# Redistributive effects of pension reforms: Who are the winners and losers?⋆

Miguel Sanchez-Romero[a,*], Philip Schuster[b], Alexia Prskawetz[c,a]

[a] *VID, Wittgenstein Centre for Demography and Global Human Capital (IIASA, ÖAW, Univ. Vienna), Austria*
[b] *Office of the Austrian Fiscal Advisory Council c/o Oesterreichische Nationalbank*
[c] *Institute of Statistics and Mathematical Methods in Economics, Research Unit Economics, TU Wien, Austria*

---

**Abstract**

As the heterogeneity in life expectancy by socioeconomic status increases, pension systems become more regressive implying wealth transfers from short to long lived individuals. Various pension reforms aim to reduce these inequalities that are caused by ex-ante differences in life expectancy. However, these pension reforms may themselves induce redistribution effects since a) life expectancy is not perfectly correlated to socioeconomic status and b) pension reforms themselves will have an impact on life cycle decisions (education, consumption, health, labor supply) and ultimately also on life expectancy and the composition of the population. To account for these feedback effects of pension reforms in heterogenous aging societies we propose an OLG framework that is populated by heterogeneous individuals that initially differ by their learning ability and disutility from the effort of attending schooling. These initial heterogeneities imply differences in ex ante life expectancies. Within this framework we study two pension reforms that aim to account for these differences in ex ante life expectancies. We show that by including the feedback of pension reforms on individual behavior, new redistributions may result.

*Keywords:* Overlapping generations, Mortality and fertility differentials, Inequality, Life cycle, Pensions, Progressivity

---

## 1. Introduction

Many studies have shown a negative and increasing correlation between mortality rates and higher socioeconomic status (SES) by occupation, education, income, and even wealth (Preston

⋆This project has received funding from the Austrian National Bank (OeNB) under Grant no. 17647. We thank Stefani Rivic for collecting historical data for Austria, Bernhard Hammer for sharing NTA data for Austria, Yuliya Kulikova and participants at Viennese Vintage Workshop 2019: "Heterogeneous Dynamic Models of Economic and Population Systems", the Fiscal Policy Modelling Workshop of the European Commission, and NOeG 2020 for giving us comments that helped us improving the paper.

*Corresponding author
*Email address:* miguel.sanchez@oeaw.ac.at (Miguel Sanchez-Romero)

and Elo, 1995; Lleras-Muney, 2005; Waldron, 2007; Manchester and Topoleski, 2008; Luy et al., 2011; Olshansky et al., 2012; NASEM, 2015; Chetty et al., 2016). The results imply a widening of the difference in life expectancy between high and low SES in recent decades. One consequence of this demographic trend is that pension systems become more regressive, since through risk pooling low SES groups unexpectedly subsidize the pension benefits of high SES groups. Because individuals who have on average a higher life expectancy receive their benefits for more years compared to those who have a low life expectancy.[1] Thus, besides the necessary pension reform to cope with the increasing life expectancy at retirement and its long-run sustainability, policy makers also need to consider that individual aging is heterogeneous across SES groups and propose reforms against the increasing regressivity of pension systems.

In a pension system with a flat pension replacement rate, a reform to avoid the increasing regressivity across SES groups would imply that all SES groups receive the same return from their pension plans regardless their ex-ante life expectancy.[2] This reform can be implemented through changes in contributions or in benefits. Since different contribution rates by occupation may create unwanted labor incentives (Pestieau and Racionero, 2016), this option is generally not considered by pension systems. Instead, many proposals suggest modifying the benefit formula, given that once individuals claim their pension benefits, they cannot modify their working lives. However, it is still likely that individuals may react to changes in the benefit formula before retiring (Sánchez-Romero et al., 2020; Sánchez-Romero and Prskawetz, 2020), which may also induce an unwanted redistribution of resources. Thus, in this paper we will study the redistributive properties of such reforms within different birth cohorts not just at the time of retirement but over the whole lifecycle.

For such reforms it is important to correctly choose the SES variable(s) used for differentiating across groups. In particular, the choice of the SES variables should be based on two main criteria. First, it should capture the strength of the increase in the longevity gap by SES and, second, it should not change over the lifecycle.[3] In the literature, the most frequently suggested SES measures, that satisfy both criteria, are education and lifetime labor income. However, neither lifetime labor income nor education can by itself account for the full variance

[1] For a detailed review of the heterogeneity in life expectancy by SES and its implication on pension schemes see Ayuso et al. (2016), Auerbach et al. (2016), Lee and Sanchez-Romero (2019), Palmer and de Gosson de Varennes (2019), and Holzmann et al. (2019).

[2] Ex-ante differences in life expectancy arise from differences in the probability of death, while the ex-post difference in the length of life arises from the random process of death.

[3] See Lee and Sanchez-Romero (2019) for a discussion.

in life expectancy by SES (Bosworth et al., 2016). Consequently, any model aiming at analyzing the redistributive properties of reforms that aim to reduce the regressivity of pensions should consider a population that is at least heterogeneous with respect to life expectancy, education, and lifetime labor income. This implies the necessity of implementing a model with more than one degree of heterogeneity. The models of Fehr et al. (2013), Fehr and Uhde (2013, 2014) and more recently Laun et al. (2019), in which agents face idiosyncratic income risk, disability risk and mortality risk by skill group, and the distribution of skill groups is the same across cohorts, are potential candidates. However, in reality, education is changing across birth cohorts, which may cause that the observed increasing gap in life expectancy by educational attainment is just driven by the fact that the low educated group becomes more negatively selected over time (Goldring et al., 2016; Hendi et al., 2021). To control for selection between life expectancy and SES, our model allows individuals to (endogenously) choose their educational attainment based on their initial endowments and, similar to Pestieau and Ponthiere (2016), we link mortality and fertility to their education decision.[4] We include heterogeneity in the schooling effort (Sánchez-Romero et al., 2016; Sánchez-Romero and Prskawetz, 2020) to avoid that the inequality in life expectancy (as determined by the endogenous schooling decision) is driven by responsibility and not by circumstances (Fleurbaey, 2008). Thus, individuals with a high learning ability do not necessarily reach the highest educational attainment and thereby life expectancy, because they can face additional psychological and social circumstances that prevent them to attain their maximum educational potential.

In this paper we study, using a dynamic general equilibrium-overlapping generations model with a heterogeneous population by education, lifetime labor income, and life expectancy, the redistributive properties of two pension reforms that aim at minimizing the regressivity of the pension system induced by the ex-ante difference in longevity by SES. The two pension proposals are those suggested by Ayuso et al. (2017) and Sánchez-Romero and Prskawetz (2020). The pension proposal of Ayuso et al. (2017) (herein ABH) recommends adjusting the pension replacement rate of each retiree according to the difference between the remaining years-lived of the population subgroup of the retiree and that of the average retiree. With this proposal, it is expected that all retirees will earn at the age of retirement the same present value of benefits relative to the contributions paid. The proposal of Sánchez-Romero and Prskawetz (2020)

[4] Our model set up relies on studies that link differences in longevity between educational groups to education specific individual behavior (Preston and Elo, 1995; Doblhammer et al., 2005; Shkolnikov et al., 2006; Manchester and Topoleski, 2008; Klotz, 2010; Luy et al., 2011; Olshansky et al., 2012).

(herein SRP) suggests finding the level of progressivity in the replacement rate such that the pension program is ex-ante neither regressive nor progressive for any population subgroup.

We apply our model to Austria and study the impact of the afore mentioned pension proposals (ABH and SRP) on the Austrian pension system. Austria's pension system is an interesting case, because similar to many other non-progressive pension systems it has neither implemented any policy that corrects for the increasing life expectancy nor the diverging life expectancy by SES. However, like many other pension systems, to guarantee its long-run sustainability proposals are indispensable and should also consider the diverging trends of life expectancy across different subgroups of the population.

To study the redistributive properties of the ABH and SRP proposals, we calculate the internal rate of return (IRR) for population subgroups that differ by their educational attainment and pension points (which is a good proxy for lifetime labor income). Previous empirical studies analyzing the progressivity of pension systems using the IRR are Aaron (1977), Hurd and Shoven (1985), Duggan et al. (1993), Gustman and Steinmeier (2001) and Liebman (2002) in the US, and Schröder (2012) and Haan et al. (2020) in Germany, among others. We obtain the following results. First, under the current Austrian pension system we obtain that agents with high SES receive a higher IRR than those with low SES. The difference in IRR for all SES groups will decline from the 1960 birth cohort to the 2020 birth cohort. The fall in the IRR is explained by the fact that the future social contribution rate will increase faster than the future gains in life expectancy. Second, the decline in the IRR across cohorts is more pronounced for the highly educated workers than for the low-educated workers. However, highly educated workers will continue receiving an IRR that doubles that of low-educated workers. Consequently, third, we find that the Austrian pension system is ex-ante regressive due to the life expectancy gradient, which is a common characteristic of all non-progressive pension systems. Fourth, after implementing the ABH and SRP proposals, we find that both proposals (SRP and ABH) reduce the inequality in the IRR across agents with different educational attainment and pension points compared to the status quo, albeit the inequality reduction is stronger in the SRP proposal than in the ABH proposal. Under the SRP proposal agents with low educational attainment and with pension points in the lowest tercile get the highest increase in IRR, bringing them closer to the average IRR. However, the SRP proposal has a drawback since it also provides the highest IRR to those agents who belong to the lowest pension points tercile and are highly educated. This is because the SRP proposal compensates not only for differences in the life expectancy but also for differences in pension points, while the ABH only compensates for differences in life expectancy.

The paper is structured as follows. Section 2 is devoted to explain why it is necessary that pension proposals account for differences in life expectancy in the replacement rate when there exists a mortality gradient. In Section 3 we present the model setup. In Section 4 we discuss the parametrization of the model and the calibration strategy using the Bayesian melding method. In Section 5 we introduce the two pension proposals and present the results of the IRR for different population subgroups that differ by their educational attainment and their lifetime labor income. Section 6 concludes. We provide a detailed derivation of the economic model in the Appendix.

## 2. Intracohort redistribution of a pension system

To analyze the redistributive properties of a pension program the most frequently used measure is the internal rate of return (IRR). The IRR is the return that equalizes the present discounted value, survival weighted, of the contributions and taxes paid and benefits received for a cohort

$$\sum_{\text{age}=0}^{\text{max age}} \text{Survival}_{\text{age}} \frac{\text{benefits}_{\text{age}} - (\text{contributions+taxes})_{\text{age}}}{(1+\text{IRR})^{\text{age}}} = 0. \tag{1}$$

Eq (1) implies that the IRR increases the higher is the survival, the benefits received, and the duration of retirement, while the IRR declines the higher are the contributions and taxes paid and the longer is the duration paying contributions and taxes.[5] The IRR is preferable to the social security wealth (SSW) for analyzing redistribution because high income earners pay more contributions than low income earners and therefore their social security wealth is ex-ante by default higher. In contrast, the IRR, as opposed to the social security wealth, is a measure that is not affected by the scale of contributions paid and hence it is not affected by the labor income level.

The IRR received from the pension system by an individual who is planning to retire depends on two main components: the pension replacement rate ($\varphi$), that transforms the contributions paid to pension benefits, and the expected remaining years of life (LE). For convenience, let us

[5] The SSW is the present value of future benefits that an individual will receive less the present value of future contributions and taxes that one has to pay. The SSW can be calculated for each age and will start to be negative during working life and be highest around retirement age.

rewrite the IRR as $\text{IRR}(\varphi, \text{LE})$. Applying the total derivate to the IRR we have

$$d\text{IRR} = \underbrace{\frac{\partial \text{IRR}}{\partial \varphi}}_{+} d\varphi + \underbrace{\frac{\partial \text{IRR}}{\partial \text{LE}}}_{+} d\text{LE}, \tag{2}$$

where the positive signs show that an increase either in the replacement rate or in the expected remaining years of life has a positive effect on the IRR. Indeed, since in non-progressive pension systems $d\varphi = 0$, Eq. (2) shows that the IRR is higher for those individuals with higher expected remaining years of life, ceteris paribus all contributions and taxes paid.

To avoid that the pension system redistributes within each cohort from short-lived individuals to long-lived individuals, the pension system should give the same IRR to every individual regardless of their life expectancy. Consequently, we have from (2) that if the goal is to provide the same IRR across all groups (i.e. $d\text{IRR} = 0$) regardless their expected remaining years of life, the replacement rate should satisfy

$$d\varphi = -\left( \frac{\partial \text{IRR}}{\partial \text{LE}} \bigg/ \frac{\partial \text{IRR}}{\partial \varphi} \right) d\text{LE} < 0. \tag{3}$$

Therefore, in order that a pension system gives the same IRR to all individuals, Eq. (3) suggests that individuals with a higher expected remaining years of life should have a lower replacement rate level.

A large body of research has recently focused on studying the heterogeneity in life expectancy (or the expected remaining years of life) by SES and its impact on the pension system (for a survey, see Lee and Sanchez-Romero, 2019). To estimate the life expectancy by SES, education ($e$) and lifetime income (LI) have frequently been used as a measure of SES. Both measures have been found to account equally well for differences in life expectancy and that neither one fully captures all the covariation of life expectancy by SES (Bosworth et al., 2016). Thus, for convenience, let us assume that the expected remaining years of life is simultaneously a function of LI and $e$; i.e. $\text{LE}(\text{LI}, e)$. Totally differentiating LE gives

$$d\text{LE} = \frac{\partial \text{LE}}{\partial \text{LI}} d\text{LI} + \frac{\partial \text{LE}}{\partial e} de. \tag{4}$$

Then, substituting (4) into (3), and dividing by $d\text{LI}$ shows how the pension replacement rate should be adjusted to changes in lifetime income

$$\frac{d\varphi}{d\text{LI}} = -\left( \frac{\partial \text{IRR}}{\partial \text{LE}} \bigg/ \frac{\partial \text{IRR}}{\partial \varphi} \right) \frac{\partial \text{LE}}{\partial \text{LI}} < 0. \tag{5}$$

Notice that we have removed the term $\frac{\partial \text{LE}}{\partial e} \frac{de}{d\text{LI}}$ because education occurs before the lifetime labor income is accumulated $de/d\text{LI} = 0$. Thus, whenever there exists a positive relationship between the expected remaining years of life and lifetime income, Eq. (5) suggests that the replacement rate should decline proportionally to the increase in lifetime income.

Alternatively, the pension replacement rate can be adjusted according to the educational attainment. Substituting the total derivative of LE in (3), and dividing by $de$ gives that the pension replacement rate should be adjusted to different educational groups

$$\frac{d\varphi}{de} = -\left(\frac{\partial \text{IRR}}{\partial \text{LE}} \Big/ \frac{\partial \text{IRR}}{\partial \varphi}\right)\left(\frac{\partial \text{LE}}{\partial \text{LI}}\frac{d\text{LI}}{de} + \frac{\partial \text{LE}}{\partial e}\right) < 0. \tag{6}$$

Notice in (6) that the adjustment of the replacement rate through the educational attainment turns out to be more complex compared to an adjustment through the lifetime labor income in (5), since education has a direct and indirect impact, through LI, on the expected remaining years of life. For this reason, in this article we use the lifetime income as the SES measure to adjust the replacement rate. This same approach has been taken in many other articles (e.g., NASEM, 2015; Bosworth et al., 2016; Waldron, 2007).

As it should be expected, the two proposals analyzed (ABH and SRP) in section 5 agree on the necessity of adjusting the replacement rate by differences in LE by SES, but they differ on the degree of proportionality that is represented by the term in parenthesis in Eq. (5). Thus, we will analyze the proposals of ABH and SRP, which assume two different degrees of proportionality. In particular, the proposal of ABH implicitly assumes that the impact of the remaining years of life on the IRR is equal to the impact of the replacement rate on the IRR. Therefore, an increase of 1 percent in the remaining years of life should be compensated with a reduction of 1 percent in the pension replacement rate. Instead, SRP argues that the impact of an increase in the remaining years of life on the IRR is higher than the impact of an increase in the replacement rate on the IRR. This is because higher expected remaining years of life affect not only the retirement period, but may also influence the labor supply and the educational attainment. For this reason, in the following section we build a dynamic general equilibrium-overlapping generations model with heterogeneous agents and allow them to optimally choose their educational attainment and labor supply. Readers who are mainly interested in the policy section can go to Section 5.

## 3. The model

This section provides a description of our discrete time dynamic general equilibrium-overlapping generations model with heterogeneous households. Our model is populated by $\mathcal{Z} = 500$ generations, or birth cohorts. Each birth cohort is comprised of $\mathcal{N}$ heterogeneous representative agents.

The solution of the household problem and the equilibrium path can be found in appendices S1 and S2.

### *3.1. Households*

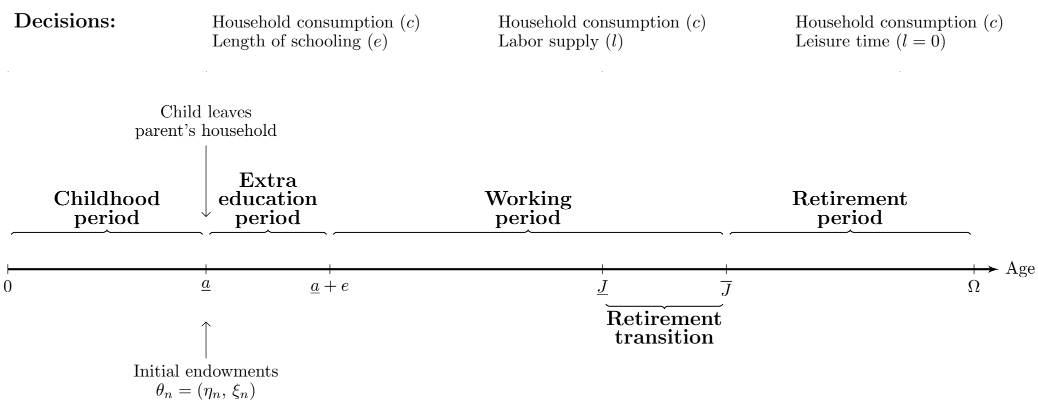



Figure 1: Agents' timeline

Households are comprised of an adult agent and dependent children. Agents give birth each year to a fraction of children according to age-specific fertility rates. This allows to incorporate changes over time in the family structure. Let the household size in equivalent adult consumers units be denoted by $H$. Agents enter the model at age 0, face mortality risk, and may live up to a maximum of $\Omega = 100$ years. See agents' timeline in Figure 1. Agents are raised by their parents from birth until the age of finishing primary schooling, denoted by $\underline{a}$. After age $\underline{a}$ agents leave their parents' home, settle their own household, and are randomly endowed with a set of initial characteristics. We denote the set of initial characteristics of an agent of type $n \in \{1, \ldots, \mathcal{N}\}$ by $\theta_n$. The set of characteristics is comprised of a level of effort of schooling $\eta_n$ and an innate learning ability $\xi_n$. Thus, the set of initial endowments of an agent of type $n$, or permanent unobserved heterogeneity, is $\theta_n = (\eta_n, \xi_n) \in \Theta_n$, where $\Theta_n$ is the set of all possible endowments for an agent of type $n$. After receiving the set of initial endowments, agents decide on the additional length of schooling (e). Agents can choose their educational attainment $e$ across three education levels $\mathbf{E}$ ={primary, secondary, college}. Each education level is associated with a different mortality and fertility pattern.[6] Moreover, agents will choose over their lifecycle

[6]For consistency reasons we assume that all education groups have the same net reproduction rate within

the total consumption of the household and the number of hours worked. Since each agent type represents a group of individuals who have similar initial characteristics (i.e. endowments), from age $\underline{J}$ (i.e, minimum retirement age) until age $\overline{J}$ (i.e., maximum retirement age) a fraction of agents of type $n$ will retire. After age $\overline{J}$ all agents are assumed to be retired and only devote time to leisure.

