CES ifo Working Papers

10684
2026

Original Version: October 2023
This Version: June 2026

Redistribution with Unequal Life Expectancy

Sebastian Koehne

CES ifo

Imprint:

CESifo Working Papers

ISSN 2364-1428 (digital)

Publisher and distributor: Munich Society for the Promotion
of Economic Research - CESifo GmbH

Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740

Email office@cesifo.de
https://www.cesifo.org

Editor: Clemens Fuest

An electronic version of the paper may be downloaded free of charge
· from the CESifo website: www.ifo.de/en/cesifo/publications/cesifo-working-papers
· from the SSRN website: www.ssrn.com/index.cfm/en/cesifo/
· from the RePEc website: https://ideas.repec.org/s/ces/ceswps.html

Redistribution with Unequal Life Expectancy

Sebastian Koehne*

First draft: September 29, 2023

This draft: June 4, 2026

Abstract

I study life-expectancy inequality in a Mirrleesian life-cycle framework with pension-linked fiscal externalities and private retirement saving. Heterogeneous longevity affects optimal redistribution by changing both the expected fiscal cost of pension benefits and the mapping from disposable income into social marginal welfare. I derive an $\mathcal{ABCD}$ optimal tax formula that incorporates these efficiency and equity effects. Calibrated to the U.S. economy, the model implies that accounting for heterogeneous longevity lowers optimal marginal tax rates, particularly at the bottom of the income distribution. These reductions generate welfare losses for the lowest-income types and welfare gains for most other earners, peaking at middle incomes.

JEL codes: D82, H21, H55

Keywords: Optimal Taxation, Redistribution, Life Expectancy, Inequality

*Kiel University and CESifo, koehne@economics.uni-kiel.de. This research was supported by the German Research Foundation (DFG), project number 454799018.

1 Introduction

Income and life expectancy are strongly correlated (Goldman, 2001; Cutler et al., 2011). In the United States, individuals in the top income decile can expect to live about ten years longer than those in the bottom decile (Chetty et al., 2016), and sizable gaps persist even in countries with lower income inequality and universal health care systems (Mackenbach et al., 2008; European Commission, 2013; Lampert et al., 2019). This gradient matters for redistribution because income is not only a measure of economic capacity, but is also correlated with the horizon over which resources are consumed and pension benefits are paid. Yet most analyses of redistribution focus narrowly on income, largely abstracting from the income-longevity gradient.[1] This raises an important policy question: how should income taxes and transfers be designed when income groups differ systematically in life expectancy?

This paper provides a theoretical and quantitative answer. I introduce heterogeneous survival probabilities into a tractable Mirrleesian life-cycle model. Individuals differ in a one-dimensional unobservable type that determines both productivity and the probability of reaching retirement. This probability can equivalently be interpreted as a measure of expected retirement duration, allowing the two-period structure to capture the income-longevity gradient in a parsimonious way. Individuals choose labor supply and allocate disposable resources between working-life consumption and private retirement assets. The price of these assets may differ from the actuarially fair value of survival-contingent payouts, thereby distinguishing the private cost of retirement resources from their expected fiscal cost.

The main analysis characterizes optimal nonlinear income taxation conditional on an existing pension system. This conditional approach isolates a policy-relevant decision problem, since income tax schedules are often adjusted around pre-existing pension rules rather than jointly redesigned with them. I also show how the characterization changes when taxes and pensions are chosen jointly. When survival probabilities are homogeneous and pension wedges are absent, the model reduces to the canonical static Mirrlees (1971) framework, providing a natural benchmark.

Unequal life expectancy modifies the optimal tax problem through two channels. On the

[1] Existing contributions show that heterogeneous survival matters for retirement incentives, pension design, and asset taxation (Cremer et al., 2004; Shourideh and Troshkin, 2017; Hosseini and Shourideh, 2019; Jang et al., 2023). Related work studies fiscal externalities between income taxation and other social policies, such as unemployment insurance (Ferey, 2022). Less is known about how unequal life expectancy changes optimal nonlinear income taxation.

equity side, a dollar of disposable income need not have the same marginal welfare value for individuals with different survival prospects and different costs of financing retirement consumption. Hence, the mechanical value of redistribution depends on life-cycle choices rather than on disposable income alone. On the efficiency side, behavioral earnings responses affect public revenues not only through income taxes, but also through the pension system. Ignoring life expectancy differences therefore mismeasures both the mechanical value of redistribution and the fiscal cost of behavioral responses.

I formalize these forces by deriving an $\mathcal{ABCD}$ optimal tax formula. The first three components parallel the familiar $\mathcal{ABC}$ logic of Diamond (1998). The additional term adjusts statutory marginal tax rates for pension-induced fiscal externalities. It reflects that earnings responses change public revenues through income taxes, pension contributions, and the expected present value of future pension benefits, so the relevant fiscal wedge generally differs from the statutory marginal tax rate faced by the individual. The mechanical redistribution term is also modified: the marginal welfare weight entering this term is evaluated at the individual intertemporal optimum. This modification reflects both a survival-probability effect, operating through the expected marginal value of retirement consumption, and an intertemporal-price effect, operating through the private cost of shifting resources to retirement.

These channels yield clear comparative statics. When life expectancy rises with income, greater life-expectancy inequality tends to lower optimal marginal tax rates at the bottom. The efficiency logic is that a wider spread makes low earners shorter-lived, so a marginal dollar of their earnings accrues only a small expected future benefit liability; its net fiscal return (contributions net of expected benefit liabilities) is correspondingly high, which makes reducing these earnings through taxation relatively more fiscally costly. The equity logic depends on the pricing of private retirement resources. I show that, under both actuarially fair annuity pricing and uniform intertemporal pricing, survival-probability heterogeneity makes social marginal welfare weights decline less steeply with income, thereby attenuating the mechanical gains from redistribution toward low earners. For higher earners, the impact of life-expectancy inequality on optimal marginal tax rates is generally ambiguous, since the fiscal-wedge channel and the modified welfare-weight channel may work in opposite directions.

