc. sec. wealth	$\mathsf{SSW}_{is}(x_0)$	$\begin{cases} <0 & \text{for } \mathsf{e}_i < \mathsf{e}_s, \\ >0 & \text{for } \mathsf{e}_i > \mathsf{e}_s. \end{cases}$	$\begin{cases} <0 & \text{for } \mathsf{e}_i < \mathsf{e}_s, \\ >0 & \text{for } \mathsf{e}_i > \mathsf{e}_s. \end{cases}$	0

		Defined Benefit (DB)		
		Non-Progressive	Progressive	Corrected-Progressive
	Symbol	DB-I	DB-II	DB-III
Indexation	r	$n+g$	$n+g$	$n+g$
Point factor	ϕ	ϱ/τ	ϱ/τ	ϱ/τ
Correction factor	E_i	1	1	$A_s(R_n, \mathsf{r})/A_i(R_n, \mathsf{r})$
Replacement rate	f_{is}	$\beta(R_n)\varphi$	$\beta(R_n)\varphi(\mathsf{pp}_i(R_n), \mathsf{r})$	$E_i(R_n)\beta(R_n)\varphi(\mathsf{pp}_i(R_n), \mathsf{r})$
Value of \$1 contributed	$\mathsf{P}_{is}(x_0)$	$\begin{cases} <1 & \text{for } \mathsf{e}_i < \mathsf{e}_s, \\ >1 & \text{for } \mathsf{e}_i > \mathsf{e}_s. \end{cases}$	$\lesseqgtr 1$	$\lesseqgtr 1$
Value of \$1 contributed	$\mathsf{P}_{is}(R_n)$	$\begin{cases} <1 & \text{for } \mathsf{e}_i < \mathsf{e}_s, \\ >1 & \text{for } \mathsf{e}_i > \mathsf{e}_s. \end{cases}$	$\lesseqgtr 1$	$\begin{cases} >1 & \text{for } \mathsf{e}_i < \mathsf{e}_s, \\ <1 & \text{for } \mathsf{e}_i > \mathsf{e}_s. \end{cases}$
Implicit tax/subsidy	t_{is}	$\begin{cases} >0 & \text{for } \mathsf{e}_i < \mathsf{e}_s, \\ <0 & \text{for } \mathsf{e}_i > \mathsf{e}_s. \end{cases}$	$\lesseqgtr 0$	$\lesseqgtr 0$
Soc. sec. wealth	$\mathsf{SSW}_{is}(x_0)$	$\begin{cases} <0 & \text{for } \mathsf{e}_i < \mathsf{e}_s, \\ >0 & \text{for } \mathsf{e}_i > \mathsf{e}_s. \end{cases}$	$\lesseqgtr 0$	$\lesseqgtr 0$

Notes: 'e_s' denotes the life expectancy 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 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 $\mathsf{r} = n + g$, (c) the pension replacement rate φ is equal to $(\tau/\varrho)/A_s(R_n, \mathsf{r})$ 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 when the effective interest rate is r.

Table 1 summarizes how each pension system may affect eqs. (5)–(16)

across individual types. Table 1 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 ($\mathsf{e}_i < \mathsf{e}_s$) and individuals with an average life expectancy above the average level ($\mathsf{e}_i > \mathsf{e}_s$). Table 1 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., $\mathsf{SSW}_{is}(x_0) < (>) 0$, they become (b) neither contributors nor beneficiaries in a NDC-III system —i.e. $\mathsf{SSW}_{is}(x_0) = 0$, 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 1 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 $r > \mathsf{r}$, the overall value of a pension point will decline in all pension systems, the average implicit tax/subsidy will increase, 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. Of course, non-actuarial corrections of the penalties/rewards for early/late retirement will have a strong effect.

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 $i \in \mathcal{I}$ across the six pension systems analyzed.

4.1 The individual problem

Let us consider an individual, who belongs to quintile $i \in \mathcal{I}$, starts making decisions after finishing the compulsory educational system at age x_0, and lives up to a maximum age ω. 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 per hour worked for an individual of type $i \in \mathcal{I}$ at age x conditional on S years of schooling be $w_i(x, S) = h_i(S)\bar{w}(x-S)$, where $h_i(S)$ is the stock of human capital of an individual with S years of schooling and $\bar{w}(x-S) > 0$ accounts for the returns to $x-S$ years of post-schooling experience. Assume the stock of human capital of an individual type $i \in \mathcal{I}$ accumulates at age $x \in [x_0, S_i)$ according to a Ben-Porath (1967) technology

$$\frac{\partial h_i(x)}{\partial x} = \theta_i h_i(x)^{\gamma} - \delta h_i(x) \text{ with } h_i(x_0) = 1, \tag{17}$$

where θ_i is the learning ability of an individual belonging to group $i \in \mathcal{I}$, $\gamma \in (0, 1)$ is the returns to scale to the time devoted to education, and δ is the human capital depreciation rate. Assume the wage rate per hour worked increases with post-schooling experience according to the following Mincer (1974) equation

$$\bar{w}(x-S) = \exp\left(\beta_0(x-S) - \beta_1(x-S)^2\right) \text{ for } x \geq S, \tag{18}$$

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 $w_i(x, S)$ and pay contributions to the pension system. Assume the individual understands that that higher contributions imply higher future benefits. Thus, individuals do not see their contributions to the pension system as a tax on labor income. This assumption connects the economic model with the general pension model introduced in Section 3. Finally, once the individual reach 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. Individuals start with zero assets, $a_i(x_0) = 0$, and in the terminal age ω they do not hold wealth, $a_i(\omega) = 0$. Using (9) the lifetime budget constraint at age $x > x_0$ of an individual of type $i \in \mathcal{I}$ is

$$\int_x^\omega e^{-r(t-x)} \frac{p_i(t)}{p_i(x)} c_i(t) dt = a_i(x) + \mathsf{P}_{is}(x) \frac{\mathsf{pp}_{is}(x)}{\phi} + \\ + \int_x^{R_i} e^{-r(t-x)} \frac{p_i(t)}{p_i(x)} (1 - \mathsf{t}_{is}(t)) w_i(t, S_i) \ell_i(t) dt, \quad (19)$$

where r is the market interest rate, $p_i(t)/p_i(x)$ is the probability of surviving to age t conditional on being alive at age x for an individual of type i, $a_i(x)$ are the assets held at age x, $\mathsf{P}_{is}(x)$ is the value of one dollar contributed to the pension system at age x for an individual type i, $\mathsf{pp}_{is}(x)$ are the pension points accumulated until age x by an individual of type i. Differentiating (4) with respect to age we have that pension points accumulate over the working life according to

$$\frac{\partial \mathsf{pp}_{is}(x)}{\partial x} = (\mathsf{r} + \mu_s(x)) \mathsf{pp}_{is}(x) + \phi \tau w_i(x, S_i) \ell_i(x) \text{ for } x \in (S_i, R_i). \quad (20)$$

r is the capitalization factor of the social security pension system, $\mu_s(x)$ is the mortality hazard rate at age x used by the pension system, ϕ is the pension point per unit of social contribution paid, τ is the social contribution rate, $\mathsf{t}_{is}(t)$ is the implicit tax/subsidy rate on labor income faced by an individual type i at age t, and $w_i(t, S_i)\ell_i(t)$ 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 $w_i(t, S_i)$.

Eq. (19) clearly shows how consumption over the remaining lifespan is financed by current assets $a_i(x)$, by the monetary value of the social security contributions paid until age x, $\mathsf{P}_{is}(x)\frac{\mathsf{pp}_{is}(x)}{\phi}$, and by the present value at age x of the remaining flow of labor income net of implicit taxes/subsidies.

Preferences. Consider an individual optimally choose 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 (17)-(18), 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

$$V_i(x) = \int_x^\omega e^{-\rho(t-x)} \frac{p_i(t)}{p_i(x)} U\left(c_i(t)\right) dt - \int_{S_i}^{R_i} e^{-\rho(t-x)} \frac{p_i(t)}{p_i(x)} \alpha_i v\left(\ell_i(t)\right) dt \\ - \int_x^{S_i} e^{-\rho(t-x)} \frac{p_i(t)}{p_i(x)} \eta_i dt + \int_{R_i}^\omega e^{-\rho(t-x)} \frac{p_i(t)}{p_i(x)} \varphi(t) dt. \quad (21)$$

The first two components on the right-hand side of (21) 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, $p_i(t)$ is the probability of surviving to age t by an individual of type $i \in \mathcal{I}$, $U(c)$ 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 $v''(\ell) > 0$), $\eta_i > 0$ is the marginal disutility from attending school and $\varphi(t) > 0$ (with $\varphi'(t) \geq 0$) is the marginal utility of leisure during retirment, which increases with age as the amount of retirement time is squeezed by later and later retirement. Similar to Bloom et al. (2014) we assume $\varphi(t)$ is proportional to the mortality rate.

The last three terms in (21) are key for (i) taking into account that the return to schooling exceeds the marginal cost of education (Heckman et al., 2006), (ii) for reproducing the supply of labor during the working life, which is hump-shaped, and (iii) for replicating actual retirement ages given that individuals would never retire because continuing 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 heterogeneous individuals. To do so, we solve the problem of maximizing the lifetime utility (21) subject to the lifetime budget constraint (19) and the laws of motion (17)–(18). The definition of the Hamiltonians and the first-order conditions are reported in the Appendix B. Then, we proceed to look at the behavioral responses of an individual of type i to the different pension systems. First, we start analyzing how each pension plan affects consumption and hours worked for a given length of schooling and retirement age. Given the optimal consumption and labor supply, we will follow analyzing the impact that each pension plan has on the length of schooling. We finalize the section studying the impact of each pension system on retirement.

In a lifecycle model behavioral responses are explained by the impact that marginal changes in values and taxes have on the decision variables. For this reason, in this section we will define a set of marginal values and taxes associated to those introduced in Section 3.2.

Consumption and hours worked. Given a length of schooling and retirement age (S_i, R_i), the optimization yields the following optimal consumption at age $x \in (x_0, \omega)$ and intensive labor supply (hours worked) at age $x \in (S_i, R_i)$:

$$c_i(x) = e^{\sigma_c(r-\rho)(x-x_0)} \left(\frac{\int_{S_i}^{R_i} D(t; x_0, -\sigma_\ell) wt_{is}(t) \left(w\tau_{is}(t)\right)^{\sigma_\ell} dt}{(\alpha_i)^{\sigma_\ell} \int_{x_0}^{\omega} D(t; x_0, \sigma_c) dt} \right)^{\frac{\sigma_c}{\sigma_\ell+\sigma_c}}, \tag{22}$$

$$\ell_i(x) = e^{\sigma_\ell(\rho-r)(x-x_0)} \left(\frac{(\alpha_i)^{-\sigma_c} \int_{x_0}^{\omega} D(t; x_0, \sigma_c) dt}{\int_{S_i}^{R_i} D(t; x_0, -\sigma_\ell) \frac{wt_{is}(t)}{(w\tau_{is}(x))^{\sigma_c}} \left(\frac{w\tau_{is}(t)}{w\tau_{is}(x)}\right)^{\sigma_\ell} dt} \right)^{\frac{\sigma_\ell}{\sigma_\ell+\sigma_c}}, \tag{23}$$

where

$$D(t; x_0, \sigma) = \frac{p_i(t)}{p_i(x_0)} e^{-[r(1-\sigma)+\sigma\rho](t-x_0)}, \tag{24}$$

$$wt_{is}(t, S_i, R_i) \equiv wt_{is}(t) = (1 - \mathsf{t}_{is}(t)) w_i(t, S_i), \tag{25}$$

$$w\tau_{is}(t, S_i, R_i) \equiv w\tau_{is}(t) = (1 - \tau_{is}(t)) w_i(t, S_i). \tag{26}$$

The term $\sigma_c = -\frac{U'(c)}{cU''(c)} > 0$ is the intertemporal elasticity of substitution for consumption, $\sigma_\ell = \frac{v'(\ell)}{\ell v''(\ell)} > 0$ is the intertemporal elasticity of substitution for labor, and $(1-\tau_{is}(t))w_i(t, S_i)$ is the net wage rate per hour worked after paying a **marginal implicit tax/subsidy rate on labor income** faced at age t by an individual of type i, who retires at age R_i and has been contributing to the system since age S_i

$$\tau_{is}(t, S_i, R_i) \equiv \tau_{is}(t) = \tau \left(1 - \mathcal{P}_{is}(t)\right). \tag{27}$$

$\mathcal{P}_{is}(t, S_i, R_i)$ is the **marginal value of one dollar of social contribution** for an individual of type i, who retires at age R_i and has been contributing to the system since age S_i.[8] Unlike P_{is}, the value of $\mathcal{P}_{is}$ is calculated as the marginal rate of substitution between pension points and assets, conditional on the retirement age R_i:

$$\mathcal{P}_{is}(t, S_i, R_i) \equiv \mathcal{P}_{is}(t) = \left. \frac{\partial V_i(t) / \partial \left(\mathsf{pp}_{is}(t)/\phi\right)}{\partial V_i(t) / \partial a_i(t)} \right|_{R=R_i} = \mathsf{P}_{is}(t) \left(1 - \varepsilon_{is}\right). \tag{28}$$

Therefore, $\mathcal{P}_{is}(t)$ can be defined as the "value of an additional unit of social contribution" at age t for an individual of type i who plans to retire at age

[8]A similar expression to (27) can be found in Auerbach and Kotlikoff (1987) and Sánchez-Romero and Prskawetz (2017).