For notational simplicity, we present in this section all control and state variables with the age subscript $a$, the educational attainment subscript $e$, and suppress the birth cohort subscript $z \in \mathcal{Z}$.

*Preferences*

Agents have preferences over household consumption ($c_a$), years of schooling ($e$), and hours worked ($l_a$). Preferences are assumed to be separable and logarithmic in consumption. The period utility function of an agent with education $e$ at age $a$ is given by

$$U(c_{a,e}, l_{a,e}) = v_C(c_{a,e}) - v_E(a,e) - v_L(l_{a,e}) + v_J(\mathrm{LE}_{a,e})\alpha_J(l_{a,e}). \tag{7}$$

Eq. (7) implies that the utility increases through household consumption ($v_C(c_{a,e}) = H_{a,e} \log(c_{a,e}/H_{a,e})$), where $H_{a,e}$ is the household size measured in equivalent adult consumption units, and through the leisure time during retirement ($v_J(\mathrm{LE}_{a,e}) = v_0 \left(\mathrm{LE}_{a,e}\right)^{v_1}$ with $v_0 > 0, v_1 < 0$). The utility from leisure is increasing with respect to the inverse of the remaining life expectancy, $\mathrm{LE}_{a,e}$, as the amount of retirement time is squeezed by delaying retirement (Sánchez-Romero et al., 2020). $\alpha_J(l_{a,e})$ is the fraction of people retired, which is inversely related to the labor supply, i.e $\alpha'_J(l_{a,e}) < 0$. Utility decreases because of the cost of attending schooling and hours worked. Agents incur a cost $v_E(a,e) = \eta_n \mathbf{1}_{\{a<\underline{a}+e\}}$ by attending schooling (Oreopoulos, 2007; Restuccia and Vandenbroucke, 2013; Le Garrec, 2015; Sánchez-Romero et al., 2016; Sánchez-Romero and Prskawetz, 2020), where $\eta_n > 0$ is the marginal cost of each additional year of schooling and $\mathbf{1}_{\{a<\underline{a}+e\}}$ is an indicator function that takes the value of one if $a < \underline{a} + e$ and zero otherwise. We consider $\eta_n$ as a proxy for the socioeconomic background. Thus, higher (resp. lower) values $\eta_n$ are associated with a lower (resp. higher) socioeconomic background. We assume a standard isoelastic disutility from working $v_L(l_{a,e})$, where the marginal disutility from working $l_{a,e}$ hours is $v'_L(l_{a,e}) = \alpha_L (l_{a,e})^{1/\sigma_L}$, with $\alpha_L, \sigma_L > 0$.

each cohort. As a consequence, agents with lower life expectancy have higher fertility rates.

*Human capital*

We denote the stock of human capital of an agent of age $a$ with $e$ years of education by $h_{a,e}$. All agent types are assumed to start at age $\underline{a}$ with the same initial stock of human capital, $h_{\underline{a},e}$, but different learning ability $\xi_n$. We assume individuals can increase their human capital by attending schooling. The accumulation of human capital is described by the following Ben-Porath function

$$h_{a+1,e} = \begin{cases} h_{a,e} + \xi_n (h_{a,e})^{\gamma_h} & \underline{a} \leq a < \underline{a} + e, \\ h_{a,e} & a \geq \underline{a} + e, \end{cases} \tag{8}$$

where the number of years of education, $e$, is a discrete choice variable. Specifically, agents choose whether to stay with compulsory education ($e = 0$), complete high school ($e = 4$), or complete college ($e = 8$).

*Budget constraint*

We assume the existence of a perfect annuity market in which agents can purchase life-insured loans, when they are in debt, and annuities in case of having positive financial wealth. Let us denote the conditional probability of surviving from age $a$ to age $a + 1$ as $\pi_{a,e}$ and the financial wealth at age $a$ as $k_{a,e}$. There are three sources of income after survival: the interests gained from the initial financial wealth annuitized $(R_{a,e} - 1)k_{a,e}$, the labor income earned net of contributions and taxes $(1-\tau^l)(1-\tau^s)y_{a,e}$, and the pension benefits (net of taxes) $(1-\tau^l)b_{a,e}\alpha_J(l_{a,e})$. The term $\alpha_J(l_{a,e})$ represents the fraction of agents with similar endowments that are already retired. We assume $\alpha_J(l)$ is inversely related to the labor supply. The income is used for consuming market goods $(1 + \tau^c)c_{a,e}$ and savings $k_{a+1,e} - k_{a,e}$. We assume agents start with zero financial wealth $k_{\underline{a},e} = 0$. The budget constraint at age $a$ of an agent with $e$ additional years of education is

$$k_{a+1,e} - k_{a,e} + (1 + \tau^c)c_{a,e} = (R_{a,e} - 1)k_{a,e} + (1 - \tau^l)[(1 - \tau^s)y_{a,e} + b_{a,e}\alpha_J(l_{a,e})], \tag{9}$$

where $R_{a,e} = (1 + r(1 - \tau^k))/\pi_{a,e}$ is the capitalization factor of the annuity, $\{\tau^k, \tau^c, \tau^l, \tau^s\}$ is the set of tax rates on capital, consumption, and labor income and the social contribution rate, respectively. Labor income $y_{a,e}$ is given by the product of the wage rate $w_{a,e}$ and the labor supply $l_{a,e}$, which is normalized between zero and one. The wage rate $w_{a,e}$ consists of three components: (1) the effective wage rate $w$, (2) the efficiency of an individual with $a - \underline{a} - e$ years of experience after $e$ years of schooling, and (3) the human capital stock $h_{a,e}$; i.e. $w_{a,e} = w\epsilon_a(e)h_{a,e}$.

*Pensions*

The model replicates the Austrian social security pension system. The Austrian pension system is an unfunded and defined benefit system. The general pension formula of the Austrian pension system follows the rule that after 45 years of contribution, retiring at age 65, workers will receive 80 percent of their average lifetime income (Knell et al., 2006; Sánchez-Romero et al., 2013). Nonetheless, to account for all past pension proposals and those cases outside the general pension rule, we model the pension system following the general framework of Sánchez-Romero et al. (2020). The dynamics of the pension points is given by:

$$\mathrm{pp}_{a+1,e} = \left[\alpha_J(l_{a,e}) + (1-\alpha_J(l_{a,e}))\mathcal{R}_a\right]\mathrm{pp}_{a,e} + \phi^p \mathrm{PBI}(y_{a,e}), \tag{10}$$

where $\mathcal{R}_a = (1+i_a)/\bar{\pi}_a$ is the capitalization factor of the pension system, which depends on a capitalization index ($i_a$) that is set by the social security system and on the average conditional survival probability of the cohort ($\bar{\pi}_a$). Note that the average conditional survival probability of the cohort, $\bar{\pi}_a$, does not necessarily coincide with the conditional survival probability of the individual, $\pi_{a,e}$. $\phi^p$ is the conversion factor of wage income to pension points and $\mathrm{PBI}(y_{a,e})$ is the pension base increment.[7] Pension points are capitalized until all individuals of the cohort retire or $\alpha_J(l) = 1$. Since the pension benefit is calculated based on the $n$ best years or the so-called pensionable income years, we create an ordered vector $\mathbf{p}_a$ for each agent comprised of the $n$ best earnings years until age $a$; i.e. $\mathbf{p}_a = \{(p_1, p_2, \ldots, p_n) \in \mathbb{R}^n_+ : p_1 > p_2 > \ldots > p_n\}$. Thus, the pension base increment is calculated as the difference between the current earnings and the lowest earnings stored in $\mathbf{p}_a$; i.e. $\mathrm{PBI}(y_{a,e}) = \max(y_{a,e} - p_n, 0)$.

Agents can retire after the minimum retirement age $\underline{J}$ and no later than a maximum retirement age $\overline{J}$. We denote the normal retirement age by $J^N$. The amount of pension benefits claimed depends on three components: (i) a pension replacement rate $\boldsymbol{\varphi}(\mathrm{pp})$, which might depend on the pension points accumulated (notice that this variable has a direct relationship with lifetime labor income), (ii) a replacement rate adjustment factor $\lambda_a$, which is a function of the average years contributed and the average retirement age, and (iii) the pension points accumulated $\mathrm{pp}_{a,e}$. When pension benefits fall below a minimum threshold $b^{\min}$, there is a supplementary benefit until the minimum pension benefit is reached. On top of these three components, we consider the possibility that the government introduces in the future a sustainability factor $\rho_a \leq 1$, which will reduce pension benefits when total pension claims exceed total

[7] We detail the historical changes of the main parametric components of the Austrian pension system in the supplementary material.

contributions. The average pension benefits received at age $a$ by an individual with education $e$ is given by

$$b_{a,e} = \max\left\{\lambda_a \cdot \boldsymbol{\varphi}(\mathrm{pp}_{a,e}) \cdot \mathrm{pp}_{a,e}, b^{\min}\right\} \cdot \rho_a. \tag{11}$$

Through $\lambda_a$ pension benefits are reduced (resp. increased) when (i) individuals retire before (resp. after) $J^N$ and when (ii) individuals do not reach the minimum of years of contribution. The pension replacement rate is assumed to have a fixed average replacement rate $\varphi$, which will be adjusted through $\mathcal{I}(\mathrm{pp}_{a,e})$ according to the difference in life expectancy by number of pension points

$$\boldsymbol{\varphi}(\mathrm{pp}_{a,e}) = \varphi\left(1 - \zeta \cdot \mathcal{I}(\mathrm{pp}_{a,e})\right), \tag{12}$$

where $\zeta \in [0,1]$ accounts for the phase-in/out period in which the adjustment factor is introduced ($\zeta = 0$ before the phase-in and $\zeta = 1$ after the phase-out). To account for the negative impact that the minimum pension benefit has on the labor supply, the minimum pension benefit is modeled assuming that individuals start with a minimum pension points; i.e. $b^{\min} = \boldsymbol{\varphi}(\mathrm{pp}^{\min}) \cdot \mathrm{pp}^{\min}$. See Section S5 in the supplementary material for more information on the evolution of the parametric components of the Austrian pension system.

*Recursive household problem*

Households choose the optimal consumption path ($c$), labor supply ($l$), and education ($e$) in two steps. First, agents determine the consumption path and hours of work conditional on a particular educational attainment $e \in \mathbf{E}$. Hence, given a set of endowments $\theta_n = (\eta_n, \xi_n) \in \Theta_n$, an educational level $e \in \mathbf{E}$, and the set of state variables $\mathbf{x}_{a,e} = \{k_{a,e}, \mathrm{pp}_{a,e}, h_{a,e}\}$, an agent chooses consumption ($c$) and labor ($l$) that maximizes from $a = \Omega$ to $a = \underline{a}$ the following Bellman equation:

$$V(\mathbf{x}_{a,e};\theta_n) = \max_{c_{a,e},\, l_{a,e}} \left\{U(c_{a,e}, l_{a,e}) + \beta\pi_{a+1,e}V(\mathbf{x}_{a+1,e};\theta_n)\right\} \tag{13}$$

subject to eqs. (7)-(12) and the boundary conditions $k_{\underline{a},e} = 0$, $h_{\underline{a},e} = h_{\underline{a}}$. See the derivation of the optimality conditions in Section S1 in the Appendix.

Second, given the optimal paths of consumption, labor supply, and the vector of state variables $\mathbf{x}^*_{\underline{a},e}(\theta_n)$ for each educational attainment $e \in \mathbf{E}$, the agent chooses the optimal level of education according to

$$\mathbf{e}(\theta_n) = \arg\max_{e \in \mathbf{E}} V(\mathbf{x}^*_{\underline{a},e}(\theta_n);\theta_n). \tag{14}$$

Notice that given the stream of prices and demographic information each representative agent is uniquely characterized according to her initial endowments. Therefore, we denote from now on the optimal policy function of variable '$X$' of a representative agent born in year $z$ at age $a$ and with initial endowments $\theta_n$ as $\mathbf{X}_{z,a}(\theta_n)$.

### *3.2. Production*

We assume one representative firm that produces a final good by combining capital ($K$) and effective labor ($L$). Final goods can either be saved or consumed. The production function, that exhibits constant returns to scale, takes the following form

$$Y_t = (K_t)^{\alpha_Y} (A_t L_t)^{1-\alpha_Y}, \tag{15}$$

where $Y_t$ is output, $\alpha_Y$ is the capital share, and $A_t$ is labor-augmenting technology, whose law of motion is $A_{t+1} = (1 + g_t^A)A_t$ and $g_t^A$ is the productivity growth rate. Aggregate capital stock evolves according to the law of motion $K_{t+1} = K_t(1 - \delta_K) + I_t$, where $\delta_K$ is the depreciation rate of capital and $I_t$ is aggregate gross investment.

We assume our representative firm maximizes the net cash flow by renting capital and hiring labor from households in competitive markets at the rates $r_t$ and $w_t$, respectively. Capital and labor inputs are chosen by firms according to the first-order conditions:

$$r_t + \delta_K = \alpha_Y (Y_t/K_t), \tag{16}$$

$$w_t = (1 - \alpha_Y)(Y_t/L_t). \tag{17}$$

### *3.3. Government*

The government provides public goods and services, denoted by $G_t$, and transfers all retirement pension benefits claimed, which are denoted by $\mathcal{S}_t$. The total amount of pension benefits claimed is

$$\mathcal{S}_t = \sum\nolimits_{a=0}^{\Omega} N_{t,a} \left[ \sum_{n=1}^{\mathcal{N}} \int_{\Theta_n} \mathbf{b}_{t-a,a}(\theta_n) \, \mathrm{d}P_{t-a}(\theta_n) \right]. \tag{18}$$

$N_{t,a}$ is the population size of age $a$ in year $t$, $\mathcal{N}$ is the number of heterogeneous agents, and $P_{t-a}(\theta_n)$ is the probability of having the initial endowments $\theta_n \in \Theta_n$. For simplicity, we assume the government does not hold debt. Following the Austrian pension system, social security contributions finance 70 percent of all retirement benefits claimed. Thus,

$$0.70\mathcal{S}_t = \tau_t^s w_t L_t, \tag{19}$$

where $\tau_t^s$ is the social security contribution rate. To finance $G_t$ and the remaining 30 percent of $\mathcal{S}_t$, the government levies taxes on labor income ($\tau^l$), on capital income ($\tau^k$), and on consumption ($\tau^c$). The budget of the government in period $t$ is

$$G_t + 0.30\mathcal{S}_t = \tau_t^l(w_t L_t + 0.30\mathcal{S}_t) + \tau_t^k r_t K_t + \tau_t^c C_t, \tag{20}$$

where $C_t$ is the total final goods consumed. Notice that the total tax base of labor income has to be augmented by the fraction of total pension benefits that are not financed by social contributions. We have not included in the model the progressivity of the Austrian tax system, since many pension systems are not financed through the general budget. Consequently, for the case of Austria our results will overestimate the differences in the IRR across agents with different SES. Albeit this effect is expected to be small given that the total labor income tax only represents 16.5 percent of the total revenues of the pension system. In contrast, our results will underestimate the inequality in the IRR for the same mortality gradient across agents with different SES, in all other countries in which the pension system is not financed through contributions.

## 4. Parametrization

The basic purpose of this section is to replicate the inflows and outflows of the Austrian pension system so as to correctly calculate the IRR of our heterogeneous agents. To do so we fit the model to historical economic and demographic data of Austria for the period 1890–2010. Before introducing the calibration of our model in section 4.5 we briefly summarize the reconstruction of education specific demographic parameters, the age profiles of labor income and taxes and the required social contributions to finance the pension system.

### *4.1. Demographics by education*

We extend the historically reconstructed population estimates for Austria implemented in the AGENTA project (www.agenta-project.eu) by introducing differential fertility and mortality by educational attainment. Demographic data before 2010 is taken from historical records (Rivic, 2019), while the demographic data after 2010 is based on Eurostat's projections. Based on existing literature (Lutz et al., 2007, 2014; Goujon et al., 2016), we assume a fixed difference in life expectancy at age 15. We consider agents can attain any of the following three educational groups: primary, secondary, and college. To be consistent with the ISCED classification and taking into account that agents start making decisions at age 14, we set the length of schooling,

$e \in \mathbf{E}$, at 0 years for primary education, at 4 years for secondary education, and at 8 years for college. Agents with primary or less education are assumed to have a life expectancy, $\mathrm{LE}_{15,0}$, five years lower than those with college, while agents with secondary education have a life expectancy, $\mathrm{LE}_{15,4}$, one year lower than those with college. The evolution across cohorts of the life expectancy (at birth) for the three educational groups is presented in Fig. 2, panel A. Moreover, we assume that the population of each educational group grows at the same rate. This assumption implies that in order to overcome the lower proportion of agents surviving through the reproductive ages, fertility is slightly higher for lower educated than for more educated agents. See the evolution across cohorts of the total fertility rate for the three educational groups in Fig. 2, panel B. As a result, agents with different educational attainment will also face a different household size consistent with their mortality and fertility profiles. The derivation of age-specific mortality rates and age-specific fertility rates is provided in section S4 in the supp. material.

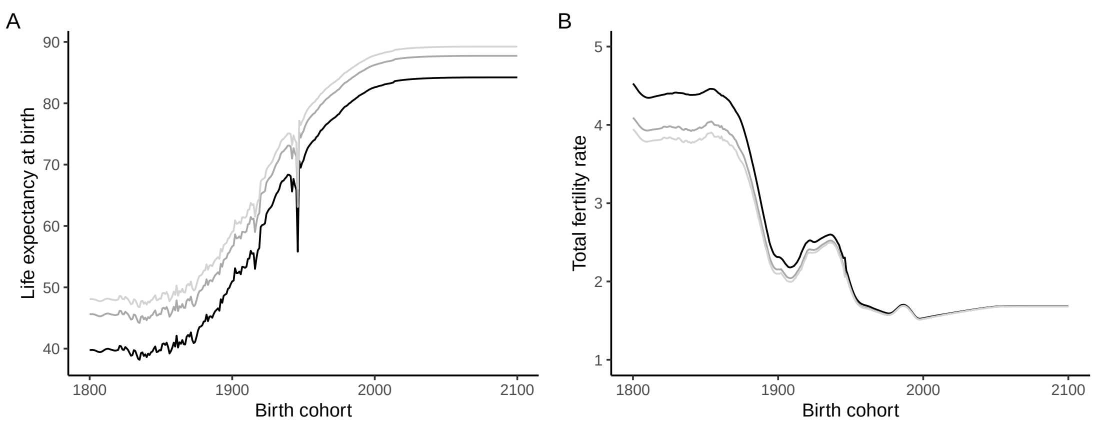



Figure 2: Estimated vital rates by educational attainment in Austria for birth cohorts born between 1800 and 2100: Primary or less (black), secondary (dark gray), and tertiary (light gray). *Source*: Own calculations. *Notes*: Panel A shows the life expectancy at birth by educational attainment. Panel B shows the total fertility rate (TFR) by educational attainment.

*4.2. Life cycle earnings by education*

The wage rate per hour worked of our agents depends on (i) the wage rate per efficient unit of labor, on (ii) the age-specific labor productivity $\epsilon_a(e)$, which is a function of the experience, and on (iii) the stock of human capital $h_{a,e}$. The age-specific productivity of an agent with $e$ additional years of education and $a - \underline{a} - e$ years of experience is assumed to follow a standard

Mincerian equation $\log \epsilon_a(e) = \beta_1(a - \underline{a} - e) + \beta_2(a - \underline{a} - e)^2$, where $\underline{a}$ is the age at finishing primary education and the parameters $(\beta_1, \beta_2)$ reflect the importance of experience on the wage rate, which are set to match EU-SILC 2011 data.