To assess the quantitative importance of these channels, I calibrate the model to the U.S.

economy and compare optimal tax schedules under two benchmark assumptions about private retirement asset markets. Accounting for life-expectancy inequality lowers optimal marginal tax rates throughout the income distribution, with effects concentrated at the bottom. The reductions are larger under actuarially fair annuity pricing, where survival-dependent annuity prices reinforce the survival-probability effect. As a consequence, optimal marginal tax rates fall by up to 4.9 percentage points relative to a benchmark with homogeneous survival probabilities, compared with 1.4 percentage points under uniform intertemporal pricing. These policy adjustments imply consumption-equivalent welfare gains reaching about 1 percent for middle earners, while the lowest-income group experiences losses exceeding 2 percent.

Finally, I clarify the role of pension design for the redistribution framework. If the policy maker jointly chooses the income tax schedule and the pension system, the separate marginal rates of taxes, contributions, and benefits are not uniquely pinned down by the optimality condition. What matters for individual behavior is their composition into the private effective marginal tax rate, whereas what matters for public revenues is their composition into the fiscal effective marginal tax rate. The corresponding optimality condition combines these statistics in a close analogue to the main formula. I then extend the optimality logic to a stylized multi-period life-cycle environment with age-dependent wages, preferences, and survival probabilities, showing that the core mechanisms are not specific to the two-period framework.

Related literature. This paper connects and extends multiple strands of the literature, particularly in the areas of optimal nonlinear income taxation, dynamic taxation, and pension design.

First, the paper's main contribution is to the Mirrleesian literature on optimal nonlinear income taxation (Mirrlees, 1971), as surveyed by Mankiw et al. (2009), Diamond and Saez (2011), and Piketty and Saez (2013). To my knowledge, this is the first paper to characterize optimal nonlinear income taxation in a Mirrleesian framework in which life expectancy varies with skill. The analysis extends the optimal tax formula by modifying the mechanical redistribution term and by incorporating pension-induced fiscal externalities from earnings responses. The latter channel connects to recent findings by Ferey (2022) on the interaction between redistribution and unemployment insurance.

Second, by adopting a life-cycle perspective, the paper relates to the literature on dynamic

optimal taxation, which typically abstracts from heterogeneity in life expectancy.[2] The closest dynamic contributions are Mirrleesian life-cycle frameworks with deterministic productivity paths and retirement decisions, as developed by Michau (2014), Shourideh and Troshkin (2017), and Hosseini and Shourideh (2019). These works study retirement incentives, intertemporal wedges, and pension design. The present paper complements them by taking asset taxation and pension rules as given and deriving how longevity heterogeneity modifies the optimal nonlinear income tax schedule conditional on those institutions.

Third, the paper contributes to the broader literature on the implications of longevity heterogeneity for public policy.[3] Cremer et al. (2004) study a discrete-type life-cycle model with heterogeneity in skill and health and show that, when first-best instruments are unavailable, optimal policy combines income taxation with distortions to the retirement decision. Sheshinski (2008) analyzes annuity pricing when lifespans differ across individuals. Bommier et al. (2011) and Leroux and Ponthiere (2013) study redistribution under longevity heterogeneity while abstracting from labor supply. Relative to these contributions, the present paper embeds longevity heterogeneity in a Mirrleesian income-tax framework with continuous skill heterogeneity, endogenous earnings, and pension-linked fiscal externalities.

The analysis also relates to recent quantitative work on life-cycle models with heterogeneous mortality, including Bagchi (2019), Abraham et al. (2023), and Jones and Li (2023). These works use calibrated models to study pension design. The present paper is complementary: it develops a tractable theoretical characterization of optimal nonlinear income taxation and derives sufficient-statistics conditions that clarify the relevant equity and efficiency margins. Finally, while the present paper adopts the utilitarian approach common in much of the optimal taxation literature, recent work by Jang et al. (2023) highlights that welfare comparisons in the presence of differential mortality can be sensitive to alternative normative criteria.

Outline. The remainder of the paper is organized as follows. Section 2 develops a Mirrleesian life-cycle model with heterogeneous life expectancy and a pension system. Section 3 presents the theoretical analysis and derives an $\mathcal{ABCD}$ formula for optimal taxation. Section 4 calibrates

[2]See for example Grochulski and Kocherlakota (2010), Farhi and Werning (2013), Koehne and Kuhn (2015), Golosov et al. (2016), Koehne (2018), and the survey by Stantcheva (2020).

[3]While not focused on longevity, Diamond and Spinnewijn (2011) analyze optimal capital taxation when individuals differ in discount rates, a conceptually related form of intertemporal heterogeneity. In the present model, high-skilled individuals place greater expected weight on retirement consumption because they have higher survival probabilities. This effect would be reinforced if high-skilled individuals were also more patient in terms of pure time preferences.

the model and quantifies the implications for optimal tax schedules under alternative market environments for private retirement assets. Section 5 concludes. All proofs are relegated to Appendix A.1.

2 Model

This section develops a tractable Mirrleesian life-cycle model with heterogeneous life expectancy and a public pension system. The basic model consists of two periods. The first period represents the individual's working life. In this period, individuals work, consume and transfer resources to retirement. The second period represents retirement. In this period, individuals have left the labor market and consume public pension benefits together with privately financed retirement resources. Individuals differ with respect to a one-dimensional, unobservable characteristic that determines both their labor productivity θ and their probability $p(\theta)$ of reaching retirement.[4] I refer to variation in $p(\theta)$ as heterogeneity in life expectancy or, more compactly, as longevity heterogeneity.

2.1 Preferences and skills

There is a continuum of individuals with identical, time-separable von Neumann–Morgenstern preferences. The individuals live for up to two periods and discount the future with the factor $\beta \in (0,1)$.

Individuals differ with respect to their skill $\theta \in \Theta := \left(\underline{\theta}, \overline{\theta}\right) \subset \mathbb{R}_{++}$. An individual with skill θ and labor hours l earns a gross income of

$$y = \theta l. \tag{1}$$

Skill and labor hours are private information, whereas gross income y is publicly observable. The distribution of skill types in the economy is defined by a smooth probability density $f : \Theta \to \mathbb{R}_{++}$ with full support. The associated cumulative distribution function is denoted by $F : \Theta \to [0,1]$.