R_i. The term ε_{is} is the elasticity between the replacement rate and pension points at retirement of an individual of type i, or

$$\varepsilon_{is} = -\frac{\mathsf{pp}_{is}(R_i)}{\mathsf{f}_{is}(R_i, \mathsf{pp}_{is}(R_i))} \frac{\partial \mathsf{f}_{is}(R_i, \mathsf{pp}_{is}(R_i))}{\partial \mathsf{pp}_{is}(R_i)} \geq 0. \tag{29}$$

From (28)-(29) we have that the marginal value of one dollar of social contribution is lower than the value of one dollar of social contribution because, as a result of paying additional contributions to the pension system, the pension replacement rate falls when $\varepsilon_{is} > 0$. The condition $\varepsilon_{is} > 0$ is satisfied not only in progressive pension systems, but also in pension systems with a flat pension benefit (or Beveridge pension systems). Hence, we can distinguish two cases:

(i) A flat pension system in which the replacement rate is invariant to the pension points accumulated; i.e., $\varepsilon_{is} = 0$.[9] In this pension system, $\mathcal{P}_{is} = \mathsf{P}_{is}$ and $\tau_{is} = \mathsf{t}_{is}$.

(ii) A progressive pension system in which the replacement rate decreases with the number of pension points accumulated; i.e., $\varepsilon_{is} > 0$. In this pension system, $\mathcal{P}_{is} < \mathsf{P}_{is}$ and $\tau_{is} > \mathsf{t}_{is}$.

Considering a stable, mature pension PAYG pension system, Fig. 4 shows the marginal value of a unit of social contribution at the normal retirement age R_n for the six different pension systems analyzed in Section 3.3.[10] Fig. 4(a) shows the value of $\mathcal{P}_{is}$ regardless of the gap in life expectancy across the different income groups. Fig. 4(b) shows the change in $\mathcal{P}_{is}$ that would result from the difference in life expectancy at age 65 when the implicit rate of return of the pension system r (i) equals the market interest rate r (solid line) or (ii) is lower than the market interest rate (dashed line). Thus, the value of $\mathcal{P}_{is}$ is obtained by multiplying the value in Fig. 4(a) by the corresponding value by longevity gap in Fig. 4(b). Looking at Fig. 4 we can see that in all NDC plans and in the DB-I plan the value of $\mathcal{P}_{is}$ and the value of P_{is} coincide and is equal to $q = \frac{A_i(R_n,r)}{A_s(R_n,\mathsf{r})}$. The value of q reflects (i) differences in longevity between each income group and the life table used by the social security system and (ii) the fact that the benefit received from the public pension system differs from the benefit the individual would receive if she/he transforms the social security wealth into a private annuity. In contrast, in DB-II and DB-III plans, the value of $\mathcal{P}_{is}$ is substantially lower than P_{is} when the total pension points is higher than one-sixth of the average labor income.

[9] We 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.

[10] We assume the normal retirement age so that no other penalties/rewards for early/late retirement are considered.

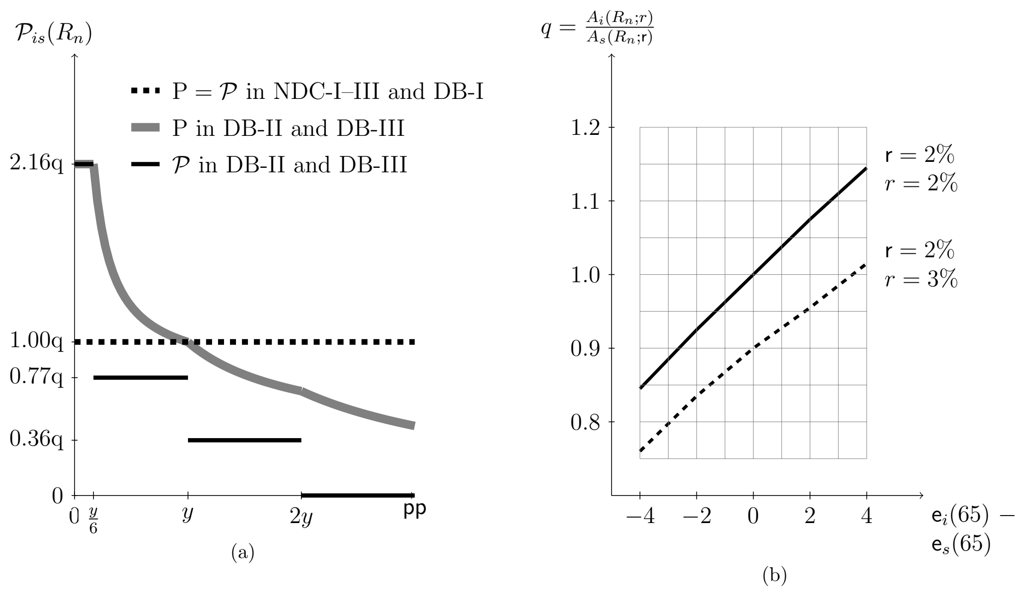


Figure 4: Marginal value of one dollar of social contribution at the normal retirement age R_n.

Note: Calculation based on a normal retirement age R_n of 65 and an average life expectancy at age 65 of 21.4 years. The term y represents the average labor income.

The difference in P_{is} and $\mathcal{P}_{is}$ and the evolution of P_{is} over the working life —see Eq. (12)— determine the impact of the alternative pension system on consumption and the labor supply. Looking at (22), using (13) and (27), and assuming that $\mathsf{r} = r$, we find that a pension system has an unambiguous positive (resp. negative) impact on consumption when the condition

$$(1 - \tau + \tau \mathsf{P}_{is}(t))(1 - \tau + \tau \mathcal{P}_{is}(t))^{\sigma_\ell} > (\text{resp.} <) 1 \, \forall t \in (S_i, R_i) \tag{30}$$

is satisfied. This occurs for individuals with a life expectancy above (resp. below) the average level in a NDC system and in the DB-I. However, in the progressive pension systems DB-II and DB-III, the impact on consumption is ambiguous. Note in Fig. 4 that the positive effect of a $\mathsf{P} > 1$ when $\mathsf{pp} < y$ is offset by $\mathcal{P} < 1$ and the value of $q = \frac{A_i(R_n;r)}{A_s(R_n;\mathsf{r})}$, which would be lower than one due to the positive correlation between income and life expectancy. The opposite occurs for individuals with $\mathsf{pp} \in (y, 2y)$, while it is negative for individuals with $\mathsf{pp} > 2y$. Thus, it is likely that only for those with a total pension points (pp) below one-sixth of the average labor income, the consumption will increase in a DB-II and DB-III systems.

The total number of hours worked are also affected by the alternative pension system when $\sigma_c \neq 1$. Nevertheless, following Chetty (2006) and assuming σ_c is equal to one, Eq. (23) suggests that pension systems with a constant replacement rate do not significantly affect the total number of hours worked over the lifecycle. Instead, Eq. (23) shows that pension systems will increase/decrease the labor supply early in the working life and compensate it with decreases/increases later in the working life. This is because the pension system influences the labor supply at the intensive margin through the evolution, and not the level, of the implicit taxes/subsidies —see the ratios of net wages rates in the denominator of Eq. (23). Thus, from Fig. 3 we can see that the intensive labor supply (hours worked) early in the working life decreases (resp. increase) for those individuals with a $\mathcal{P}_{is}$ increasing (resp. decreasing) with age, while the intensive labor supply late in the working life increases (resp. decrease) for those individuals with a $\mathcal{P}_{is}$ increasing (resp. decreasing) with age. A similar restructure of the hours worked over the working life occurs in progressive pension systems (DB-II and DB-III). However, given that in a progressive pension system $(1-\mathsf{t}_{is}(t)) \geq (1-\tau_{is}(t))$ for all $t \in (S_i, R_i)$, the DB-II and DB-III plans also reduce the intensive labor supply for individuals with a $\mathsf{pp} \in (\frac{y}{6}, 2y)$, while these pension systems do no affect the intensive labor supply for those with a $\mathsf{pp} < \frac{y}{6}$ or with a $\mathsf{pp} > 2y$.

Length of schooling. In the previous section we have considered the length of schooling fixed. Now, we relax this condition and study the impact of each pension system on the optimal length of schooling. Given a retirement age R_i, the optimal length of schooling $S^* = S_i$ satisfies:

$$r_i^h(S^*) = \bar{r}_i(S^*, R_i) + \frac{\eta_i - \alpha_i v(\ell(S^*))}{U'(c_i(S^*))W_i(S^*, R_i)}. \tag{31}$$

See the derivation of (31) in Appendix B.2. Eq. (31) states that the return to education at the S^*th unit of schooling is equal to the sum of the marginal cost of the S^*th unit of schooling expressed in terms of foregone earnings and the nonpecuniary cost/benefit of schooling.[11] $r_i^h(S_i^*)$ is the return to education at age S_i^* for an individual of type i

$$r_i^h(S^*) = \frac{1}{h_i(S^*)}\frac{\partial h_i(S^*)}{\partial S}. \tag{32}$$

[11]See 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.

$\bar{r}_i(S^*, R_i)$ is the rate of return lost from an additional S^*th unit of schooling, which is given by

$$\bar{r}_i(S^*, R_i) = \int_{S^*}^{R_i} \frac{1}{\bar{w}(t-S^*)} \frac{\partial \bar{w}(t-S^*)}{\partial t} \psi_{is}(t) dt + \psi_{is}(S^*), \tag{33}$$

and where

$$\psi_{is}(t) = \frac{D(t; x_0, -\sigma_\ell)\left(w\tau_{is}(t)\right)^{1+\sigma_\ell}}{\int_{S^*}^{R_i} D(u; x_0, -\sigma_\ell)\left(w\tau_{is}(u)\right)^{1+\sigma_\ell} du}. \tag{34}$$

The first term in (33) accounts for the rate of return lost from not accumulating further labor experience, while the second term in (33) accounts for the foregone returns that the individual would have generated from the additional savings at age S^*. Eq. (34) represents the weight of the (net) income earned at age t relative to the (net) human capital wealth at age S^*. The last term in (31) is the nonpecuniary cost/benefit of schooling. For expositional simplicity, if we assume that $\sigma_c = 1$ and $\rho = 0$, we can rewrite (31) as

$$\frac{\frac{p_i(S^*)}{p_i(x_0)}\eta_i}{\mathsf{e}_i(x_0)} \frac{\int_{S^*}^{R_i} D(t; x_0, -\sigma_\ell) w \mathsf{t}_{is}(t)(w\tau_{is}(t))^{\sigma_\ell} dt}{\int_{S^*}^{R_i} D(t; x_0, -\sigma_\ell) w \tau_{is}(t)(w\tau_{is}(t))^{\sigma_\ell} dt} - \frac{\sigma_\ell}{1+\sigma_\ell}\psi_{is}(S^*). \tag{35}$$

The first component is the schooling effort measured in monetary terms, while the second component accounts for additional utility the individual gets by not being working at age S^*.

