Journal of Monetary Economics 134 (2023) 35–52

Contents lists available at ScienceDirect

Journal of Monetary Economics

journal homepage: www.elsevier.com/locate/jmoneco

Rising earnings inequality and optimal income tax and social security policies☆

Pavel Brendler

University of Bonn, Institute for macroeconomics, Adenaueralle 24–42, Bonn 53113, Germany

ARTICLE INFO

Article history:
Received 19 March 2021
Revised 25 October 2022
Accepted 28 October 2022
Available online 31 October 2022

JEL classification:
D3
E6
H2
H3

Keywords:
Optimal tax
Progressive taxation
Social security
Earnings inequality
Aging population

ABSTRACT

How did the US government preferences over income redistribution across generations and within generations change during 1980–2010? Using a rich quantitative model in which a Ramsey government chooses income taxation and Social Security, I decompose the total change in the actual policies into the impact of new economic and demographic conditions and government preferences. I find that the US government preferences have shifted toward more educated and older households since the 1980s. Preferences over income redistribution within and across generations interact and, therefore, must be analyzed jointly.

1. Introduction

The recent macroeconomic literature has extensively studied the optimal *income* tax-and-transfer system, largely ignoring the role played by Social Security, a publicly provided pension system in the US. Since the two programs redistribute incomes both across generations and within generations, their joint analysis allows to study how government preferences over income redistribution have changed over time, which is the paper's main focus.

One of the paper's key innovations is to relax the commonly used assumption in the optimal income taxation literature that the Ramsey government maximizes the welfare of *newborns*. Instead, the policymaker cares about *all* agents who are alive at the time when the policy is implemented. This departure is crucial because newborn workers in the calibrated model economy prefer to shut down the public pension system altogether, regardless of their characteristics, so the model is unable to rationalize why Social Security and the income tax-and-transfer program coexist. The second significant deviation from

☆ I am particularly thankful to Yongsung Chang and an anonymous referee for very useful comments and suggestions. I benefited a lot from insightful discussions with Pawel Doligalski, Emanuel Hansen, Thomas Hintermaier, Byeongju Jeong, Moritz Kuhn, and Benjamin Larin. I also received very valuable feedback from Arpad Abraham, Marco Bassetto, Gregor Boehl, Klaus Gruendler, Fatih Guvenen, Jonathan Heathcote, Urban Jermann, Joachim Jungherr, Sagiri Kitao, Winfried Koeniger, Tim Krieger, Dirk Krueger, Patrick Macnamara, Alisdair McKay, Christian Moser, Roberto Pancrazi, Facundo Piguillem, Dominik Sachs, Almuth Scholl, Ctirad Slavik, Kjetil Storesletten, Arnau Valladares-Esteban, Chunzan Wu as well as the seminar participants at several seminars and conferences. I thank Oliver Freyermuth at the Physics Department of the University of Bonn for his excellent IT support. As always, all mistakes are mine.

E-mail address: pavel.brendler@uni-bonn.de

https://doi.org/10.1016/j.jmoneco.2022.10.004
0304-3932/

the existing studies is the introduction of age- and education-specific Pareto weights, which allow the model to account for the actual sizes of both programs in the data. By quantifying the change in the joint distribution of Pareto weights by age and education during 1980–2010, this paper is able to analyze the evolution of the US government preferences over income redistribution across generations *and* within generations, which is the paper's main contribution.

The government has access to two policy instruments in the model: the income tax progressivity and the average replacement rate in the pension system. The government sets both policies jointly, internalizing how the policy will affect all currently alive agents along the transition. Workers accumulate skills and make retirement decisions, so both income taxes and payroll contributions exert a distortionary effect on their decisions. In line with the literature, the policy reform is unanticipated and permanent.

Through the lens of the model, two distinct forces are responsible for the change in the observed policies during 1980–2010. The government might have found it necessary to adjust both policies because it faced new economic and demographic conditions in the 2010s compared to the 1980s. To quantify this force, I compute the optimal response to the new conditions, applying the identified Pareto weight distribution from the 1980s. Then, by construction, the residual difference between the actual policy in the 2010s and the optimal response to the economic and demographic changes must be attributed to the shift in the Pareto weight distribution, which is the second force central to this paper.

The main findings are as follows. First, when a *utilitarian* government who cares about *newborn* workers is restricted to set the income tax policy only, while the replacement rate is fixed at the calibrated level, the optimal progressivity in the 1980s almost coincides with the data, consistent with Heathcote et al. (2020). But when the policymaker is allowed to set both programs jointly, it prefers to shut down the pension system altogether. Hence, this approach fails to explain why income taxation and Social Security coexist in the data. By contrast, a utilitarian government who maximizes the welfare of *all* agents alive, chooses counterfactually high replacement rates and income tax progressivity below the empirical value. Therefore, to rationalize the data, the Pareto weight distribution in the 1980s must be skewed toward younger, uneducated workers.

Second, in response to the new economic and demographic conditions, both income tax progressivity and the average replacement rate rise relative to the 1980s. These results highlight the importance of the identified Pareto weight distribution, under which the optimal solution is computed. In response to inequality, mainly driven by the growing ex-ante heterogeneity among workers consistent with the empirical evidence by Guvenen et al. (2017), young, uneducated workers demand higher transfers. Since the government cares relatively more about the young, it makes income taxes more progressive. As the level of the income tax schedule declines through the government budget constraint, the policymaker is able to simultaneously increase the average replacement rate because earnings inequality during the working stage propagates into income inequality during retirement. As opposed to this optimal solution, actual income tax progressivity declined and the average replacement rate increased three times more between 1980 and 2010. To rationalize these developments, the model predicts that the Pareto weight distribution must shift toward more educated and older agents, indicating that the US government has become less willing to tolerate redistribution from educated people toward uneducated people and more willing to tolerate income redistribution from workers toward retirees.[1] Although the model is silent regarding the forces behind the identified shift in Pareto weights, empirical evidence on the trend in voter turnout rates corroborates the model predictions.

Finally, the paper uncovers an important channel, previously ignored in the literature. As the relative weight on educated agents increases over time, the government reduces income tax progressivity, similar to the mechanism in Heathcote and Tsujiyama (2021) and Wu (2021). However, the policymaker also raises the average replacement rate to better cater to the preferences of college graduates: due to the education-specific longevity gap, they receive pensions, on average, for a longer period than uneducated agents. Similarly, as the age distribution of weights shifts toward older agents, the replacement rates fall, as in Brendler (2020). Simultaneously, however, the government also reduces income tax progressivity. Hence, the government preferences over income redistribution within and between generations are interconnected and must be studied jointly, which has not been done before.

This paper relates to three strands in the literature. The first strand applies the inverse optimum approach to recover social preferences for income redistribution by looking at the progressivity of the income tax-and-transfer system in the US. Chang et al. (2018) and Heathcote and Tsujiyama (2021) study the progressivity of the tax-and-transfer system using cross-sectional data. Wu (2021) finds that a rising Pareto weight on high-ability households accounts for a significant portion of the drop in the income tax progressivity during 1980–2010. By contrast, Heathcote et al. (2020) find that a utilitarian solution rationalizes the observed income tax progressivity. While this literature endogenizes income taxation, it takes other policies, including Social Security, as given.[2] The second strand in the literature asks a *normative* question of how an income tax or a public pension system should optimally look like (e.g., Huggett and Parra (2010), Hosseini and Shourideh (2019), Moser and Olea de Souza e Silva (2019), Ndiaye (2020)). However, these studies primarily focus on the decentralization of the first-best policies. Finally, the third strand analyzes different retirement financing reforms. Among many others, Conesa and Krueger (1999), Nishiyama and Smetters (2007), Kitao (2014), and McGrattan and Prescott (2017) have made significant

[1] The rising Pareto weight on older and more educated households is obtained *conditional* on population aging and the increased share of college graduates in the data. This implies that aging and college participation alone are *insufficient* to rationalize the occurred change in the policies.

[2] In my previous work (Brendler, 2020), I recover social preferences over inter-generational income redistribution in the US by endogenizing Social Security, while taking income taxation as given.

progress in this field. These studies focus on the macroeconomic and welfare consequences of various policy reforms taking them as given, whereas my work attempts to understand how existing policies arise in the first place.

2. The model

This section lays out the quantitative model employed in the optimal policy analysis.[3]

2.1. Overview

The economy is populated by overlapping generations of agents. Each period a continuum of agents is born. Agents enter the economy as workers at age $j=1$ and continue working until they reach mandatory retirement age J^R. At age 1 and before any decision is made, agents draw an educational level z from the invariant distribution Π_z. Agents can be either college graduates ($z=H$) or non-college graduates ($z=L$), and their education remains constant throughout life. Agents live up to a maximum of $J>J^R$ periods, but may die earlier due to stochastic mortality. Denote by $\psi_{z,j}$ the probability that the type z agent survives up to age $j+1$, conditional on surviving up to age j. Population grows at an exogenous rate n. At any point in time, the total population size is normalized to 1.

The financial markets are incomplete in that there is no insurance against idiosyncratic mortality and labor productivity shocks. Agents enter without any assets but can self-insure against these shocks by accumulating shares in a representative firm and government bonds. Since both assets bear no risk and generate the same return, agents are indifferent between either asset, and their asset holdings are denoted by a. Borrowing is ruled out.

2.2. Households

Preferences are time-separable, with a constant discount factor β. Agents are endowed with one unit of productive time each period, which they allocate among three activities: leisure ℓ, learning s, and work l. The utility from consumption and leisure in each period is given by the function $u(c,\ell)$. Agents also derive warm-glow utility from bequeathing wealth a denoted by $\phi(a)$. Bequests are equally redistributed among all agents alive.

The worker's wage per unit of time worked is given by

$$w_{z,t}h_{j,z}\nu_z y_{j,z}. \tag{1}$$

The deterministic part of the agent's hourly wage is governed by the wage rate $w_{z,t}$ (determined in equilibrium), the skill level $h_{j,z}$, and the fixed effect ν_z, while the stochastic part is given by the idiosyncratic productivity shock $y_{j,z}$. Each component is education-specific.

Upon entering the model, workers draw their educational level z which comes with a fixed initial amount of skills denoted by $h_{1,z}$ and an immutable learning ability denoted by θ_z. During the working stage, the agent's skills evolve according to the following deterministic law of motion borrowed from Guvenen et al. (2014):

$$h_{j+1,z}=(1-\delta^h)h_{j,z}+\theta_z(h_{j,z}s)^{\gamma^h}. \tag{2}$$

$h_{j,z}s$ is the total effective time the agent spends acquiring new skills, $\gamma^h\in(0,1)$ is the speed at which the agent builds up human capital, and δ^h is a rate at which skills depreciate.