In the first period, individual utility depends on consumption $c_1 \in \mathbb{R}_+$ and labor hours

[4] Equivalently, $p(\theta)$ can be interpreted as a measure of retirement duration.

$l \in \mathbb{R}_+$ and is given by

$$u\left(c_1 - v\left(l\right)\right). \tag{2}$$

The utility function $u : \mathbb{R} \to \mathbb{R}$ is strictly increasing, strictly concave and twice continuously differentiable. The labor disutility function $v : \mathbb{R}_+ \to \mathbb{R}_+$ is strictly increasing, strictly convex, twice continuously differentiable and satisfies $v\left(0\right) = 0$. Note that this specification eliminates income effects on labor supply.

With probability $p\left(\theta\right) \in \left(0, 1\right)$, an individual with skill θ survives to the second period. In this period, the individual does not work and her utility is determined exclusively by her consumption:

$$u(c_2) = u\left(c_2 - v\left(0\right)\right). \tag{3}$$

The model treats the survival probability $p\left(\theta\right)$ as exogenous. Thus, the probability of reaching retirement can be correlated with skill, but it is not affected by individual choices or government policy within the model. This assumption reflects that current evidence does not support a credible "production function" for longevity (Cutler et al., 2006, 2011). Moreover, panel and quasi-experimental studies suggest that the causal effect of income on health in advanced economies is typically small or negligible (Adams et al., 2003; Meer et al., 2003; Frijters et al., 2005; Banerjee et al., 2010; Cesarini et al., 2016; Gerdtham et al., 2016; Erixson, 2017).

2.2 Individual problem

At the beginning of the first period, individuals observe their skill realization θ. They then choose their labor supply $l\left(\theta\right)$ and, as a consequence, their gross income

$$y\left(\theta\right) = \theta \cdot l\left(\theta\right). \tag{4}$$

Individuals pay income taxes $T\left(y\left(\theta\right)\right)$ and public pension contributions $S\left(y\left(\theta\right)\right)$. They allocate their disposable income between contemporaneous consumption $c_1\left(\theta\right)$ and privately financed retirement resources $a\left(\theta\right)$. One unit of such retirement resources costs $q\left(\theta\right)$ units of first-period consumption and pays one unit in the second period conditional on survival.

The intertemporal price $q(\theta)$ summarizes the private cost of shifting resources to retirement. It is a reduced-form representation of private insurance and saving opportunities in the presence

of survival risk. Two benchmark cases are useful for the analysis below. Under *actuarially fair annuity pricing*, the price equals the expected discounted payout, $q(\theta) = \beta p(\theta)$. At the opposite extreme, individuals may face a *uniform intertemporal price* $q(\theta) = \bar{q}$, as would arise under pooled annuity pricing or a non-annuitized storage technology whose return is independent of individual survival probabilities. More generally, $q(\theta)$ may differ from $\beta p(\theta)$ because annuity prices are regulated, partially pooled, or otherwise imperfectly linked to individual survival probabilities.

In the second period, which represents retirement and is reached with probability $p(\theta)$, individuals do not work. If alive, they consume the sum

$$c_2(\theta) = a(\theta) + B(y(\theta)) \tag{5}$$

of privately financed retirement resources $a(\theta)$ and public pension benefits $B(y(\theta))$. If an individual dies before reaching retirement, second-period utility is normalized to zero. Because survival probabilities are exogenous, this normalization has no effect on the optimization problem of the individual or the government.

Formally, individuals solve the following optimization problem:

$$\begin{aligned} V(\theta) := \max_{c_1, c_2, a, y} \; & u\left(c_1 - v\left(\frac{y}{\theta}\right)\right) + \beta p(\theta)\, u(c_2) \\ \text{subject to} \quad & c_1 + q(\theta)\, a = y - T(y) - S(y) \quad \text{and} \quad c_2 = a + B(y). \end{aligned} \tag{6}$$

The two budget constraints can be combined into the intertemporal budget constraint:

$$c_1 + q(\theta)\, c_2 = y - T(y) - S(y) + q(\theta)\, B(y). \tag{7}$$

Using this constraint, the individual problem can be written as follows:

$$V(\theta) = \max_{c_2, y} u\left(y - T(y) - S(y) + q(\theta)\, B(y) - v\left(\frac{y}{\theta}\right) - q(\theta)\, c_2\right) + \beta p(\theta)\, u(c_2). \tag{8}$$

This decision problem can be decomposed into two steps. First, individuals choose earnings

to maximize net lifetime income minus labor disutility:

$$\max_{y} y - T(y) - S(y) + q(\theta) B(y) - v\left(\frac{y}{\theta}\right). \tag{9}$$

Given the solution of this problem, individuals allocate resources intertemporally according to:

$$\max_{c_2} u\left(y - T(y) - S(y) + q(\theta) B(y) - v\left(\frac{y}{\theta}\right) - q(\theta) c_2\right) + \beta p(\theta) u(c_2). \tag{10}$$

Throughout the analysis, I assume that individuals can transfer resources across periods at price $q(\theta)$ without binding borrowing or saving constraints, so that the intertemporal first-order condition holds with equality.[5]

From the point of view of the individual, the labor supply problem (9) is equivalent to that of the typical static Mirrlees problem, except that the statutory income tax is replaced by the *private effective* income tax

$$\hat{T}(y;\theta) := T(y) + S(y) - q(\theta) B(y), \tag{11}$$

which accounts for public pension contributions and the private value of public pension benefits. The corresponding private effective marginal tax rate is

$$\hat{T}'(y;\theta) = T'(y) + S'(y) - q(\theta) B'(y). \tag{12}$$

It equals the statutory marginal tax rate T' plus the *privately valued pension wedge,* defined as

$$\omega(y,\theta) := S'(y) - q(\theta) B'(y). \tag{13}$$

Thus, $\hat{T}' = T' + \omega$. This private effective marginal tax rate is the object that governs individual labor-supply incentives.

[5]Provided that public pension benefits are not overly generous, individuals allocate strictly positive private resources to retirement.