Fig. 5 illustrates the equilibrium condition for the optimal length of schooling that results from (31). The black solid curve represents the rate of return to Sth years of education $r_i^h(S)$, which is a decreasing function with respect to S. The black dashed curve represents the marginal pecuniary cost of schooling $\bar{r}_i(S, R_i)$, which is an increasing function with respect to S. While the gray solid curve is the sum of the marginal pecuniary and non-pecuniary costs of schooling. Thus, due to the existence of non-pecuniary costs of schooling, individuals do not maximize the returns to education, which occurs at $\tilde{S}_i$, and instead decide to have $S^* < \tilde{S}_i$ years of schooling. Some of the underlying mechanisms explaining the gap between $\tilde{S}_i$ and S^* have been explained by Oreopoulos and Savanes (2011).

How does each pension system impact on the length of schooling across the different individual types? Looking at (31), we can observe that the returns to schooling do not depend on the pension system, whereas the marginal cost of education is influenced by the evolution of the marginal implicit tax/subsidy rate, τ_{is}, throughout (34). Recalling that the evolution of τ_{is} is given by (12)

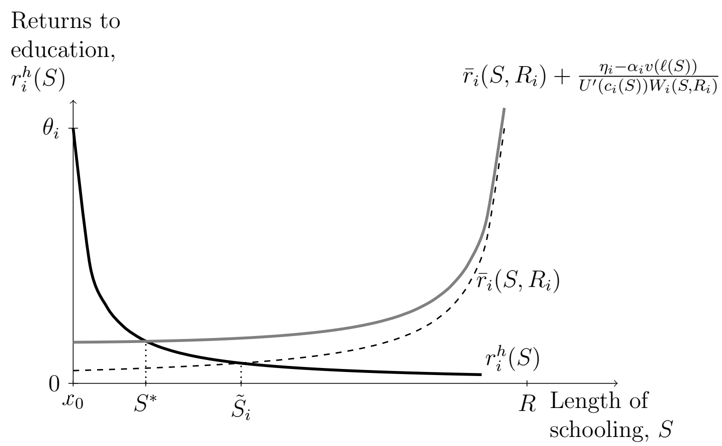


Figure 5: Optimal length of schooling decision.

and assuming $\sigma_c = 1$, we have that pension systems reduce (resp. increase) the marginal pecuniary costs of schooling when $\mathcal{P}_{is}$ increase (resp. decrease) with age. This is equivalent to a downward (resp. upward) shift of the gray curve in Fig. 5. The intuition is simple. The monetary lost of paying contributions to the system is compensated with higher wage rates, which are only possible through longer investments in education. When the pension replacement rate is progressive, however, the decline in the marginal pecuniary cost of schooling is offset by an increase in the marginal non-pecuniary cost of schooling —see Eq. (35). Hence, in DB-II and DB-III plans the net effect of the pension system is a priori ambiguous when $\mathsf{pp} \in (\frac{y}{6}, 2y)$. In the special case that $\tau_{is} = 0$, which occurs in DB-II and DB-III plans when $\mathsf{pp} > 2y$, the pension system does not have an impact on schooling.

Optimal retirement age. Consider now that the length of schooling is fixed. Under this setting the optimal retirement age $R^* = R_i$ satisfies the following condition

$$U'(c_i(R^*))w_i(R^*, S_i)\ell_i(R^*)\left(1 - \tau_{is}^{GW}(R^*)\right) = \alpha_i v\left(\ell_i(R^*)\right) + \varphi(R^*). \qquad (36)$$

See the proof in Appendix B.3. Eq. (36) states the standard optimal retirement age condition in which the marginal benefit of continue working (left-hand side) must be equal to the marginal cost of continue working (right-hand side). The new term, $\tau_{is}^{GW}(R^*)$, located on the left-hand side of (36) is the effective tax/subsidy rate on work calculated in Gruber and Wise (1999). The marginal cost of continue working is given by the sum of the marginal cost of continue working, which depends on the disutility of labor (or the lost of current leisure time)

$$\alpha_i v\left(\ell_i(R^*)\right) = \frac{\sigma_\ell}{1+\sigma_\ell} w \tau_{is}(R^*, S_i)\ell_i(R^*)U'(c_i(R^*)), \tag{37}$$

plus the lost of utility of leisure from retirement, $\varphi(R^*)$. Plugging (37) in (36) notice that without the marginal utility of leisure from retirement, individuals would only retire when the marginal implicit tax/subsidy rate on work would be null.

In our general pension setting, the effective tax/subsidy rate on work, $\tau_{is}^{GW}(R^*)$, is given by

$$\tau_{is}(R^*) + \left(\frac{1}{A_i(R^*, r)} - (\mathsf{r} + \mu_s(R^*))(1-\varepsilon_{is}) - \frac{1}{\mathsf{f}_{is}}\frac{\partial \mathsf{f}_{is}}{\partial R^*}\bigg|_{\mathsf{pp}}\right)\frac{\mathsf{P}_{is}(R^*)\frac{\mathsf{pp}_{is}(R^*)}{\phi}}{w_i(R^*, S)\ell_i(R^*)}. \tag{38}$$

The first term in (38) is the marginal implicit tax/subsidy rate on labor income. The second term is the ratio of pension benefits to labor income. The third term, or the sum of the negative terms, is the increase in social security wealth from delaying retirement, which is a function of the increase in pension points and the increase in the value of one dollar of social contribution through the percentage change in replacement rate for each year of delayed retirement. To explain the difference in retirement age across the pension plans analyzed, the percentage change in replacement rate for each year of delayed retirement is key; i.e., $\frac{1}{\mathsf{f}_{is}}\frac{\partial \mathsf{f}_{is}}{\partial R}\Big|_{\mathsf{pp}}$. Thus, we show in Figure 6 the percentage change in replacement rate for each year of delayed retirement for the six pension plans under both mortality regimes. For the NDC-II, NDC-III, and DB-III cases, which correct for differences in mortality, we only plot the value of $\frac{1}{\mathsf{f}_{is}}\frac{\partial \mathsf{f}_{is}}{\partial R}\Big|_{\mathsf{pp}}$ for the top (q5) and bottom (q1) quintiles. All other cases lie within the top and bottom quintiles. The most important result from Fig. 6 is that DB plans, which use the credits for delayed retirement of the actual US system, provide different retirement age incentives than NDC plans. In particular, under the mortality regime of the 1930 birth cohort, DB plans give a higher incentive to delay the retirement age than the NDC plans when the optimal retirement age is below age 65. Under the mortality regime of the 1960 birth cohort, DB plans give a higher incentive to delay the retirement age than NDC plans.

To conclude this section, we should bear in mind that we have studied the effect of each pension system on each variable independently. However, in reality, a delay in the retirement age will change the optimal length of schooling. Since the the length of schooling and the retirement age are interweaved. The

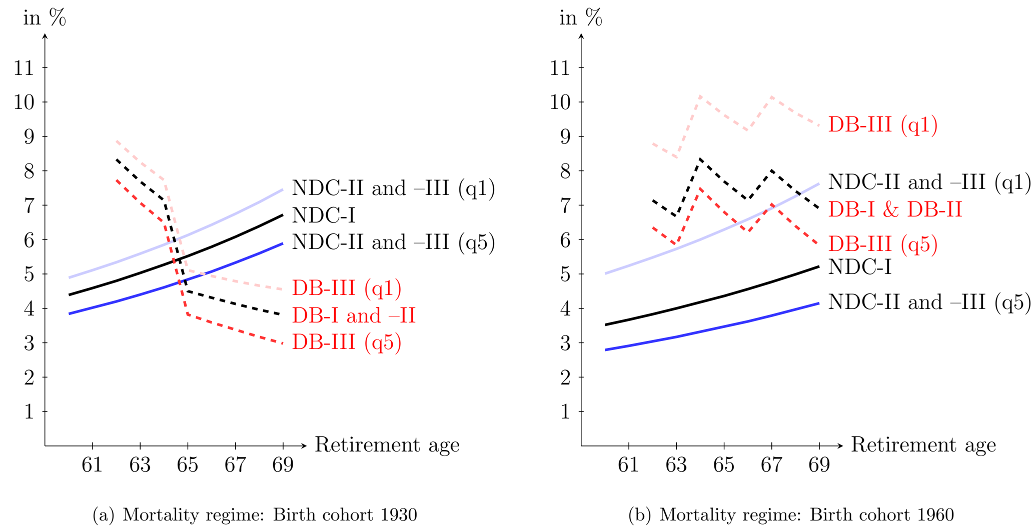


Figure 6: Percentage change in replacement rate for each year of delayed retirement by pension system and mortality regime; i.e., $\frac{1}{\mathsf{f}_{is}} \frac{\partial \mathsf{f}_{is}}{\partial R}\bigg|_{\mathsf{pp}}$.

Note: Calculations for the NDC plans are based on a notional interest rate of 2%.

sign of the relationship between the length of schooling and the retirement depends on the risk aversion coefficient and the effort from attending schooling (see Sánchez-Romero et al., 2016).

5 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). 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. Third, the annual capitalization factor of the unfunded pension system (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 (12) we know that this assumption implies that the marginal value of a unit of social contribution 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 $\mathcal{P}_{is}$, 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, $p_s(x) := p(x) = \sum_{i\in\mathfrak{I}} p_i(x)/I$. Fifth, we assume for the NDC systems a minimum retirement age of 55 and a maximum retirement age of 70 for all $i \in \mathfrak{I}$. 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 balanced[12]

$$\sum_{i\in\mathfrak{I}} \int_{S_i}^{R_i} e^{-nt} p_i(t)\tau w_i(t,S_i)\ell_i(t)dt =$$
$$= \sum_{i\in\mathfrak{I}} e^{-nR_i} \int_{R_i}^{\omega} e^{-(n+g)(t-R_i)} p_i(t) b_{is}(R_i, \mathsf{pp}_{is}(R_i;\mathsf{r}))dt. \quad (39)$$

Using Eq. (39) 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 10.43%, under the hypothetical assumption that the population face the survival probabilities of the cohort born in 1930. While τ must be set at 10.87% 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

[12]Note in Eq. (39) 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.

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 monotonically increasing with age. Similar to Bloom et al. (2014) we proxy $\varphi(\cdot)$ to the mortality rate in 1930. 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. Finally, the weight of disability cost (α_i) and the learning ability (θ_i) for each income quintile group are simultaneously calculated in order to replicate the length of schooling and the retirement age from the Health and Retirement Survey (HRS) and the present value of lifetime benefits reported in NASEM (2015) for the cohort born in 1930.[13] See Table 6 in Appendix B.3.

Table 2 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 B.3 tables 4 and 5 report the optimal length of schooling and retirement ages for all pension scenarios. Life expectancy and total years-worked are calculated using the specific mortality rates for each income quintile. The last column in Tab. 2 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 2: Optimal length of schooling (S_i) and retirement age (R_i) and by income quintile. US males born in 1930, US pension system (DB-II)

Inc. quintile	Schooling	Retirement	Life expectancy at $S_i + 7$	Years-worked	
i	S_i	R_i	$\mathsf{e}_i(S_i + 7)$	(YW_i)	$\frac{\mathsf{e}_i(S_i) - \mathsf{YW}_i}{\mathsf{YW}_i}$
1	12	62	55,5	41,8	0.33
2	12	63	56,3	42,8	0.31
3	12	63	57,4	43,1	0.33
4	13	63	58,2	42,3	0.37
5	16	64	58,2	41,1	0.42

Table 3 summarizes the model economy parameters.