The logarithm of the fixed effect ν_z is a standard-normally distributed random variable with education-specific variance $\sigma^2_{\nu,z}$. The stochastic part of the wage is driven by the idiosyncratic productivity shock $y_{j,z}$ that follows an AR(1) process:

$$\log(y_{j,z})=\rho_z\log(y_{j-1,z})+\epsilon_{j,z} \text{ with } y_{1,z}=1 \text{ and } \epsilon_{j,z}\sim\mathcal{N}(0,\sigma^2_{\epsilon,z}). \tag{3}$$

$y_{j,z}$ follows a Markov chain with states $y\in\mathcal{Y}_{j,z}$ and transitions $\pi_{j,z}(y_{j+1,z}\mid y_{j,z})$.

2.3. Production technology

College graduate workers and non-college graduate workers are imperfectly substitutable in production, but workers of the same education are perfectly substitutable across different ages and skill levels $h_{j,z}$. Let $N_{z,t}$ denote the aggregate labor of type z at time t measured in efficiency units. Then, the total effective labor at time t is given by

$$N_t=\left(N^\rho_{L,t}+N^\rho_{H,t}\right)^{\frac{1}{\rho}}, \tag{4}$$

where $1/(1-\rho)$ is the elasticity of substitution between two worker types.

A representative firm produces the final output good according to the production function

$$Y_t=ZK_t^{\varpi}N_t{}^{1-\varpi}, \tag{5}$$

[3] In this section, all aggregate variables, prices, and policy variables are indexed by time. Individual variables are indexed only by those characteristics that appear necessary in a given context. However, it is understood that all individual variables also depend on time and the agent's characteristics.

where K_t denotes the aggregate capital stock, $\varpi \in (0, 1)$ measures the elasticity of output with respect to the input of capital services, and Z is a scaling factor. The output can be consumed or invested in capital. The firm rents capital and hires labor on competitive spot markets at prices $r_t + \delta$ and $w_{z,t}$, where r_t is the rental price of capital, δ – the capital depreciation rate, and $w_{z,t}$ – the wage per effective unit of type-z labor. The interest rate and the wage rate follow from the firm's profit maximization:

$$w_{z,t} = (1 - \varpi)Z(K_t/N_t)^{\varpi} (N_t/N_{z,t})^{1-\rho} \text{ for } z \in \{H, L\}, \tag{6}$$

$$r_t = \varpi Z(K_t/N_t)^{\varpi-1}. \tag{7}$$

The first condition implies that firms pay a wage premium $w_{H,t}/w_{L,t}$ to workers with college education. When $\rho = 1$, college graduates and non-college graduates are perfectly substitutable in the production process and, therefore, receive the same wage in equilibrium. As long as $\rho < 1$, both types of workers are imperfectly substitutable.

2.4. Government policies

The government administers a pay-as-you-go pension system, Medicare, and an income tax-and-transfer program. Each activity is described in detail below.

2.4.1. Social security

All workers pay a proportional Social Security tax $\tau_{SS,t}$ on their taxable earnings given by $\tilde{e}_{SS} = \min(e, cap_{SS})$, where cap_{SS} is a maximum taxable earnings threshold and e are the agent's total pre-tax earnings given by $e = w_{z,t} h v y l$. Workers must retire once they reach the normal retirement age J^R, but they become eligible to receive pension benefits already at age $J^E < J^R$ and, therefore, may choose to retire early.[4] Since the timing of retirement will affect the benefit amount, I introduce an individual state variable j^R that is equal to the age at which the agent has retired; for workers, the convention is to set $j^R = 0$.

During the working stage, the agent accumulates average taxable earnings denoted by $\bar{e}_j$. The worker's average lifetime earnings are equal to the simple average of her past Social Security-taxable earnings, i.e., $\bar{e}_{j+1} = [(j-1)\bar{e} + \tilde{e}_{SS}]/j$. Once the agent retires, $\bar{e}_j$ remains constant. All workers enter the model with no prior histories of earnings, i.e., $\bar{e}_1 = 0$.

During retirement, the agent receives a pension benefit denoted by $b = b(\bar{b}_t, j^R)$, which depends on two components: 1) the full pension amount $\bar{b}_t$ for which the agent would have qualified had she retired at the normal retirement age J^R and 2) the actual retirement age j^R. The specification of the penalty function $b(\cdot)$ is relegated to the calibration.

The full pension amount is determined by a replacement rate schedule $R_t(\cdot)$. The statutory schedule in the data is approximated by a simple parametric class

$$R_t(\bar{e}, j^R; \alpha_t) = \begin{cases} \alpha_t \cdot \left(\bar{e}/\tilde{E}_{SS,t-j+j^R}\right)^{\bar{\alpha}} & \text{if } \bar{e} \geq \bar{e}_{\min}, \\ \alpha_t \cdot \left(\bar{e}_{\min}/\tilde{E}_{SS,t-j+j^R}\right)^{\bar{\alpha}} & \text{otherwise,} \end{cases} \tag{8}$$

where $\alpha_t \in \mathcal{R}_+$ is a policy variable, $\bar{e}_{\min}$ is a fixed threshold and $\tilde{E}_{SS,t-j+j^R}$ are the economy-wide average taxable earnings at the time when the agent has entered retirement.[5] The policy variable α_t controls the level of the replacement rate schedule, while pension system progressivity, governed by parameter $\bar{\alpha} \in \mathcal{R}$, is assumed to be exogenous. Individual average lifetime earnings are normalized by the economy-wide average taxable earnings, $\tilde{E}_{SS,t}$, such that α_t is equal to the replacement rate of an agent whose average lifetime earnings are exactly the same as the economy-wide taxable earnings at the time when the agent has retired. The first line refers to retirees with $\bar{e}$ above the minimum threshold $\bar{e}_{\min}$, while the second line shows a constant replacement rate for the remaining retirees. Adding this latter case allows the model to fit more accurately the empirical schedule (Section 4). Fig. 1 plots the schedule for $\alpha_t = 37\%$ (solid line). As α_t increases, the schedule shifts upward, which is shown by a dashed line. Note that when $\alpha_t = 0\%$, Social Security is shut down.

Given the total amount of pension benefits denoted by B_t, the Social Security tax $\tau_{SS,t}$ adjusts to balance the government budget constraint[6]

$$\tau_{SS,t}\mu_{W,t}\tilde{E}_{SS,t} = B_t, \tag{9}$$

where $\mu_{W,t}$ is the population share of working agents and $\tilde{E}_{SS,t}$ – the economy-wide average taxable earnings, such that $\mu_{W,t}\tilde{E}_{SS,t}$ are the total taxable earnings in the economy.

[4] To save computational time, workers are not allowed to delay retirement beyond J^R.

[5] Given α_t, the full pension amount of an agent who retires at the normal retirement age J^R with average lifetime earnings $\bar{e}$ is given by $\bar{b}_t(\bar{e}; \alpha_t) = \bar{e} \times R_t(\bar{e}, J^R; \alpha_t)$.

[6] I solved a version of the model in which Social Security accumulates asset reserves, but found that this feature did not have any significant effect on the numerical results.

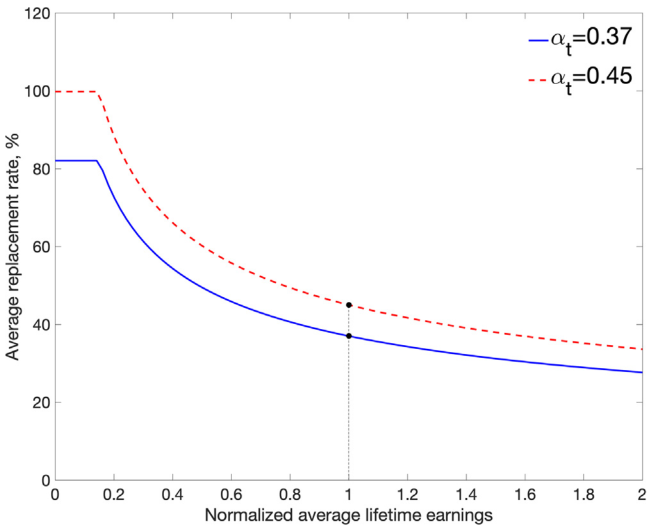

Fig. 1. Replacement rate schedule. Notes: The figure illustrates how the replacement rate schedule $R_t(\bar{e}, j^R; \alpha_t)$, specified in (8), depends on the policy variable α_t. Replacement rates are plotted as a function of the agent's average lifetime earnings normalized by the economy-wide taxable earnings at the time of retirement, $\bar{e}/\tilde{E}_{SS,t-j+j^R}$. The figure shows the effect of an increase in α_t from 37% to 45% ($\bar{\alpha}$ is fixed at -0.42). The minimum threshold $\bar{e}_{\min}$ is set to 0.15.

2.4.2. Medicare

All retired agents become eligible for Medicare benefits once they reach age $J^M > J^E$. This is also the age when retired agents start facing medical costs. Hence, the model implicitly assumes that expenses incurred during the working stage or during retirement prior to reaching age J^M are fully covered through employer-provided health insurance plans or other sources not present in the model. Agent's gross medical expenses are denoted by $\tilde{m}_j$, out of which a constant fraction $1 - \eta$ is covered by Medicare and the remaining share must be paid by the agent. Medical expenses are deterministic and age-dependent.[7]

Medicare consists of two subprograms: Hospital Insurance and Supplementary Medical Insurance. Hospital Insurance is financed by a linear hospital insurance tax τ_M on Medicare-taxable earnings given by $\tilde{e}_M = \min(cap_M, e)$ and runs a separate, balanced budget[8]

$$\tau_M \mu_{W,t} \tilde{E}_{M,t} = M_{A,t}, \tag{10}$$

where $M_{A,t}$ are the total medical expenses covered by Hospital Insurance, $\mu_{W,t}$ is the population share of workers, and $\tilde{E}_{M,t}$ are the average Medicare-taxable earnings. Supplementary Medical Insurance is part of the consolidated government budget, presented below, and is partially financed by a premium $p_{B,t}$ paid by all retirees eligible to receive Medicare.

2.4.3. Income tax-and-transfer program

Besides Social Security and Medicare, the government needs to finance an exogenous stream of wasted spending G_t and pay income transfers to working agents. The share of wasted spending in GDP is constant and given by $gy = G_t/Y_t$, where gy is a parameter. To finance these expenditures, the government issues debt $D_t > 0$, imposes a linear consumption tax τ_c, and collects income taxes. The initial stock of government debt D_0 is given, and the share of debt in output, $dy = D_t/Y_t$, remains constant at any point in time.