2.3 Government problem

The government does not observe skills or labor hours and sets the tax schedule T as a nonlinear function of income y. For the main analysis, I focus on the statutory tax schedule T as the government's design object and treat the public pension system (S, B) as exogenous. This approach isolates the redistributive role of income taxation and captures the fact that income tax policy is often adjusted conditional on existing pension rules rather than jointly redesigned with them.[6] In Section 3.3.1, I show how the characterization changes when income taxes and pension policies are chosen jointly.

The government can save and borrow at the interest rate $r = 1/\beta - 1$. Its intertemporal budget constraint is therefore

$$\int_{\underline{\theta}}^{\overline{\theta}} \left[T\left(y\left(\theta\right)\right) + S\left(y\left(\theta\right)\right) - \beta p\left(\theta\right) B\left(y\left(\theta\right)\right)\right] dF\left(\theta\right) - E \geq 0, \tag{14}$$

where E denotes fixed expenditures on other government programs. The term $\beta p(\theta)B(y(\theta))$ is the expected present-value fiscal cost of pension benefits for type θ. Thus, the government values one unit of pension benefits at the public cost factor $\beta p(\theta)$, whereas the individual values the same benefit through the private intertemporal price $q(\theta)$. These two prices coincide under actuarially fair annuity pricing but may differ when annuity prices are pooled, imperfectly individualized, or when retirement saving is not fully annuitized.

This distinction gives rise to the fiscal counterpart of the privately valued pension wedge introduced above. I define the *fiscal pension wedge* as

$$\omega^F\left(y, \theta\right) := S'\left(y\right) - \beta p\left(\theta\right) B'\left(y\right). \tag{15}$$

It measures the marginal pension contribution net of the expected present-value fiscal cost of the associated marginal pension benefit. The corresponding *fiscal effective marginal tax rate* is

$$\hat{T}^{F\prime}\left(y; \theta\right) := T'\left(y\right) + \omega^F\left(y, \theta\right) = T'\left(y\right) + S'\left(y\right) - \beta p\left(\theta\right) B'\left(y\right). \tag{16}$$

This rate records the budgetary effect of marginal earnings responses along the equilibrium

[6] In the same spirit, I treat the private intertemporal price q as given and abstract from asset-tax instruments that would endogenize this price. See Hosseini and Shourideh (2019) for a model focused on asset taxation.

allocation. Note that its dependence on θ does not relax the informational constraint. The government still chooses an income-contingent tax schedule, while $\hat{T}^{F\prime}(y;\theta)$ is an equilibrium statistic summarizing the fiscal effect of an earnings response by the type θ who chooses income y. Because the government values pension benefits at $\beta p(\theta)$, whereas individuals value them at the private intertemporal price $q(\theta)$, the fiscal effective marginal tax rate can differ from the private effective marginal tax rate $\hat{T}'$ that governs individual labor-supply incentives.

The government chooses a budget-balanced income tax schedule to maximize utilitarian welfare. This objective isolates the consequences of life expectancy that operate through resource allocation and fiscal externalities, rather than building in a direct social preference to compensate individuals for shorter lifespans.[7] Accordingly, the tax function T is *optimal* if it satisfies the government budget constraint (14) and solves:

$$\max_{T} \int_{\underline{\theta}}^{\overline{\theta}} V(\theta)\, dF(\theta), \tag{17}$$

where $V(\theta)$ is the indirect utility defined in Equation (6).

2.4 Optimal control approach

Following standard procedures in the optimal taxation literature, I reformulate the government problem (17) as an optimal control problem.

First, I map the labor supply problem to a reporting problem in which the individuals report their skill to a fictitious social planner, who allocates consumption resources and income targets based on the reports. An individual with skill θ who reports skill $\hat{\theta}$ to the planner is required to work $y\left(\hat{\theta}\right)/\theta$ hours, pays taxes and pension contributions $T\left(y\left(\hat{\theta}\right)\right) + S\left(y\left(\hat{\theta}\right)\right)$, and obtains pension benefits $B\left(y\left(\hat{\theta}\right)\right)$ if she reaches retirement. Hence, if the individual reports a skill of $\hat{\theta}$, her net lifetime income minus labor disutility is

$$\hat{x}\left(\theta,\hat{\theta}\right) := y\left(\hat{\theta}\right) - T\left(y\left(\hat{\theta}\right)\right) - S\left(y\left(\hat{\theta}\right)\right) + q(\theta)\, B\left(y\left(\hat{\theta}\right)\right) - v\left(\frac{y\left(\hat{\theta}\right)}{\theta}\right). \tag{18}$$

[7] The theoretical characterization extends directly to linear social welfare functions with nonnegative, type-dependent Pareto weights $\chi(\theta)$. If these weights depend directly on life expectancy, however, quantitative comparisons across alternative life-expectancy profiles become less transparent, since changes in optimal policy would then reflect both changes in survival probabilities and changes in normative weights.

The tax system induces truthful reporting if it satisfies

$$x(\theta) := \hat{x}(\theta, \theta) = \max_{\hat{\theta}} \hat{x}\left(\theta, \hat{\theta}\right). \tag{19}$$

Following common practice (Mirrlees, 1971, 1976; Diamond, 1998; Saez, 2001), I replace the incentive compatibility constraint (19) by the associated envelope condition:

$$\frac{dx(\theta)}{d\theta} = \frac{\partial \hat{x}(\theta, \theta)}{\partial \theta} = \frac{l(\theta)}{\theta} v'(l(\theta)) + q'(\theta) B(\theta l(\theta)). \tag{20}$$

Next, I transform the government budget constraint. Using the variables x and l, Equation (14) can be expressed as

$$\int_{\underline{\theta}}^{\overline{\theta}} [\theta l(\theta) - x(\theta) - v(l(\theta)) - [\beta p(\theta) - q(\theta)] B(\theta l(\theta))] dF(\theta) - E \geq 0. \tag{21}$$

Finally, I express the objective function in terms of variables that are convenient for the control problem. As explained above, for a given level x of net lifetime income minus labor disutility, individual welfare is obtained by the following intertemporal value function:

$$V(\theta) = \tilde{V}(x(\theta); \theta) := \max_{c_2} u(x(\theta) - q(\theta) c_2) + p(\theta) \beta u(c_2). \tag{22}$$

By the envelope theorem, the marginal welfare value of a transfer is

$$\frac{\partial \tilde{V}(x(\theta); \theta)}{\partial x(\theta)} = u'(x(\theta) - q(\theta) c_2(\theta)), \tag{23}$$

where $c_2(\theta)$ solves the intertemporal problem (22). The argument $x(\theta) - q(\theta)c_2(\theta)$ equals first-period consumption net of labor disutility, $c_1(\theta) - v(l(\theta))$, which is the composite over which first-period utility is defined. To lighten the notation, I write u' for this marginal utility throughout, with the understanding that its argument is consumption net of labor disutility rather than consumption alone.