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

Table 3: Model parameters

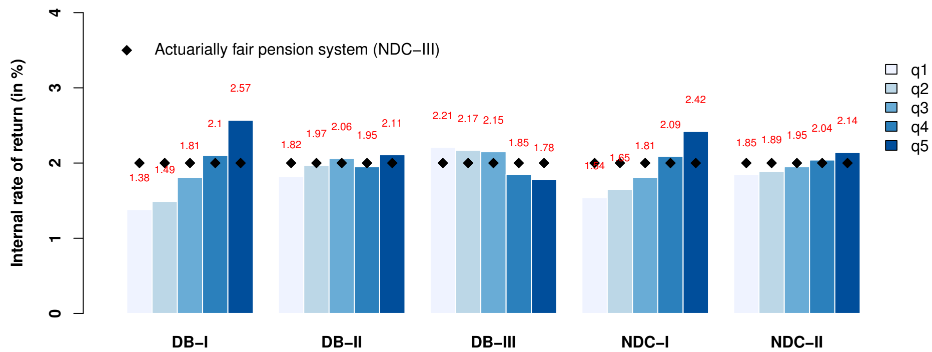
Parameter	Symbol	Value	Parameter	Symbol	Value
Demographics			**Preferences**		
First age at entrance	x_0	14	Subjective discount factor	ρ	0,005
Maximum age	ω	115	Utility cost of not being retired	$\varphi(x)$	$92e^{-8.7+0.073x}$
Annual population growth	n	0,005	Labor elasticity of substitution	σ_ℓ	0,33
Minimum length of schooling	$\underline{S}$	10	Utility weight of labor	$\alpha(q1)$	140
Maximum length of schooling	$\bar{S}$	20		$\alpha(q2)$	120
				$\alpha(q3)$	80
Technology				$\alpha(q4)$	50
Market interest rate	r	0,030		$\alpha(q5)$	85
Labor-augmenting technological progress growth rate	g	0,015			
			Education		
Social security system			Returns of scale in education	γ	0,65
Minimum retirement age	$\underline{R}$	NDC=55, DB=62	Disutility of schooling	η	3
Maximum retirement age	$\overline{R}$	70	Mincerian eq.	β_0	0,07
Capitalization factor	r	0,02		β_1	0.0011
Accrual rate in DB systems	ϕ	1/45	Learning ability	$\theta(q1)$	0,113
Avg. replacement rate in DB systems	$f(\mathsf{pp})$	0,4167		$\theta(q2)$	0,113
Social contribution rate				$\theta(q3)$	0,113
Cohort 1930	τ_{1930}	0,1043		$\theta(q4)$	0,114
Cohort 1960	τ_{1960}	0,1087		$\theta(q5)$	0,124

6 Redistributive effects of each pension system

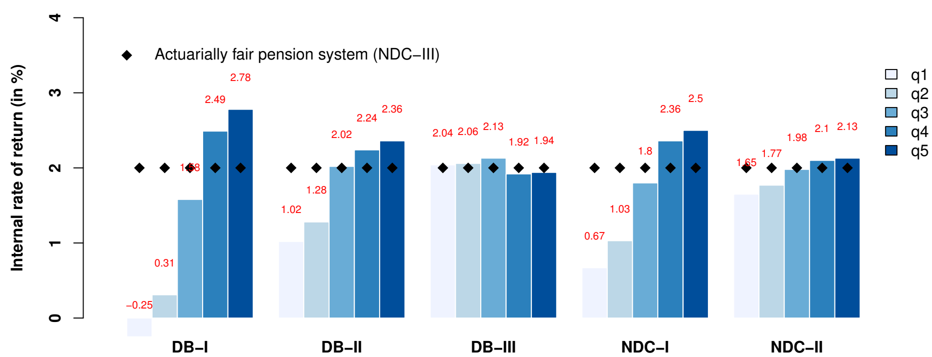
The results presented in this section are based on a small-open economy with a stable population. The pension and economic models, introduced in Sections 3 and 4, have been parametrized in order to replicate for the mortality of the 1930 birth cohort the present value of lifetime benefits at age 50 reported in the NASEM (2015) as well as the length of schooling and the retirement age by income quintile from HRS. To account for the redistributive effects of each pension scheme when the heterogeneity in life expectancy increases, we compare the results obtained using the survival probabilities of the cohort born in 1930 to those obtained using the survival probabilities of the cohort born in 1960.

6.1 Internal rate of return (irr).

One measure to analyze the redistributive characteristics of a pension system, which is not affected by the scale of contributions (i.e., pp_{is}), is the internal rate of return. Thus, we report in Figure 7 the internal rate of return values of each pension system by income quintile and birth cohort. The differences in irr across income quintiles, shown in Fig. 7, 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. (10) and Table 1).



(a) Cohort 1930



(b) Cohort 1960

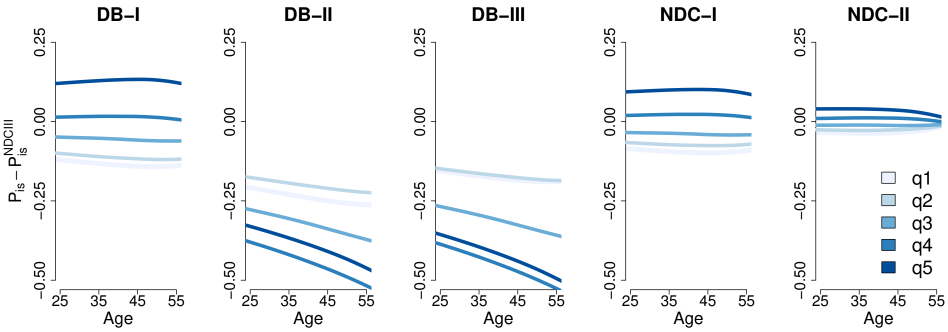
Figure 7: Internal rate of return of each pension system by income quintile (in percentage). US males, birth cohorts 1930 (Panel a) and 1960 (Panel b)

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. 7 we see that this rate of return is achieved by all income groups and mortality regimes under NDC-III —see black diamonds— 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 —see the numbers in red on top of each bar— and therefore are redistributing income to the higher income q4 and q5 who have irr>2, 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 q3 and q5 under both mortality regimes, but particularly with more unequal mortality. The q4 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.

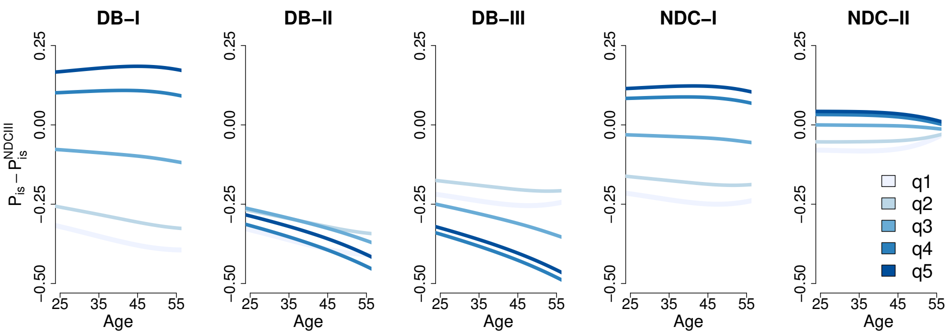
6.2 Marginal value of one dollar of social contribution ($\mathcal{P}_{is}$).

Another method to measure the redistributive effects of a pension system is to look at the value of $\mathcal{P}_{is}$. We showed in sections 3.2 and 4.2 how $\mathcal{P}_{is}$ can be used to study at each age the redistributive effect of each pension system. Hence, $\mathcal{P}_{is}$ complements the information provided in the previous section, since the irr measures the redistributive effects of each pension system over the whole life cycle, while $\mathcal{P}_{is}$ does it by age.

To understand the redistributive properties, we compare the value of $\mathcal{P}_{is}$ for each pension system to our benchmark (NDC-III). Thus, we report in Figure 8 the difference in the evolution of $\mathcal{P}_{is}$ for each pension plan and that in the NDC-III by income quintile and birth cohort. Vertical axes in Fig. 8 reflect whether the contribution to the system represents a subsidy (positive values) or tax (negative values) relative to the contribution paid to the NDC-III system. Thus, we see that NDC-I, NDC-II, and DB-I (pension plans with a



(a) Cohort 1930



(b) Cohort 1960

Figure 8: Marginal value of one dollar of social contribution, $\mathcal{P}_{is}(x)$, from age 25 to 55 by income quintile relative to NDC-III. US males, birth cohorts 1930 (Panel a) and 1960 (Panel b)

flat replacement rate) redistributive income from poor (q1, q2, and q3) to rich income groups (q4 and q5). The situation is reversed in progressive pension systems —DB-II and DB-III—, in which we see that the higher income quintiles q4 and q5 pay higher implicit taxes than lower income quintiles q1 and q2. However, in DB-II and DB-III, the marginal value of one dollar of contribution to the pension system is significantly lower than a dollar contributed to the NDC-III plan. 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.

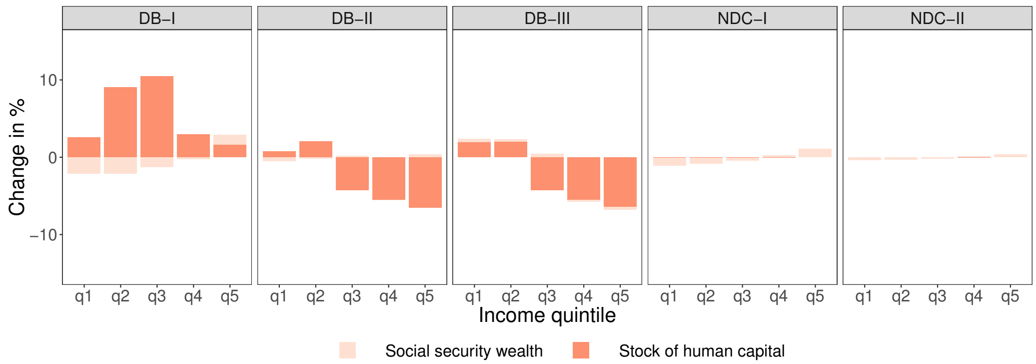
6.3 Wealth.

The fact that pension systems redistribute income across income groups leads individuals to response in order to cope with the increase/lost of wealth. The lifetime wealth measure gives the most comprehensive assessment of the effects of the different pension designs on economic wellbeing, because it includes the general equilibrium responses. These responses are reflected in the lifetime wealth (LW) —see Eq. (19)— through changes in the social security wealth at age x_0, $\mathsf{SSW}_i(x_0)$, and through changes in the stock of human capital (HK) at age x_0

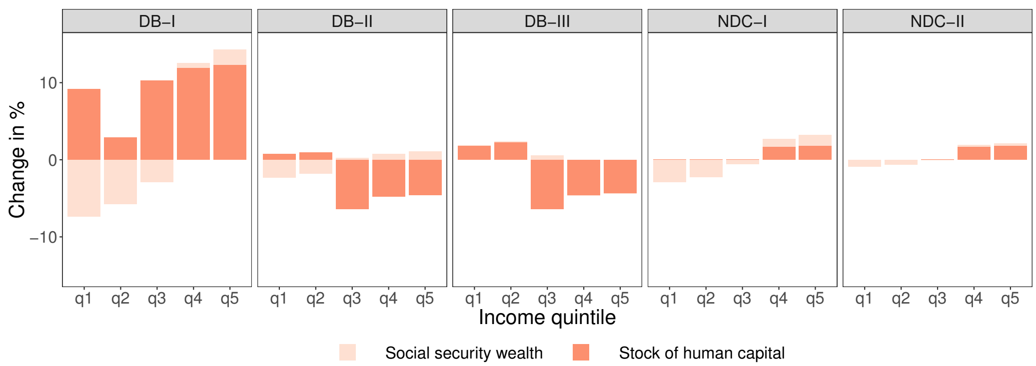
$$\int_{S_i}^{R_i} e^{-r(t-x_0)} \frac{p_i(t)}{p_i(x_0)} w_i(t, S_i)\ell_i(t)dt. \tag{40}$$

Note that changes in the SSW have a one-to-one relation with the internal rate of return.