The income tax system discriminates between the income sources. More specifically, capital interest income, $r_t a$, is taxed at a proportional rate τ_a, and the after-tax return on saving is denoted by $\tilde{r}_t = (1 - \tau_a) r_t$. Labor income is taxed according to the income tax function Λ. Following Heathcote et al. (2017), the net tax liability is specified as

$$\Lambda(\tilde{\iota}; \tau_{I,t}, \bar{\tau}_{I,t}) = \tilde{\iota} - (1 - \bar{\tau}_{I,t}) \cdot (\tilde{\iota})^{1-\tau_{I,t}}, \tag{11}$$

where $\tilde{\iota} = \iota/\iota^m$ denotes the agent's taxable labor income ι normalized by the median taxable labor income ι^m.[9] The policy variables are $(\tau_{I,t}, \bar{\tau}_{I,t})$. The variable $\tau_{I,t}$ determines the progressivity of the income tax schedule. When $\tau_{I,t} > 0$, marginal income tax rates exceed average rates, and the income tax system is *progressive*. Conversely, the tax system is *regressive*

[7] Following Imrohoroglu and Kitao (2012), I exploit the Medical Expenditure Panel Survey and find no significant variation in medical expenses by income or education (Appendix C.5).

[8] The Hospital Insurance Trust Fund ran a deficit of 0.03% of GDP in the 1980s and 0.2% in the 2010s, which is why it is ignored in the model. See Table V.H5 in the 2022 Medicare Trustees Report accessible at https://www.cms.gov/files/document/2022-medicare-trustees-report.pdf.

[9] This normalization allows me to use the empirical estimate of $\bar{\tau}_{I,t}$ obtained by Wu (2021) who applies the same normalization when estimating $\bar{\tau}_{I,t}$ in the data.

when $\tau_{I,t} < 0$. The variable $\bar{\tau}_{I,t}$ is the tax rate of an agent whose taxable income is equal to the median taxable income at time t.

The agent's taxable labor income ι is given by

$$\iota = e - 0.5\tau_{SS,t}\tilde{e}_{SS} - 0.5\tau_M\tilde{e}_M, \tag{12}$$

where $e = w_{z,t}h\upsilon yl$ are the agent's pre-tax earnings and $(\tilde{e}_{SS}, \tilde{e}_M)$ are the agent's earnings taxable for the Social Security and Medicare purposes. Since the part of labor income that is paid by the employer as Social Security and Medicare contributions is not subject to income taxes, a corresponding portion is deducted from the worker's taxable labor income.

The income tax-and-transfer program is part of a consolidated government budget:

$$G_t + (1 + r_t)D_t + M_{B,t} = \mathcal{I}_t + P_{B,t} + \tau_a r_t A_t + \tau_c C_t + D_{t+1}, \tag{13}$$

where $\mathcal{I}_t$, A_t and C_t denote the aggregate net income taxes, the aggregate asset holdings, and the aggregate consumption, respectively. $M_{B,t}$ denotes the total medical expenses covered by Supplementary Medical Insurance and $P_{B,t}$ is the total amount of health insurance premium.

2.5. Dynamic programming problem of households

This subsection spells out the agent's dynamic programming problems at different stages of her lifecycle.

Problem of workers of age $1, \ldots, J^E - 2$

Workers become eligible for early retirement benefits once they reach age J^E. The retirement decision must be made one period in advance, which implies that agents must remain workers during the first $J^E - 2$ periods. During this stage, workers choose how much to consume, how much to save, and how to split the unitary time endowment between leisure, skill acquisition, and work. In each period, agents draw the labor productivity shock y before making these decisions. The agent's dynamic programming problem is given by

$$V_t(z, j, j^R, \upsilon, y, \bar{e}, h, a) = \max_{\substack{a' \geq 0, \\ l \in [0,1], \\ s \in [0,1]}} u(c, 1 - s - l) + \beta\psi_{z,j}\mathbb{E}_{y'|y}\left[V_{t+1}(z, j+1, j^{R\prime}, \upsilon, y', \bar{e}', h', a')\right] + (1 - \psi_{z,j})\phi(a') \tag{14}$$

subject to the budget constraint

$$a' + (1 + \tau_c)c = (1 + \tilde{r}_t)a + e - \tau_{SS,t}\tilde{e}_{SS} - \tau_M\tilde{e}_M - \Lambda(\tilde{\iota}) + T_t. \tag{15}$$

V_t denotes the agent's value function at time t. Since the agent must remain in the labor force in the next period, her retirement status is unchanged: $j^{R\prime} = j^R = 0$. If the agent dies (which occurs with probability $1 - \psi_{z,j}$), she receives an instantaneous warm-glow utility $\phi(a')$. Bequests are equally split among alive agents as a private transfer T_t.

Problem of workers of age $J^E - 1, \ldots, J^R - 1$

Once agents reach age $J^E - 1$, they may choose to retire in the next period. Before making the retirement decision, agents draw their labor productivity y. If the agent chooses to retire in the current period, her welfare is given by:

$$V_t^R(z, j, j^R, \upsilon, y, \bar{e}, h, a) = \max_{a' \geq 0, l \in [0,1]} u(c, 1 - l) + \beta\psi_{z,j}V_{t+1}(z, j+1, j^{R\prime}, \bar{e}', a') + (1 - \psi_{z,j})\phi(a')$$

subject to budget constraint (15). Since the agent retires in the next period, $j^{R\prime} = j + 1$. The stock of human capital h, the fixed effect υ, and the productivity shock y do not affect the household's decisions and welfare during retirement and, therefore, become redundant. For the same reason, the agent optimally chooses not to spend any time on learning, i.e., $s = 0$. I continue to keep track of the agent's education level z because it affects the survival probability rate $\psi_{z,j}$. If instead the agent chooses to continue working in the next period, her welfare is given by the right-hand side of (14). The agent decides whether to retire or continue working by comparing the welfare associated with each choice. If they choose to retire, $j^{R\prime} = j + 1$; otherwise, $j^{R\prime}$ continues to equal zero.

Problem of retirees of age $J^E, \ldots, J$

Retired agents devote their unitary time endowment to leisure. At age J^M, they start facing medical expenses. Their maximization problem reads

$$V_t(z, j, j^R, \bar{e}, a) = \max_{c, a' \geq 0} u(c, \ell = 1) + \beta\psi_{z,j}V_t(z, j+1, j^R, \bar{e}, a') + (1 - \psi_{z,j})\phi(a') \tag{16}$$

subject to the budget constraint

$$a' + (1 + \tau_c)c = (1 + \tilde{r}_t)a + b + T_t - \mathbb{1}_{j \geq J^M} m_j, \tag{17}$$

where m_j denotes out-of-pocket medical expenditures.[10] Observe that the retirement age indicator remains unchanged ($j^{R\prime} = j^R$).

[10] m_j is defined as the sum of the health insurance premium and the payment net of the part of the expenditure paid by Medicare.

2.6. Definition of equilibrium

Appendix A defines the competitive equilibrium. Appendix B describes the algorithm to solve the model numerically.

3. Quantitative experiment

Section 3.1 introduces the government's problem and Section 3.2 lays out the experiment.

3.1. Government's maximization problem

At time t, the economy is in a steady-state equilibrium with an initial policy denoted by $\boldsymbol{\Upsilon}_t^0 = (\tau_{l,t}^0, \alpha_t^0)$, where $\tau_{l,t}^0$ is the income tax progressivity and α_t^0 – the average replacement rate. Given the stationary distribution F_t implied by the initial policy, the government makes an unanticipated and credible announcement that it will adopt a constant future policy $\boldsymbol{\Upsilon}_t^\star$, which becomes effective the following period and solves

$$\boldsymbol{\Upsilon}_t^\star = \arg\max_{\boldsymbol{\Upsilon}_t} SWF(\boldsymbol{\Upsilon}_t, \boldsymbol{\Upsilon}_t^0; \boldsymbol{\kappa}_t), \tag{18}$$

where SWF denotes the social welfare function given by

$$SWF(\boldsymbol{\Upsilon}_t, \boldsymbol{\Upsilon}_t^0; \boldsymbol{\kappa}_t) = \sum_j \int \omega(z, j; \boldsymbol{\kappa}_t) V_t(z, j, j^R, \nu, y, \bar{e}, h, a; \boldsymbol{\Upsilon}_t, \boldsymbol{\Upsilon}_t^0) dF_{t,j}. \tag{19}$$

The term $\omega(j, z; \boldsymbol{\kappa}_t)$ denotes age- and education-specific Pareto weights that depend on parameter vector $\boldsymbol{\kappa}_t$. $V_t(.; \boldsymbol{\Upsilon}_t, \boldsymbol{\Upsilon}_t^0)$ is the agent's value function in the period of reform announcement on the economy's transition path from the stationary equilibrium associated with policy $\boldsymbol{\Upsilon}_t^0$ to a new stationary equilibrium associated with constant policy $\boldsymbol{\Upsilon}_t$. The new policy is implemented once-and-for-all. A change in the average replacement rate affects all agents in the economy, including those who have already retired at time t, which is consistent with the fact that there is no legal entitlement to receive a certain pension benefit in the US.

The Pareto weight function is specified as

$$\omega(j, z; \boldsymbol{\kappa}_t) = \exp(-\kappa_{1,t} \cdot j + \kappa_{2,t} \cdot \mathbb{1}_{z=H}), \tag{20}$$

where $\mathbb{1}_{z=H}$ is an indicator function that equals one for agents with college education. The parameter $\kappa_{1,t} \in \mathcal{R}$, referred to as the *age bias* below, determines the extent to which the government discriminates agents with the same educational background by age. When $\kappa_{1,t} = 0$, the policymaker trades off the utility of different age groups at the same rate. When $\kappa_{1,t} > 0$, the relative Pareto weight on newborns is above one, while the opposite is the case when $\kappa_{1,t} < 0$. Besides age, agents may be also discriminated by their education. The weight assigned to a college graduate relative to a non-college graduate of the same age is given by $\exp(\kappa_{2,t})$, where $\kappa_{2,t} \in \mathcal{R}$ is a parameter referred to as the *educational bias* below. When $\kappa_{2,t} = 0$, both educational groups are treated equally. For $\kappa_{2,t} > 0$, the government favors college graduates over non-college graduates; the opposite is the case when $\kappa_{2,t} < 0$.

3.2. Design of the quantitative experiment

Denote by $\boldsymbol{\Psi}_t$ with $t \in \{1980, 2010\}$ the set of all model parameters at time t, except for the calibrated income tax policy and the average replacement rate captured separately in $\tilde{\boldsymbol{\Upsilon}}_t = (\tilde{\tau}_{l,t}, \tilde{\alpha}_t)$. Furthermore, let $\boldsymbol{\Upsilon}_t^\star(\boldsymbol{\Psi}_t, \boldsymbol{\kappa}_t; \boldsymbol{\Upsilon}_t^0)$ denote the solution to the government's maximization problem in (18) when the government faces model parameters $\boldsymbol{\Psi}_t$, a set of Pareto weights induced by $\boldsymbol{\kappa}_t$, and initial policy $\boldsymbol{\Upsilon}_t^0$. The experiment proceeds in 3 steps.