With these preparations, the optimal control problem is formulated as follows:

$$\max_{x,l} \int_{\underline{\theta}}^{\overline{\theta}} \tilde{V}\left(x\left(\theta\right);\theta\right) dF\left(\theta\right) \tag{24}$$

$$\text{subject to} \quad \int_{\underline{\theta}}^{\overline{\theta}} \left[\theta l\left(\theta\right) - x\left(\theta\right) - v\left(l\left(\theta\right)\right) - \left[\beta p\left(\theta\right) - q\left(\theta\right)\right] B\left(\theta l\left(\theta\right)\right)\right] dF\left(\theta\right) - E \geq 0,$$

$$\text{and} \quad \frac{dx}{d\theta} = \frac{l\left(\theta\right)}{\theta} v'\left(l\left(\theta\right)\right) + q'\left(\theta\right) B\left(\theta l\left(\theta\right)\right).$$

3 Theoretical characterization of optimal redistribution

This section characterizes optimal income taxation under heterogeneous life expectancy. I proceed in two steps. I begin with a simplified problem that abstracts from labor-supply responses in order to isolate how longevity heterogeneity affects the equity value of redistribution. I then return to the full model and derive an $\mathcal{ABCD}$ optimal tax formula. Relative to the classic $\mathcal{ABC}$ formula of Diamond (1998), heterogeneous life expectancy affects optimal taxation through two distinct mechanisms: on the equity side, it changes the mapping from disposable income to social marginal welfare; on the efficiency side, it introduces a correction for pension-induced fiscal externalities. I conclude by discussing optimal pension design and multi-period life cycles.

3.1 Mechanical redistributive effect

Before deriving the full optimal tax formula, it is useful to isolate how longevity heterogeneity affects the mechanical value of redistribution. I therefore start with a simplified problem in which labor-supply responses are absent.

Consider the two-period environment of Section 2, but suppose that gross income $y(\theta)$ is fixed. Individuals pay income taxes $T\left(y\left(\theta\right)\right)$ and public pension contributions $S\left(y\left(\theta\right)\right)$, and receive public pension benefits $B\left(y\left(\theta\right)\right)$. They can transfer private resources to retirement at the intertemporal price $q\left(\theta\right)$. Since labor supply is exogenous, labor disutility v is no longer a choice-relevant margin; I therefore normalize it to zero to simplify the exposition.

Let

$$x^{M}\left(\theta\right) := y\left(\theta\right) - T\left(y\left(\theta\right)\right) - S\left(y\left(\theta\right)\right) + q\left(\theta\right) B\left(y\left(\theta\right)\right) \tag{25}$$

denote net lifetime income in this mechanical redistribution model. The individual problem is

$$\tilde{V}^M\left(x^M(\theta);\theta\right) := \max_{c_2^M} u\left(x^M(\theta) - q(\theta)c_2^M\right) + \beta p(\theta) u\left(c_2^M\right). \tag{26}$$

By the envelope theorem,

$$\frac{\partial \tilde{V}^M\left(x^M(\theta);\theta\right)}{\partial x^M(\theta)} = u'\left(c_1^M(\theta)\right), \tag{27}$$

where

$$c_1^M(\theta) := x^M(\theta) - q(\theta)c_2^M(\theta) \tag{28}$$

denotes working-life consumption at the optimum. Hence, the planner values a marginal transfer to type θ according to $u'\left(c_1^M(\theta)\right)$. In contrast to a static model, the social value of a transfer therefore depends not only on disposable income, but also on how this income is allocated over the life cycle.

The intertemporal allocation is characterized by the first-order condition

$$q(\theta)u'\left(c_1^M(\theta)\right) = \beta p(\theta) u'\left(c_2^M(\theta)\right). \tag{29}$$

This condition highlights two channels through which longevity heterogeneity shapes the consumption profile. First, the survival probability $p(\theta)$ affects the expected marginal value of retirement consumption. Second, the private intertemporal price $q(\theta)$ determines the cost, in working-life resources, of privately financed retirement consumption.

The following proposition summarizes these two effects separately, without imposing a specific pricing rule for retirement resources.

Proposition 1 (Survival and price effects on marginal welfare) *Holding $y(\theta)$, $T(y(\theta))$, $S(y(\theta))$, and $B(y(\theta))$ fixed, the following results hold. First, a higher survival probability increases retirement consumption, reduces working-life consumption, and raises marginal welfare:*

$$\frac{\partial c_2^M(\theta)}{\partial p(\theta)} > 0, \qquad \frac{\partial c_1^M(\theta)}{\partial p(\theta)} < 0, \qquad \frac{\partial u'\left(c_1^M(\theta)\right)}{\partial p(\theta)} > 0. \tag{30}$$

Second, a higher private intertemporal price has an ambiguous effect on retirement consumption, working-life consumption, and marginal welfare. If privately financed retirement consumption is

positive, i.e., if $c_2^M(\theta) > B(y(\theta))$*, then a higher private intertemporal price reduces retirement consumption:*

$$\frac{\partial c_2^M(\theta)}{\partial q(\theta)} < 0. \tag{31}$$

Its effect on working-life consumption satisfies:

$$\frac{\partial c_1^M(\theta)}{\partial q(\theta)} < 0 \qquad \Longleftrightarrow \qquad 1 - \frac{B(y(\theta))}{c_2^M(\theta)} > \psi\left(c_2^M(\theta)\right), \tag{32}$$

where $\psi(c) := -u'(c)/(cu''(c))$ *is the intertemporal elasticity of substitution.*

Proposition 1 shows how longevity heterogeneity shapes the mechanical value of redistribution by changing the intertemporal consumption profile. The survival-probability channel is unambiguous. A higher probability of reaching retirement raises the expected marginal value of retirement consumption. Individuals therefore shift resources toward retirement, which lowers working-life consumption and raises the marginal welfare value of a transfer.