To analyze the changes in lifetime wealth across pension systems by income quintiles, we again use the NDC-III system as a benchmark against which we can assess these changes. Figure 9 shows the percentage change in social security wealth (light red bars) and stock of human capital (dark red bars) by income quintile between the alternative pension systems and the NDC-III system. We can see in Fig. 9 that the DB-I system leads to an increase in HK for all income quintiles. Under the mortality regime of the cohort 1930 the increase is higher for quintiles q2 and q3, while for the cohort 1960 the DB-I system benefits mostly the higher 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. A similar rise in human capital investment under the DB-I system is also derived in Sánchez-Romero and Prskawetz (2017). Consistent with the results shown for irr and $\mathcal{P}$, the DB-I system reduces the social security wealth of lower income quintiles q1–q3 and increases that of higher income quintiles q4–q5, especially under the mortality regime of 1960. Instead, the actual US pension system reduces the stock of human capital of higher income quintiles (see Fig. 9) due to the implicit tax on labor income, which is not offset by an increase in human capital of the low income quintiles. Moreover, under the mortality regime of the cohort born in 1960 the overall effect on the social security wealth is negative for the lower income quintiles and positive for the higher income quintiles. This is because individuals in the higher income quintiles response through changes in schooling and retirement so as to improve their lifetime wealth. The impact of the DB-III system on HK is similar to that in DB-I but now the progressivity of the pension system is re-introduced.



(a) Cohort 1930



(b) Cohort 1960

Figure 9: Redistributive effects of each pension system on lifetime wealth by income quintile (measured in percentage change with respect to the results in the NDC-III system). US males, birth cohorts 1930 and 1960.

Note: 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.

It is striking that in DB plans the indirect effects on HK (lifetime earnings) arising from incentives for school, work, and retirement, usually are far larger than the direct redistributive effects through the pension system. As shown in Fig. 6, this is because DB plans produce an incentive to retire at later ages, which do not necessarily coincide with those of the NDC plans. As a consequence, the increase in the retirement age leads to an increase in the length schooling and in the number of hours worked.

Finally, we can see in Fig. 9 that the size of the impact of the NDC systems is small in magnitude and mainly concentrated on the social security wealth. Under the NDC-I and NDC-II plans, lower income quintiles q1–q3 experience a reduction in their social security wealth, whereas the social security wealth of higher income quintiles increase. Under the mortality regime of the 1960 birth cohort, the increase in human capital under the NDC-I and NDC-II is due to a postponement of the retirement age compared to the NDC-III plan.

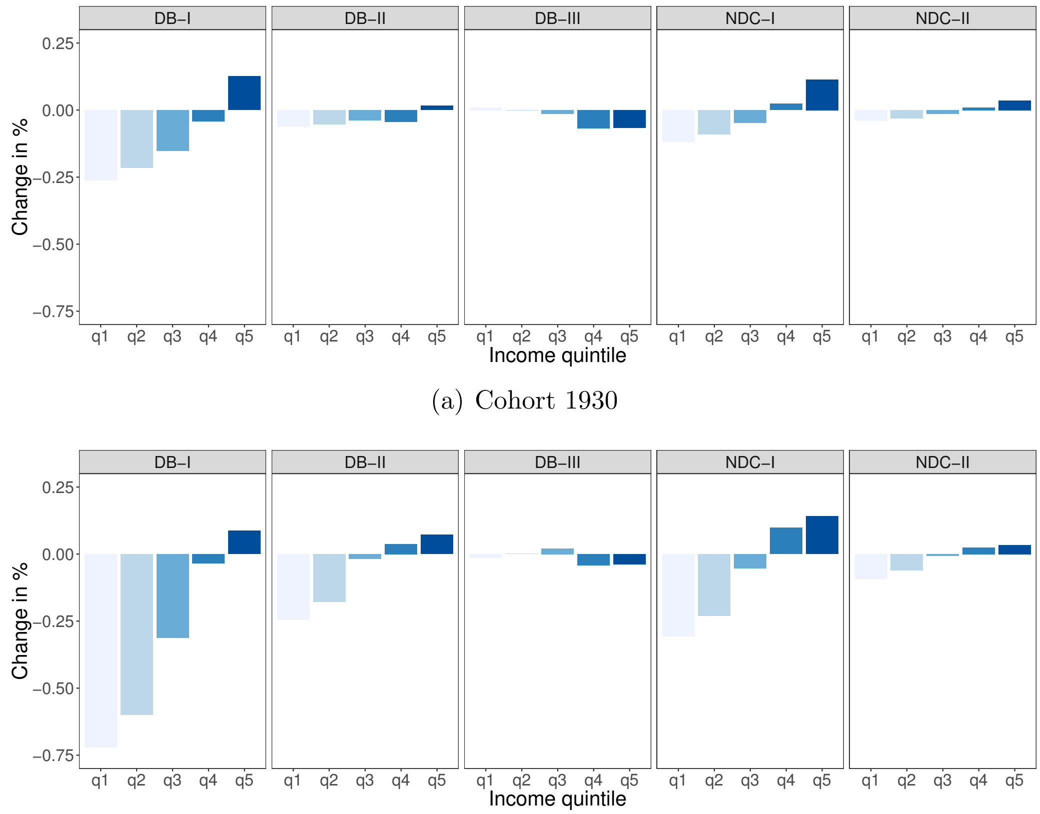
6.4 Welfare.

Above we have shown that differences in lifetime wealth come with changes in schooling and leisure time, through age at retirement and amount of labor supplied while working. Given that our lifetime utility (see Eq. 21) includes disutility of schooling, labor, and the utility from retirement, we can provide a comprehensive assessment of the impact of the different pension plans on lifetime welfare by income quintile. This exercise is key for understanding the results in Fig. 9, which shows the impact of each pension system on lifetime wealth.

In Figure 10 we show the impact on lifetime welfare of each pension system by income quintile relative to the NDC-III plan.[14] The first important result can be seen by comparing in Fig. 10 the differences between the NDC plans and DB plans. Recall looking at Fig. 6 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 64 in NDC plans, whereas individuals retire at older ages under the DB systems —see Table 5 in Appendix B.3. 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. 9), and the welfare lost through the heavy cost in leisure.

In NDC plans we do not observe significant differences in the length of schooling and the retirement age. Thus, the sign of the redistribution across income quintiles in the irr (Fig. 7), the P (Fig. 8), the lifetime wealth (Fig. 9), and the lifetime welfare all coincide, which is the same as that directly induced by the social security system. Moreover, given that the NDC-II is closer than the NDC-I to the NDC-III, the NDC-II plan creates less welfare differences across income quintiles than the NDC-I.

[14]Note that 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. Therefore, we can abstract from whether our baseline NDC-III is better or worse than no program at all.



(a) Cohort 1930

(b) Cohort 1960

Figure 10: Impact of each pension system on welfare by income quintile (relative to the NDC-III system). US males, birth cohorts 1930 and 1960.

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. 9), which is more pronounced for q1–q3 under the mortality regime of the 1930 cohort, and for q3–q4 under the mortality regime of the 1960 cohort. However, the increase in human capital comes at the expense of facing a higher disutility from schooling, longer working hours, 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. 7). 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 for q1–q3, the number of hours worked are reduced (decreasing the marginal benefit of education), and the length of schooling is shortened. The combination of these three behavioral reactions explains the reduction in human capital (see Fig 8) 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-II 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 PAYGO, 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 PAYGO.

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 not only have direct effects on systems and their progressivity given the current behavior 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 full effect of a variety of policy adjustments to DB and NDC pension programs operating in environments of more or less mortality heterogeneity (reflecting the mortality regimes of the 1930 vs 1960 birth cohorts), including both direct and indirect effects of these adjustments. We do this in a general equilibrium context 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 (no 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, policy adjustments have both direct and indirect effects. One way to assess these is through their impact on lifetime wealth. Under all NDC versions, the indirect effects of policy adjustments are quite small, and the relatively small direct effects are slightly regressive. For the DB systems, the indirect effects strongly dominate the direct ones, which is interesting in itself. The indirect effects of adding progressive benefits are slightly positive at lower incomes and 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, but not surprisingly the gains in direct effects from moving to group-specific life tables are larger under the more heterogeneous mortality regime.

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. In most pension systems, we find welfare losses

for lower incomes and small welfare gains for higher incomes. This pattern arises from 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 is achieved with the progressivity of benefits. However, only when correcting the benefits using life tables for each group, we observe similar lifetime utilities.

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 value of one dollar of social contribution, 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.

It is important to note that we are comparing outcomes across programs that are assumed to have existed over the long term. We have not yet attempted to investigate transitions from one program to another, although that is the situation that policy makers must face.

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., 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.

Bloom, D. E, Canning, D., Moore, M., 2014. Optimal retirement with increasing longevity. *The Scandinavian Journal of Economics* 116 (3), 838–858.

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.

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.

Feldstein, M., 1974. Social security, induced retirement, and aggregate capital accumulation. *Journal of Political Economy*, 82 (5), 905–926.

Gruber, J., Wise, D., 1999. Social security programs and retirement around the World. *Research in Labor Economics*, 18, 1–40.

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.

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.

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.

Appendix

A Derivation of Eq. (9)

Eq. (9) can easily be derived from (8) by elementary algebraic manipulations. Substituting (4) and (5) in (8) gives

$$\begin{aligned}\mathsf{SSW}_{is}(x) = e^{-r(R_i-x)}\frac{p_i(R_i)}{p_i(x)} f_{is}(R_i, \mathsf{pp}_{is}(R_i))A_i(R_i,r)\phi \int_{x_0}^{R_i} e^{\mathsf{r}(R_i-t)}\frac{p_s(t)}{p_s(R_i)}\tau y_i(t)dt- \\ - \int_x^{R_i} e^{-r(t-x)}\frac{p_i(t)}{p_i(x)}\tau y_i(t)dt. \qquad (A.1)\end{aligned}$$

Splitting in two the integral in the first term of the right-hand side of (A.1) gives

$$\begin{aligned}\mathsf{SSW}_{is}(x) = e^{-r(R_i-x)}\frac{p_i(R_i)}{p_i(x)} f_{is}(R_i, \mathsf{pp}_{is}(R_i))A_i(R_i,r)\phi \int_{x_0}^{x} e^{\mathsf{r}(R_i-t)}\frac{p_s(t)}{p_s(R_i)}\tau y_i(t)dt+ \\ + e^{-r(R_i-x)}\frac{p_i(R_i)}{p_i(x)} f_{is}(R_i, \mathsf{pp}_{is}(R_i))A_i(R_i,r)\phi \int_{x}^{R_i} e^{\mathsf{r}(R_i-t)}\frac{p_s(t)}{p_s(R_i)}\tau y_i(t)dt- \\ - \int_x^{R_i} e^{-r(t-x)}\frac{p_i(t)}{p_i(x)}\tau y_i(t)dt. \qquad (A.2)\end{aligned}$$

Multiplying and dividing the first two terms on the right-hand side of (A.2) by $e^{\mathsf{r}(x-R_i)}\frac{p_s(R_i)}{p_s(x)}$, using (4) and (10), gives

$$\begin{aligned}\mathsf{SSW}_{is}(x) = \mathsf{P}_{is}(x,R_i)\frac{\mathsf{pp}_{is}(x)}{\phi} + \mathsf{P}_{is}(x,R_i)\int_x^{R_i} e^{\mathsf{r}(x-t)}\frac{p_s(t)}{p_s(x)}\tau y_i(t)dt- \\ - \int_x^{R_i} e^{-r(t-x)}\frac{p_i(t)}{p_i(x)}\tau y_i(t)dt. \qquad (A.3)\end{aligned}$$

Next, multiplying and dividing the inner part of the first integral on the right-hand side of (A.3) by $e^{-r(t-x)}\frac{p_i(t)}{p_i(x)}$, and using the fact that $\mathsf{P}_{is}(t)$ is equal to $e^{(\mathsf{r}-r)(x-t)}\frac{p_i(x)}{p_i(t)}\frac{p_s(t)}{p_s(x)}\mathsf{P}_{is}(x)$, we have

$$\begin{aligned}\mathsf{SSW}_{is}(x) = \mathsf{P}_{is}(x,R_i)\frac{\mathsf{pp}_{is}(x)}{\phi} + \int_x^{R_i} e^{-r(t-x)}\frac{p_i(t)}{p_i(x)}\tau \mathsf{P}_{is}(t) y_i(t)dt- \\ - \int_x^{R_i} e^{-r(t-x)}\frac{p_i(t)}{p_i(x)}\tau y_i(t)dt. \qquad (A.4)\end{aligned}$$

Adding both integrals on the right-hand side of (A.4) and using (13) gives Eq. (9).