### *4.3. Private sector*

Our choices for capital share and depreciation of capital are $\alpha_Y = 0.375$ and $\delta_K = 0.05$, respectively. The values of these two parameters imply an interest rate of 3.3% for an average capital-to-output ratio of 4.5. We assume no productivity growth before year 1800. From 1800 to 2070 the exogenous productivity growth rate is taken from two main sources. For the period 1890–2018 we take Austrian historical productivity estimates from Bergeaud et al. (2016). For the period 2018–2070 we rely on the productivity assumptions from the European Commission (2018). After year 2070 we take the last productivity growth rate assumed by the European Commission (2018) and assume that it stays constant until the end of the simulation period. For the intermediate period 1800–1890 we linearly extrapolate the productivity growth rate. See the productivity growth rate in panel B, Figure 3.

### *4.4. Public sector*

To account for the differential impact of capital taxes, labor income taxes, and consumption taxes on the age profiles of labor income and pension benefits, and given that taxes also finance thirty percent of the total public pensions claimed (see Eq. 20), we collected historical information from Statistisches Handbuch Österreichs (1966, 1991) on the public consumption spending from 1913 to 2018. Before 1913 and after 2018 we assume that public consumption represents 8 percent and 20 percent of the total output, respectively, which coincides with the first and the last public consumption to output ratio from the time series taken from Statistisches Handbuch Österreichs (1966, 1991). See the ratio of public consumption to output in panel A Figure 3.

Based on National Accounts data from Statistics Austria for the period 1995–2018 we consider that labor income taxes finance 55 percent, consumption taxes finance 35 percent, and capital income taxes finances the remaining 10 percent of the total budget. The implementation of the evolution of all the historical parametric components of the Austrian pension system is taken from the General Law on Social Security (ASVG) and the General Pensions Act (APG).[8] We detail the values of the parametric components in Section S5 in the supplementary material.

[8] All the historical proposals can be found in the historic law database `www.sozdok.at`.

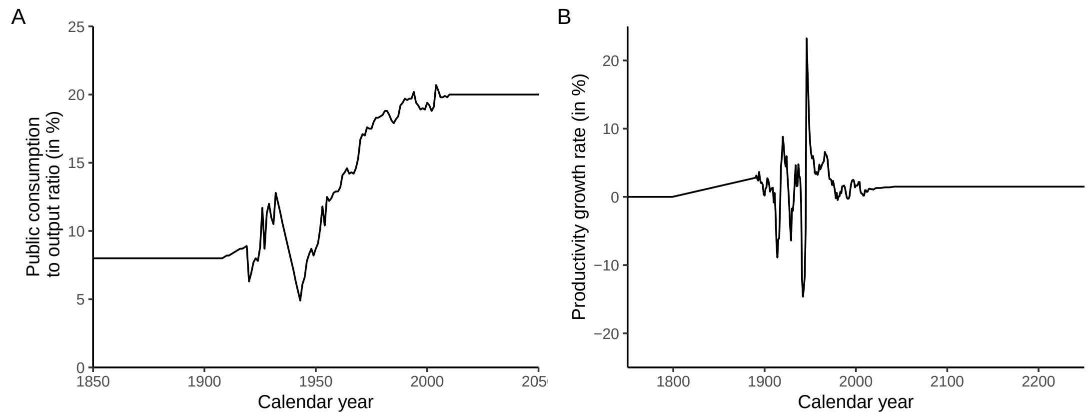



Figure 3: Public consumption to output ratio (**A**) and exogenous productivity growth rate (**B**). *Source*: Data on public consumption to output ratio comes from Statistisches Handbuch Österreichs (1966, 1991) and the National Accounts from Statistics Austria. The exogenous productivity growth rate is taken from Bergeaud et al. (2016) and European Commission (2018).

Under the current law of the Austrian pension system, we estimate that pension spending will represent more than 20% of the total output by year 2100, which is 5 percent higher than the current pension spending. The social contribution rate, measured over the total cost of a worker, is expected to reach 25 percent by year 2100, as compared to 19.1% in 2010.[9] To reduce the expected increasing cost of the pension system due to population aging and the longer life expectancy of retirees, we introduce a pension sustainability factor, which guarantees a maximum social security contribution rate, denoted by $\overline{\tau_t^s}$, of 22 percent. When the maximum social security contribution rate is reached, the government will adjust downwards the pension replacement rate by reducing the pension sustainability factor, denoted by $\rho_t$, until the system is balanced

$$\begin{cases} \rho_t = 1 \text{ and } 0.70\mathcal{S}_t = \tau_t^s w_t L_t & \text{if } \tau_t^s < \overline{\tau_t^s}, \\ \rho_t < 1 \text{ and } 0.70\mathcal{S}_t = \overline{\tau_t^s} w_t L_t & \text{if } \tau_t^s \geq \overline{\tau_t^s}. \end{cases} \tag{21}$$

This policy will transform the DB system to a DC system once that the maximum social security rate is reached (see Sánchez-Romero and Prskawetz, 2019a). In addition to the sustainability factor, in this paper, we analyze two pension proposals that aim at correcting the regressivity

[9] The social contribution rate is calculated as the total pension spending financed through contributions divided by the total wage bill of the economy.

of the pension system when there is an ex-ante difference in life expectancy by socio-economic status. For the sake of comparability, we assume that the two pension proposals are introduced by cohort and have a similar phase-in/out period of 20 years

$$\zeta_z = \begin{cases} 0 & \text{for } z \leq 1960, \\ \frac{z-1960}{20} & \text{for } 1960 < z \leq 1980, \\ 1 & \text{for } z > 1980, \end{cases} \tag{22}$$

where $z$ denotes the birth cohort. Eq. (22) implies that the cohort that is currently retiring ($z = 1960$) is the last cohort without any correction ($\zeta_z = 0$) and that this policy is fully implemented ($\zeta_z = 1$) for all cohorts born after year 1980. From the 1960 birth cohort to the 1980 birth cohort, the proposal is gradually introduced, increasing the importance of the replacement rate adjustment factor by 5%(=1/20) per year. The minimum pension points $\text{pp}^{\min}$ are set to match the minimum pension benefits $b^{\min}$, which in Austria is close to 1/3 of the average income ($\bar{y}$) at the age of retirement.[10]

*4.5. Calibration*

We follow a two-stage process to replicate the evolution of the Austrian economy. In the first stage, we assign values using the literature on the parameters governing the human capital accumulation and preferences. In a second stage, we estimate using the Bayesian melding method the permanent unobserved heterogeneity and the number of heterogeneous representative agents that best fit the evolution of the educational attainment in Austria. See the Appendix, Section S3, for a detailed explanation of the Bayesian melding method.

We assume the same initial stock of human capital ($h_{\underline{a},e}$) for all agents and normalize it to one. The returns-to-education is set at $\gamma_h = 0.65$, similar to Cervellati and Sunde (2013). The parameters governing the behavior of agents are set to replicate specific features of the labor supply. Specifically, we assume an intertemporal elasticity of substitution (IES) of consumption ($\sigma_c$) of 1.0, which coincides with the upper range values for $\sigma_c$ suggested by Chetty (2006) and guarantees a steady-state equilibrium. We assume an IES of labor supply, $\sigma_l$, equal to 0.40. Notice from Eq. (7) that $\sigma_l$ coincides with the Frisch elasticity, which is between the lower bound of 0.1 and upper bound of 2.0 (Keane and Rogerson, 2012). The value of the weight of the disutility of labor ($\alpha_L = 866.28$) is chosen so as to obtain that prime aged agents work

[10]In year 2010, the minimum pension ("Ausgleichszulage") was 10 976 Euros, which was 35 percent of the average income of the age group 56-60 (31 673 Euros) in Austria, see §293 ASVG in year 2010.

33.0 percent of their available time in year 2010. This is equivalent to an average of 37 hours of work per week and year. The preferences for retirement $v_0$ and $v_1$ are set at 77.06 and -1.94, respectively, to guarantee an average retirement age between 57 and 58 for the cohort born in year 1950. The subjective discount factor $\beta$ is calibrated to have a (real) interest rate between 3 and 4 percent along the XXI century. This interest rate should be thought as the opportunity cost of contributing to the pension system.

Table 1: Model parameters

| Parameter | Symbol | Value | Parameter | Symbol | Value |
|---|---|---|---|---|---|
| **Preferences** | | | **Human capital** | | |
| Marginal schooling cost† | $\eta_n$ | [0,40] | Learning ability† | $\xi_n$ | [0.00,0.30] |
| Labor elasticity | $\sigma_L$ | 0.40 | Initial human capital | $h_{\underline{a}}$ | 1.00 |
| Labor weight | $\alpha_L$ | 866.28 | Returns to education | $\gamma_h$ | 0.65 |
| Max. labor supply before retirement | $\overline{L}$ | 0.4 | Experience | | |
| Leisure in retirement | $v_0$ | 77.0552 | Age | $\beta_1$ | 0.070 |
| | $v_1$ | -1.9425 | Age-squared | $\beta_2$ | 0.00092 |
| Subjective discount factor | $\beta$ | 1.02 | | | |
| | | | **Production** | | |
| | | | Capital depreciation rate | $\delta_K$ | 0.05 |
| | | | Capital share | $\alpha_Y$ | 0.375 |
| | | | Productivity growth rate | $g_t^A$ | see Fig. 3 |

† Parameter calibrated using the Bayesian melding method.

The last set of parameters corresponds to the initial endowments of our heterogeneous individuals $\theta \in \Theta$: the effort of attending schooling, $\eta_n$, and the innate learning ability, $\xi_n$. These two parameters are estimated using the Bayesian melding method (Poole and Raftery, 2000; Raftery and Bao, 2010), which provides an inferential framework that takes into account both model's inputs and outputs. In particular, we apply the Bayesian melding method to derive the distribution of the set of initial endowments that best replicate the educational distribution for Austrian cohorts born between 1890 and 1980. Thus, by following this strategy our model accounts for selectivity within educational groups, since agents with different initial endowments endogenously choose their educational attainment and, consequently, the characteristics (schooling effort and learning ability) of each educational group changes across cohorts.

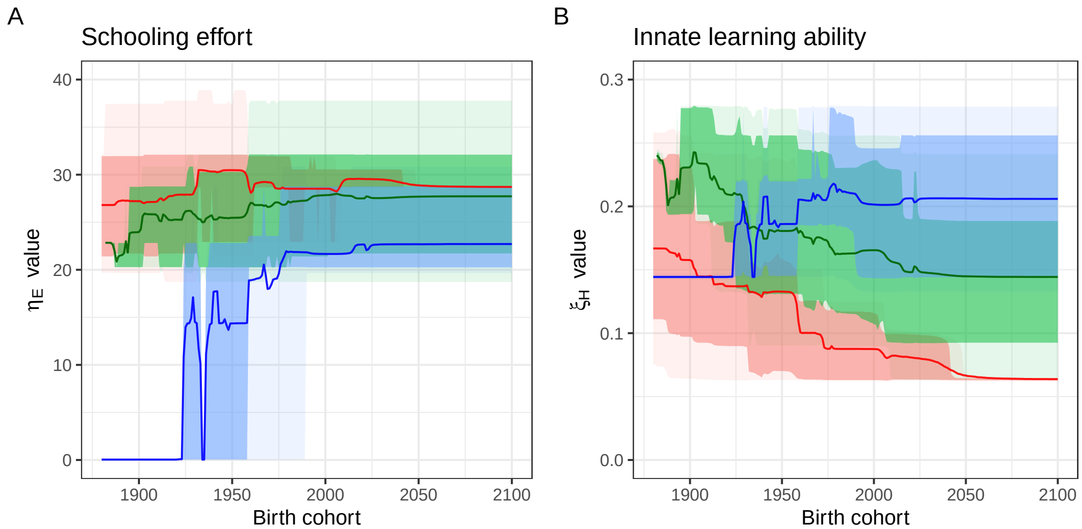



Figure 4: Evolution across cohorts of the initial endowments of each educational group, birth cohorts 1880–2100. *Notes*: Each panel shows for each education group (primary -**red**-, secondary -**green**-, and college -**blue**-) the evolution of the mean (solid line), the 50% confidence interval (darker area), and the 75% confidence interval (lighter area) of each initial endowment: schooling effort (**A**) and innate learning ability (**B**).

The best fit to the educational distribution is obtained with $\mathcal{N}$=25 heterogeneous agents.[11] Figure 4 shows the evolution across cohorts of the initial endowments —generated with the model— that characterize agents belonging to each educational group. Panel **A**, Fig. 4, shows the evolution of the schooling effort (we use this endowment as a proxy for the SES of parents), which is the most important endowment for choosing the educational attainment. Agents with high schooling effort are likely to stay with primary education, while agents with low schooling effort are more likely to attain college. The positive slope of the schooling effort across cohorts suggests that the relative importance of schooling effort for attaining higher education has decreased over time due to the rise in technological progress and in the length of life. As explained by Sánchez-Romero et al. (2016), our agents consider schooling effort as a fixed cost of education. Hence, as the lifetime income raises, the marginal cost of education decreases and it becomes more interesting for the agents to invest in education. Panel **B** shows that the innate learning ability is also an important endowment for choosing the educational attainment. Indeed, as the importance of schooling effort diminishes for younger generations, the innate learning ability

[11]We fitted the model to the educational distribution for $\mathcal{N}$ equal to 5, 15, 25, 35, 50, and 100.

becomes a key determinant of the educational attainment of agents. Consequently, agents with primary education are becoming more selected and have lower learning ability levels (see the red area). Instead, agents with college become more heterogeneous and have a higher innate learning ability level (see the blue area).

Figure 4 has shown that given the set of initial conditions for our heterogeneous agents our model can control for selectivity within educational groups, since our agents optimally choose their educational attainment and hence the characteristics of each educational group endogenously change in the model. Besides the choice of education, our representative agents also choose their consumption of final goods, their labor supply, their financial wealth, their social security wealth, their conditional survival probabilities, and the number of children. Moreover, Fig 5 shows that the model is capable of replicating well the educational distribution for cohorts born between 1890 and 1980 in Austria, which we have targeted with the Bayesian melding. The fit our the model (solid lines) to the educational distribution data (dots), taken from the Wittgenstein Centre Database (2015), is shown in Fig. 5, panel A. Panels from B to E, in Fig. 5, are used as an external validation of the calibration procedure. For instance, the model reproduces well the time series of pension spending to output ratio from 1950 to 2010 (panel **B**), the evolution of per capita income (panel **C**), and the average labor income profile in 2010 (panel **E**) taken from the AGENTA database. Besides, the model can replicate the number of pensioners for years 2011 to 2017 (panel **D**). Hence, the five panels in Fig 5 imply that our model is capable of matching key variables necessary to replicate well the cost of the evolution of the Austrian pension system. At the aggregate level, the evolution of the total pension spending and the evolution of per capita income. At the micro level, the profiles of labor income and pension benefits. Therefore, the model is validated to calculate the past and future cost and benefits of the pension system and its redistributive consequences on heterogeneous individuals by educational attainment.

## 5. Policy analysis

After solving the dynamic general equilibrium-overlapping generations model with the existing pension system (status quo), in this section we study the redistributive properties within each cohort that results from implementing either the ABH or the SRP proposal, which aim at reducing the regressivity of the pension system caused by the ex-ante difference in life expectancy. Consequently, we run two additional simulations. As explained in section 2, both proposals account for the relationship between the remaining years of life and lifetime labor in-

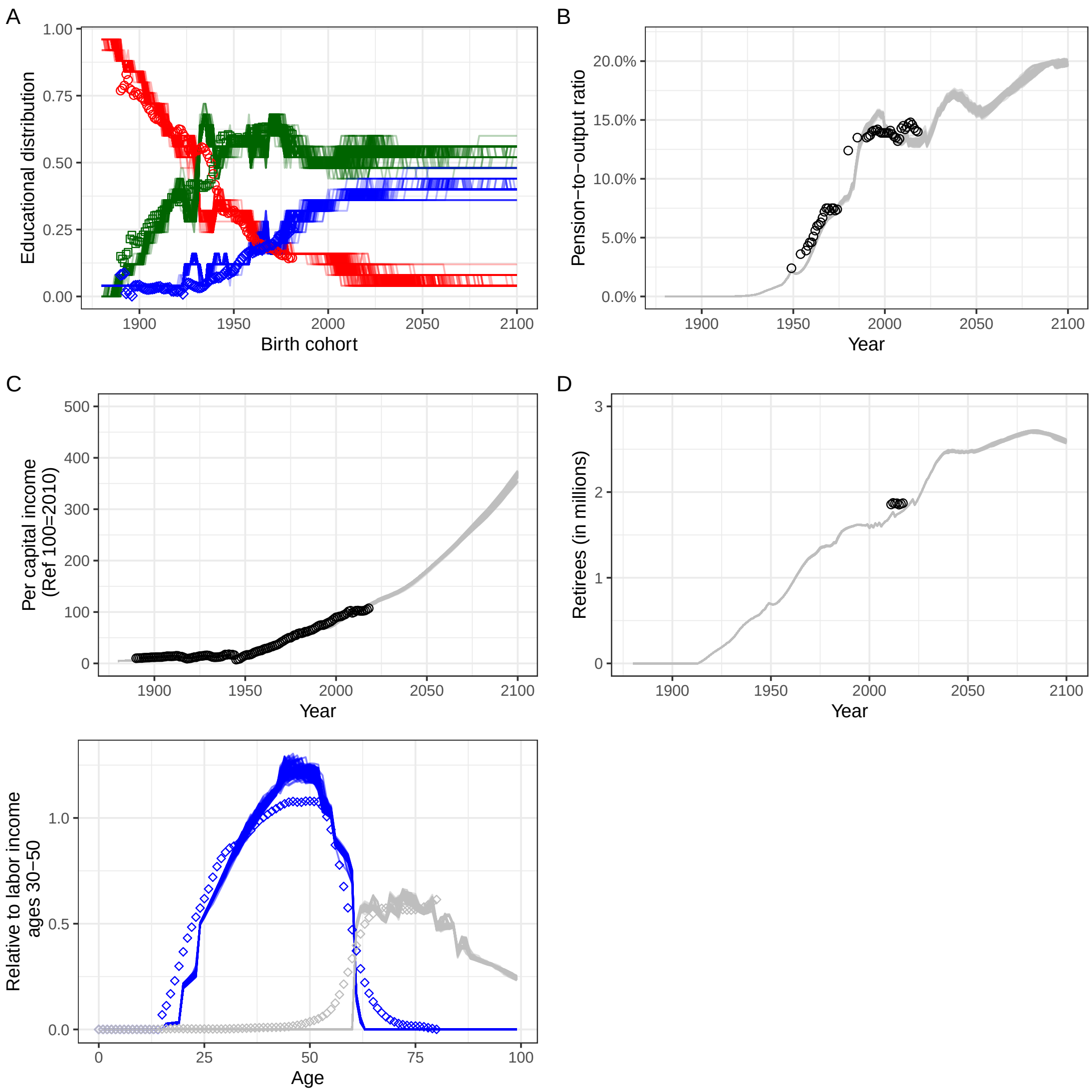



Figure 5: Model fit, 1880–2100. *Notes and sources*: Panel **A** shows the fit of the model (solid lines) to the educational distribution data (shaped lines) from Wittgenstein Centre Database (2015). Panel **B** shows the fit of the model (solid line) to the total pension spending to output ratio (dots). Panel **C** shows the model fit (solid line) to the per capita income (dots) taken from Bergeaud et al. (2016). Panel **D** shows the model fit (solid lines) to the number of pensioners (gray dots) from Statistics Austria. Panel **E** shows the model fit (solid lines) of labor income and pension benefits profiles in year 2010 to AGENTA data (dots) (see `http://dataexplorer.wittgensteincentre.org/nta/`).

come, through the pension replacement rate. Nevertheless, since the number of pension points is already a measure of lifetime labor income, we assume the social security system uses the number of pension points, $\mathbf{pp}$, as the SES measure. Moreover, we also assume, yet realistically, that the pension system does not have information about the exact life expectancy of each agent.