The intertemporal-price channel is more subtle. A higher private intertemporal price makes retirement consumption more expensive relative to working-life consumption. This substitution effect pushes consumption away from retirement. At the same time, a higher private intertemporal price reduces real lifetime income to the extent that retirement consumption is privately financed. This income effect pushes consumption downward in both periods. Pension benefits partly offset this income effect because a higher private intertemporal price also raises the working-life value of the pension claim B.

For retirement consumption, the substitution effect and the income effect work in the same direction when privately financed retirement consumption is positive. For working-life consumption, they work in opposite directions when privately financed retirement consumption is positive: the substitution effect raises working-life consumption, whereas the negative income effect lowers it. Equation (32) reflects the relative strength of these two effects by comparing the intertemporal elasticity of substitution, ψ, to the privately financed share of retirement consumption, $1 - B/c_2^M$. Working-life consumption falls with the private intertemporal price only when the privately financed share of retirement consumption is sufficiently large relative to the willingness to substitute consumption across periods. As a result, the role of the private intertemporal price for marginal welfare remains ambiguous.

In summary, higher survival probabilities make retirement consumption more valuable, while higher private intertemporal prices make retirement consumption more costly and thereby generate substitution and income effects. Because these forces can point in different directions, the composite effect of longevity heterogeneity on the marginal welfare value of a transfer is not signed in general.

To determine this composite effect, one must specify how private intertemporal prices vary with survival probabilities. The most transparent benchmark is actuarially fair annuity pricing:

$$q(\theta) = \beta p(\theta). \tag{33}$$

In this case, a higher survival probability proportionally raises the private intertemporal price. The next result characterizes the resulting composite comparative static.

Proposition 2 (Composite marginal welfare effect under fair annuity pricing) *Suppose annuities are actuarially fairly priced, $q(\theta) = \beta p(\theta)$. Holding $y(\theta)$, $T(y(\theta))$, $S(y(\theta))$, and $B(y(\theta))$ fixed, the allocation satisfies*

$$c_1^M(\theta) = c_2^M(\theta) = \frac{z(\theta) + \beta p(\theta) B(y(\theta))}{1 + \beta p(\theta)}, \tag{34}$$

where $z(\theta) := y(\theta) - T(y(\theta)) - S(y(\theta))$ denotes net income before pension benefits. Hence, if privately financed retirement consumption is positive, i.e., if $c_2^M(\theta) > B(y(\theta))$, then

$$\frac{\partial c_1^M(\theta)}{\partial p(\theta)} < 0 \qquad \textit{and} \qquad \frac{\partial u'\left(c_1^M(\theta)\right)}{\partial p(\theta)} > 0. \tag{35}$$

Under actuarially fair annuity pricing, a higher survival probability therefore raises marginal welfare whenever individuals privately finance part of retirement consumption.

Proposition 2 clarifies how the two separate comparative-static channels combine under fair annuity pricing. When p and q vary independently, a higher survival probability changes the implicit weight attached to retirement consumption, while a higher private intertemporal price affects the allocation through substitution and income effects. These forces need not point in the same direction. Under actuarially fair pricing, however, the intertemporal price equals the survival-weighted discount factor attached to retirement consumption. The intertempo-

ral first-order condition therefore reduces to perfect consumption smoothing, $c_1^M(\theta) = c_2^M(\theta)$, irrespective of the level of longevity.

This condition gives the comparative statics a particularly simple interpretation. Longevity no longer changes the relative allocation between working-life and retirement consumption. Instead, it affects welfare solely by changing the amount of consumption that can be sustained in each expected period of life. If retirement consumption is partly privately financed, a higher survival probability spreads any given amount of private resources over a longer expected lifetime. Consumption therefore falls, and the marginal welfare value of a transfer rises.

Assuming a positive income-longevity gradient, the key takeaway is that heterogeneous life expectancy weakens the equity motive for redistribution. Under actuarially fair annuity pricing, unless pension benefits fully cover retirement consumption, a transfer to a longer-lived individual has higher marginal welfare value than the same transfer would have in an otherwise identical environment with homogeneous life expectancy. Heterogeneous life expectancy consequently attenuates the usual decline of marginal welfare weights with income when higher-income individuals are longer lived.

Although this argument was derived under actuarially fair annuity pricing, a related conclusion arises under the second benchmark of uniform intertemporal pricing,

$$q(\theta) = \bar{q}. \tag{36}$$

This benchmark captures environments in which individuals face a common price for retirement resources, as under pooled annuity pricing or non-annuitized storage technologies. In this case, the intertemporal-price channel is shut down because q no longer varies with longevity. The survival-probability channel remains operative: a higher probability of reaching retirement raises the expected marginal value of retirement consumption, shifts resources toward retirement, lowers working-life consumption, and raises the marginal welfare value of a transfer. Thus, when higher earners live longer, this channel again dampens the usual decline of marginal welfare weights with income. Unlike in the benchmark with fair annuities, however, this conclusion does not require retirement consumption to be partly privately financed.[8]

[8]Note that the survival-probability channel operates even when individuals borrow against future pension benefits. Ceteris paribus, a higher survival probability reduces such borrowing, lowers working-life consumption, and thereby raises the marginal welfare value of a transfer.

3.2 The $\mathcal{ABCD}$ formula

I now return to the full two-period model and derive the optimal tax formula from the first-order conditions of the planner's control problem (24).[9] The next lemma collects the behavioral elasticities needed to interpret the optimality condition.