B Economic problem

We solve the problem of maximizing the lifetime utility (21) subject to the constraints (17)–(20) 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 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 $i \in \mathcal{I}$ before working $(t \leq S_i)$ is defined as

$$\begin{aligned} H^1 = {} & e^{-\rho(t-x_0)} \frac{p_i(t)}{p_i(x_0)} \left[U(c_i(t)) - \eta_i\right] + \lambda_a^1(t) \left[(r + \mu_i(t))a_i(t) - c_i(t)\right] + \\ & + \lambda_h[\theta_i h_i(t)^\gamma - \delta h_i(t)] \end{aligned} \tag{B.1}$$

where $\lambda_a^1(t)$ and $\lambda_h(t)$ are the adjacent 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:

$$H_c^1 = e^{-\rho(t-x_0)} \frac{p_i(t)}{p_i(x_0)} U'(c_i(t)) - \lambda_a^1(t) = 0. \tag{B.2}$$

The dynamic laws of the adjacent variables are:

$$\frac{\partial \lambda_a^1(t)}{\partial t} = -\lambda_1^1(t)(r + \mu_i(t)), \tag{B.3}$$

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

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

$$\begin{aligned} H^2 = {} & e^{-\rho(t-x_0)} \frac{p_i(t)}{p_i(x_0)} \left[U(c_i(t)) - \alpha_i v(\ell_i(t))\right] + \\ & + \lambda_a^2(t) \left[(r + \mu_i(t))a_i(t) + (1 - \tau)h_i(S_i)\bar{w}(t - S_i)\ell_i(t) - c_i(t)\right] + \\ & + \lambda_p^2(t) \left[(\mathsf{r} + \mu_s(t))\mathsf{pp}_{is}(t) + \phi\tau h_i(S_i)\bar{w}(t - S_i)\ell_i(t)\right], \end{aligned} \tag{B.5}$$

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

$$H_c^2 = e^{-\rho(t-x_0)} \frac{p_i(t)}{p_i(x_0)} U'(c_i(t)) - \lambda_a^2(t) = 0, \tag{B.6}$$

$$\begin{aligned} H_\ell^2 = &-e^{-\rho(t-x_0)} \frac{p_i(t)}{p_i(x_0)} \alpha_i v'(\ell_i(t)) + \lambda_a^2(t)(1-\tau)h_i(S_i)\bar{w}(t-S_i) + \\ &+ \lambda_p^2(t)\phi\tau h_i(S_i)\bar{w}(t-S_i) = 0. \end{aligned} \tag{B.7}$$

The dynamic laws of the adjacent variables are:

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

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

$$\frac{\partial \lambda_p^2(t)}{\partial t} = -\lambda_p^2(t)(\mathsf{r} + \mu_s(t)). \tag{B.10}$$

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

$$\begin{aligned} H^3 = &e^{-\rho(t-x_0)} \frac{p_i(t)}{p_i(x_0)} \left[U(c_i(t)) + \varphi(t)\right] \\ &+ \lambda_a^3(t) \left[(r + \mu_i(t))a_i(t) + \mathsf{f}_{is}(R_i, \mathsf{pp}_i(R_i))\mathsf{pp}_i(R_i) - c_i(t)\right], \end{aligned} \tag{B.11}$$

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

$$H_c^3 = e^{-\rho(t-x_0)} \frac{p_i(t)}{p_i(x_0)} U'(c_i(t)) - \lambda_a^2(t) = 0. \tag{B.12}$$

Using the definition (29), the dynamic laws of the adjacent variables are:

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

$$\frac{\partial \lambda_p^3(t)}{\partial t} = -\lambda_a^3(t)\mathsf{f}_{is}(R_i, \mathsf{pp}_i(R_i))(1 - \varepsilon_{is}). \tag{B.14}$$

Moreover, the following matching conditions hold at the switching ages S_i and R_i for the adjacent variables

$$\lambda_a(S_i) := \lambda_a^1(S_i) = \lambda_a^2(S_i), \tag{B.15a}$$
$$\lambda_h(S_i) := \lambda_h^1(S_i) = \lambda_h^2(S_i), \tag{B.15b}$$
$$\lambda_a(R_i) := \lambda_a^2(R_i) = \lambda_a^3(R_i), \tag{B.15c}$$
$$\lambda_p(R_i) := \lambda_p^2(R_i) = \lambda_p^3(R_i), \tag{B.15d}$$

and for the Hamiltonians

$$H(S_i) := H^1(S_i) = H^2(S_i), \tag{B.16a}$$
$$H(R_i) := H^2(R_i) = H^3(R_i). \tag{B.16b}$$

Taking into account the above matching conditions, let us define the marginal rate of substitution of assets for social contributions $\mathcal{P}(t) = \phi\lambda_p(t)/\lambda_a(t)$ for periods $\{2, 3\}$, and the marginal rate of substitution of assets for human capital $\mathcal{H}(t) = \lambda_h(t)/\lambda_a(t)$ for periods $\{1, 2\}$. Differentiating $\mathcal{P}(t)$ and $\mathcal{H}(t)$ with respect to time t, and using the dynamics of the adjoint variables, gives

$$\frac{\partial \mathcal{P}(t)}{\partial t} = \begin{cases} \mathcal{P}(t)(r - \mathsf{r} + \mu_i(t) - \mu_s(t)) & \text{for } S_i < t < R_i, \\ \mathcal{P}(t)(r + \mu_i(t)) - \phi \mathsf{f}_{is}(R_i, \mathsf{pp}_i(R_i))(1 - \varepsilon_{is}) & \text{for } t \geq R_i \end{cases} \tag{B.17}$$

$$\frac{\partial \mathcal{H}(t)}{\partial t} = \begin{cases} \mathcal{H}(t)(r + \mu_i(t) + \delta - \gamma\theta_i h_i(t)^{\gamma-1}) & \text{for } t \leq S_i, \\ \mathcal{H}(t)(r + \mu_i(t)) - (1 - \tau(t))\, \bar{w}(t - S_i)\ell_i(t) & \text{for } S_i < t < R_i, \end{cases} \tag{B.18}$$

where $\tau(t)$ is the marginal implicit tax/subsidy rate on labor income defined in (27). Solving (B.17) and (B.18) and using the fact that $\mathcal{P}(\omega) = 0$ and $\mathcal{H}(R_i) = 0$, the marginal rate of substitution of assets for a unit of social contribution and the marginal rate of substitution of assets for human capital are:

$$\mathcal{P}(t) = e^{(\mathsf{r}-r)(R_i-t)} \frac{p_i(R_i)}{p_i(t)} \frac{p_s(t)}{p_s(R_i)} \phi \mathsf{f}_{is}(R_i, \mathsf{pp}_i(R_i))(1 - \varepsilon_{is}) A_i(R_i; r) \tag{B.19}$$

and

$$\mathcal{H}(t) = \frac{h_i(t)}{h_i(S_i)} \int_{S_i}^{R_i} e^{-r(x-t)} \frac{p_i(x)}{p_i(t)} (1 - \tau_{is}(x)) \bar{w}(x - S_i)\ell_i(x) dx. \tag{B.20}$$

B.1 Optimal consumption (c) and hours worked (ℓ)

Now, using the budget constraint (19) at age x_0, we have

$$\int_{x_0}^{\omega} e^{-r(t-x_0)} \frac{p_i(t)}{p_i(x_0)} c_i(t) dt = \int_{S_i}^{R_i} e^{-r(t-x_0)} \frac{p_i(t)}{p_i(x_0)} (1 - \mathsf{t}_i(t)) w_i(t, S_i) \ell_i(t) dt. \tag{B.21}$$

Using (25)–(26) and the FOCs on consumption and labor supply along the three periods, we have

$$\begin{aligned} &\int_{x_0}^{\omega} e^{-r(t-x_0)} \frac{p_i(t)}{p_i(x_0)} \left(e^{-\rho(t-x_0)} \frac{p_i(t)}{p_i(x_0)} \frac{1}{\lambda_a(t)} \right)^{\sigma_c} dt = \\ &\qquad = \int_{S_i}^{R_i} e^{-r(t-x_0)} \frac{p_i(t)}{p_i(x_0)} w \mathsf{t}_i(t) \left(\frac{1}{\alpha_i} \frac{\lambda_a(t)}{e^{-\rho(t-x_0)} \frac{p_i(t)}{p_i(x_0)}} w \uptau_i(t) \right)^{\sigma_\ell} dt. \end{aligned} \tag{B.22}$$

Solving the dynamics of the adjoint variables for the capital stock, we have

$$\lambda_a(t) = \lambda_a(x_0) e^{-r(t-x_0)} \frac{p_i(t)}{p_i(x_0)}. \tag{B.23}$$

Using (24), substituting (B.23) in (B.22), and after rearranging terms, we obtain that $\lambda_a(x_0)$ is given by

$$\lambda_a(x_0) = \left(\frac{(\alpha_i)^{\sigma_\ell} \int_{x_0}^{\omega} D(t; x_0, \sigma_c) dt}{\int_{S_i}^{R_i} D(t; x_0, -\sigma_\ell) w \mathsf{t}_i(t) \left(w \uptau_i(t) \right)^{\sigma_\ell} dt} \right)^{\frac{1}{\sigma_\ell + \sigma_c}}. \tag{B.24}$$

Therefore, using the FOCs, the optimal consumption and hours worked at age x are given by

$$c_i(x) = e^{\sigma_c (r-\rho)(x-x_0)} \frac{1}{\lambda_a(x_0)^{\sigma_c}}, \tag{B.25}$$

$$\ell_i(x) = e^{\sigma_\ell (\rho-r)(x-x_0)} \left(\frac{w \uptau_i(x)}{\alpha_i} \lambda_a(x_0) \right)^{\sigma_\ell}. \tag{B.26}$$

Note that substituting (B.24) in (B.25)–(B.26) gives (22)–(23), respectively.

B.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

$$\int_{x_0}^{\omega} e^{-\rho(t-x_0)} \frac{p_i(t)}{p_i(x_0)} U'\left(c_i(t)\right) \frac{\partial c_i(t)}{\partial S} dt - \int_{S}^{R_i} e^{-\rho(t-x_0)} \frac{p_i(t)}{p_i(x_0)} \alpha_i v'\left(\ell_i(t)\right) \frac{\partial \ell_i(t)}{\partial S} dt$$
$$= e^{-\rho(S-x_0)} \frac{p_i(S)}{p_i(x_0)} \left(\eta_i - \alpha_i v(\ell(S))\right). \tag{B.27}$$

Substituting the FOCs in the previous equation gives

$$\int_{x_0}^{\omega} \lambda_a(t) \frac{\partial c_i(t)}{\partial S} dt - \int_{S}^{R_i} \lambda_a(t)(1-\tau_{is}(t)) w_i(t, S_i) \frac{\partial \ell_i(t)}{\partial S} dt$$
$$= e^{-\rho(S-x_0)} \frac{p_i(S)}{p_i(x_0)} \left(\eta_i - \alpha_i v(\ell(S))\right). \tag{B.28}$$

Using (B.23) and rearranging terms gives

$$\int_{x_0}^{\omega} e^{-r(t-x_0)} \frac{p_i(t)}{p_i(x_0)} \frac{\partial c_i(t)}{\partial S} dt - \int_{S}^{R_i} e^{-r(t-x_0)} \frac{p_i(t)}{p_i(x_0)} (1-\tau_{is}(t)) w_i(t, S_i) \frac{\partial \ell_i(t)}{\partial S} dt$$
$$= e^{-\rho(S-x_0)} \frac{p_i(S)}{p_i(x_0)} \frac{\eta_i - \alpha_i v(\ell(S))}{\lambda_a(x_0)}. \tag{B.29}$$