*5.1. Pension proposals: Adjustment factors of the pension replacement rates*

The ABH proposal suggests adjusting the replacement rate of an agent of age $a$ according to the difference between the average life expectancy of the cohort at age $a$ and the life expectancy of the agent at age $a$. Since the government does not have perfect information about the life expectancy of each agent, we consider that the pension system follows Holzmann et al. (2019) and regresses the remaining years of life on the log of the number of pension points (see section S6 in the supplementary material). The adjustment factor of the replacement rate in the ABH proposal is given by

$$\mathcal{I}_{\text{ABH}}(\mathbf{pp}_{z,a}(\theta_n)) = 1 - \frac{\widehat{\text{LE}}(\overline{\mathbf{pp}}_{z,a})}{\widehat{\text{LE}}(\mathbf{pp}_{z,a}(\theta_n))}, \tag{23}$$

where $\widehat{\text{LE}}(\mathbf{pp})$ denotes the estimated life expectancy as a function of the number of pension points $\mathbf{pp}$ and $\overline{\mathbf{pp}}_{z,a}$ is the average number of pension points for the cohort $z$ at age $a$. From Eq. (23) and assuming that the ABH proposal is fully implemented (i.e. $\zeta = 1$), the pension replacement rate becomes

$$\varphi_{\text{ABH}}(\mathbf{pp}_{z,a}(\theta_n)) = \varphi \frac{\widehat{\text{LE}}(\overline{\mathbf{pp}}_{z,a})}{\widehat{\text{LE}}(\mathbf{pp}_{z,a}(\theta_n))}. \tag{24}$$

Thus, the ABH pension replacement rate will be equal to $\varphi$ (i.e., the same as the status quo) for agents with pension points equal to the average pension points ($\mathbf{pp}_{z,a}(\theta_n) = \overline{\mathbf{pp}}_{z,a}$) and lower (resp. higher) than $\varphi$ for those agents with pension points higher (resp. lower) than the average pension points. Notice that since life expectancy is positively related to the number of pension points, agents with higher (resp. lower) life expectancy would receive a lower (resp. higher) replacement rate than agents with lower (resp. higher) life expectancy, ceteris paribus the retirement age.

The second pension proposal by SRP suggests that the short-lived and poor workers are compensated by finding the level of progressivity of the replacement rate that makes the pension system ex-ante neither regressive nor progressive. This proposal implies the following adjustment factor of the replacement rate

$$\mathcal{I}_{\text{SRP}}(\mathbf{pp}_{z,a}(\theta_n)) = \underbrace{\left[\left(1 - \frac{\widehat{\text{LE}}(\mathbf{pp}^{\min}_{z,a})}{\widehat{\text{LE}}(\mathbf{pp}^{\max}_{z,a})}\right) \Big/ \left(1 - \frac{\mathbf{pp}^{\min}_{z,a}}{\mathbf{pp}^{\max}_{z,a}}\right)\right]}_{=\phi^{\text{SRP}}} \left(1 - \frac{\overline{\mathbf{pp}}_{z,a}}{\mathbf{pp}_{z,a}(\theta_n)}\right) \tag{25}$$

where $\phi^{\text{SRP}}$ measures the average percentage change of life expectancy in response to a change in the number of pension points, or average elasticity between life expectancy and pension points, and $(\mathbf{pp}_{z,a}^{\min}, \mathbf{pp}_{z,a}^{\max})$ are respectively the minimum and maximum pension points of the cohort $z$ at age $a$. From Eq. (25) and assuming that the SRP proposal is fully implemented (i.e. $\zeta = 1$), the pension replacement rate in the SRP proposal becomes

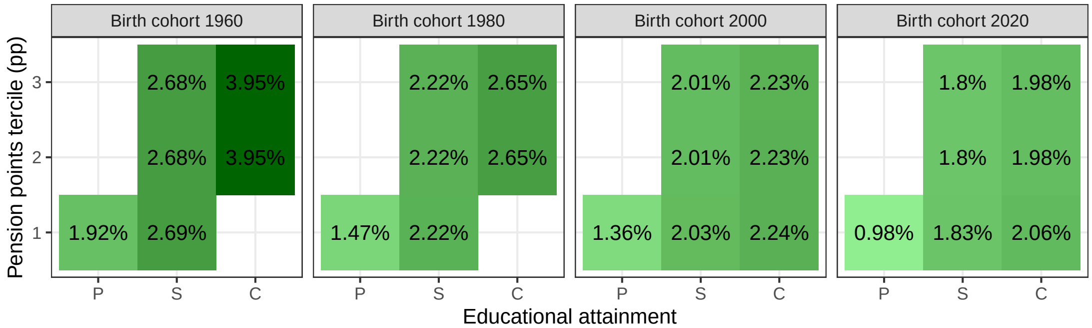
$$\boldsymbol{\varphi}_{\text{SRP}}(\mathbf{pp}_{z,a}(\theta_n)) = \varphi \frac{(1-\phi^{\text{SRP}})\mathbf{pp}_{z,a}(\theta_n) + \phi^{\text{SRP}}\overline{\mathbf{pp}}_{z,a}}{\mathbf{pp}_{z,a}(\theta_n)}. \tag{26}$$

Notice in Eq. (26) that the positive relationship between life expectancy and lifetime labor income, explained in section 2, is contained in the elasticity term $\phi^{\text{SRP}}$. Using the simulation results from the benchmark case we set the value of $\phi^{\text{SRP}}$ at 0.1348 (see section S6 in the supplementary material). Thus, Eq. (26) implies that those agents with pension points equal to the average number of pension points ($\mathbf{pp}_{z,a}(\theta_n) = \overline{\mathbf{pp}}_{z,a}$) will also receive a replacement rate equal to $\varphi$ (as in the ABH proposal and the status quo). While agents with pension points lower (resp. higher) than $\overline{\mathbf{pp}}_{z,a}$ will have a replacement rate that is $\phi^{\text{SRP}} \frac{\overline{\mathbf{pp}}_{z,a}}{\mathbf{pp}_{z,a}(\theta_n)}$ times higher (resp. lower) than that of the average replacement rate $\varphi$.

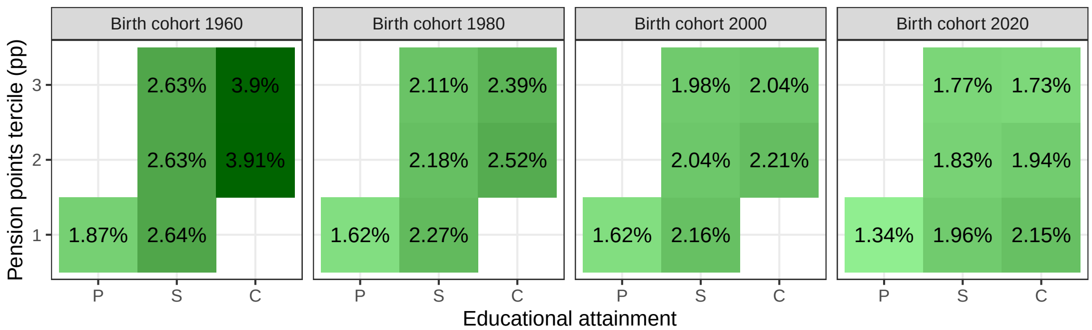
### *5.2. Pension proposals: Redistributive effects*

To study the redistributive properties of the ABH and SRP proposals we use the internal rate of return (IRR), which reflects the average return received by each contribution paid to the pension system. See section 2 for more details. An interesting characteristic of the Austrian pension system is that 30 percent of the total pension spending is also financed through the general budget. Thus, our calculations of the IRR include the social contributions paid by each representative agent and the fraction of labor income tax, capital income tax, and consumption tax devoted to finance the pension system. To assess the redistributive effects of ABH and SRP proposals, we divide our $\mathcal{N}$(=25) heterogeneous agents into 9 groups according to their educational attainment and the pension points accumulated until age 65. We choose age 65 in order to guarantee that almost all agents of a cohort are retired. Moreover, each agent type belongs to one out of three possible educational groups $\mathbf{E}$ = (primary, secondary, college) and one out of three possible levels of pension points accumulated (low, middle, high). Given that an absolute number of pension points will not mean the same for different cohorts and that its value will not lay over time in the same position of the pension points distribution, we split our heterogeneous agents by pension points terciles.

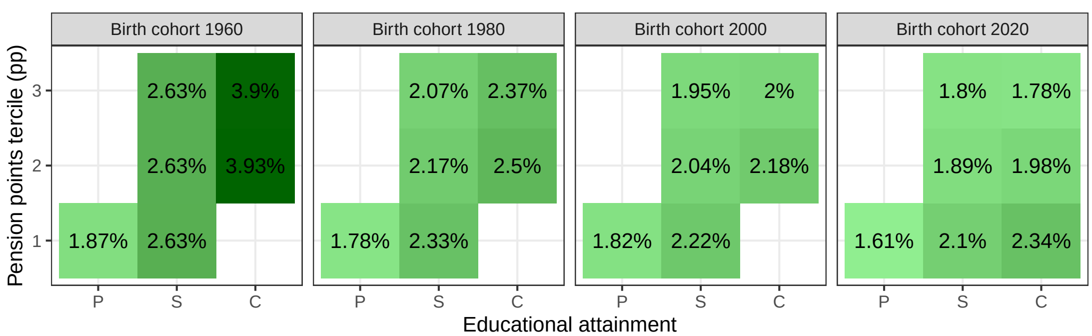
Figure 6 shows the IRR by educational attainment and pension points tercile for four selected birth cohorts (1960, 1980, 2000, and 2020). For each of the nine groups, the IRR is calculated as




(a) Benchmark




(b) Experiment 1: Ayuso et al. (2017)




(c) Experiment 2: Sánchez-Romero and Prskawetz (2020)

Figure 6: Internal rate of returns (IRR) of the Austrian pension system by pension points tercile (1=low, 2=middle, 3=high) and educational attainment (P=primary, S=secondary, C=college): selected birth cohorts under three different pension proposals. *Source*: Authors' calculations. *Note*: The reported IRR in each cell is the mean IRR across the 1 000 random simulations drawn from the posterior distribution.

the mean IRR across the 1 000 random simulations drawn from the posterior distribution of the initial endowments (learning ability and schooling effort). We include in Figure 6 the 1960 birth cohort in order to show the IRR before the pension proposals are introduced. See the phase-in/out period in Eq. (22). Figure 6 is divided in three panels: (a) the benchmark (or status quo), (b) the ABH pension proposal, and (c) the SRP pension proposal. The benchmark case shows the model results without the implementation of any policy correcting for the difference in life expectancy by pension points. Thus, we use the benchmark case to assess the marginal effects of the ABH and SRP policy proposals by comparing them to the status quo. In addition, to better read Fig. 6, for each cohort, the cell(s) with a lighter color represents the group(s) of agents who have the lowest IRR, while cell(s) with a darker color highlights the group(s) of agents who have the highest IRR.

We obtain the following results from panel (a) in Fig. 6. First, agents with college education receive a higher IRR than that received by agents with lower education. For instance, for the 1960 birth cohort, the average IRR is 3.95 percent for those with college education, 2.68 percent for those with secondary education, and 1.92 percent for those with primary education. Second, the difference across cohorts in the IRR between those with college education and those with primary education will diminish from 2.03%(=3.95%-1.92%) to 1%(=1.98%-0.98%) from the 1960 birth cohort to the 2020 birth cohort. This is because the decline in the IRR from the 1960 to the 2020 birth cohort is more pronounced for those with college education, from 3.95 percent to 1.98 percent, than for those with primary education, from 1.92 percent to 0.98 percent. However, despite the more pronounced decline among the highly-educated, third, agents with college education will continue receiving an IRR that doubles the IRR of those with primary education (3.95/1.92$\approx$ 2.06 for the 1960 birth cohort and 1.98/0.98$\approx$ 2.02 for the 2020 birth cohort). As a result, we can conclude that the Austrian pension system is *ex-ante* regressive.

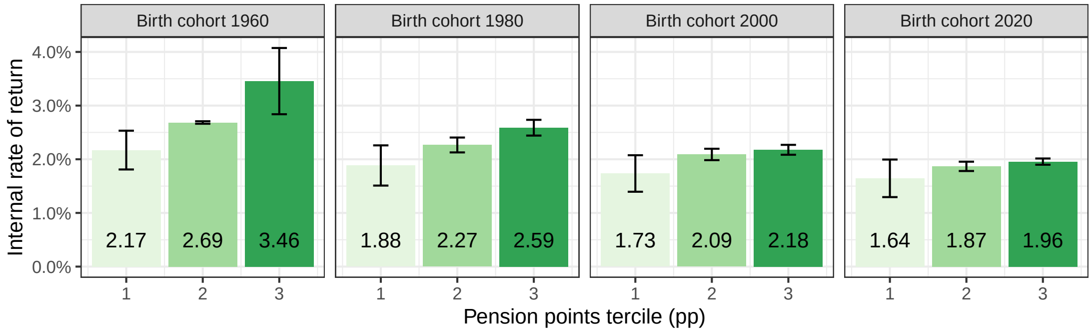
Panels (b) and (c), in Fig. 6, show the IRRs in the ABH and SRP pension proposals across the nine different subgroups. In both pension proposals we observe the same IRR pattern by education and by pension points tercile. On the one side, since in both pension proposals the replacement rate is inversely related to the number of pension points, ceteris paribus the educational attainment, agents in lower pension points terciles receive a higher IRR than those in the highest pension points tercile. Thus, for agents born in 2020 who have college education, we find in the ABH proposal that the IRR is 1.73% for those in the highest pension point tercile and 2.15% in the lowest pension point tercile and, similarly, in the SRP proposal the IRR is 1.78% for those in the highest pension point tercile and 2.34% in the lowest pension point tercile. On the other side, since education is positively related to life expectancy, ceteris paribus the number of pension points, agents who are highly-educated receive a higher IRR than those agents who are low-educated. Looking at agents born in 2020 who belong to the lowest pension points tercile, we find in the ABH proposal that the IRR is 1.34% for those with primary education and 2.15% for those with college education. Similarly, in the SRP proposal the IRR is 1.61% for those with primary education and 2.34% for those with college education. As a result, these two patterns reduce the difference in the IRR across educational groups but, conditional on a specific education level, both proposals also increase the difference in the IRR across pension points terciles as compared to the benchmark case. An advantage of the SRP proposal, relative to the other cases, is that low-educated agents in the lowest pension points tercile receive in this pension system the highest IRR (=1.61% for the 2020 birth cohort). However, the SRP proposal has as a drawback that it also provides the highest IRR (=2.34% for the 2020 birth cohort) to those agents who belong to the lowest pension points tercile and are highly educated. On the other hand, the ABH proposal generates an IRR that is between the benchmark case and the SRP proposal for those agents who belong to the lowest pension points tercile and are either low-educated (=1.33% for the 2020 birth cohort) or highly-educated (=2.2% for the 2020 birth cohort). This is because the SRP proposal compensates not only for differences in the life expectancy but also for differences in pension points, while the ABH only compensates for differences in life expectancy.

An alternative way of analyzing the ABH and SRP proposals is to look at the IRR only across pension points terciles (see Fig. 7) and only across educational groups (see Fig. 8). First, looking only across pension points terciles, Figure 7 shows in the benchmark case that agents that belong to the lower pension points terciles receive a lower IRR than those that belong to the highest pension points tercile. For the 1960 birth cohort, the average IRR of agents in the lowest pension points tercile is 2.17% and 3.46% for those in the highest pension points tercile. For the 2020 birth cohort, the average IRR of agents in the lowest pension points tercile is 1.64% and 1.96% for those in the highest pension points tercile. The reduction in the regressivity of the pension system as the rates of returns decline, was already found in the US pension system by Steuerle and Bakija (1994). This is because a reduction in the IRR generates a larger fall in the net benefits of high income earners than in the net benefits of low income earners (Feldstein and Liebman, 2002). Under the ABH and SRP proposals (see panels (b) and (c) in Fig. 7), we find for the 2020 birth cohort that the ABH proposal will make the pension system more favorable to the average retiree (i.e., those agents in the middle pension points tercile). Agents in the lowest pension points tercile have an IRR of 1.83%, which is higher than the IRR of 1.74% received by agents in the highest pension points tercile. But those agents with the medium pension points tercile receive the highest IRR, which is equal to 1.87%. Instead, the SRP proposal gives the highest IRR to agents in the lowest pension points tercile (1.99%), followed by those with medium pension points (1.92%), and the lowest IRR is given to agents in the highest pension points tercile (1.79%). Hence, only looking across pension points, we can see for the 2020 birth cohort that the SRP proposal will transform the pension system into a progressive system. Nonetheless, it should be noticed that for earlier cohorts (i.e., 1980 and 2000), all the analyzed pension systems are still regressive across pension point terciles.

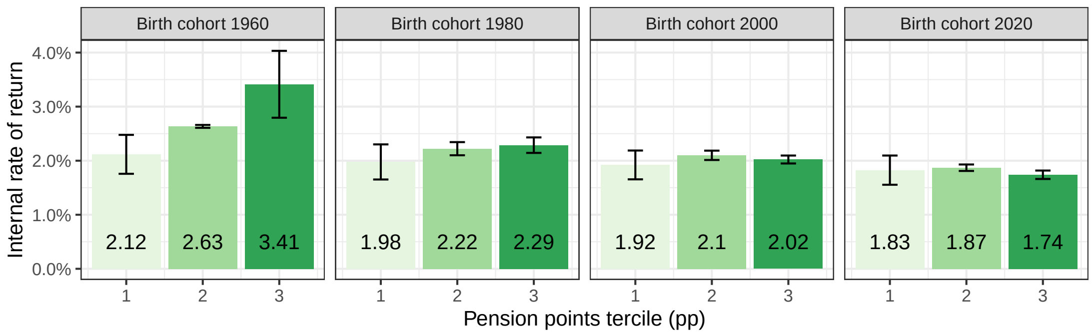
To compare the IRR results using an alternative SES measure, Figure 8 shows the IRR only across educational groups. Panel (a), in Fig. 8, reports the same IRR values as those shown in Figure 6, panel (a). This is because we have assumed that life expectancy depends on the educational attainment. However, contrary to Fig. 7 we find that neither the ABH proposal nor the SRP proposal makes the pension system progressive once that we analyze the IRR only across educational groups. Indeed, even for the 2020 birth cohort, in both proposals agents with secondary education receive the highest IRR (1.89% in the ABH proposal and 1.98% in the SRP proposal) and agents with primary education receive the lowest IRR (1.34% in the ABH proposal and 1.61% in the SRP proposal). The fact that the IRR is the lowest for those with primary education is mainly due to the fact that both replacement rate adjustment factors, $\mathcal{I}(\cdot)$, do not fully compensate for the difference in life expectancy. Moreover, when the overall IRR is higher, we can see for the 1980 birth cohort that both pension proposals are regressive.