Lemma 1 (Elasticities) *Assuming a linearized tax and pension system, i.e., $T'' = S'' = B'' = 0$, the elasticities of taxable income are as follows:*

$$\varepsilon_{y,\theta} = 1 + \frac{v'(l)}{lv''(l)} + \frac{\theta^2 q'(\theta) B'}{lv''(l)} \quad , \tag{37}$$

$$\varepsilon_{y,(1-T')} = \frac{v'(l)}{lv''(l)} + \frac{\theta(S' - q(\theta) B')}{lv''(l)} \quad , \tag{38}$$

$$\varepsilon_{y,(1-\hat{T}')} = \frac{v'(l)}{lv''(l)} \quad . \tag{39}$$

In particular, if the privately valued pension wedge is zero:

$$\varepsilon_{y,(1-T')}\big|_{\omega=0} = \frac{v'(l)}{lv''(l)}. \tag{40}$$

I can now state the optimality condition. Equation (41) is written in terms of the skill distribution and uses the fiscal effective marginal tax rate, $T' + \omega^F$, to value the budgetary effect of marginal earnings responses. Equation (42) rearranges this condition into an $\mathcal{ABCD}$ form whose left-hand side matches the familiar static expression $T'/(1-T')$. This statutory-tax representation preserves comparability with the classic Diamond (1998) formula, while the correction factor $\mathcal{D}$ accounts for the fact that the fiscal wedge differs from the statutory tax wedge. Equation (43) rewrites the same condition in terms of the endogenous income distribution and connects to the sufficient-statistics literature (Saez, 2001; Chetty, 2009).

Proposition 3 ($\mathcal{ABCD}$ formula) *If the tax function T is optimal, taking the public pension system (S, B) as given, then it satisfies*

$$\frac{T' + \omega^F}{1 - T'} = \frac{\varepsilon_{y,\theta}}{\varepsilon_{y,(1-T')}} \left(1 - \frac{\int_\theta^{\bar{\theta}} u' \, dF}{\lambda(1 - F)}\right) \frac{1 - F}{\theta f} \tag{41}$$

[9] Appendix A.2 provides a complementary derivation using a tax perturbation approach in the spirit of Saez (2001).

where λ is the multiplier of the government budget constraint (21)*, $\omega^F = S' - \beta p B'$ is the fiscal pension wedge, and u' is the marginal utility of working-life consumption net of labor disutility. Equivalently, this condition can be rearranged into the following $\mathcal{ABCD}$ form:*

$$\frac{T'}{1-T'} = \frac{\varepsilon_{y,\theta}}{\varepsilon_{y,(1-T')}} \left(1 - \frac{\int_\theta^{\overline{\theta}} u' \, dF}{\lambda (1-F)}\right) \frac{1-F}{\theta f} \frac{T'}{T' + \omega^F}. \tag{42}$$

In terms of the endogenous income distribution H, with density $h = H'$, the same condition can be expressed as

$$\frac{T'}{1-T'} = \frac{1}{\varepsilon_{y,(1-T')}} \left(1 - \frac{\int_y^{\overline{y}} u' \, dH}{\lambda (1-H(y))}\right) \frac{(1-H(y))}{y \cdot h(y)} \frac{T'}{T' + \omega^F}. \tag{43}$$

Relative to the classic $\mathcal{ABC}$ formula of Diamond (1998), Proposition 3 contains two new effects. First, as discussed above, longevity heterogeneity modifies the mechanical value of redistribution,

$$\mathcal{B} := 1 - \frac{\int_\theta^{\overline{\theta}} u' \, dF}{\lambda (1-F)}, \tag{44}$$

where $u' = u'(x(\theta) - q(\theta) c_2)$ is the marginal utility of working-life consumption net of labor disutility, evaluated at the individual's intertemporal optimum. Second, the formula includes the fiscal pension wedge correction,

$$\mathcal{D} := \frac{T'}{T' + \omega^F}, \tag{45}$$

which captures the fact that marginal earnings responses affect not only income tax revenue, but also pension contributions and expected pension liabilities. The remaining terms,

$$\mathcal{A} := \frac{\varepsilon_{y,\theta}}{\varepsilon_{y,(1-T')}} \quad \text{and} \quad \mathcal{C} := \frac{1-F}{\theta f}, \tag{46}$$

retain their standard roles and capture, respectively, the elasticity of the tax base and the proportions between marginal and inframarginal taxpayers. For completeness, I briefly review their roles in Appendix A.3.

Like every characterization of this kind, Proposition 3 provides implicit optimality conditions rather than closed-form solutions. The right-hand sides of all formulations depend on the tax system through endogenous incomes, marginal utilities, and behavioral responses, so $\mathcal{A}$, $\mathcal{B}$, and

$\mathcal{D}$ are endogenous objects, as is $\mathcal{C}$ in its income-based formulation. The factor $\mathcal{D} = T'/(T' + \omega^F)$ shows this endogeneity most explicitly, since it refers directly to the marginal tax rate, whereas in the remaining terms the dependence on the marginal tax rate enters only indirectly. In the discussion that follows, I focus on the $\mathcal{ABCD}$ form because it preserves comparability with the classic $\mathcal{ABC}$ formula and isolates the pension-induced correction as a single transparent term.

The term $\mathcal{B}$ is the full-model counterpart of the equity mechanism isolated in Section 3.1. It captures the mechanical welfare effect of a marginal tax increase, abstracting from behavioral responses, by comparing the average marginal welfare weight, u'/λ, among taxpayers above skill level θ to the population-wide average (normalized to one). The standard redistributive logic is unchanged: a marginal tax increase is more attractive when the inframarginal taxpayers who finance it have relatively low marginal welfare weights.

What changes in the life-cycle model is the mapping from disposable income into marginal welfare weights. The planner evaluates a transfer through its effect on working-life consumption at the individual intertemporal optimum. This optimum depends on survival probabilities and on the private intertemporal price of retirement resources. The benchmark cases in Section 3.1 show that both actuarially fair annuity pricing and uniform intertemporal pricing generate the same qualitative implication: longer-lived individuals tend to have lower working-life consumption and hence a higher marginal welfare value of a transfer. Under actuarially fair annuity pricing, this result requires that retirement consumption be at least partly privately financed. Under uniform intertemporal pricing, the result follows from the survival-probability channel alone and therefore does not require privately financed retirement consumption. Thus, if life expectancy rises with income, longevity heterogeneity tends to attenuate the usual decline of marginal welfare weights with income and thereby weakens the mechanical redistributive gain from higher marginal tax rates.