Second, we differentiate the budget constraint (19) at age x_0 w.r.t. S

$$\begin{aligned}\int_{x_0}^{\omega} e^{-r(t-x_0)} \frac{p_i(t)}{p_i(x_0)} \frac{\partial c_i(t)}{\partial S} dt & \\ = \int_{S}^{R_i} e^{-r(t-x_0)} \frac{p_i(t)}{p_i(x_0)} (1-\mathsf{t}_{is}(t)) w_i(t, S) \frac{\partial \ell_i(t)}{\partial S} dt + & \\ + \int_{S}^{R_i} e^{-r(t-x_0)} \frac{p_i(t)}{p_i(x_0)} (1-\mathsf{t}_{is}(t)) \frac{\partial w_i(t, S)}{\partial S} \ell_i(t) dt - & \\ - \int_{S}^{R_i} e^{-r(t-x_0)} \frac{p_i(t)}{p_i(x_0)} \frac{\partial \mathsf{t}_{is}(t)}{\partial S} w_i(t, S) \ell_i(t) dt - & \\ - e^{-r(S-x_0)} \frac{p_i(S)}{p_i(x_0)} (1-\mathsf{t}_i(S)) w_i(S, S) \ell_i(S).\end{aligned} \tag{B.30}$$

From (10), (13), and using the definition of ε_{is} in (29), we have

$$\frac{\partial \mathsf{t}_{is}(t)}{\partial S} = \tau \varepsilon_{is} \mathsf{P}_{is}(t) \frac{1}{\mathsf{pp}_{is}(R_i)} \frac{\partial \mathsf{pp}_{is}(R_i)}{\partial S}. \tag{B.31}$$

Substituting (B.31) on the third term on the right-hand side of Eq. (B.30), and using $\mathsf{P}_{is}(t) = \mathsf{P}_{is}(R_i)e^{(\mathsf{r}-r)(R_i-t)}\frac{p_i(R_i)}{p_i(t)}\frac{p_s(t)}{p_s(R_i)}$ gives

$$\frac{1}{\mathsf{pp}_{is}(R_i)}\frac{\partial \mathsf{pp}_{is}(R_i)}{\partial S}\varepsilon_{is}\int_S^{R_i} e^{-r(t-x_0)}\frac{p_i(t)}{p_i(x_0)}\tau \mathsf{P}_{is}(t)w_i(t,S)\ell_i(t)dt =$$
$$= \frac{1}{\mathsf{pp}_{is}(R_i)}\frac{\partial \mathsf{pp}_{is}(R_i)}{\partial S}\mathsf{P}_{is}(R_i)\varepsilon_{is}e^{-r(R_i-x_0)}\frac{p_i(R_i)}{p_i(x_0)}\int_S^{R_i} e^{\mathsf{r}(R_i-t)}\frac{p_s(t)}{p_s(R_i)}\tau w_i(t,S)\ell_i(t)dt =$$
$$= \frac{\partial \mathsf{pp}_{is}(R_i)}{\partial S}\frac{\mathsf{P}_{is}(R_i)}{\phi}\varepsilon_{is}e^{-r(R_i-x_0)}\frac{p_i(R_i)}{p_i(x_0)} \quad \text{(B.32)}$$

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

$$\frac{\partial \mathsf{pp}_{is}(R_i)}{\partial S} = \phi\tau\int_S^{R_i} e^{\mathsf{r}(R_i-t)}\frac{p_s(t)}{p_s(R_i)}w_i(t,S)\frac{\partial \ell_i(t)}{\partial S}dt +$$
$$+ \phi\tau\int_S^{R_i} e^{\mathsf{r}(R_i-t)}\frac{p_s(t)}{p_s(R_i)}\frac{\partial w_i(t,S)}{\partial S}\ell_i(t)dt +$$
$$- \phi\tau e^{\mathsf{r}(R_i-S)}\frac{p_s(S)}{p_s(R_i)}w_i(S,S)\ell_i(S) \quad \text{(B.33)}$$

Then, plugging (B.33) in (B.32), and using the fact that $\mathsf{P}_{is}(R_i)$ can be rewritten as $\mathsf{P}_{is}(t)e^{-(\mathsf{r}-r)(R_i-t)}\frac{p_i(t)}{p_i(R_i)}\frac{p_s(R_i)}{p_s(t)}$, gives

$$\frac{\partial \mathsf{pp}_{is}(R_i)}{\partial S}\frac{\mathsf{P}_{is}(R_i)}{\phi}\varepsilon_{is}e^{-r(R_i-x_0)}\frac{p_i(R_i)}{p_i(x_0)} =$$
$$= \int_S^{R_i} e^{-r(t-x_0)}\frac{p_i(t)}{p_i(x_0)}\tau \mathsf{P}_{is}(t)\varepsilon_{is}w_i(t,S)\frac{\partial \ell_i(t)}{\partial S}dt +$$
$$+ \int_S^{R_i} e^{-r(t-x_0)}\frac{p_i(t)}{p_i(x_0)}\tau \mathsf{P}_{is}(t)\varepsilon_{is}\frac{\partial w_i(t,S)}{\partial S}\ell_i(t)dt +$$
$$- e^{-r(S-x_0)}\frac{p_i(S)}{p_i(x_0)}\tau \mathsf{P}_{is}(S)\varepsilon_{is}w_i(S,S)\ell_i(S) \quad \text{(B.34)}$$

Now, substituting (B.34) in (B.30) gives

$$\int_{x_0}^{\omega} e^{-r(t-x_0)}\frac{p_i(t)}{p_i(x_0)}\frac{\partial c_i(t)}{\partial S}dt - \int_S^{R_i} e^{-r(t-x_0)}\frac{p_i(t)}{p_i(x_0)}(1-\tau_{is}(t))w_i(t,S)\frac{\partial \ell_i(t)}{\partial S}dt =$$
$$= \int_S^{R_i} e^{-r(t-x_0)}\frac{p_i(t)}{p_i(x_0)}(1-\tau_{is}(t))\frac{\partial w_i(t,S)}{\partial S}\ell_i(t)dt -$$
$$- e^{-r(S-x_0)}\frac{p_i(S)}{p_i(x_0)}(1-\tau_{is}(S))w_i(S,S)\ell_i(S). \quad \text{(B.35)}$$

Using the fact that the left-hand side of (B.2) is equal to (B.36), then

$$e^{-\rho(S-x_0)}\frac{p_i(S)}{p_i(x_0)}\frac{\eta_i-\alpha_i v(\ell(S))}{\lambda_a(x_0)}=\int_S^{R_i}e^{-r(t-x_0)}\frac{p_i(t)}{p_i(x_0)}(1-\tau_{is}(t))\frac{\partial w_i(t,S)}{\partial S}\ell_i(t)dt- \\ -e^{-r(S-x_0)}\frac{p_i(S)}{p_i(x_0)}(1-\tau_{is}(S))w_i(S,S)\ell_i(S). \quad \text{(B.36)}$$

Differentiating $w_i(t,S)$ w.r.t. S gives

$$\begin{aligned}\frac{\partial w_i(t,S)}{\partial S}&=\frac{\partial\bar{w}(t-S)}{\partial S}h_i(S)+\bar{w}(t-S)\frac{\partial h_i(S)}{\partial S}\\&=-\frac{1}{\bar{w}(t-S)}\frac{\partial\bar{w}(t-S)}{\partial t}w_i(t,S)+\frac{1}{h_i(S)}\frac{\partial h_i(S)}{\partial S}w_i(t,S)\end{aligned} \quad \text{(B.37)}$$

Third, we use (B.37) in (B.36)

$$\begin{aligned}&e^{-\rho(S-x_0)}\frac{p_i(S)}{p_i(x_0)}\frac{\eta_i-\alpha_i v(\ell(S))}{\lambda_a(x_0)}=\\&\quad=\frac{\frac{\partial h_i(S)}{\partial S}}{h_i(S)}\int_S^{R_i}e^{-r(t-x_0)}\frac{p_i(t)}{p_i(x_0)}(1-\tau_{is}(t))w_i(t,S)\ell_i(t)dt+\\&\quad-\int_S^{R_i}e^{-r(t-x_0)}\frac{p_i(t)}{p_i(x_0)}\frac{\frac{\partial\bar{w}(t-S)}{\partial t}}{\bar{w}(t-S)}(1-\tau_{is}(t))w_i(t,S)\ell_i(t)dt-\\&\qquad-e^{-r(S-x_0)}\frac{p_i(S)}{p_i(x_0)}(1-\tau_{is}(S))w_i(S,S)\ell_i(S).\end{aligned} \quad \text{(B.38)}$$

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

$$\begin{aligned}&\frac{\frac{\partial h_i(S)}{\partial S}}{h_i(S)}\int_S^{R}e^{-r(t-x_0)}\frac{p_i(t)}{p_i(x_0)}(1-\tau_{is}(t))w_i(t,S)\ell_i(t)dt=\\&\quad=\int_S^{R_i}e^{-r(t-x_0)}\frac{p_i(t)}{p_i(x_0)}\frac{\frac{\partial\bar{w}(t-S)}{\partial t}}{\bar{w}(t-S)}(1-\tau_{is}(t))w_i(t,S)\ell_i(t)dt+\\&\qquad+e^{-r(S-x_0)}\frac{p_i(S)}{p_i(x_0)}(1-\tau_{is}(S))w_i(S,S)\ell_i(S)+\\&\qquad\quad+e^{-\rho(S-x_0)}\frac{p_i(S)}{p_i(x_0)}\frac{\eta_i-\alpha_i v(\ell_i(S))}{\lambda_a(x_0)}\end{aligned} \quad \text{(B.39)}$$

Defining the net human capital wealth at age S out of effective labor income tax of an individual of type i as

$$W(S,R)=\int_S^{R}e^{-r(t-S)}\frac{p_i(t)}{p_i(S)}(1-\tau_{is}(t))w_i(t,S)\ell_i(t)dt. \quad \text{(B.40)}$$

and dividing both sides of (B.39) by $W(S, R_i)$, multiplying by $e^{r(S-x_0)}\frac{p_i(x_0)}{p_i(S)}$, we obtain the optimal length of schooling condition

$$r_i^h(S) = \bar{r}_i(S, R_i) + \frac{\eta_i - \alpha_i v(\ell_i(S))}{U'(c_i(S))W_i(S, R_i)}. \tag{B.41}$$

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

$$r_i^h(S) = \frac{1}{h_i(S)}\frac{\partial h_i(S)}{\partial S}, \tag{B.42}$$

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

$$\bar{r}_i(S, R_i) = \int_S^{R_i} \frac{\frac{\partial \bar{w}(t-S)}{\partial t}}{\bar{w}(t-S)}\psi_{is}(t)dt + \psi_{is}(S), \tag{B.43}$$

where

$$\psi_{is}(t) = \frac{D(t; x_0, -\sigma_\ell)\left(w\tau_{is}(t)\right)^{1+\sigma_\ell}}{\int_S^{R_i} D(u; x_0, -\sigma_\ell)\left(w\tau_{is}(u)\right)^{1+\sigma_\ell} du}. \tag{B.44}$$

Note from (B.43) and (B.44) we have $\int_S^{R_i} \psi_{is}(t)dt = 1$ and $\lim_{S\to R_i} \bar{r}_i(S, R_i) = 1$. The last term in (B.41), which represents the non-pecuniary cost of schooling, is

$$\begin{aligned} &\frac{\eta_i - \alpha_i v(\ell_i(S))}{U'(c_i(S))W_i(S, R_i)} = \\ &= e^{-\rho(S-x_0)}\frac{p_i(S)}{p_i(x_0)}\eta_i \frac{(\alpha_i)^{\sigma_\ell}\lambda_a(x_0)^{-1-\sigma_\ell}}{\int_S^{R_i} D(t; x_0, -\sigma_\ell)(w\tau_{is}(t))^{1+\sigma_\ell}dt} - \frac{\sigma_\ell}{1+\sigma_\ell}\psi_{is}(S). \end{aligned} \tag{B.45}$$

Note that assuming $\sigma_c = 1$ and $\rho = 0$, we get (35).