### *5.3. Pension proposals: The impact on inequality*

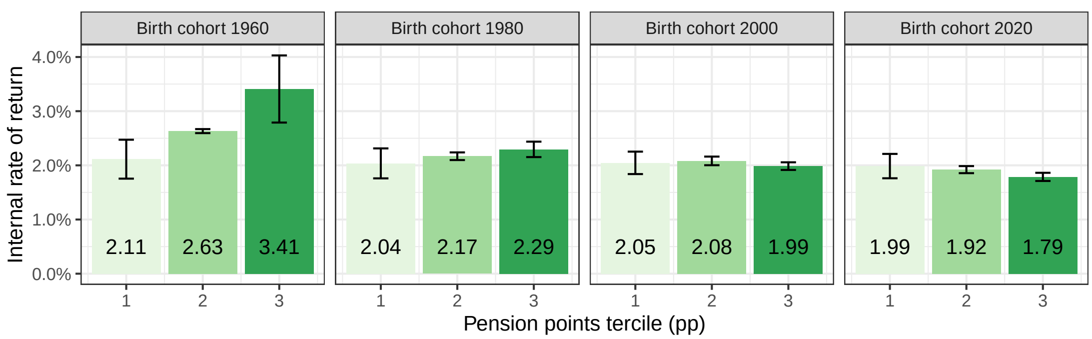
We have seen in Section 5.2 that both the ABH and the SRP proposals reduce the difference in IRR across educational groups, giving both pension proposals quite similar results. To provide an additional insight on the advantages and disadvantages of each proposal, Figure 9 shows for the three pension cases the coefficient of variation (CV) of the IRR. The CV is frequently used as a measure of inequality. The CV is calculated with all the $\mathcal{N}(=25)$ heterogeneous agents in order to maximize the variance in the data. For each birth cohort, the height of the bar gives the CV, which is also displayed as a number at the bottom of the bar. We associate a blue color to the benchmark, a light blue color to the ABH proposal, and a dark blue to the SRP proposal. Fig. 9 shows for the 1980 birth cohort that the highest IRR inequality is found in the benchmark case (0.18), followed by the ABH proposal (0.12) and then the SRP proposal (0.10). The same inequality gradient is observed for the 2020 birth cohort, in which the highest CV is 0.14 for the benchmark, followed by the ABH proposal with a CV of 0.10 and the SRP




(a) Benchmark



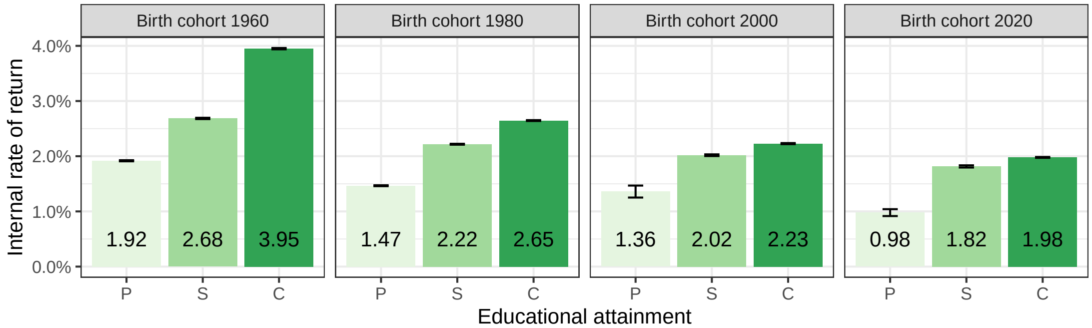

(b) Experiment 1: Ayuso et al. (2017)



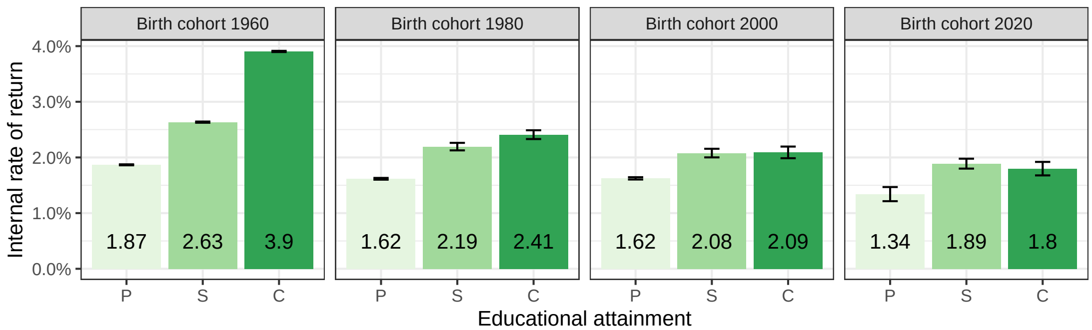

(c) Experiment 2: Sánchez-Romero and Prskawetz (2020)

Figure 7: Internal rate of returns (IRR) of the Austrian pension system by pension points tercile (1=low, 2=middle, 3=high): selected birth cohorts under three different pension proposals. *Source*: Authors' calculations. *Note*: The reported IRR in each cell is the mean IRR across the 1 000 random simulations drawn from the posterior distribution.

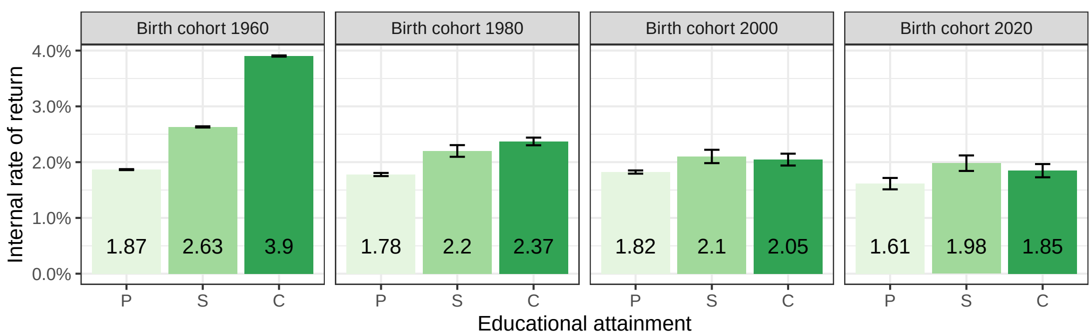
proposal with a CV of 0.09. Therefore, Fig. 9 clearly shows that the SRP proposal provides a more equal IRR across all agent types. The ABH proposal also reduces the inequality in the




(a) Benchmark




(b) Experiment 1: Ayuso et al. (2017)




(c) Experiment 2: Sánchez-Romero and Prskawetz (2020)

Figure 8: Internal rate of returns (IRR) of the Austrian pension system by educational attainment (P=primary, S=secondary, C=college): selected birth cohorts under three different pension proposals. *Source*: Authors' calculations. *Note*: The reported IRR in each cell is the mean IRR across the 1 000 random simulations drawn from the posterior distribution.

IRR, compared to the benchmark case, but to a lesser extent.

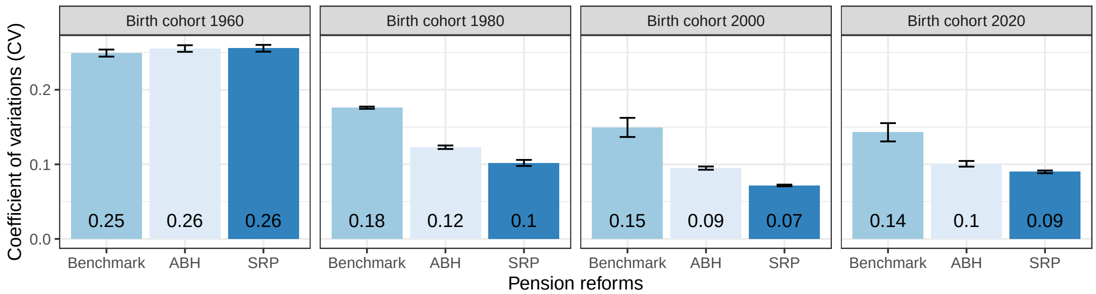



Figure 9: Coefficient of variation of the internal rate of return (IRR) for selected birth cohorts under three different pension proposals. *Source*: Authors' calculations. *Note:* The estimated CV is calculated as the mean CV across the 1 000 random simulations drawn from the posterior distribution.

### *5.4. Pension proposals: The impact on labor supply and education*

In this subsection we analyze whether the ABH and SRP proposals lead to additional behavioral reactions that will change the labor supply and the educational attainment found in the benchmark model (or status quo). Comparing the underlying variables that determine the IRR calculations, our results indicate that the pension points and the retirement ages are almost the same in the three pension proposals (see figs S14–S15 in the supp. material). Agents with primary education accumulate 50% of the pension points gained by agents in the reference group (ref=secondary education), while agents with college education accumulate 188% of the pension points of the reference group, which is almost three times greater than the pension points of those with primary. These two values are relatively close to the penalties and advantages of education on salaries reported by the OECD (2014) for Austria, that suggests values of 70%(less than upper secondary) and 171% (tertiary education), respectively. Notice, however, that the OECD estimates are derived from a period perspective, while our calculations are done from a cohort perspective. For those cohorts born between 1980 and 2020, agents with primary education tend to retire at age 58, with second education at 59, and slightly above 60 for those with college education. Thus, given that we are controlling for the years of education, these results imply that the ABH and SRP proposals have the same incentives and disincentives for working and retiring as the benchmark case. In contrast, we detect a marginal disincentive, relative to the benchmark, in both pension proposals to attain higher education (see fig S16 in the supp. material). Nonetheless, this effect is small and does not significantly affect on the IRRs reported. In sum, we obtain that the ABH proposal and the SRP proposals have similar incentives as the current pension system for education, labor supply, and retirement.

## 6. Conclusion

Population aging, as caused by low fertility levels and increasing life expectancy, challenges any social security system that is based on the redistribution of resources from the employed towards the dependent older population. The persistent population aging observed in most developed countries prompts governments to introduce pension reforms that guarantee the long-run sustainability of their social security systems. Such proposals are, among others, delaying the effective retirement age, introducing penalties and rewards for early and late retirement, and linking the pension replacement rate to the remaining life expectancy, among others. However, in many countries, the difference in life expectancy between the high and low socioeconomic groups have widened in recent decades. Ignoring this heterogeneity might jeopardize any proposal, as pension schemes become highly regressive. The introduction of any pension proposal needs to take into account that individual aging is heterogeneous across socioeconomic groups. Therefore, it is necessary to investigate how pension proposals that correct for ex-ante differences in life expectancy impact on the decisions of heterogeneous individuals by SES and on the degree of regressivity of the system across socioeconomic groups. This task implies developing models that account for the behavioral response of heterogeneous individuals with different life expectancies to changes in the pension system.

To account for potential behavioral responses and to control for the implications of changes in the educational distribution on the life expectancy gradient, this paper builds a computable overlapping generation model of labor supply with endogenous length of schooling and life expectancy. Agents are heterogeneous by their learning ability, life expectancy, and their effort of attending schooling. The model is applied to Austria and analyzes the redistributive characteristics of implementing two pension proposals: Ayuso et al. (2017) (ABH) and Sánchez-Romero and Prskawetz (2020) (SRP), which aim at reducing the regressivity of the pension program caused by the difference in life expectancy by SES.

Our simulations suggest the following results. Under the current Austrian pension system we obtain that agents with high SES receive a higher IRR than those with low SES. The difference in IRR for all SES groups will decline from the 1960 birth cohort to the 2020 birth cohort. The decline in the IRR will be more pronounced for the highly-educated workers than for the low-educated workers. Nonetheless, highly-educated workers will continue receiving from the pension system an IRR that doubles that of low-educated workers. Once that the ABH and SRP pension proposals are implemented, we obtain a reduction of the inequality in the IRR across agents with different educational attainment and pension points compared to the status quo. Comparing the effects of both proposals, we find that the SRP proposal provides a more equal IRR across population subgroups than the ABH proposal. The main advantage of the SRP proposal is to provide the highest IRR to agents who are short-lived and belong to the lowest pension points tercile. Its main disadvantage is that this proposal also provides the highest IRR to those agents who belong to the lowest pension points tercile and are highly educated. This result is due to the fact that the ABH only compensates for differences in life expectancy, while the SRP proposal compensates for differences in life expectancy as well as for differences in pension points. Comparing the ABH and SRP proposals to the current Austrian pension system, we did not find any significant distortion on lifecycle decisions of education, labor supply, and retirement.

Our simulation results also show that is crucial to introduce heterogeneity in several dimensions when analyzing the degree of progressivity/regressivity of pension proposals, since partial analyses might lead to contradictory results. In particular, if we report the IRR only by pension points, we find that the SRP proposal will transform the pension system into a progressive system for the cohort born in 2020. In contrast, if we report the IRR only by educational attainment, we find that the pension system will become regressive across educational groups, even when both proposals (ABH and SRP) are implemented.

This model is the first step to evaluate from a lifecycle perspective different social programs. In future work we plan to incorporate in this model an additional layer of heterogeneity by health status in order to analyze the IRR of other social benefits such as health care, disability, and other family benefits.

## References

Aaron, H.J. 1977. Demographic effects on the equity of social security benefits. In *The economics of public services*, 151–173. Palgrave Macmillan, London.

Auerbach, A.J., Charles, K.K., Coile, C.C. et al. 2017. How the growing gap in life expectancy may affect retirement benefits and proposals. *Geneva Papers on Risk and Insurance–Issues and Practice* 42, 475–499. https://doi.org/10.1057/s41288-017-0057-0

Ayuso, M., Bravo, J. M., Holzmann, R. 2016. On the Heterogeneity in Longevity Among Socioeconomic Groups: Scope, Trends, and Implications for Earnings-Related Pension Schemes. IZA Discussion Paper No. 10060.

Ayuso, M., Bravo, J. M., Holzmann, R. 2017. Addressing longevity heterogeneity in pension scheme design. *Journal of Finance and Economics*, 6 (1), 1–21.

Bergeaud, A., Cette, G., Lecat, R. 2016. Productivity trends in advanced countries between 1890 and 2012. *Review of Income and Wealth*, 62 (3), 420–444. `https://doi.org/10.1111/roiw.12185`

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)

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.

Doblhammer, G., Rau, R., Kytir, J., 2005. Trends in educational and occupational differentials in all-cause mortality in Austria between 1981/82 and 1991/92. *Wiener Klinische Wochenschrift*, 117 (13-14), 468–479.

Duggan, J.E., Gillingham, R., Greenlees, J.S. 1993. Returns paid to early social security cohorts. *Contemporary Economic Policy*, 11(4), 1–13.

European Commission 2018. The 2018 ageing report: underlying assumptions & projections. Institutional paper 065.

Fehr, H., Kallweit, M., Kindermann, F., 2013. Should pensions be progressive? *European Economic Review*, 63, 94–116.

Fehr, H., Uhde, J., 2013. On the optimal design of pension systems. *Empirica*, 40 (3), 457–482.

Fehr, H., Uhde, J. (2014). Means-testing and economic efficiency in pension design. *Economic Modelling*, 44, S57–S67.

Feldstein, M., Liebman, J. B., 2002. Social security. *Handbook of Public Economics*, in: A. J. Auerbach, M. Feldstein (ed.), Handbook of Public Economics, edition 1, volume 4, chapter 32, 2245–2324 Elsevier.

Fleurbaey, M., 2008. *Fairness, responsibility, and welfare*. Oxford University Press.

Goldring, T., Lange, F., Richards-Shubik, S., 2016. Testing for changes in the SES-mortality gradient when the distribution of education changes too. *Journal of Health Economics*, 46, 120–130. `https://doi.org/10.1016/j.jhealeco.2015.12.002`

Goujon, A., K.C., S., Speringer, M., Barakat, B., Potancoková, M., Eder, J.,... Lutz, W., 2016. A harmonized dataset on global educational attainment between 1970 and 2060 –an analytical window into recent trends and future prospects in human capital development–. *Journal of Demographic Economics*, 82 (3), 315–363. `doi:10.1017/dem.2016.10`

Gustman, A.L., Steinmeier, T.L. 2001. How effective is redistribution under the social security benefit formula? *Journal of Public Economics*, 82 (1), 1–28.

Haan, P., Kemptner, D., Lüthen, H., 2020. The rising longevity gap by lifetime earnings–Distributional implications for the pension system. *The Journal of the Economics of Ageing*, 100199.

Hendi, A. S., Elo, I. T., Martikainen, P., 2021. The implications of changing education distributions for life expectancy gradients. *Social Science & Medicine*, 272, 113712. `https://doi.org/10.1016/j.socscimed.2021.113712`

Holzmann, R., Alonso-García, J., Labit-Hardy, H., Villegas, A.M., 2019. NDC schemes and heterogeneity in longevity: proposals for redesign. In Holzmann R, Palmer E, Palacios R, Sacchi S (eds). *Progress and challenges of nonfinancial defined pension schemes: volume 1. Addressing marginalization, polarization, and the labor market*. World Bank, Washington,DC

Hurd, M.D., Shoven, J.B. 1985. The Distributional Impact of Social Security. In Wise, D.A. (ed). *Pensions, Labor, and Individual Choice*. Chicago: University of Chicago Press, 193–221.

Keane, M., Rogerson, R., 2012. Micro and macro labor supply elasticities: A reassessment of conventional wisdom. *Journal of Economic Literature*, 50 (2), 464–76.

Klotz, J., 2010. Convergence or divergence of educational disparities in mortality and morbidity? The evolution of life expectancy and health expectancy by educational attainment in Austria in 1981-2006. *Vienna Yearbook of Population Research*, 8, 139–174.

Knell, M., Köhler-Töglhofer, W., Prammer, D., 2006. The Austrian pension system –how recent proposals have changed fiscal sustainability and pension benefits. *Monetary Policy and the Economy* 2, 69–93.

Laun, T., Markussen, S., Vigtel, T.C., Wallenius, J. 2019. Health, longevity and retirement proposal. *Journal of Economic Dynamics and Control*, 103, 123–157.

Lee, R., Sanchez-Romero, M., (2019). Overview of the relationship of heterogeneity in life expectancy to pension outcomes and lifetime income. In Holzmann R, Palmer E, Palacios R, Sacchi S (eds). *Progress and challenges of nonfinancial defined pension schemes: volume 1. Addressing marginalization, polarization, and the labor market.* World Bank, Washington, DC.

Le Garrec, G., 2015. Increased longevity and social security proposal: questioning the optimality of individual accounts when education matters. *Journal of Population Economics* 28, 329–352.

Liebman, J.B. 2002. Redistribution in the Current U.S. Social Security System. In Feldstein, M., Liebman, J.B. (eds) *The Distributional Aspects of Social Security and Social Security Reform*, 11–48. University of Chicago Press, .

Lleras-Muney, A., 2005. The relationship between education and adult mortality in the United States. *The Review of Economic Studies*, 72(1), 189–221.

Lutz, W., Butz, W. P., Samir, K. E. (Eds.), 2014. *World population and human capital in the twenty-first century.* OUP Oxford.

Lutz, W., Goujon, A., KC, S., Sanderson, W. ,2007. Reconstruction of populations by age, sex and level of educational attainment for 120 countries for 1970–2000. *Vienna Yearbook of Population Research*, 193–235.

Luy, M., Di Giulio, P., Caselli, G., 2011. Differences in life expectancy by education and occupation in Italy, 1980-94: Indirect estimates from maternal and paternal orphanhood. *Population Studies*, 65 (2), 137–155.

Manchester, J., Topoleski, J., 2008. Growing disparities in life expectancy. Washington, DC: Congressional Budget Office.

OECD, 2014. *Education at a Glance 2014: OECD Indicators.* OECD Publishing. http://dx.doi.org/10.1787/eag-2014-en

Olshansky, S.J., Antonucci, T., Berkman, L., Binstock, R.H., Boersch-Supan, A., Cacioppo, J.T., Carnes, B.A., Carstensen, L.L., Fried, L.P., Goldman, D.P., Jackson, J., Kohli, M., Rother, J., Zheng, Y., Rowe, J., 2012. Differences in life expectancy due to race and educational differences are widening, and many may not catch up. *Health Affairs* 31 (8),1803–1813.

Oreopoulos, P., 2007. Do dropouts drop too soon? Wealth, health and happiness from compulsory schooling. *Journal of Public Economics* 91, 2213–2229.

Palmer, E., de Gosson de Varennes, Y. Z. 2019. Annuities in (N)DC pension schemes: Design, heterogeneity, and estimation issues. In Holzmann R, Palmer E, Palacios R, Sacchi S (eds). *Progress and challenges of nonfinancial defined pension schemes: volume 1. Addressing marginalization, polarization, and the labor market.* World Bank, Washington,DC

Pestieau, P., Ponthiere, G., 2016. Longevity variations and the welfare state. *Journal of Demographic Economics*, 82 (2), 207–239.