In the first step, the Pareto weight parameter $\boldsymbol{\kappa}_{1980}$ is recovered by assuming that, given parameters $\boldsymbol{\Psi}_{1980}$, the calibrated policy $\tilde{\boldsymbol{\Upsilon}}_{1980}$ from the 1980s steady-state equilibrium, and (to be found) Pareto weight parameter $\boldsymbol{\kappa}_{1980}$, the government optimally chooses to remain in the status quo. In this case, $\boldsymbol{\kappa}_{1980}$ is implicitly defined by $\tilde{\boldsymbol{\Upsilon}}_{1980} = \boldsymbol{\Upsilon}_{1980}^\star(\boldsymbol{\Psi}_{1980}, \boldsymbol{\kappa}_{1980}; \tilde{\boldsymbol{\Upsilon}}_{1980})$.

In the second step, the economy is in an interim stationary equilibrium, in which the model parameters have been updated to $\boldsymbol{\Psi}_{2010}$ but the government has not yet responded to the new environment: the income tax progressivity and the replacement rate remain the same as in the 1980s and are given by $\tilde{\boldsymbol{\Upsilon}}_{1980}$. I ask how the government would optimally respond to the new conditions and denote the solution by $\boldsymbol{\Upsilon}_{\text{int}}^\star = \boldsymbol{\Upsilon}_{\text{int}}^\star(\boldsymbol{\Psi}_{2010}, \boldsymbol{\kappa}_{1980}; \tilde{\boldsymbol{\Upsilon}}_{1980})$. The difference $\boldsymbol{\Upsilon}_{\text{int}}^\star - \tilde{\boldsymbol{\Upsilon}}_{1980}$ quantifies the impact of the new conditions on the optimal policy.

In the final step, the economy is in a stationary equilibrium in the 2010s characterized by parameters $\boldsymbol{\Psi}_{2010}$ and optimal policy $\boldsymbol{\Upsilon}_{\text{int}}^\star$ computed in the previous step. The Pareto weight parameter $\boldsymbol{\kappa}_{2010}$ is identified such that the government chooses to adopt the actual policy from the 2010s, $\tilde{\boldsymbol{\Upsilon}}_{2010}$. In this case, $\boldsymbol{\kappa}_{2010}$ is implicitly defined by $\tilde{\boldsymbol{\Upsilon}}_{2010} = \boldsymbol{\Upsilon}_{\text{int}}^\star(\boldsymbol{\Psi}_{2010}, \boldsymbol{\kappa}_{2010}; \boldsymbol{\Upsilon}_{\text{int}}^\star)$. The difference $\boldsymbol{\kappa}_{2010} - \boldsymbol{\kappa}_{1980}$ quantifies the required shift in the government preferences necessary to rationalize the residual change in the policy $\tilde{\boldsymbol{\Upsilon}}_{2010} - \boldsymbol{\Upsilon}_{\text{int}}^\star$.

Table 1
Model parameters.

Parameter	Interpretation	Value in 1980s	Value in 2010s
Exogenously calibrated parameters			
Demographics			
J	Maximum life span (real-life age 100)	76	76
J^E	Early retirement age (real-life age 62)	38	38
J^R	Normal retirement age	41 (real-life age 65)	42
J^M	Medicare eligibility age (real-life age 65)	41	41
$\psi_{z,j}$	Education-specific age profile of survival probabilities	Estimates	
Π_z	Distribution of non-college and college graduates, %	(75,25)	(56,44)
n	Population growth rate, %	0.9	1.2
Preferences			
(β, σ)	Discount factor and coefficient of relative risk aversion	(0.997,2)	(0.997,2)
(ϕ_1, ϕ_2)	Bequest function	(−9.5, 11.6)	(−9.5, 11.6)
Labor productivity			
γ^h	Elasticity of human capital production	0.7	0.7
(ρ_H, ρ_L)	Persistence parameter	(0.989,1.0)	(0.993, 0.977)
$(\sigma^2_{\epsilon,H}, \sigma^2_{\epsilon,L})$	Variance of persistent shock	(0.016,0.006)	(0.015,0.012)
Production			
(ϖ, δ)	Capital share and capital depreciation, %	(43.0,8.0)	(46.0,6.0)
ρ	Elasticity of substitution is $1/(1-\rho)$	0.75	0.285
Government policies			
$\tilde{\alpha}$	Pension system progressivity	−0.42	−0.43
τ_l	Income tax progressivity	0.187	0.137
(τ_c, τ_a)	Consumption and capital income tax, %	(5.3,38.4)	(4.1,33.0)
dy	Debt-to-GDP ratio, %	35.0	100.0
m_j	Age profile of gross medical expenses	Estimates	
η	Share of out-of-pocket medical expenses, %	51.0	31.0
τ_M	Hospital Insurance tax, %	2.6	2.9
pm	Share of premia in Supplementary Insurance expenses, %	25.0	25.0
Parameters calibrated in equilibrium (targets in brackets)			
γ	Weight on consumption (average hours worked)	0.485	0.485
(θ_H, θ_L)	Learning ability (age profile of wages)	(0.097,0.096)	(0.138,0.104)
$(h_{1,H}, h_{1,L})$	Initial skill level (age profile of wages)	(0.77,0.77)	(1.35,0.60)
δ^h	Skill depreciation, % (age profile of wages)	0.50	0.49
$(\sigma^2_{\upsilon,H}, \sigma^2_{\upsilon,L})$	Variances of fixed effect (earnings Gini)	(0.0,0.009)	(0.018,0.025)
Z	Scaling factor in production (change in average wage)	0.758	0.323
α	Average replacement rate, % (Social Security tax)	36.0	39.4
$\bar{e}_{\min}$	Bend point (bend point-to-earnings ratio)	0.05	0.11
δ^p	Penalty for early retirement (share of retired at age 62)	0.07	0.14
gy	Wasted spending-to-GDP, % (income tax level $\bar{\tau}_l$)	16.6	12.3
(cap_{SS}, cap_M)	Taxable earnings threshold (share of workers above cap)	(0.62,0.62)	(1.32, ∞)

4. Calibration

The model parameters are calibrated to two time periods: 1980s and 2010s. In both periods, the economy is assumed to be in a steady-state equilibrium, so the time index t is dropped throughout this section. One model period equals one year. A large set of parameters is calibrated using the Current Population Survey (CPS). An agent in the model corresponds to a household head in the CPS. The agent's educational type z corresponds to the household head's educational level (college graduate or high-school graduate). Appendix C.1 describes the sample selection criteria. Table 1 shows the parameters calibrated outside the model and the parameters calibrated inside the model with empirical targets shown in brackets. Below I explain the calibration strategy in detail.

Demographics. The maximum possible age is $J = 76$ (real-life age 100). Consistent with the data, agents qualify for early retirement at age $J^E = 38$ (real-life age 62) and for Medicare at age $J^M = 41$ (real-life age 65). The values of (J, J^E, J^M) remain the same in both steady states. The normal retirement age, J^R, is equal to 41 (real-life age 65) in the 1980s and rises by one year in the 2010s. The share of college graduates, Π_H, is equal to 25% in the 1980s and 44% in the 2010s. The education- and age-specific mortality rates, $1 - \psi_{z,j}$, are estimated outside the model by matching two targets: 1) the life expectancy profile by age for an average worker and 2) the life expectancy profiles by age and education. According to the estimates, the life expectancy of a 25-year-old worker rose by 7 years during since 1980, while the life expectancy gap between the two educational groups at age 25 increased from 4 to 6 years. The birth rate, n, is estimated outside the model matching the dependency ratio, defined as the ratio of household heads aged 65–100 to household heads aged 25–64, of 22% in the 1980s and 27% in the 2010s. Appendix C.2 explains further details.

Preferences. The agent's utility u is a constant relative risk aversion function given by

$$u(c, \ell) = \frac{\left[c^{\gamma} \ell^{1-\gamma}\right]^{1-\sigma} - 1}{1-\sigma}, \tag{21}$$

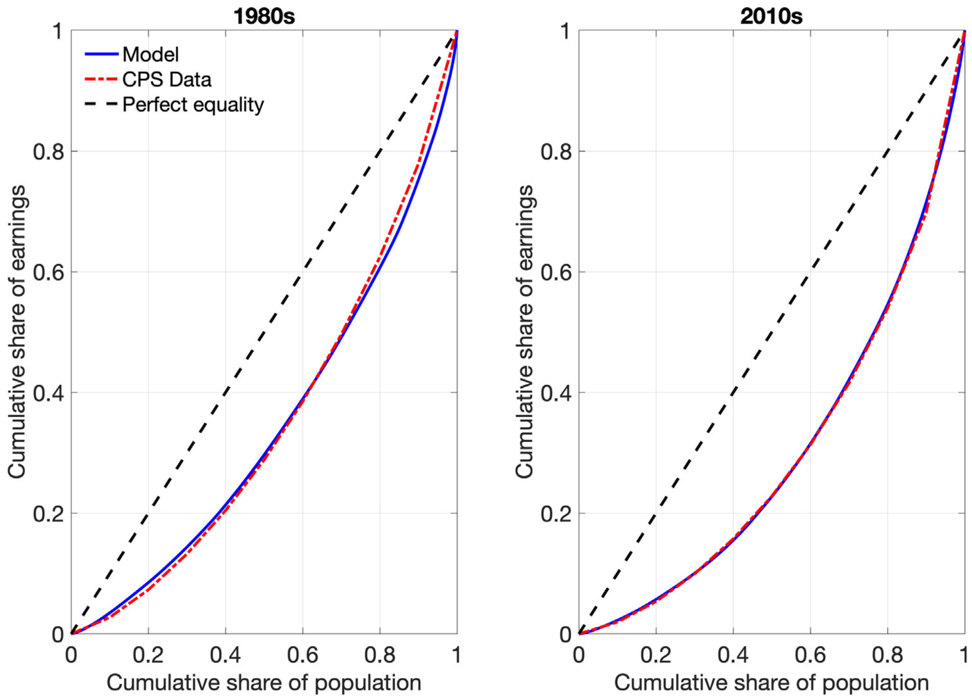

Fig. 2. Earnings inequality in the model and data. *Notes:* The figure shows how the model fits the Lorenz curves for the earnings distribution in the 1980s (left panel) and 2010s (right panel). Earnings are measured before taxes and government transfers. In the model, the Lorenz curve is constructed using the cross-sectional distribution of earnings in the respective steady state. The empirical curves are computed using the CPS data.

where σ controls the degree of relative risk aversion and γ is the relative weight on consumption. All preference parameters remain constant across the steady states.[11] In particular, σ is fixed at 2.0 and the discount factor β is set to 0.997. The bequest function is specified and parameterized as in De Nardi (2004) using $\phi(a) = \phi_1(1 + a/\phi_2)^{1-\sigma}$, where $\phi_1 = -9.5$ reflects the agent's concern about leaving bequest a, while $\phi_2 = 11.6$ measures the extent to which bequests are a luxury good. Given these values, the model generates the capital-to-GDP ratio, K/Y, equal to 3.1 in the 1980s and 3.8 in the 2010, which is fairly close to the empirical counterparts of 3.2 and 3.7, respectively.[12] Finally, γ is set to 0.485. Given this value, the model calibrated to the 2010s matches exactly the fraction of a unitary time endowment that workers devote to work (40.7%). For the 1980s calibration, the model predicts 41.5% compared to 42.4% in the CPS.