B.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

$$\int_{x_0}^{\omega} e^{-\rho(t-x_0)} \frac{p_i(t)}{p_i(x_0)} U'(c_i(t)) \frac{\partial c_i(t)}{\partial R} dt - \alpha_i \int_{S}^{R} e^{-\rho(t-x_0)} \frac{p_i(t)}{p_i(x_0)} v'(\ell_i(t)) \frac{\partial \ell_i(t)}{\partial R} dt$$
$$= e^{-\rho(R-x_0)} \frac{p_i(R)}{p_i(x_0)} \left(\alpha_i v(\ell_i(R)) + \varphi(R)\right). \tag{B.46}$$

Substituting the FOCs in the previous equation gives

$$\int_{x_0}^{\omega} \lambda_a(t) \frac{\partial c_i(t)}{\partial R} dt - \int_{S}^{R} \lambda_a(t)(1-\tau_{is}(t)) w_i(t,S_i) \frac{\partial \ell_i(t)}{\partial R} dt$$
$$= e^{-\rho(R-x_0)} \frac{p_i(R)}{p_i(x_0)} \left(\alpha_i v(\ell_i(R)) + \varphi(R)\right). \tag{B.47}$$

Using (B.23) and rearranging terms gives

$$\int_{x_0}^{\omega} e^{-r(t-x_0)} \frac{p_i(t)}{p_i(x_0)} \frac{\partial c_i(t)}{\partial R} dt - \int_{S_i}^{R} e^{-r(t-x_0)} \frac{p_i(t)}{p_i(x_0)} (1-\tau_{is}(t)) w_i(t,S_i) \frac{\partial \ell_i(t)}{\partial R} dt$$
$$= e^{-\rho(R-x_0)} \frac{p_i(R)}{p_i(x_0)} \frac{\alpha_i v(\ell_i(R)) + \varphi(R)}{\lambda_a(x_0)}. \tag{B.48}$$

Second, we differentiate the budget constraint (19) at age x_0 w.r.t. R

$$\int_{x_0}^{\omega} e^{-r(t-x_0)} \frac{p_i(t)}{p_i(x_0)} \frac{\partial c_i(t)}{\partial R} dt =$$
$$= \int_{S_i}^{R} e^{-r(t-x_0)} \frac{p_i(t)}{p_i(x_0)} (1-\mathsf{t}_{is}(t)) w_i(t,S) \frac{\partial \ell_i(t)}{\partial R} dt -$$
$$- \int_{S_i}^{R} e^{-r(t-x_0)} \frac{p_i(t)}{p_i(x_0)} \frac{\partial \mathsf{t}_{is}(t)}{\partial R} w_i(t,S) \ell_i(t) dt +$$
$$+ e^{-r(R-x_0)} \frac{p_i(R)}{p_i(x_0)} (1-\mathsf{t}_{is}(R)) w_i(R,S) \ell_i(R). \tag{B.49}$$

From (4), (10), (13), and using the definition of ε_{is} in (29), we have

$$\frac{\partial \mathsf{t}_{is}(t)}{\partial R} = -\tau \mathsf{P}_{is}(t) \frac{1}{\mathsf{P}_{is}(t)} \frac{\partial \mathsf{P}_{is}(t)}{\partial R} =$$
$$= -\tau \mathsf{P}_{is}(t) \left(\mathsf{r} + \mu_s(R) + \frac{1}{\mathsf{f}_{is}} \frac{\partial \mathsf{f}_{is}}{\partial R} - \frac{1}{A_i(R,r)}\right) +$$
$$+ \tau \mathsf{P}_{is}(t) \varepsilon_{is} \frac{1}{\mathsf{pp}_{is}(R)} \frac{\partial \mathsf{pp}_{is}(R)}{\partial R}. \tag{B.50}$$

Substituting (B.50) on Eq. (B.49), and using $\mathsf{P}_{is}(t) = \mathsf{P}_{is}(R_i) e^{(\mathsf{r}-r)(R_i - t)} \frac{p_i(R_i)}{p_i(t)} \frac{p_s(t)}{p_s(R_i)}$ gives

$$\begin{aligned}\int_{x_0}^{\omega} e^{-r(t-x_0)} \frac{p_i(t)}{p_i(x_0)} \frac{\partial c_i(t)}{\partial R} dt = \\ = \int_{S_i}^{R} e^{-r(t-x_0)} \frac{p_i(t)}{p_i(x_0)} (1 - \mathsf{t}_{is}(t)) w_i(t,S) \frac{\partial \ell_i(t)}{\partial R} dt + \\ + \left(\mathsf{r} + \mu_s(R) + \frac{1}{\mathsf{f}_{is}} \frac{\partial \mathsf{f}_{is}}{\partial R} - \frac{1}{A_i(R,r)}\right) e^{-r(R-x_0)} \frac{p_i(R)}{p_i(x_0)} \mathsf{P}_{is}(R) \frac{\mathsf{pp}_{is}(R)}{\phi} - \\ - \frac{\partial \mathsf{pp}_{is}(R)}{\partial R} e^{-r(R-x_0)} \frac{p_i(R)}{p_i(x_0)} \frac{\mathsf{P}_{is}(R) \varepsilon_{is}}{\phi} + \\ + e^{-r(R-x_0)} \frac{p_i(R)}{p_i(x_0)} (1 - \mathsf{t}_{is}(R)) w_i(R,S) \ell_i(R). \quad \text{(B.51)}\end{aligned}$$

Differentiating the total pension points at age R with respect to R gives

$$\begin{aligned}\frac{\partial \mathsf{pp}_{is}(R)}{\partial R} = \phi \int_{S_i}^{R} e^{\mathsf{r}(R-t)} \frac{p_s(t)}{p_s(R)} \tau w_i(t,S) \frac{\partial \ell_i(t)}{\partial R} dt + \\ + (\mathsf{r} + \mu_s(R)) \mathsf{pp}_{is}(R) + \phi \tau w_i(R, S_i) \ell_i(R) \quad \text{(B.52)}\end{aligned}$$

Plugging (B.52) in (B.51) gives, after rearranging terms,

$$\begin{aligned}\int_{x_0}^{\omega} e^{-r(t-x_0)} \frac{p_i(t)}{p_i(x_0)} \frac{\partial c_i(t)}{\partial R} dt - \int_{S_i}^{R} e^{-r(t-x_0)} \frac{p_i(t)}{p_i(x_0)} (1 - \tau_{is}(t)) w_i(t,S) \frac{\partial \ell_i(t)}{\partial R} dt = \\ = \left((\mathsf{r} + \mu_s(R))(1 - \varepsilon_{is}) + \frac{1}{\mathsf{f}_{is}} \frac{\partial \mathsf{f}_{is}}{\partial R} - \frac{1}{A_i(R,r)}\right) e^{-r(R-x_0)} \frac{p_i(R)}{p_i(x_0)} \mathsf{P}_{is}(R) \frac{\mathsf{pp}_{is}(R)}{\phi} + \\ + e^{-r(R-x_0)} \frac{p_i(R)}{p_i(x_0)} (1 - \tau_{is}(R)) w_i(R,S) \ell_i(R). \quad \text{(B.53)}\end{aligned}$$

Using the fact that the left-hand side of (B.3) is equal to (B.53), then

$$\begin{aligned}e^{-\rho(R-x_0)} \frac{p_i(R)}{p_i(x_0)} \frac{\alpha_i v(\ell_i(R)) + \varphi(R)}{\lambda_a(x_0)} = \\ = \left((\mathsf{r} + \mu_s(R))(1 - \varepsilon_{is}) + \frac{1}{\mathsf{f}_{is}} \frac{\partial \mathsf{f}_{is}}{\partial R} - \frac{1}{A_i(R,r)}\right) e^{-r(R-x_0)} \frac{p_i(R)}{p_i(x_0)} \mathsf{P}_{is}(R) \frac{\mathsf{pp}_{is}(R)}{\phi} + \\ + e^{-r(R-x_0)} \frac{p_i(R)}{p_i(x_0)} (1 - \tau_{is}(R)) w_i(R,S) \ell_i(R). \quad \text{(B.54)}\end{aligned}$$

Multiplying both sides of the equation by $e^{r(R-x_0)}\frac{p_i(x_0)}{p_i(R)}$ gives

$$
\begin{aligned}
&e^{(r-\rho)(R-x_0)}\frac{\alpha_i v\left(\ell_i(R)\right)+\varphi(R)}{\lambda_a(x_0)} = \\
&\quad = \left((\mathsf{r}+\mu_s(R))(1-\varepsilon_{is})+\frac{1}{\mathsf{f}_{is}}\frac{\partial \mathsf{f}_{is}}{\partial R}-\frac{1}{A_i(R,r)}\right)\mathsf{P}_{is}(R)\frac{\mathsf{pp}_{is}(R)}{\phi}+ \\
&\qquad + (1-\tau_{is}(R))w_i(R,S)\ell_i(R).
\end{aligned} \tag{B.55}
$$

Using (38) and the fact that $U'(c_i(R)) = e^{(r-\rho)(R-x_0)}\lambda_a(x_0)$ we get the optimal retirement age condition

$$
\alpha_i v\left(\ell_i(R)\right)+\varphi(R) = U'(c_i(R))w_i(R,S)\ell_i(R)(1-\tau_{is}^{GW}(R)), \tag{B.56}
$$

which coincides with (36).

■

Additional simulated data

Length of schooling S_i.

Table 4: Optimal length of schooling by income quintile (S_i), US male birth cohorts 1930 and 1960

	Defined Contribution (NDC)			Defined Benefit		
	Avg. LT NDC-I	Corrected Avg. LT NDC-II	i–th LT NDC-III	Non–progressive DB-I	Progressive DB-II	Progressive Corrected DB-III
Cohort 1930						
Quintile 1	12	12	12	12	12	12
Quintile 2	12	12	12	13	12	12
Quintile 3	13	13	13	14	12	12
Quintile 4	14	14	14	14	13	13
Quintile 5	17	17	17	17	16	16
Cohort 1960						
Quintile 1	11	11	11	12	11	11
Quintile 2	13	13	13	13	13	13
Quintile 3	15	15	15	16	14	14
Quintile 4	17	17	17	18	16	16
Quintile 5	19	19	19	20	18	18

Retirement ages R_i.

Table 5: 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	61	61	61	63	62	63
Quintile 2	61	61	61	63	63	63
Quintile 3	61	61	61	64	63	63
Quintile 4	62	62	62	64	63	63
Quintile 5	64	64	64	65	64	64
Cohort 1960						
Quintile 1	60	60	60	62	61	62
Quintile 2	61	61	61	63	62	63
Quintile 3	63	63	63	66	63	63
Quintile 4	65	65	64	68	65	65
Quintile 5	66	66	65	69	66	66

Present value of lifetime benefits.

Lifetime wealth.

Table 7 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. 7 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.

Table 6: 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	115	126	133	135	128	142
Quintile 2	125	134	138	145	141	148
Quintile 3	154	161	163	190	166	170
Quintile 4	210	208	205	272	193	185
Quintile 5	269	249	223	337	229	209
Cohort 1960						
Quintile 1	94	124	137	97	105	134
Quintile 2	117	144	154	128	128	158
Quintile 3	186	196	197	252	188	189
Quintile 4	290	268	261	419	282	224
Quintile 5	327	293	283	525	293	259

Table 7: Distribution of the lifetime wealth (LW) at age 14 by income quintile and mortality regime in the NDC-III system

	Cohort 1930			Cohort 1960†		
	SSW	HK	LW	SSW	HK	LW
	I	II	III=I+II	I	II	III=I+II
Quintile 1	-2,2	78,2	76,0	-2,2	72,1	70,0
Quintile 2	-2,5	86,5	84,0	-2,8	92,7	89,9
Quintile 3	-2,9	102,9	100,0	-3,8	124,4	120,6
Quintile 4	-3,5	125,9	122,4	-5,0	163,5	158,4
Quintile 5	-4,0	145,2	141,2	-5,2	173,1	167,9

Notes: SSW stands for the social security wealth, HK denotes the stock of human capital, and LW is the lifetime wealth. †Values for the cohort 1960 are detrended by an annual productivity growth of 1.5 percent.