Pestieau, P., Racionero, M., 2016. Harsh occupations, life expectancy and social security. *Economic Modelling*, 58, 194–202.

Poole, D., Raftery, A.E., 2000. Inference for deterministic simulation models: the Bayesian melding approach. *Journal of the American Statistical Association*, 95 (452), 1244–1255.

Preston, S.H., Elo, I.T., 1995. Are educational differentials in adult mortality increasing in the United States? *Journal of Aging and Health*, 7 (4), 476–496.

Raftery, A.E., Bao, L., 2010. Estimating and projecting trends in HIV/AIDS generalized epidemics using incremental mixture importance sampling. *Biometrics*, 66, 1162–1173.

Restuccia, D., Vandenbroucke, G., 2013. A century of human capital and hours. *Economic Inquiry* 51 (3), 1849–1866.

Rivic, S., 2019. Konstruktion einer demographischen und ökonomischen Datenbank für Österreich zwischen 1820 und 2010. Bachelor Thesis. TU Wien.

Sánchez-Romero, M., d'Albis, H., Prskawetz, A., 2016. Education, lifetime labor supply, and longevity improvements. *Journal of Economic Dynamics and Control* 73, 118–141.

Sánchez-Romero, M., Lee, R. D., Prskawetz, A., 2020. Redistributive effects of different pension systems when longevity varies by socioeconomic status. *The Journal of the Economics of Ageing*, 17, 100259. `https://doi.org/10.1016/j.jeoa.2020.100259`

Sánchez-Romero, M., Prskawetz, A., 2019. Social Security: Long-term financing and proposal. In: Gu D., Dupre M. (eds) *Encyclopedia of Gerontology and Population Aging*. Springer, Cham. https://doi.org/10.1007/978-3-319-69892-2_535-1

Sánchez-Romero, M., Prskawetz, A., 2020. The impact of reducing the pension generosity on inequality and schooling. *De Economist*, 168 (2), 279–304.https://doi.org/10.1007/s10645-020-09359-w

Sánchez-Romero, M., Sambt, J., Prskawetz, A., 2013. Quantifying the role of alternative pension proposals on the Austrian economy. *Labour Economics*, 22 (C), 94–114.

Schröder, C. 2012. Profitability of pension contributions?evidence from real-life employment biographies. *Journal of Pension Economics & Finance*, 11 (3), 311–336.

Shkolnikov, V.M., Andreev, E.M., Jasilionis, D., Leinsalu, M., Antonova, O.I., McKee, M., 2006. The changing relation between education and life expectancy in central and eastern Europe in the 1990s. *Journal of Epidemiology and Community Health*, 60 (10), 875–881.

Steuerle, C. E., Bakija, J. M. (1994). *Retooling Social Security for the 21st century: Right and wrong approaches to proposal.* The Urban Insitute.

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.

Wittgenstein Centre for Demography and Global Human Capital, 2015. Wittgenstein Centre Data Explorer Version 1.2. Available at: http://www.wittgensteincentre.org/dataexplorer

## S. Supplementary material

*S1. Solution: Household problem*

Given the set of endowments $\theta_n \in \Theta_n$ we solve the household problem of maximizing the lifetime utility (13) and the educational decision (14) subject to the constraints (7)-(12) and the boundary conditions $k_{\underline{a},e} = 0$ and $h_{\underline{a},e} = h_{\underline{a}}$. For notational convenience, let us define the marginal rate of substitution between pension points and assets for an agent of age $a$ with education $e$ as

$$\mathcal{P}_{a,e} = \frac{\partial V(\mathbf{x}_{a,e};\theta_n)}{\partial \mathrm{pp}_{a,e}} \Big/ \frac{\partial V(\mathbf{x}_{a,e};\theta_n)}{\partial k_{a,e}}$$

and the marginal rate of substitution between human capital and assets for an agent of age $a$ with education $e$ as

$$\mathcal{H}_{a,e} = \frac{\partial V(\mathbf{x}_{a,e};\theta_n)}{\partial h_{a,e}} \Big/ \frac{\partial V(\mathbf{x}_{a,e};\theta_n)}{\partial k_{a,e}} .$$

Each marginal rate of substitution measures the value, assigned by an agent with endowments $\theta_n$, of investing in each state (pension points and human capital) relative to investing in assets.

The first-order conditions (FOCs) of this problem are:

$$U_c(c_{a,e}, l_{a,e}) = \beta\pi_{a+1,e}\frac{\partial V(\mathbf{x}_{a+1,e};\theta_n)}{\partial k_{a+1,e}}(1+\tau_a^c), \tag{S.1}$$

$$-U_l(c_{a,e}, l_{a,e}) = U_c(c_{a,e}, l_{a,e})\left(1-\tau_{a,e}^L\right) w_{a,e}, \tag{S.2}$$

where $\tau_{a,e}^L = \frac{\tau_a^c+\tau_a^l+\tau_{a,e}^S+\tau_{a,e}^J(-\alpha_J'(l_a))}{1+\tau_a^c}$ is the effective labor income tax. Notice that the effective labor income tax includes the effective social security tax rate at the intensive margin, denoted by $\tau_{a,e}^S$, and the retirement tax/subsidy rate, denoted by $\tau_{a,e}^J$, which are given by

$$\tau_{a,e}^S = \tau_a^s(1-\tau_a^l) - \mathcal{P}_{a+1,e}\phi^p \mathrm{PBI}'(y_{a,e}), \tag{S.3}$$

$$\tau_{a,e}^J = (1-\tau_a^l)\left(1+\varepsilon_{b,\alpha_J,e}\right)\frac{b_{a,e}}{w_{a,e}} - (\mathcal{R}_a - 1)\frac{\mathrm{pp}_{a,e}\mathcal{P}_{a+1,e}}{w_{a,e}}. \tag{S.4}$$

The term $\varepsilon_{b,\alpha_J,e}$ is the retirement-elasticity of pension benefit; i.e. $\frac{1}{b_{a,e}}\frac{\partial b_{a,e}}{\partial l_{a,e}}\frac{\alpha_J(l_{a,e})}{\alpha_J'(l_{a,e})}$. Eqs. (S.3)-(S.4) coincide with the effective social security tax rate and the retirement tax/subsidy rate in Sánchez-Romero et al. (2020).

The envelope conditions (ECs) imply that:

$$U_c(c_{a,e}, l_{a,e}) = R_{a+1,e}\beta\pi_{a+1,e}\frac{1+\tau_a^c}{1+\tau_{a+1,e}^c}U_c(c_{a+1,e}, l_{a+1,e}), \tag{S.5}$$

$$R_{a,e}\mathcal{P}_{a.e} = (1-\tau_a^l)\frac{\partial b_{a,e}}{\partial \mathrm{pp}_{a,e}}\alpha_J(l_{a,e}) + \mathcal{P}_{a+1,e}\frac{\partial \mathrm{pp}_{a+1,e}}{\partial \mathrm{pp}_{a,e}}, \tag{S.6}$$

$$R_{a,e}\mathcal{H}_{a,e} = \left(1-\tau_a^l-\tau_{a,e}^S\right)\frac{y_{a,e}}{h_{a,e}} + \mathcal{H}_{a+1,e}\frac{\partial h_{a+1,e}}{\partial h_{a,e}}, \tag{S.7}$$

Combining FOCs and ECs we have that the total expenditure on final goods not only changes with age because of the difference between the market and the subjective time discount factors, but also because of changes in the household size

$$\frac{(1+\tau^c_{a+1})c_{a+1,e}}{(1+\tau^c_a)c_{a,e}} = \beta(1+r_a(1-\tau^k_a))\frac{H_{a+1,e}}{H_{a,e}}. \tag{S.8}$$

The labor supply of our representative agents is given by

$$l_{a,e} = \begin{cases} \left(\frac{1}{\alpha_L}\frac{(1-\tau^L_{a,e})w_{a,e}}{c_{a,e}/H_{a,e}}\right)^{\sigma_L} & \text{if } a < J, \\ \left(\frac{1}{\alpha_L}\frac{(1-\tau^L_{a,e})w_{a,e}}{c_{a,e}/H_{a,e}} - \frac{1}{\alpha_L}\frac{v_0(LE_{a,e})^{-v_1}}{\overline{L}}\right)^{\sigma_L} & \text{if } a \geq J. \end{cases} \tag{S.9}$$

Eq. (S.9) implies that agents who have lower effective labor income tax and higher wage rates, relative to the average consumption of the household, supply more labor. Once that retirement is allowed, those agents with longer life expectancy, lower effective labor income tax, and lower wage rates, relative to the average consumption of the household, will retire later, ceteris paribus the initial endowments.

The value of $\mathcal{H}_a$, which is also calculated backwards, gives

$$\mathcal{H}_{a,e}h_{a-1,e} = \sum_{s=a}^{\Omega-1}\left(\prod_{z=a}^{s}\frac{1}{R_{z,e}}\right)\left(1-\tau^l_s-\tau^S_{s,e}\right)y_{s,e}. \tag{S.10}$$

The value of human capital times the stock of human capital is the present value of the remaining lifetime income, which includes the present value of future pension benefits through the stream of $\{\tau^S_{a,e}y_{a,e}\}_{a=\underline{a}}^{\overline{J}}$ values.

*S2. Equilibrium conditions*

Given initial time, cohort, and age sets $\{\mathcal{T},\mathcal{Z},\mathcal{A}\}$, the set of education levels $\mathbf{E}$, the probability space of initial endowments $(\Theta,\boldsymbol{\theta},\mathrm{P})$, the number of heterogeneous agents $\mathcal{N}$ in each birth cohort, the model parameters (see Table 1), exogenous economic data $\{A_t\}_{t\in\mathcal{T}}$, and demographic data $\{N_{t,a},\pi_{z,a,e},\mathrm{fer}_{z,a,e},H_{z,a,e},\Delta_e\}_{t\in\mathcal{T},z\in\mathcal{Z},a\in\mathcal{I},e\in\mathbf{E}}$, a recursive competitive equilibrium is a sequence of a set of household policy functions $\{\mathbf{c}_{z,a}(\theta_n),\ \mathbf{l}_{z,a}(\theta_n),\mathbf{k}_{z,a}(\theta_n),\mathbf{pp}_{z,a}(\theta_n),\ \mathbf{h}_{z,a}(\theta_n)\}$ for $z\in\mathcal{Z},a\in\mathcal{A},\boldsymbol{\theta}\in\Theta,n\in 1,\ldots,\mathcal{N}$, government policy functions $\{G_t,\ \tau^c_t,\ \tau^l_t,\tau^k_t,\ \tau^s_t\}_{t\in\mathcal{T}}$ and factor prices $\{w_t,r_t\}_{t\in\mathcal{T}}$ such that

i. Given the factor prices and government policy functions, household policy functions satisfy (7)-(14).

ii. Factor prices $w_t$, $r_t$ equal their marginal productivities.

iii. The government's budget constraints (18) and (20) are satisfied.

iv. The stock of capital and the effective labor input are given by:

$$K_t = \sum_{a=0}^{\Omega} \sum_{n=1}^{\mathcal{N}} N_{t,a} \int_{\Theta_n} \mathbf{k}_{t-a,a}(\theta_n) \mathrm{d}\,\mathrm{P}_{t-a}(\theta_n), \tag{S.11}$$

$$L_t = \sum_{a=0}^{\Omega} \sum_{n=1}^{\mathcal{N}} N_{t+1,a+1} \int_{\Theta_n} \epsilon_a(\mathbf{e}_{t-a}(\theta_n)) \mathbf{h}_{t-a,a}(\theta_n) \mathbf{l}_{t-a,a}(\theta_n) \mathrm{d}\,\mathrm{P}_{t-a}(\theta_n). \tag{S.12}$$

v. The market of final goods clears

$$Y_t = C_t + G_t + I_t. \tag{S.13}$$

Aggregate consumption of final goods is given by

$$C_t = \sum_{a=0}^{\Omega} \sum_{n=1}^{\mathcal{N}} N_{t,a} \int_{\Theta_n} \mathbf{c}_{t-a,a}(\theta_n) \mathrm{d}\,\mathrm{P}_{t-a}(\theta_n). \tag{S.14}$$

*S3. Bayesian melding method*

We use the Bayesian melding method to derive in our dynamic general equilibrium-overlapping generations model the unobserved initial heterogeneity of our heterogeneous agents, while keeping consistency between the micro- and the macroeconomic information. To implement the Bayesian melding we initially used the sampling importance resampling (SIR) algorithm (Poole and Raftery, 2000). However, after running the algorithm thousands of times the number of unique points was very low, which is a signal of poor performance and suggests that the algorithm is not suitable for finding the most likely parameters. To cope with this problem, Raftery and Bao (2010) suggest to use a more sophisticated algorithm such as the incremental mixture importance sampling (IMIS) algorithm, which outperforms the Markov chain Monte Carlo (MCMC) algorithm. We modify the IMIS algorithm of Raftery and Bao (2010) in order to allow for heterogeneous agents.

Let our large scale dynamic general equilibrium-overlapping generations model be $M(\cdot)$. Let us assume each cohort is represented by a set of $\mathcal{N}$ heterogeneous agents whose endowments are randomly assigned at birth. Let the set of endowments characterizing the $n$-th agent be $\theta_n = (\xi_n, \eta_n)$ or permanent unobserved heterogeneity. Let $\Theta$ be the the product set of $\Theta_1, \ldots, \Theta_{\mathcal{N}}$ that consists of all $\mathcal{N}$-tuples $(\theta_1, \theta_2, \ldots, \theta_n, \ldots, \theta_{\mathcal{N}})$ where $\theta_n \in \Theta_n$ for each $n$. Let a realization of $\Theta$ be $\boldsymbol{\theta}$. The initial endowments $\boldsymbol{\theta}$ are random variables with a joint prior distribution denoted by $q_1(\Theta)$. We assume independent uniform priors for the distribution on the inputs

$$q_1(\Theta) = \mathcal{U}\left([0, 0.3] \times [0, 0.4]\right).$$

Let $\Phi = (\mathbf{e}_{z_0}(\boldsymbol{\theta}), \ldots, \mathbf{e}_{z_T}(\boldsymbol{\theta}))$ be the set of outputs of the dynamic general equilibrium-overlapping generations model given the model inputs $\boldsymbol{\theta}$; i.e., $M(\boldsymbol{\theta}) = \Phi$. We assume the likelihood of the model's output is given by

$$\mathcal{L}(\Phi|\text{data}) \propto -\frac{1}{2}\sum_{z=z_0}^{z_T}(\mathbf{m}_z(\boldsymbol{\theta}) - \widehat{\mathbf{m}}_z)'\widehat{W}^{-1}(\mathbf{m}_z(\boldsymbol{\theta}) - \widehat{\mathbf{m}}_z) \tag{S.15}$$

where $\mathbf{m}_z(\boldsymbol{\theta}) = (\mathrm{E}[\mathbf{e}_z(\boldsymbol{\theta})]; \sigma[\mathbf{e}_z(\boldsymbol{\theta})])$ is the vector with the model mean and standard deviation of the additional years of education for cohort $z$, $\widehat{\mathbf{m}}_z$ is the vector with the estimated mean and standard deviation of the additional years of education for cohort $z$, and $\widehat{W} = \operatorname{diag}(\sigma[\boldsymbol{\mu}^e], \sigma[\boldsymbol{\sigma}^e])$ is the weighting matrix with the standard deviations of the estimated mean and standard deviation of the additional years of education across all cohorts.

**IMIS algorithm (Raftery and Bao, 2010)**

1. Initial Stage:

   (a) Run $B_0$ samples of $\boldsymbol{\theta} \in \Theta$ realizations from the joint prior distribution on inputs $q_1(\Theta)$ obtained with the SIR algorithm

   (b) For each $\boldsymbol{\theta}_i$ sampled, run the model to obtain the set of output $M(\boldsymbol{\theta}_i) = \Phi_i$,

   (c) Calculate the likelihood of each model output

   $$\mathcal{L}(\Phi_i|\text{data}) \text{ for } i = \{1, \ldots, B_0\}$$

   (d) Construct the importance sampling weights (ISW)

   $$w_0(\boldsymbol{\theta}_i) \propto \frac{\mathcal{L}(\Phi_i|\text{data})}{\sum_{i=1}^{B_0}\mathcal{L}(\Phi_i|\text{data})}$$

2. Importance Sampling Stage: for $k = 1, 2, \ldots,$ until a stopping criteria is satisfied

   (a) Compute $\mathcal{N}$ multivariate Gaussian distribution $H_n^{(k)}$ with center $\mu_n^{(k)}$ and covariance $\Sigma_n^{(k)}$ for $n \in \{1, \ldots, \mathcal{N}\}$. Choose the input set $\boldsymbol{\theta}_i = (\theta_{i1}, \ldots, \theta_{i\mathcal{N}})$ with maximum weight, $w_{k-1}(\boldsymbol{\theta}_i)$. Choose as the center $\mu_n^{(k)}$ the set of parameters $\theta_{in}^{(k)}$. Calculate the weighted covariance matrix $\Sigma_n^{(k)}$ with the $(B)$ agents, one for each sampled $\vartheta$, with the smallest Mahalanobis distance to $\theta_{in}^{(k)}$ and the weights are the average between the importance weight and $1/B_k$.

   (b) Sample $B$ new inputs $\theta_{jn}$, with $j \in \{1, \ldots, B\}$, from $H_n^{(k)}$ for each $n$-th agent and form inputs $\boldsymbol{\theta}_j$ and combine them with the previous realizations.

(c) Compute steps 1(b)–(c) and calculate the new importance sampling weights as follows

$$w_k(\boldsymbol{\theta}_i) \propto \mathcal{L}(M(\boldsymbol{\theta}_i)|\text{data}) \times \prod_{n=1}^{\mathcal{N}} \frac{q_1(\theta_{in})}{q_n^{(k)}(\theta_{in})},$$

where $q_n^{(k)}(\theta_{in})$ is the mixture sampling distribution for the $n$-th agent, with $q_n^{(k)}(\theta_{in}) = \frac{B_0}{B_k} q_1(\theta_{in}) + \frac{B}{B_k} \sum_{s=1}^{k} H_n^{(s)}(\theta_{in})$ and $B_k = B_0 + Bk$ is the total number of inputs up to iteration $k$.

3. Resample Stage: For $J$ equal to 1 000, if the expected fraction of unique points after resampling $\hat{Q}(w) = \frac{1}{J} \sum_{i=1}^{B_k} (1 - (1 - w_i)^J)$ is less than 63%, go to Step 2.; otherwise, resample ($J$) 1 000 inputs with replacement from $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_{B_K}$ with weights $w_1, \ldots, w_{B_K}$, where $K$ is the number of iterations at step 2.

After running the IMIS algorithm we have obtained the 1 000 most likely inputs (i.e., initial endowments). Figure S10 shows how the two initial endowments (learning ability and schooling effort) are positively correlated. Table S2 reports the mean and standard deviation across the 1 000 initial endowments for each of the $\mathcal{N}$ clusters.