Human capital accumulation. The speed at which agents acquire skills, γ^h, is constant and equals 0.7 following Badel et al. (2020). The remaining parameters $(\theta_z, h_{1,z}, \delta^h)$ are calibrated inside the model by targeting the profiles of hourly wages by age and education in the CPS. The results for the 1980s imply no heterogeneity in the entry skill levels and abilities between the educational groups but the disparity becomes significant in the 2010s. Appendix C.3 shows that the model matches accurately the empirical hourly wage profiles.

Idiosyncratic risk. Following Storesletten et al. (2004), the parameters of the idiosyncratic productivity shock $y_{j,z}$ are estimated outside the model by fitting the profiles of residual wage dispersion by age and education to the data. According to the estimates, the variance of the shocks, $\sigma^2_{\epsilon,z}$, remains almost unchanged for college graduates and doubles for non-college graduates during 1980–2010. Appendix C.3 explains the procedure.

Fixed effect. The logarithm of the fixed effect υ_z is assumed to be a standard-normally distributed variable with an education-specific variance $\sigma^2_{\upsilon,z}$. The fixed effect is calibrated inside the model by matching the residual cross-sectional earnings inequality by education (measured by the Gini index), conditional on the calibrated differences in the persistent shock and skills. The obtained estimates imply that a substantial portion of the residual earnings inequality in the 2010s arises due to larger dispersion in the initial conditions at labor market entry, consistent with the recent evidence by Guvenen et al. (2017). Fig. 2 shows the fit of the Lorenz curves for pre-government earnings in each steady state.

[11] This assumption will allow me to attribute any change in the agent's welfare to the underlying change in the fundamental parameters that govern inequality, aging, etc. and not the preference parameters themselves.

[12] The share of capital stock in GDP is computed from the National Income and Product Accounts (NIPA) following the procedure in Hosseini and Shourideh (2019). See their online Supplement, Section S5.

Production. The capital share ϖ and the depreciation rate δ are set outside the model to match the average ratio of capital income in GDP and the average ratio of investment in GDP, respectively.[13] The scaling factor Z is calibrated inside the model such that the average hourly wage is equal to 1.0 in the 1980s; for the 2010s, Z is recalibrated to match the change in the average real hourly wage equal to 24% in line with the CPS. Following Heathcote et al. (2017) and Abbott et al. (2019), my theoretical framework attributes the rise in the wage premium since the late 1980s in the data to a stronger complementarity between college graduate and high-school graduate workers in production governed by parameter ρ. In line with Katz and Murphy (1992), ρ is set to 0.75 for the 1980s and 0.285 for the 2010s, which implies that the elasticity of substitution declines from 4.0 to 1.4. Given these values, the model matches the wage premium of 36% in the 1980s and 54% in the 2010s.

Social Security. I estimate the replacement rate level α and progressivity $\bar{\alpha}$ by fitting the replacement rate schedule in (8) to its empirical counterpart. Appendix C.4 explains the estimation and shows that the fit is very accurate. The obtained estimate of α is then adjusted inside the model, keeping $\bar{\alpha}$ fixed, such that the Social Security tax, τ_{SS}, matches its empirical counterpart. More specifically, α is reduced from 39.0% to 36.0% to match $\tau_{SS} = 8.9\%$ in the 1980s and from 42.0% to 39.4% to match $\tau_{SS} = 10.6\%$ in the 2010s. cap_{SS} is calibrated inside the model by matching the fraction of households in the CPS whose pre-tax earnings exceed the maximum taxable earnings threshold (13% in the 1980s and 8% in the 2010s). Finally, the penalty function for early retirement is specified as:

$$b(\bar{b}, j^R) = (1-\delta^p)\cdot \bar{b}(\bar{e};\alpha) + \left(\frac{j^R - J^E}{J^R - J^E}\right)\cdot \delta^p \cdot \bar{b}(\bar{e};\alpha). \tag{22}$$

The penalty factor δ^p is estimated inside the model targeting the share of retired households at age 62 in the CPS (39% in the 1980s and 26% in the 2010s).

Medicare. Following Imrohoroglu and Kitao (2012), I exploit the Medical Expenditure Panel Survey (MEPS) to obtain the age profile of gross medical expenses m_j and the share of out-of-pocket medical expenses η. Appendix C.5 describes the data and explains the estimation. According to the results, η drops from 51% in the old calibration to 31% in the new calibration, while real medical expenditures faced during retirement rise, on average, by 43%. The Hospital Insurance tax, τ_M, is equal to 2.6% in the 1980s and 2.9% in the 2010s. cap_M and cap_{SS} are equalized in the initial calibration, in line with the data. Since there has been no limitation on taxable earnings in the Medicare program starting from 1994, cap_M is set to a large number in the 2010s, which practically eliminates the cap. The individual premium p_B is pinned down by assuming that the share of aggregate premium income P_B in the aggregate Supplementary Medical Insurance expenses M_B is held constant at $pm = P_B/M_B$, where pm is calibrated at 25% in both steady states.[14]

Income tax-and-transfer program. The estimates of the income tax progressivity τ_I are borrowed from Wu (2021) who utilizes very similar CPS samples and applies the same definition of taxable income. According to his results, income tax progressivity declines from 0.187 in 1978–1980 to 0.137 in 2014–2016.[15] The share of wasted spending in GDP, $gy = G/Y$, is calibrated inside the model such that the equilibrium value of the income tax level $\bar{\tau}_I$, which balances the consolidated government budget (13), coincides with the empirical estimates obtained by Wu (2021).[16] The share of federal government debt in GDP, dy, is set to 35% in the 1980s and 100% in the 2017 calibration.[17] The estimates for consumption tax, τ_c, and capital income tax τ_a are borrowed from Wu (2021).

5. Findings

Section 5.1 explains the identification of Pareto weights. Section 5.2 presents the results of the major decomposition exercise. Section 5.3 zooms onto the impact of economic and demographic changes, while Section 5.4 elaborates on the impact of government preferences.

5.1. Pareto weight distribution and optimal policy in the 1980s

This subsection provides intuition for the Pareto weight identification and highlights the significance of the paper's contribution. The results presented below are based on the 1980s calibration, but all arguments also hold qualitatively for the 2010s.[18]

Consider first the case of a utilitarian government who maximizes the welfare of *newborn* agents only (upon observing their state $\boldsymbol{x}$), which is the prevailing assumption in the optimal income taxation literature. Table 2 (column *Newborns*)

---

[13] The empirical targets were computed using the NIPA data. See Footnote [12] for the source.

[14] See Table V.H7 (Operations of the Part B Account) in the 2022 Medicare Trustees Report. Footnote [8] shows the data source.

[15] Heathcote et al. (2020) estimate progressivity using the Congressional Budget Office data and find that it has remained unchanged between the 1980s and the 2010s. Since their definition of pre-government income comprises both labor *and* capital income, while disposable income includes Social Security and Medicare, which my model accounts for separately, I choose not to use their estimates.

[16] See Appendix C.6 for details.

[17] The data is provided by the Office of Management and Budget at the White House and is publicly accessible at https://www.whitehouse.gov/omb/historical-tables/ (Table 7.1).

[18] The time index t is dropped in this subsection to simplify notation.

Table 2
Utilitarian policies versus actual policies in the 1980s.

	Equal weights		Data
	Newborns	All alive	(1980s)
Optimal policy:			
– Progressivity $\tau_I^\star$	0.141	0.048	0.187
– Replacement rate $\alpha^\star$, %	0.0	70.0	36.0
Equilibrium variables (in %):			
– Income tax level $\bar{\tau}_{I,\infty}$	11.42	11.76	9.30
– Social Security tax $\tau_{SS,\infty}$	0.0	19.53	8.90

Notes: The values of equilibrium variables are taken from the final steady state associated with the corresponding solution.

shows the optimal income tax progressivity $\tau_I^\star$ and the optimal average replacement rate $\alpha^\star$ in this scenario. To facilitate comparison, the last column (*Data*) records the calibrated policy. As one can see, the government chooses to shut down the public pension system altogether by setting $\alpha^\star$ to 0%, as opposed to 36.0% in the data. To see why, consider the effect of a marginal increase in the replacement rate. As the replacement rate schedule shifts permanently upward, all current and future retirees enjoy higher pensions, which leads to an upward adjustment of the Social Security tax. Hence, workers must weigh welfare gains from receiving a larger pensions during retirement against welfare losses induced by falling after-tax wages. The table suggests that the welfare cost dominates for young workers.[19] Overall, this approach fails to explain why income taxation and Social Security coexist in the data.

In the previous experiment, the government chooses a less progressive system ($\tau_I^\star = 0.141$) than in the data (0.187). Interestingly, when the policymaker is restricted to set progressivity only, while the replacement rate level is fixed at the calibrated level (not shown in the table), $\tau_I^\star$ becomes equal to 0.189, which almost coincides with the data. This special case is, therefore, consistent with Heathcote et al. (2020) who show, in a similar setting, that a "utilitarian equal-weights objective function for the planner offers a reasonable description of US societys taste for redistribution in1980."

Assume next that the policymaker cares equally about *all* agents who are alive at the time when the optimal policy is implemented. This scenario arises as a special case of the Pareto weight function in (20) with $\kappa_1 = \kappa_2 = 0$. The optimal solution, documented in column *All alive*, reveals that the model does a very poor job matching both policies in the data. First, the model predicts an almost twice as high average replacement rate (70.0% versus 36.0%). Counterfactually high replacement rates lead to an excessively high payroll tax of 19.53%, compared to only 8.90% in the data.[20] The government chooses to expand the pension system to better cater to the preferences of retired and middle-aged agents who constitute the largest share of the population. Retired agents do not bear the cost of Social Security taxation, so they unanimously opt to raise α above the status quo level. Similarly, middle-aged workers also opt for a larger pension system because the welfare cost of falling after-tax wages during the remainder of their working careers is relatively small compared to the welfare benefit of receiving larger pensions during the rest of their lives.

Second, the model predicts a substantially lower income tax progressivity than in the data (0.048 versus 0.187). Middle-aged workers do not rely on income transfers as they have had enough time to accumulate precautionary buffer stocks of assets. Moreover, retired agents are not subject to the income tax-and-transfer system altogether. However, both groups favor lowering the distortionary pressure of labor income taxation as to increase aggregate earnings and, therefore, their future pension benefits. This explains why these agents opt for income tax progressivity significantly below the empirical value.

Despite matching the data poorly, this special case has one important qualitative difference compared to the first scenario: the government chooses a *positive* size of the pension system now. Augmenting the government's objective function with the age- and education-specific Pareto weights will allow the model to match both policies in the data, which is shown next. Figs. 3–4 plot the optimal policy as a function of the Pareto weight parameter $\boldsymbol{\kappa}$. Both figures also display the equilibrium values of the Social Security tax $\tau_{SS,\infty}$ and the income tax level $\bar{\tau}_{I,\infty}$ computed in the final steady-state associated with a given solution. All variables, except $\tau_I^\star$, are shown in percentage point deviations from their respective calibrated values, while $\tau_I^\star$ is plotted in absolute deviations.

Consider first Fig. 3, which shows how the optimal policy depends on the age bias κ_1, while the educational bias κ_2 is fixed at 0. For values of κ_1 closer to 0, the policymaker chooses a counterfactually large average replacement rate, which is consistent with the previously discussed case of a utilitarian government. As κ_1 rises and the age distribution of Pareto weights becomes more skewed toward the young, $\alpha^\star$ declines approaching its calibrated value, accompanied by a reduction in the Social Security tax. Hence, one can expect κ_1 to be positive for the model to match the average replacement rate in the data. Also observe that κ_1 affects not only income redistribution between workers and retirees, but also income

[19] This result is consistent with Conesa and Krueger (1999) who find that the political support for Social Security abolishment is the strongest among young workers.

[20] The optimal policy induces an entire transitional path of all endogenous variables, so the table displays the values of $\tau_{SS,\infty}$ and $\bar{\tau}_{I,\infty}$ from the *final* steady state associated with the corresponding solution.

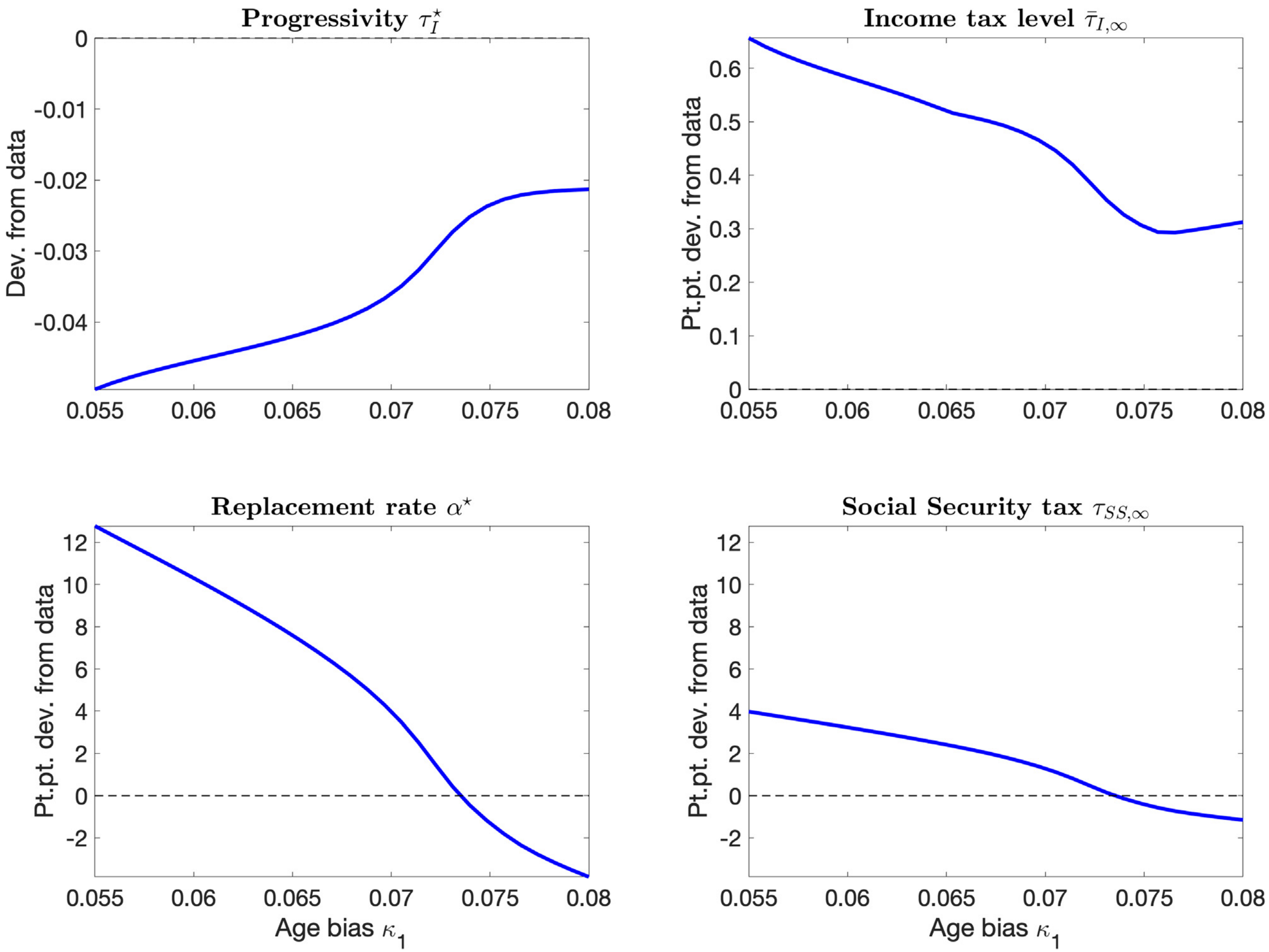

Fig. 3. Optimal policy in the 1980s as a function of the age bias κ_1. *Notes*: The values of the Social Security tax $\tau_{SS,\infty}$ and the income tax level $\bar{\tau}_{I,\infty}$ are taken from the final steady state associated with the corresponding optimal policy. Progressivity $\tau_I^\star$ is shown in absolute deviations from the calibrated value in the 1980s, while other variables are depicted in percentage point deviations. Educational bias parameter κ_2 is held fixed at 0.

Journal of Monetary Economics 134 (2023) 35–52

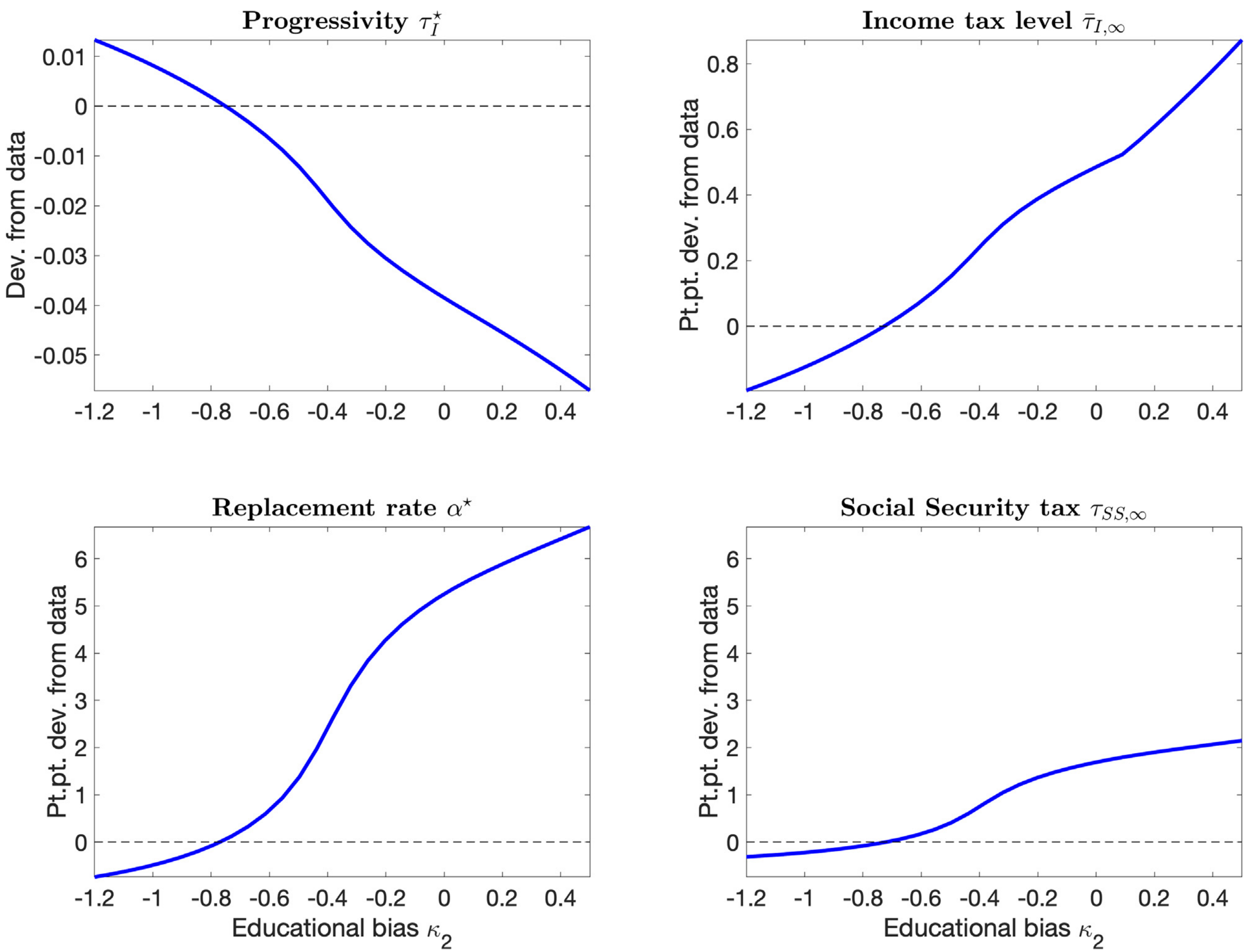

Fig. 4. Optimal policy in the 1980s as a function of the educational bias κ_2. *Notes*: The values of the Social Security tax $\tau_{SS,\infty}$ and the income tax level $\bar{\tau}_{I,\infty}$ are taken from the final steady state associated with the corresponding optimal policy. Progressivity $\tau_I^\star$ is shown in absolute deviations from the calibrated value in the 1980s, while other variables are depicted in percentage point deviations. Age bias parameter κ_1 is held fixed at 0.069, which is the calibrated value for the 1980s.

Table 3
Decomposing the total change in the observed policies during 1980–2010.