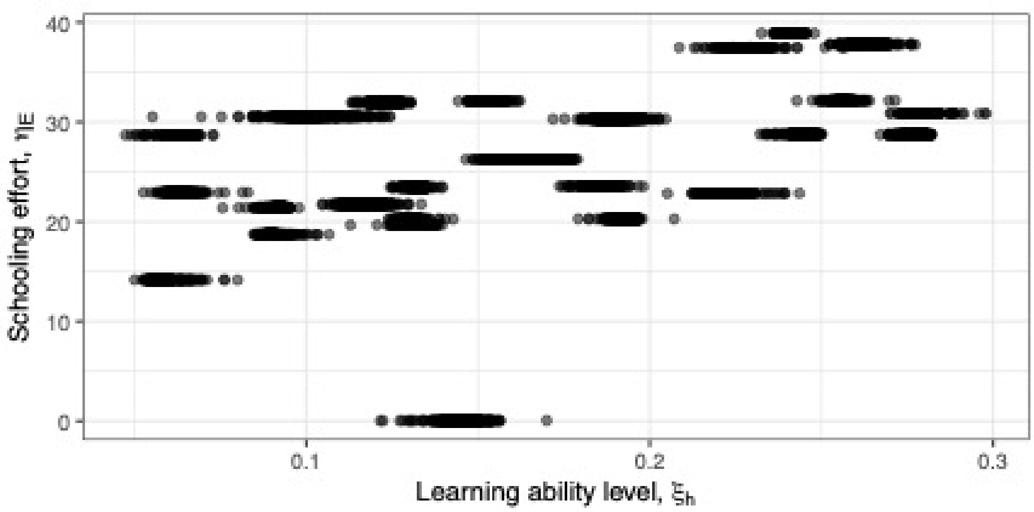



Figure S10: Correlation matrix of the initial endowments $\vartheta$ for the $\mathcal{N} = 25$ agents of each cohort. *Notes*: Dots represent the initial endowments of the most likely set of parameters obtained from the posterior distribution.

*S4. Introducing differential fertility and mortality in the model*

*Mortality.* We use standard mortality differentials by education based on existing literature (Lutz et al., 2007, 2014; Goujon et al., 2016). The next table shows the difference in life expectancy at age 15 between agents with education $e$ and those with college (reference group).

To include the differential mortality by educational group across cohorts, we first calculate the life expectancy of the reference group (=college). Let us denote by $\Delta_e$ the difference in life expectancy at age 15 between agents with education $e$ and those with college. Thus, the life expectancy at age 15 of an agent born in year $z$ with educational attainment $e$ can be written as $\mathrm{LE}_{z,e} = \mathrm{LE}_{z,8} - \Delta_e$. Let the average life expectancy at age 15 of the cohort born in year $z$ be denoted by $\overline{\mathrm{LE}}_z$, which can be expressed as

$$\overline{\mathrm{LE}}_z = \sum\nolimits_{e\in\mathbf{E}} \frac{N_{z,e}\mathrm{LE}_{z,e}}{N_z} = \sum\nolimits_{e\in\mathbf{E}} \frac{N_{z,e}}{N_z}(\mathrm{LE}_{z,8} - \Delta_e) = \mathrm{LE}_{z,8} - \sum\nolimits_{e\in\mathbf{E}} \frac{N_{z,e}\Delta_e}{N_z}, \tag{S.16}$$

Table S2: Mean and standard deviation of the initial endowments across the 1 000 parameter sets withdrawn from the posterior distribution

| Cluster | Learning ability, $\xi_n$ | | Schooling effort, $\eta_n$ | |
|---|---|---|---|---|
| $\mathcal{N}$ | $\mathrm{E}[\xi_n]$ | $\mathrm{sd}[\xi_n]$ | $\mathrm{E}[\eta_n]$ | $\mathrm{sd}[\eta_n]$ |
| 1 | 0.144 | 0.003 | 0.040 | 0.003 |
| 2 | 0.058 | 0.002 | 14.162 | 0.002 |
| 3 | 0.090 | 0.002 | 18.726 | 0.001 |
| 4 | 0.133 | 0.002 | 19.653 | 0.003 |
| 5 | 0.131 | 0.001 | 20.253 | 0.001 |
| 6 | 0.193 | 0.002 | 20.277 | 0.003 |
| 7 | 0.092 | 0.002 | 21.404 | 0.002 |
| 8 | 0.117 | 0.003 | 21.716 | 0.008 |
| 9 | 0.221 | 0.003 | 22.819 | 0.005 |
| 10 | 0.063 | 0.002 | 22.912 | 0.003 |
| 11 | 0.133 | 0.001 | 23.460 | 0.001 |
| 12 | 0.186 | 0.002 | 23.563 | 0.004 |
| 13 | 0.170 | 0.005 | 26.250 | 0.004 |
| 14 | 0.063 | 0.002 | 28.675 | 0.002 |
| 15 | 0.278 | 0.002 | 28.724 | 0.002 |
| 16 | 0.245 | 0.002 | 28.764 | 0.002 |
| 17 | 0.189 | 0.003 | 30.291 | 0.004 |
| 18 | 0.100 | 0.006 | 30.548 | 0.008 |
| 19 | 0.280 | 0.002 | 30.859 | 0.002 |
| 20 | 0.124 | 0.002 | 31.955 | 0.001 |
| 21 | 0.151 | 0.002 | 32.102 | 0.001 |
| 22 | 0.256 | 0.002 | 32.128 | 0.002 |
| 23 | 0.229 | 0.003 | 37.456 | 0.006 |
| 24 | 0.262 | 0.003 | 37.780 | 0.004 |
| 25 | 0.241 | 0.001 | 38.894 | 0.002 |

| Education level, $e$ | Primary or less ($e=0$) | Secondary ($e=4$) | College (Ref.) ($e=8$) |
|---|---|---|---|
| Life expectancy differential, $\Delta_e$ | -5.0 | -1.5 | 0 |

where $\frac{N_{z,e}}{N_z}$ is the fraction of people of cohort $z$ with education $e$. Rearranging terms in (S.16) we have that the life expectancy of cohort $z$ at age 15 with education $e$ is given by

$$\mathrm{LE}_{z,e} = \overline{\mathrm{LE}}_z + \sum\nolimits_{e\in\mathbf{E}} \left(\frac{N_{z,e}}{N_z} - 1\right) \Delta_e. \tag{S.17}$$

Second, assuming that $\pi_{z,a,e} = (\pi_{z,a})^{\kappa_e^m}$, we calculate the age-specific conditional survival probabilities by education $e$ of each cohort $z$ (i.e. $\pi_{z,a,e}$) as

$$\min_{\kappa_e^m\in\mathbb{R}} \left(\mathrm{LE}_{z,e} - \sum\nolimits_{a=15}^{\Omega} \left[\Pi_{s=15}^{a} (\pi_{z,s})^{\kappa_e^m}\right]\right) \text{ for } e \in \mathbf{E}. \tag{S.18}$$

*Fertility.* We introduce the fertility differential by education assuming that the net reproduction rate (NRR) is the same across educational group; i.e. $\mathrm{NRR}_{z,e} = \mathrm{NRR}_z$ for all $e \in \mathbf{E}$. The total number of daughters born from the birth cohort $z$, or net reproduction rate, is

$$\mathrm{NRR}_z = \sum\nolimits_{a=0}^{\Omega} \left[\Pi_{s=0}^{a}\pi_{z,s}\right] \mathrm{fer}_{z,a} f_{\mathrm{fab}}, \tag{S.19}$$

where $\mathrm{fer}_{z,a}$ is the age-specific fertility rates for the cohort $z$ and $f_{\mathrm{fab}}$ is the fraction of females at birth. Let us now consider that the birth cohort is comprised of individual with different educational attainment. Thus, we can rewrite the previous equation as

$$\mathrm{NRR}_z = \sum\nolimits_{e\in\mathbf{E}} \frac{N_{z,e}}{N_z} \underbrace{\left(\sum\nolimits_{a=0}^{\Omega} \left[\Pi_{s=0}^{a}\pi_{z,s,e}\right] \mathrm{fer}_{z,a,e} f_{\mathrm{fab}}\right)}_{\mathrm{NRR}_{z,e}}, \tag{S.20}$$

where $\mathrm{fer}_{z,a,e}$ is the age-specific fertility rate for agents that belong to cohort $z$ with education $e$.

To minimize the change in the age distribution of the population caused by the introduction of heterogeneity by education, we assume that fertility profiles across the different education groups are given by $\mathrm{fer}_{z,a,e} = \kappa_e^f \mathrm{fer}_{z,a}$, where $\kappa_e^f$ is calculated as

$$\kappa_e^f = \mathrm{NRR}_z \Big/ \left(\sum\nolimits_{a=0}^{\Omega} \left[\Pi_{s=0}^{a}\pi_{z,s,e}\right] \mathrm{fer}_{z,a} f_{\mathrm{fab}}\right). \tag{S.21}$$

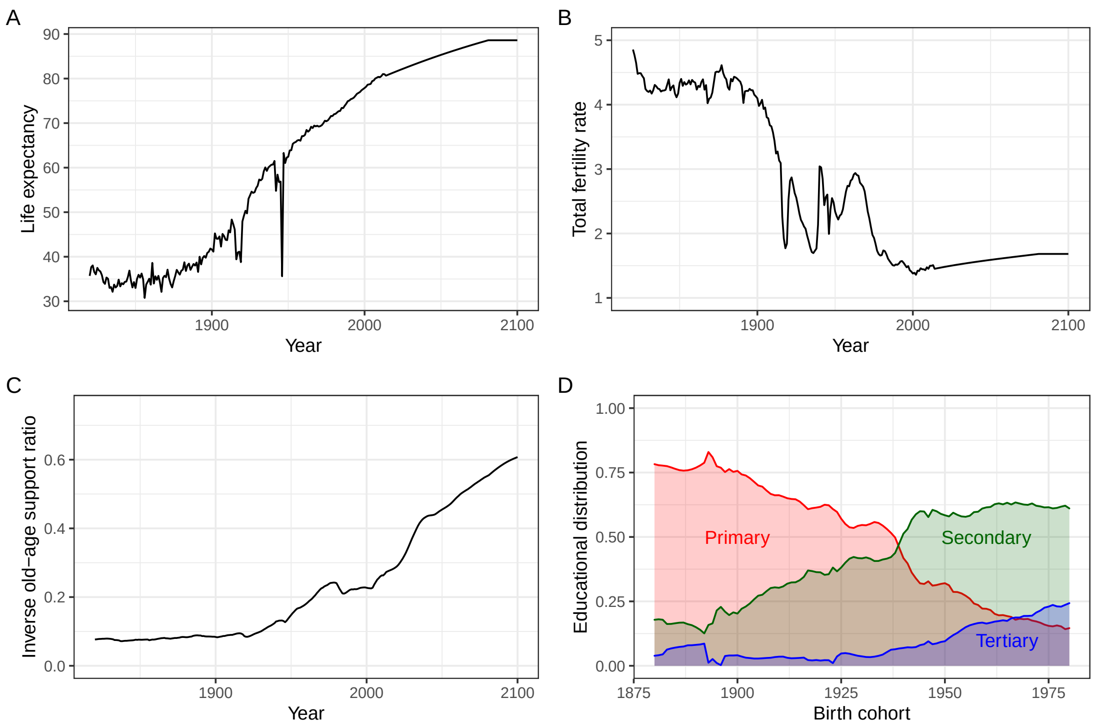



Figure S11: Austrian demographics, 1880–2100. *Source*: . *Notes*: Panel A shows the life expectancy at birth, Panel B shows the total fertility rate, Panel C shows the inverse of the old-age support ratio, and Panel D is the educational distribution by birth cohort. The inverse of the old-age support rate is the ratio of the population aged 65+ to the economically active population (ages 15-64).

*S5. Parametric components of the Austrian Pension System*

The model has been designed to reflect at each age the average pension points accumulated and the average pension benefit of the members of a cohort and not of a single individual. In what follows we explain how we have introduced in the model the cohort perspective.

Following Sánchez-Romero et al. (2020) the contribution period of the Austrian pension system can be summarized through a pension point system according to the following equation

$$\mathrm{pp}_{a+1} = \left[\alpha_J(l_a) + (1-\alpha_J(l_a))\mathcal{R}_a\right]\mathrm{pp}_a + \phi^p \mathrm{PBI}(y_{a,e}), \tag{S.22}$$

where $\mathrm{pp}_a$ are the pension points, $\alpha_J(l_a)$ is the proportion of individuals retired at age $a$ within the cohort, $\mathcal{R}_a$ is the capitalization factor, $\phi^p$ is the conversion factor of wages to pension points or accrual rate, and $\mathrm{PBI}(y)$ is the increase of the pension base. The accrual rate $\phi^p$ is defined as the inverse of the pensionable income years $n$, i.e. $\phi^p = 1/n$. The second column in Tab. S3 shows the evolution of the pensionable income years across cohort.

The pension benefit at age $a \geq \underline{J}$ of the Austrian pension system can be summarized through the following equation

$$b_a = \max\left\{\lambda_a \cdot \varphi \cdot \mathrm{pp}_a, b^{\min}\right\}. \tag{S.23}$$

where $b_a$ is the pension benefit, $\lambda_a$ is the adjustment factor of the replacement rate, $\varphi$ is the full pension replacement rate, $\mathrm{pp}_a$ are the pension points, and $b_{\min}$ is the minimum pension benefit. The replacement rate $\varphi$ of the Austrian pension system is 80% when retiring at the normal retirement age $J^N$ and after having contributed $\mathrm{wy}_z$ years (see the evolution of $\mathrm{wy}_z$ in Tab. S3); otherwise, the replacement rate is adjusted according to $\lambda_a$. This adjustment factor takes into account the average number of years contributing to the system and the average retirement age of the cohort. Since not all individuals of the cohort retire simultaneously, we write $\lambda_a$ recursively as follows[12]

$$\lambda_a = \frac{\bar{L} - l_{a-1}}{\bar{L} - l_a}\left(\mathbf{f}_{a-1} + \lambda_a^{\mathrm{yc}}\lambda_a^{\mathrm{ra}}\frac{l_{a-1} - l_a}{\bar{L} - l_{a-1}}\right) \text{ with } \lambda_{\underline{J}} = \lambda_{\underline{J}}^{\mathrm{yc}}\lambda_{\underline{J}}^{\mathrm{ra}}. \tag{S.24}$$

The constant term $\bar{L}$ is set at 0.40, which is the fraction of time devoted by an individual who works full time, i.e. $\bar{L} = \frac{(52-5-3)\times 40}{52\times 7\times(24-12)}$. We calculate the penalty factor of the pension benefit

[12] Notice from Eq. (S.24) that the average replacement rate adjustment factor is the sum across age of the replacement rate adjustment factor applied to individuals retiring at age $a$ times the fraction of individuals retired at age a, i.e. $\lambda_a = \sum_{i=\underline{J}}^{a}\left(\lambda_i^{\mathrm{yc}}\lambda_i^{\mathrm{ra}}\right)\frac{l_{i-1}-l_i}{\bar{L}-l_a}$ with $l_{\underline{J}-1} = \bar{L}$.

for not having contributed enough years before retiring as

$$\lambda_a^{\mathrm{yc}} = \left[\mathrm{ly}_{\underline{J}} + (a - \underline{J})\right] / \mathrm{wy}. \tag{S.25}$$

where $\mathrm{ly}_a$ is the average number of years worked until the minimum retirement age is reached

$$\mathrm{ly}_{a+1} = \mathrm{ly}_a + \left(l_a / \bar{L}\right) \;\text{ for } \underline{a} < a \leq \underline{J} \tag{S.26}$$

Therefore, Eq. (S.25) implies that an additional year of work after the minimum retirement age increases the pension benefit by 1/wy or 2.2%. We also adjust the pension pension formula by taking into account the penalties and rewards for early and late retirement, respectively,

$$\lambda_a^{\mathrm{ra}} = \begin{cases} 1 + \mathrm{pen} \cdot (a - J^N) & \text{if } \underline{J} \leq a \leq J^N, \\ 1 + \mathrm{rew} \cdot (a - J^N) & \text{if } J^N < a \leq \overline{J}. \end{cases} \tag{S.27}$$

Both the penalty rate (pen) and the reward rate (rew), introduced in a sequence of pension proposals that started in the early 2000s, are

$$\mathrm{pen}_t = \begin{cases} 0 & \text{for } t < 2000, \\ 0.042 & \text{for } 2000 \leq t < 2013, \\ 0.051 & \text{for } t \geq 2013, \end{cases} \qquad \mathrm{rew}_t = \begin{cases} 0 & \text{for } t < 2000, \\ 0.042 & \text{for } t \geq 2000. \end{cases} \tag{S.28}$$

*S6. Pension proposals*

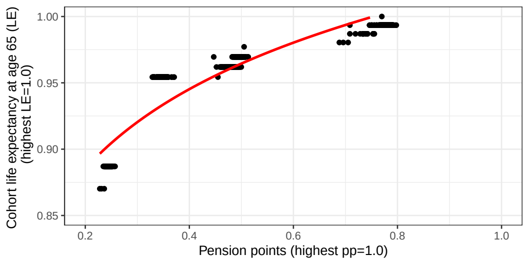
*Life expectancy by pension point level.* The two pension proposals by ABH and SRP correct the pension replacement rate for differences in life expectancy at retirement. However, we assume, yet realistically, that the social security system has no information on the life expectancy of each agent. To make this calculation we consider that the social security system uses the information on the number of pension points, which is known by the social security, to estimate the average remaining years of life at age 65 for all agents belonging to the same pension points quintile. Note that the number of pension points is a good proxy for lifetime labor income, which is frequently used to calculate the difference in life expectancy by SES (see, for instance, Chetty et al., 2016; Holzmann et al., 2019). Thus, we follow the literature and regress the relative average remaining years of life at age 65 to the logarithm of the relative number of pension points

$$le_{ij} = a + b\log(p_{ij}) + u_{ij}, \tag{S.29}$$

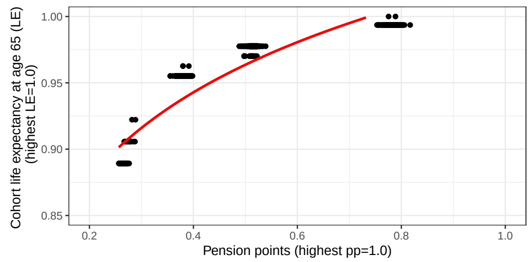
where $le_{ij} \in [0,1]$ is the relative remaining year of life at age 65 in quintile $i$ and model $j$ with respect to the highest life expectancy at age 65 in model $j$ (i.e., $le_{ij} = \mathrm{LE}_{65,ij}/\max(\mathrm{LE}_{65,j})$), $p_{ij} \in [0,1]$ is the relative number of pension points at age 65 in quintile $i$ in model $j$ with respect to the maximum number of pension points at age 65 in model $j$ (i.e., $p_{ij} = \mathrm{pp}_{65,ij}/\max(\mathrm{pp}_{65,j})$) and $u_{ij}$ is the error term.

Table S4 shows the estimated parameters ($\hat{a}$, $\hat{b}$) for a group of selected cohorts (1980, 2000, 2020, and those living in the final steady-state $\infty$). We obtain that an increase of 1% in the relative number of pension points is associated with an increase between 7.33% (cohort 2020) and 9.33% (cohort 2000) in the remaining years of life at age 65 relative to the highest life expectancy at age 65. Since the paper focuses on the impact of both proposals on cohorts 1980–2020, we use the intermediate parameters values of the 1980 birth cohort ($\hat{a} = 1.0244$, $\hat{b} = 8.64\%$) to calculate $\widehat{le}_{ij}$ for each representative agent.

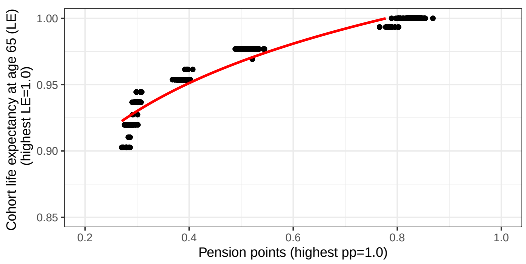
The difference in the estimated value of $\hat{b}$ across cohorts reflects the variance of the educational distribution of each cohort. In particular, the smaller is the variance of the educational distribution of a cohort, the smaller is the difference in life expectancy across groups and hence the smaller is the value of $\hat{b}$. Figure S12 shows the relationship between the relative remaining years of life at age 65 and the relative number of pension points at age 65. The red line represents the fit of model (S.29) to the simulated data, where the value of $\hat{b}$ is the slope of the red curve.