	Total change	Decomposition:	
		Economic	Political
	(Data)	forces	forces
Optimal policy:			
– Progressivity $\tau^{\star}_{l,t}$	−0.050	+0.042	−0.092
– Replacement rate $\alpha^{\star}_{t}$	+3.40	+1.15	+2.25
Equilibrium variables:			
– Income tax level $\bar{\tau}_{l,\infty}$	−1.50	−4.82	+3.32
– Social Security tax $\tau_{SS,\infty}$	+1.70	+1.04	+0.66

Notes: All variable, except $\tau^{\star}_{l,t}$, are reported in percentage point deviations, while $\tau^{\star}_{l,t}$ is given in absolute deviations. The total change and the impact of economic forces are shown relative to the 1980s. The impact of the shift in Pareto weights is computed relative to the impact of economic forces. The values of equilibrium variables are taken from the final steady state associated with the corresponding solution. Section 3.2 explains formally the decomposition exercise.

redistribution among workers.[21] Indeed, as κ_1 increases, so does income tax progressivity. Young workers who enter the model without assets and face idiosyncratic labor productivity risk while being excluded from borrowing rely heavily on transfers. As the policymaker becomes more concerned about these agents, it chooses higher progressivity to raise their net transfers. Since college graduates are, on average, subject to higher average tax rates than high-school graduates, income redistribution from educated to uneducated agents increases.

Consider next Fig. 4, which shows the optimal policy as a function of the educational bias κ_2 ($\kappa_1 = 0.069$, which is its identified value). As κ_2 increases and the relative weight on college graduates rises, income tax progressivity goes down approaching its empirical value. Note that if the government was restricted to control progressivity only, κ_2 would be solely in charge of income redistribution between high-skilled and low-skilled workers, similar to Chang et al. (2021), Heathcote and Tsujiyama (2021), and Wu (2021). In the present model, however, κ_2 positively impacts the optimal average replacement rate and, therefore, affects income redistribution between workers and retirees. This finding is surprising because one would expect college graduates to prefer a smaller size of the public pension system than high-school graduates. Recall, however, that educated workers receive pensions, on average, for a longer period due to the education-specific survival rates $\psi_{z,j}$.[22]

Overall, the figure indicates that the age- and education-specific distribution of Pareto weights jointly affects the optimal income tax and the average replacement rate, and therefore, jointly determines the inter- and intra-generational income redistribution in the model. In contrast to the literature that has backed out government preferences over each redistribution type in isolation, this paper identifies them jointly and highlights their interaction.

After providing this intuition, I estimate $\boldsymbol{\kappa}$ formally and find that $\kappa_1 = 0.069$ implying that Pareto weights decrease in age. In particular, the weight on a newborn agent (real-life age 25) relative to an agent of real-life age 65 with the same education, is equal to 15.8. This finding is in stark contrast to the optimal income taxation literature, in which the age distribution locates its entire mass on newborn agents. Furthermore, $\kappa_2 = -0.731$, indicating that high-school graduates receive a 2.1 times larger weight compared to college graduates of the same age. Below, the identified distribution of Pareto weights is applied to compute the optimal response of the government to the economic and demographic changes.

5.2. Accounting for the total change in the policies during 1980–2010

Through the lens of the model, two forces are responsible for the change in the calibrated income tax and replacement rate policies since the 1980s. First, the government might have adjusted the policies because the economic environment has changed. Second, conditional on the economic changes, the government might have amended both programs because its preferences over income redistribution have changed over time. Table 3 quantifies each force.

Consider first the second column which reports the total change in the calibrated policies that the model must account for: income tax progressivity declines by 0.05 (from 0.187 to 0.137), while the average replacement rate increases by 3.4% points (from 36.0% to 39.4%). The third column (*Economic forces*) disentangles the cumulative impact of the economic and demographic forces on the total change. As one can see, the government optimally chooses to raise income tax progressivity by 0.042, contrary to the development in the data, and to increase the average replacement rate by 1.15% points, which is

[21] If the government would be allowed to set the average replacement rate only, then κ_1 would control the inter-generational income redistribution only, as in Brendler (2020).

[22] When both educational groups face homogeneous mortality rates by age, $\alpha^{\star}$ falls in κ_2 (Appendix D).

Table 4
Impact of demographic and economic changes on the optimal policy.

Experiment	Parameters updated	Optimal policies $\tau^{\star}_{I,t}$	Optimal policies $\alpha^{\star}_t$	Equilib. variables $\bar{\tau}_{I,\infty}$	Equilib. variables $\tau_{SS,\infty}$
1. Aging	$(\psi_{z,j}, n)$	−0.010	+9.83	+0.46	+5.34
2. Production	(ϖ, δ)	−0.005	−6.55	−2.60	−1.70
3. Social Security	$(J^R, \bar{\alpha}, \bar{e}_{\min}, \delta^p, cap_{SS})$	−0.060	−0.26	+1.02	−1.31
4. Medicare	$(m_j, \eta, \tau_M, cap_M)$	−0.050	−1.62	+1.79	−0.31
5. Other policies	(τ_c, τ_a, gy, dy)	−0.048	−9.24	+0.82	−2.41
6. Earnings inequality:					
– Supply of college graduates	Π_z	−0.046	−4.10	+1.10	−0.71
– Human capital accumulation	$(\theta_z, h_{1,z}, \delta^h)$	+0.063	+9.67	−5.72	+2.81
– Fixed effects	$\sigma^2_{v,z}$	+0.064	+4.41	−4.26	+1.03
– Skill complementarity	(ρ, Z)	+0.014	+9.20	−2.96	+3.45
– Idiosyncratic labor risk	$(\rho_z, \sigma^2_{\epsilon,z})$	−0.030	−2.24	+0.69	−0.59
7. Total impact	All listed above	+0.042	+1.15	−4.82	+1.04

Notes: All variables, except $\tau^{\star}_{I,t}$, are shown in percentage point deviations from their respective values in the 1980s, while $\tau^{\star}_{I,t}$ is displayed in absolute deviations. The optimal policies are computed under the Pareto weights from the 1980s. *Total impact* shows the cumulative effect of all parameter changes. The values of the income tax level $\bar{\tau}_{I,\infty}$ and Social Security tax $\tau_{SS,\infty}$ are taken from the final steady state associated with the corresponding optimal solution.

less than half of the total increase in $\alpha^{\star}_t$ in the data. Conditional on the economic and demographic changes, the impact of government preferences on the policy is determined residually. According to the last column, $\tau_{I,t}$ must decline by 0.092, while α_t must increase by 2.25% points.

The next sections discuss these results in more detail. Section 5.3 zooms onto the impact of economic forces, while Section 5.4 quantifies the shift in government preferences.

5.3. First force: Economic and demographic changes

To investigate the effect of the economic forces on the optimal policy, I divide all model parameters whose values vary across the two calibration periods into six groups (Aging, Production, Social Security, Medicare, Other policies, and Earnings inequality) and then quantify the marginal impact of each parameter group on the optimal solution. With respect to Earnings inequality, I further split the corresponding parameters into five subgroups: Supply of college graduates, Human capital accumulation, Fixed effects, Skill complementarity, and Idiosyncratic labor risk. In all experiments, the economy is assumed to be in a steady-state equilibrium, in which the parameters from a given group take the calibrated values from the 2010s, while all remaining parameters and the initial policies are yet the same as in the 1980s, and the government faces the set of old preferences given by $\boldsymbol{\kappa}_{1980}$.

Table 4 presents the findings. To ease comparison across experiments, all variables, except $\tau^{\star}_{I,t}$, are shown in percentage point deviations from their respective calibrated values in the 1980s, while $\tau^{\star}_{I,t}$ is reported in absolute deviations. Column *Total impact* reproduces the cumulative effect of demographic and economic changes previously shown in Table 3.

Several findings emerge from the table. First, Skill complementarity, Human capital accumulation, and Fixed effects turn out to be the sole contributors to a more progressive income tax schedule, with the strongest impact (+0.064) generated by Fixed effects. One common feature across these experiments is that earnings inequality rises substantially. Poor young workers opt for higher income tax progressivity because they rely heavily on transfer income. The government optimally chooses to raise progressivity because the Pareto weight distribution, under which the optimal policies are computed, is skewed toward younger, uneducated workers.[23] Since the cumulative impact of the economic and demographic changes is to make the income tax system more progressive (by 0.042) and since all remaining parameter changes contribute to a fall in progressivity, these three experiments must dominate quantitatively.[24] Second, the three experiments mentioned above account for the total rise in the optimal average replacement rate. As the marginal income tax rates at the upper tail of the income distribution rise, the government collects higher tax revenues. Through the government budget constraint (13), the income tax level $\bar{\tau}_{I,\infty}$ declines, which allows the policymaker to simultaneously increase $\alpha^{\star}_t$ because earnings inequality

[23] To check how results depend on the education-specific Pareto weights, I recompute the optimal policy assuming $\kappa_{2,1980} = 0$, while maintaining $\kappa_{1,1980}$ at 0.069. Even though the impact of Human capital accumulation and Fixed effects on $\tau^{\star}_{I,t}$ becomes much smaller (+0.027 and +0.045, respectively) and even negative (−0.031) for Skill complementarity, the cumulative effect of all changes on progressivity remains positive.

[24] Social Security and Medicare exert the most pronounced mitigating effect on $\tau^{\star}_{I,t}$ (−0.06 and −0.05, respectively). Recall that cap_{SS} doubles during 1980–2010, while cap_M is eliminated altogether. Since the labor income tax burden on young, productive workers increases, the government reduces $\tau^{\star}_{I,t}$.

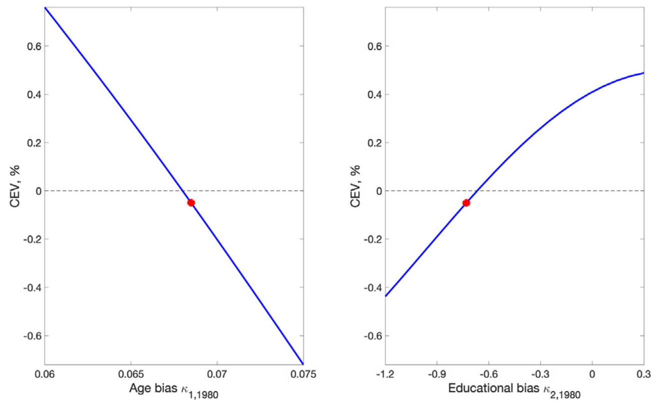

Fig. 5. Welfare effects of the optimal policy as a function of Pareto weight parameter $\boldsymbol{\kappa}_{1980}$. *Notes*: The figure shows the welfare effect of the optimal policy in response to the economic and demographic changes as a function of the age bias $\kappa_{1,1980}$ (left panel) and the educational bias $\kappa_{2,1980}$ (right panel). Welfare is computed by weighting all households using the pre-reform stationary distribution of agents and is measured as consumption equivalent variation (in %). In the left panel, $\kappa_{2,1980} = -0.731$; in the right panel, $\kappa_{1,1980} = 0.069$. The dot in each panel corresponds to the welfare effect of the optimal policy in the baseline model (−0.05%).

during the working stage in these experiments propagates into income inequality during retirement. Finally, notice that all remaining experiments, except Aging, contribute to a reduction in $\alpha_t^\star$.[25]

Next, I explore the welfare implications of the optimal policy under all parameter changes by asking: How much percent does the consumption of the agents who are alive in the economy with the new economic and demographic conditions but the old policy have to change in all future periods and contingencies (keeping their leisure and bequests unchanged) so that their expected utility equals that which would arise if the government implemented the optimal solution. When aggregating welfare across households, I apply the utilitarian social welfare measure and weigh agents using the stationary distribution obtained from solving the model for a pre-reform steady-state equilibrium.

The government's response to the economic and demographic conditions, computed in the previous subsection and recorded in column *Economic forces* (Table 3), worsens the aggregate welfare, although the magnitude of the welfare loss is fairly small with 0.05% in consumption equivalent terms. Intuitively, the government adopts a policy that benefits greatly young, uneducated workers because progressivity rises and retired agents because the replacement rate schedule shifts upward. However, the same policy hurts productive young and middle-aged workers who suffer substantial welfare losses due to a more progressive income tax schedule. The results indicate that these losses dominate.

Fig. 5 explores the reform's welfare implications in more detail. The left panel shows how welfare varies with the age bias $\kappa_{1,1980}$ ($\kappa_{2,1980}$ is fixed at −0.731). A dot corresponds to the welfare impact induced by the optimal policy in the baseline model (i.e., −0.05%). Relative to the baseline model, a small reduction in $\kappa_{1,1980}$ reverses the welfare implications of the optimal policy making it welfare improving. As the relative Pareto weight on older agents increases, the government raises Social Security taxes to finance a more generous pension system in exchange for a less progressive income tax system (recall Fig. 3). When $\kappa_{1,1980}$ is sufficiently small, welfare gains enjoyed by productive workers and retired agents dominate welfare losses suffered by young, uneducated workers.

The right panel of Fig. 5 plots welfare as a function of the education bias $\kappa_{2,1980}$, where $\kappa_{1,1980}$ is fixed at 0.069. Starting from the welfare effect in the baseline model, the reform becomes socially desirable as the relative weight on educated agents slightly increases. On the one hand, income tax progressivity declines (recall Fig. 4), which causes welfare losses to young, uneducated agents, but greatly benefits productive workers. On the other hand, the government chooses to raise the average replacement rate to better cater to the preferences of the educated young workers who receive pension annuities, on average, for a longer period. Note that this policy measure also benefits all currently alive middle-aged and retired agents.[26] Overall, the figure indicates that welfare gains dominate quantitatively.

To sum up, the welfare implications of the optimal policy crucially depend on the age- and education-specific distribution of Pareto weights. The next subsection analyzes how this distribution has evolved over time in the US.

5.4. Second force: Evolution of social welfare preferences

So far, I have analyzed the optimal government response to the new economic and demographic conditions using the identified Pareto weight distribution from the 1980s. This step allowed me to extract an unexplained portion of the total

[25] Aging increases the relative size of retirees and the cross-sectional share of educated households due to education-specific mortality rates. Both effects contribute significantly (9.83% points) to a rise in $\alpha_t^\star$.

[26] Assuming that reform does not affect benefits of already retired agents lowers the magnitude of welfare effects shown in the right panel of Fig. 5, but does not change the welfare sign.

Table 5
Estimated parameters $\boldsymbol{\kappa}_t$ of the Pareto weight function and the implied weights.

	1980s	2010s
Age bias $\kappa_{1,t}$	0.069	0.060
– Implied weight on age 25 / 65	15.80	11.02
Educational bias $\kappa_{2,t}$	−0.731	1.260
– Implied weight on col. / non-col.	0.48	3.53

Notes: The table reports the implied weight on a newborn agent (real-life age 25) relative to the agent of (real-life) age 65 with the same educational level, and the implied weight on a college graduate relative to a non-college graduate of the same age. Section 3.2 explains the experimental design to estimate $\boldsymbol{\kappa}_t$.

change in both policies: according to the last column of Table 3, income tax progressivity declines by 0.092 and the average replacement rate increases by 2.25% points, conditional on the economic and demographic changes. This subsection exploits this residual change in the policies to identify $\boldsymbol{\kappa}_{2010}$.[27]

Table 5 (last column) shows the obtained estimates of $\boldsymbol{\kappa}_{2010}$, while the central column replicates the estimates of $\boldsymbol{\kappa}_{1980}$. The age bias $\kappa_{1,t}$ declines over time from 0.069 to 0.060, indicating that the distribution of Pareto weights becomes less skewed toward young workers. More specifically, the implied Pareto weight on a 25-year-old worker relative to a 65-year-old agent with the same education drops from 15.80 to 11.02. At the same time, the educational bias $\kappa_{2,t}$ increases from −0.731 to 1.26, meaning that the relative weight on college graduates in the new calibration rises to 3.53. Overall, these findings suggest that the US government has become less willing to tolerate redistribution from educated people toward uneducated people and more willing to tolerate income redistribution from workers toward retirees.[28] Note that the model attributes only a portion of the total increase in α_t during 1980–2010s to stronger government's preferences to redistribute incomes from young to old, while the remaining portion is explained by the fact that incentives to redistribute incomes from educated to uneducated workers have become weaker over time. When the weight attached to college graduates increases (so as to account for the drop in progressivity), the government raises $\alpha^\star_{2010}$ to better cater to these agents' preferences who receive pension benefits, on average, for a longer period than uneducated agents due to the education-specific mortality.[29] Similar intuition implies to income taxation. A portion of the decline in progressivity $\tau^\star_{l,2010}$ is accounted for by a shift of Pareto weights toward more educated workers. This channel is similar to Heathcote and Tsujiyama (2021) and Wu (2021), though in those studies the entire portion of the drop in progressivity must be attributed to $\kappa_{2,2010}$. But as the age distribution of weights shifts toward older agents, the government optimally chooses to reduce $\tau^\star_{l,2010}$, so the shift of weights toward older agents reduces incentives to redistribute income among workers and, therefore, exerts an offsetting effect on $\kappa_{2,2010}$.

Although the model is silent regarding the underlying forces behind the identified shift in the Pareto weight distribution, Appendix E provides some supporting evidence by studying the evolution of voter turnout rates during Congressional elections in the US.

6. Conclusions

The current analysis opens exciting paths for future research. Below I suggest one of them. A large strand of the literature on retirement has analyzed the macroeconomic and welfare consequences of various financing reforms targeted to make public pension system sustainable. Numerous quantitative studies have concluded that a public pension system reform may have very different welfare implications for households. Then, the following question arises: Which policy proposal is the policymaker likely to implement? This is an important question because expectations about future government policies impact the households' decisions today. The identified distribution of Pareto weights may be helpful in this analysis because it captures feasibility constraints imposed on the political process. Hence, the Pareto weights can be applied in policy analyses to reduce the set of all economically feasible proposals to those that are also implementable from the political standpoint.

[27] Section 3.2 explains the procedure to estimate κ_{2010} (see the last step).

[28] Heathcote et al. (2020) estimate that income tax progressivity has remained roughly unchanged during 1980–2010 in the data. As a robustness check, I recompute $\boldsymbol{\kappa}_{2010}$ assuming that $\tau_{l,2010} = 0.187$. Qualitatively, this does not change the identified trend in Pareto weights toward older and more educated agents.

[29] In Brendler (2020), I recover social preferences over inter-generational income redistribution by endogenizing α, while taking income taxation as given and ignoring mortality differences. By construction, the paper attributes the entire change in the actual policy to a shift in the age distribution of weights.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jmoneco.2022.10.004

References

Abbott, B., Gallipoli, G., Meghir, C., Violante, G.L., 2019. Education policy and intergenerational transfers in equilibrium. Journal of Political Economy 127 (6), 2569–2624.
Badel, A., Huggett, M., Luo, W., 2020. Taxing top earners: a human capital perspective. The Economic Journal 130 (629), 1200–1225.
Brendler, P., 2020. Why hasn't social security changed since 1977? Rev Econ Dyn 36, 134–157.
Chang, B.H., Chang, Y., Kim, S.-B., 2018. Pareto weights in practice: a quantitative analysis across 32 OECD countries. Rev Econ Dyn 28, 181–204.
Chang, B.H., Chang, Y., Kim, S.-B., 2021. Equilibrium tax rates under ex-ante heterogeneity and income-dependent voting. The B.E. Journal of Macroeconomics. 2020-0274
Conesa, J.C., Krueger, D., 1999. Social security reform with heterogeneous agents. Rev Econ Dyn 2 (4), 757–795.
De Nardi, M., 2004. Wealth inequality and intergenerational links. Rev Econ Stud 71 (3), 743–768.
Guvenen, F., Kaplan, G., Song, J., Weidner, J., 2017. Lifetime incomes in the United States over six decades. NBER Working Papers. National Bureau of Economic Research, Inc.
Guvenen, F., Kuruscu, B., Ozkan, S., 2014. Taxation of human capital and wage inequality: across-country analysis. Rev Econ Stud 81 (2 (287)), 818–850.
Heathcote, J., Storesletten, K., Violante, G.L., 2017. Optimal tax progressivity: an analytical framework. Q J Econ 132 (4), 1693–1754.
Heathcote, J., Storesletten, K., Violante, G.L., 2020. How should tax progressivity respond to rising income inequality? J Eur Econ Assoc 18 (6).
Heathcote, J., Tsujiyama, H., 2021. Optimal income taxation: Mirrlees meets Ramsey. Journal of Political Economy 129 (11), 3141–3184.
Hosseini, R., Shourideh, A., 2019. Retirement financing: an optimal reform approach. Econometrica 87 (4), 1205–1265.
Huggett, M., Parra, J.C., 2010. How well does the U.S. social insurance system provide social insurance? Journal of Political Economy 118 (1), 76–112.
Imrohoroglu, S., Kitao, S., 2012. Social security reforms: benefit claiming, labor force participation, and long-run sustainability. American Economic Journal: Macroeconomics 4 (3), 96–127.
Katz, L.F., Murphy, K.M., 1992. Changes in relative wages, 1963–1987: supply and demand factors. Q J Econ 107 (1), 35–78.
Kitao, S., 2014. Sustainable social security: four options. Rev Econ Dyn 17 (4), 756–779.
McGrattan, E.R., Prescott, E.C., 2017. On financing retirement with an aging population. Quant Econom 8 (1), 75–115.
Moser, C., Olea de Souza e Silva, P., 2019. Optimal paternalistic savings policies. Working Paper.
Ndiaye, A., 2020. Flexible retirement and optimal taxation. Federal Reserve Bank of Chicago Working Paper.
Nishiyama, S., Smetters, K., 2007. Does social security privatization produce efficiency gains? Q J Econ 122 (4), 1677–1719.
Storesletten, K., Telmer, C.I., Yaron, A., 2004. Consumption and risk sharing over the life cycle. J Monet Econ 51 (3), 609–633.
Wu, C., 2021. More unequal income but less progressive taxation. J Monet Econ 117 (C), 949–968.