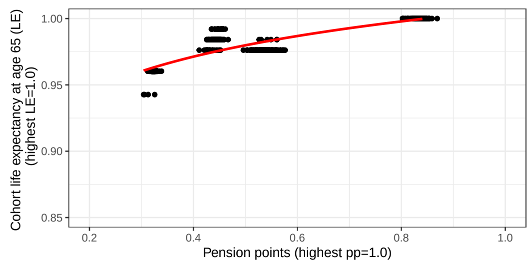

(a) Cohort 1980




(b) Cohort 2000




(c) Cohort 2020




(d) Cohort 2150

Figure S12: Relationship between the cohort life expectancy and the relative number of pension points at age 65. *Source*: Authors' calculations using the results of the benchmark model.

*Pension replacement rate progressivity.* Both pension proposals imply that the pension replacement rate varies according to the life expectancy. Hence, we have from (S.29) that under the two pension proposals the pension replacement rate becomes a function of the number of pension points accumulated. Substituting (S.29) in (24) the penalties and rewards by life expectancy in the model of ABH is given by

$$\frac{\widehat{\mathrm{LE}}(\mathbf{pp})-\widehat{\mathrm{LE}}(\overline{\mathbf{pp}})}{\widehat{\mathrm{LE}}(\mathbf{pp})}=\frac{\frac{\widehat{\mathrm{LE}}(\mathbf{pp})}{\widehat{\mathrm{LE}}(\mathbf{pp}^{\max})}-\frac{\widehat{\mathrm{LE}}(\overline{\mathbf{pp}})}{\widehat{\mathrm{LE}}(\mathbf{pp}^{\max})}}{\frac{\widehat{\mathrm{LE}}(\mathbf{pp})}{\widehat{\mathrm{LE}}(\mathbf{pp}^{\max})}}=\frac{\hat{b}[\log(\overline{p})-\log(p)]}{\hat{a}+\hat{b}\log(p)}, \tag{S.30}$$

while plugging (S.29) in (25) the penalties and rewards by pension point in the model of SRP is given by

$$\underbrace{\left[\left(1-\frac{\widehat{\mathrm{LE}}(\mathbf{pp}^{\min})}{\widehat{\mathrm{LE}}(\mathbf{pp}^{\max})}\right)\Big/\left(1-\frac{\mathbf{pp}^{\min}}{\mathbf{pp}^{\max}}\right)\right]}_{=\phi^{\mathrm{SRP}}}\left(\frac{\frac{\mathbf{pp}}{\mathbf{pp}^{\max}}-\frac{\overline{\mathbf{pp}}}{\mathbf{pp}^{\max}}}{\frac{\mathbf{pp}}{\mathbf{pp}^{\max}}}\right)=\phi^{\mathrm{SRP}}\frac{p-\overline{p}}{p}, \tag{S.31}$$

where $\overline{p}$ is the average relative number of pension points and $\phi^{\mathrm{SRP}}$ is the degree of progressivity (with $\phi^{\mathrm{SRP}}=\frac{1-\hat{a}-\hat{b}\log(p_{\min})}{1-p_{\min}}\approx 0.1348$). Figure S13 shows the increase and the reduction in the pension replacement rates across agents with different pension points that results from applying the pension proposal of ABH (see red dots) and SRP (see blue diamonds). Notice that under both pension proposals fig. S13 shows that the agents with pension points below the average pension points have a higher replacement rate, while agents with pension points above the average pension points have a lower replacement rate. It is also important to notice that the penalties and rewards are more pronounced in the proposal proposed by SRP than in that proposed by ABH. This is because the proposal of SRP not only compensates for the difference in life expectancy as in ABH, but also through the difference in pension points.

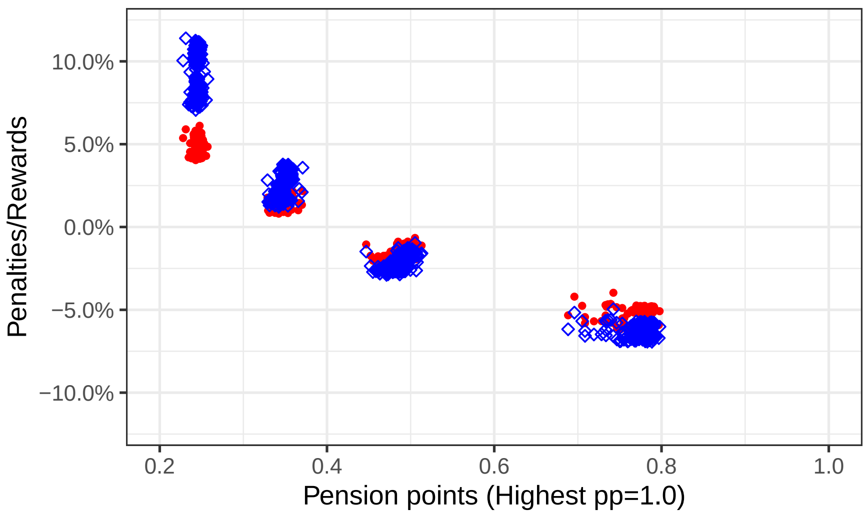
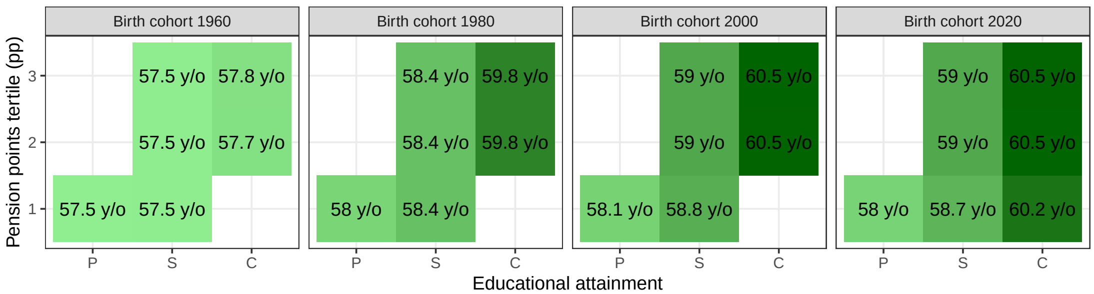
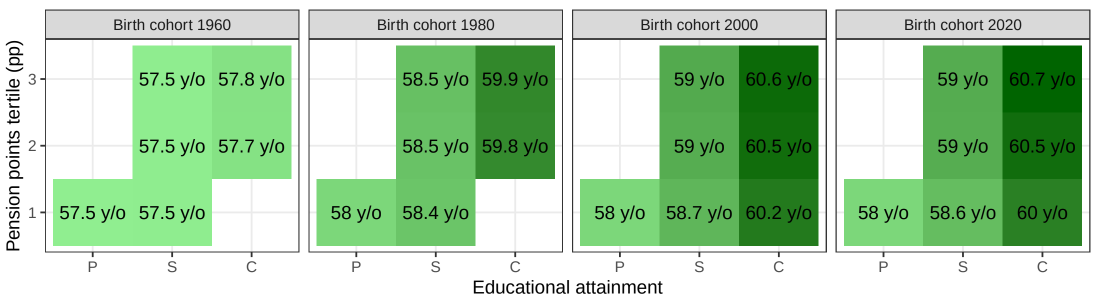


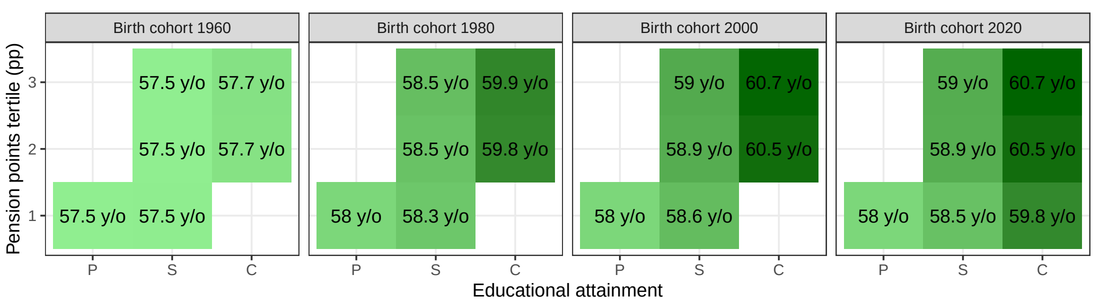

Figure S13: Penalties and rewards by relative number of pension points and policy proposal. *Source*: Authors' calculations using the model results for the cohort born in the final steady-state. *Notes:* Blue diamonds correspond to the proposal of Sánchez-Romero and Prskawetz (2020) and red dots correspond to the proposal of Ayuso et al. (2017).




(a) Benchmark




(b) Experiment 1: Ayuso et al. (2017)




(c) Experiment 2: Sánchez-Romero and Prskawetz (2020)

Figure S14: Average retirement age by pension points tertile (1=low, 2=middle, 3=high) and educational attainment (P=primary, S=secondary, C=college): selected birth cohorts under three different pension proposals. *Source*: Authors' calculations.

Table S3: Parametric components of the Austrian pension system by birth cohort

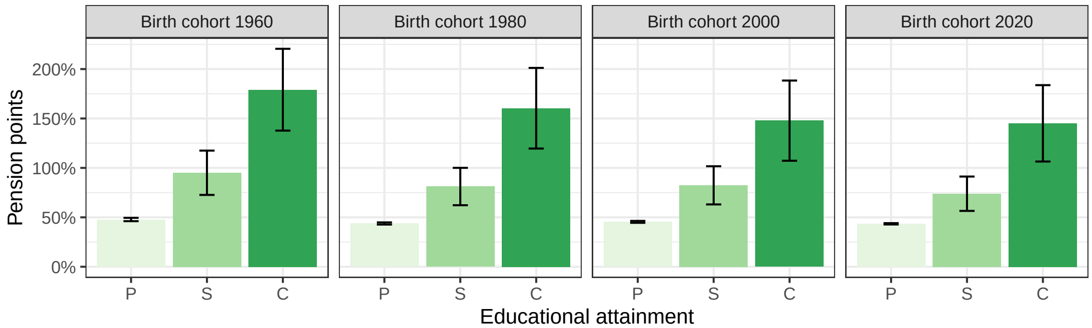
| Birth cohort $z$ | Pensionable income years $n_z$ | Working years $\mathrm{wy}_z$ | Early retirement $\underline{J}_z$ | Normal retirement $J_z^N$ | Late retirement $\overline{J}_z$ | Replacement rate $\varphi_z$ |
|---|---|---|---|---|---|---|
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 1875 | 5 | 45 | 57.0 | 63.0 | 68 | 0.00 |
| 1880 | 5 | 45 | 57.0 | 63.0 | 68 | 0.01 |
| 1885 | 5 | 45 | 57.0 | 63.0 | 68 | 0.06 |
| 1890 | 5 | 45 | 57.0 | 63.0 | 68 | 0.40 |
| 1895 | 5 | 45 | 57.0 | 63.0 | 68 | 0.74 |
| 1900 | 5 | 45 | 57.0 | 63.0 | 68 | 0.79 |
| 1905 | 5 | 45 | 57.0 | 63.0 | 68 | 0.80 |
| 1910 | 5 | 45 | 57.0 | 63.0 | 68 | 0.80 |
| 1915 | 5 | 45 | 57.0 | 63.0 | 68 | 0.80 |
| 1920 | 5 | 45 | 57.0 | 63.0 | 68 | 0.80 |
| 1925 | 5 | 45 | 57.0 | 63.0 | 68 | 0.80 |
| 1930 | 12 | 45 | 57.0 | 63.0 | 68 | 0.80 |
| 1935 | 15 | 45 | 57.0 | 63.0 | 68 | 0.80 |
| 1940 | 15 | 45 | 57.0 | 63.0 | 68 | 0.80 |
| 1945 | 15 | 45 | 59.0 | 63.0 | 68 | 0.80 |
| 1950 | 15 | 45 | 60.4 | 63.0 | 68 | 0.80 |
| 1955 | 15 | 45 | 60.9 | 63.0 | 68 | 0.80 |
| 1960 | 20 | 45 | 61.4 | 63.0 | 68 | 0.80 |
| 1965 | 25 | 45 | 61.9 | 63.8 | 68 | 0.80 |
| 1970 | 30 | 45 | 62.0 | 65.0 | 68 | 0.80 |
| 1975 | 35 | 45 | 62.0 | 65.0 | 68 | 0.80 |
| 1980 | 40 | 45 | 62.0 | 65.0 | 68 | 0.80 |
| 1985 | 45 | 45 | 62.0 | 65.0 | 68 | 0.80 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

*Notes*: Men and women combined.

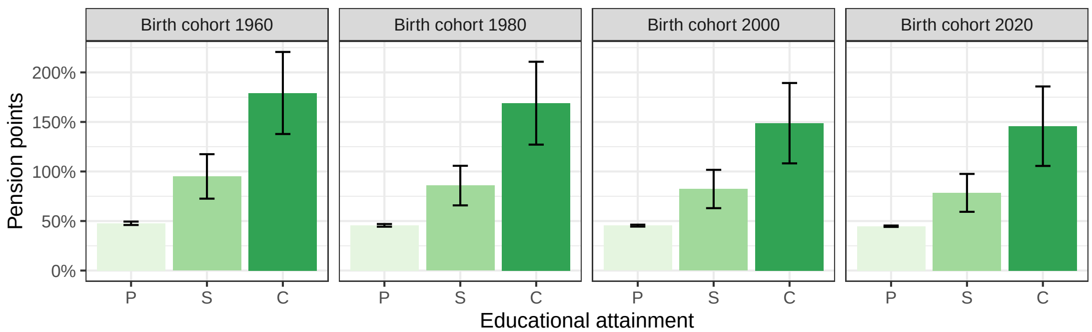
Table S4: OLS regression of the relative remaining years of life at age 65 by the relative number of pension points

| Cohort: | | 1980 | 2000 | 2020 | $\infty$ |
|---|---|---|---|---|---|
| (Intercept) | $\hat{a}$ | 1.0244*** | 1.0283*** | 1.0183*** | 1.0064*** |
| | | (0.0005) | (0.0005) | (0.0002) | (0.0002) |
| $\log p$ | $\hat{b}$ | 0.0864*** | 0.0933*** | 0.0733*** | 0.0384*** |
| | | (0.0005) | (0.0006) | (0.0002) | (0.0003) |
| $R^2$ | | 0.8753 | 0.8773 | 0.9567 | 0.8402 |
| Adj. $R^2$ | | 0.8752 | 0.8773 | 0.9567 | 0.8401 |
| Num. obs. | | 4000 | 4000 | 4000 | 4000 |

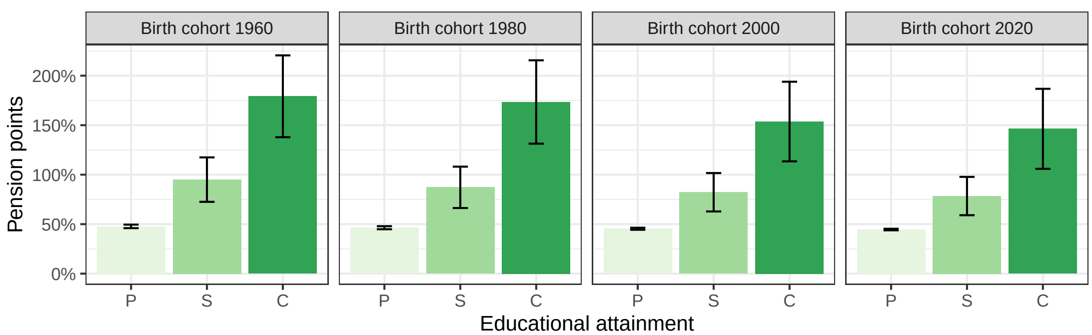
$^{***}p < 0.001, ^{**}p < 0.01, ^{*}p < 0.05$




(a) Benchmark

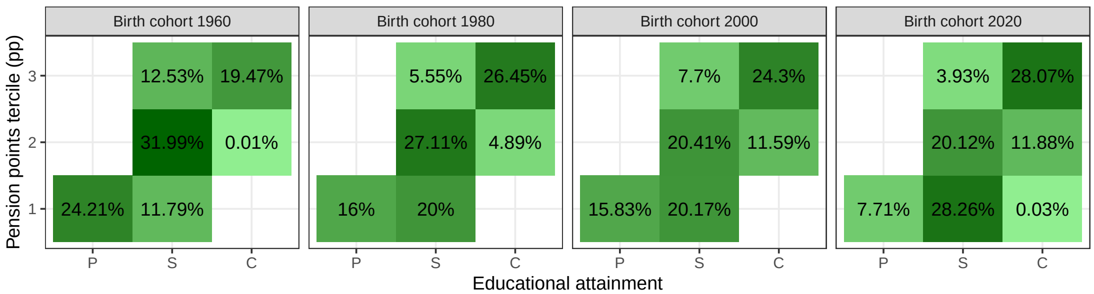



(b) Experiment 1: Ayuso et al. (2017)

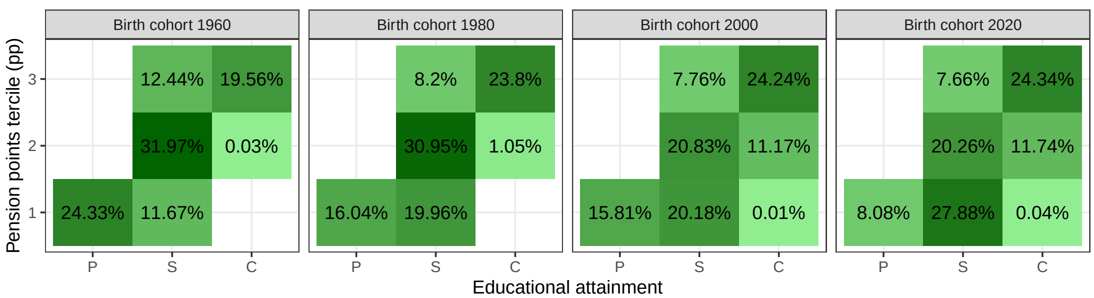



(c) Experiment 2: Sánchez-Romero and Prskawetz (2020)

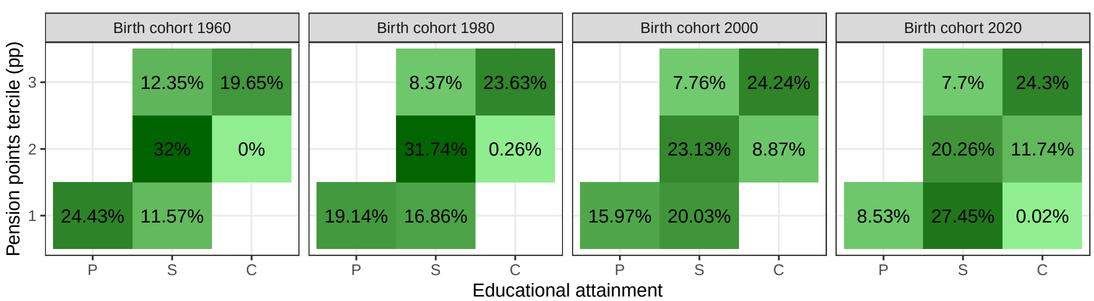
Figure S15: Average pension points relative to the mean by educational attainment (P=primary, S=secondary, C=college): selected birth cohorts under three different pension proposals. *Source*: Authors' calculations.




(a) Benchmark




(b) Experiment 1: Ayuso et al. (2017)




(c) Experiment 2: Sánchez-Romero and Prskawetz (2020)

Figure S16: Distribution of agent across pension points terciles and educational attainment (P=primary, S=secondary, C=college): selected birth cohorts under three different pension proposals. *Source*: Authors' calculations.