While the $\mathcal{B}$-term captures the welfare value of transfers generated by a marginal tax increase, the $\mathcal{D}$-term concerns the fiscal value of associated earnings responses. Changes in the statutory marginal tax rate T' affect incomes and, with them, government revenues not only through taxes but also through pension contributions and future benefit entitlements. This total fiscal effect is summarized by the fiscal effective marginal tax rate, $T' + \omega^F$. The $\mathcal{D}$-term therefore serves as a correction factor that maps the total fiscal wedge into the statutory tax

formula. Its implication is intuitive. When the pension system generates a positive local fiscal externality, i.e., when $\omega^F > 0$ or equivalently $S' > \beta p B'$, an increase in income raises government revenue by more than the statutory tax component. For a given statutory marginal tax rate, behavioral reductions in earnings are then more costly in fiscal terms, because they reduce not only tax revenue but also the fiscal surplus generated by the pension system. In that case, $\mathcal{D}$ lowers the right-hand side of the optimality condition, which tends to depress the optimal statutory marginal tax rate.

In income ranges where the fiscal pension wedge, $\omega^F = S' - \beta p B'$, increases with income, for example because the benefit schedule is progressive, this mechanism implies relatively higher marginal tax rates at lower income levels and relatively lower rates at higher income levels. However, the income-longevity gradient modifies this wedge profile. For any given marginal pension benefit $B' > 0$, a higher survival probability lowers the fiscal pension wedge ω^F. Hence, when life expectancy rises with income, survival heterogeneity tends to compress the fiscal wedge, dampening the slope that would otherwise arise from a progressive benefit schedule.

To formalize the role of life-expectancy inequality, consider a mean-preserving spread of a baseline survival-probability schedule $p_0(\theta)$ with mean $\bar{p}$, given by

$$p(\theta) = p_0(\theta) + \eta\left(p_0(\theta) - \bar{p}\right), \tag{47}$$

where $\eta > 0$ parameterizes the extent of life-expectancy inequality. The effect of this spread on the $\mathcal{D}$-term is characterized by the following result.

Proposition 4 (Comparative statics of $\mathcal{D}$) *Let $p(\theta) = p_0(\theta) + \eta\left(p_0(\theta) - \bar{p}\right)$ and define $\mathcal{D} = T'(y(\theta)) \,/\, \left(T'(y(\theta)) + \omega^F(y(\theta), \theta)\right)$. Holding $y(\theta)$, $T'(y(\theta))$, $S'(y(\theta))$, and $B'(y(\theta))$ fixed, and assuming $T'(y(\theta)) > 0$ and $B'(y(\theta)) > 0$, the comparative statics of $\mathcal{D}$ with respect to the spread parameter η are as follows:*

$$\frac{\partial \mathcal{D}}{\partial \eta} \gtreqless 0 \quad \Longleftrightarrow \quad p_0(\theta) \gtreqless \bar{p}. \tag{48}$$

Moreover, $\frac{\partial \mathcal{D}}{\partial \eta} = 0$ if $B'(y(\theta)) = 0$.

Proposition 4 shows that life-expectancy inequality alters optimal marginal tax rates in a skill-dependent manner through the fiscal-wedge channel. For individuals with below-average

life expectancy, an increase in inequality lowers $p(\theta)$. This raises the fiscal pension wedge, thereby increasing the fiscal externality associated with behavioral earnings responses. As a result, $\mathcal{D}$ falls, which tends to lower the statutory marginal tax rate implied by the $\mathcal{ABCD}$ formula. For individuals with above-average life expectancy, the effect is reversed. For them, the spread raises $p(\theta)$, lowers the fiscal pension wedge, raises $\mathcal{D}$, and thereby tends to increase the optimal statutory marginal tax rate.

This channel disappears when the marginal pension replacement rate is zero, that is, when additional income no longer affects future pension benefits. In that case, the fiscal pension wedge is unaffected by life-expectancy inequality, and the $\mathcal{D}$-term no longer transmits longevity differences into optimal marginal tax rates. Note that this invariance is specific to the fiscal-wedge channel; longevity heterogeneity may still affect optimal marginal tax rates through the $\mathcal{B}$-term.

Summing up, the theory identifies two channels through which unequal life expectancy affects optimal taxation, even in the absence of any direct normative concern for longevity. The $\mathcal{D}$-term captures how longevity modifies the efficiency cost of taxation through pension-linked fiscal externalities, while the $\mathcal{B}$-term captures how it alters the mechanical value of redistribution by changing the mapping from disposable income into marginal welfare weights. The quantitative analysis below assesses the empirical importance of these channels.

3.3 Extensions

Next, I extend the $\mathcal{ABCD}$ characterization to two alternative environments. The first extension allows the policy maker to jointly determine income taxation and pension policy. The second extension generalizes the model to a multi-period horizon.

3.3.1 Optimal pension system

If the policy maker jointly chooses the income tax schedule T and the public pension system (S, B), the separate marginal rates T', S', B' are not uniquely pinned down. What matters for individual labor-supply incentives is their composition into the private effective marginal tax rate $\hat{T}'$. By contrast, what matters for public revenues is their composition into the fiscal effective marginal tax rate $\hat{T}^{F\prime}$.

Proceeding as in the proof of Proposition 3, and using the elasticity of income with respect to the private effective net-of-tax rate from Equation (39), the optimality condition can be formulated as

$$\frac{\hat{T}^{F\prime}}{1-\hat{T}'} = \frac{\varepsilon_{y,\theta}}{\varepsilon_{y,\left(1-\hat{T}'\right)}} \left(1 - \frac{\int_{\theta}^{\overline{\theta}} u'\, dF}{\lambda\,(1-F)}\right) \frac{1-F}{\theta f}. \tag{49}$$

Equivalently, this condition can be rearranged into an $\mathcal{ABCD}$-type formula for the private effective marginal tax rate:

$$\frac{\hat{T}'}{1-\hat{T}'} = \frac{\varepsilon_{y,\theta}}{\varepsilon_{y,\left(1-\hat{T}'\right)}} \left(1 - \frac{\int_{\theta}^{\overline{\theta}} u'\, dF}{\lambda\,(1-F)}\right) \frac{1-F}{\theta f} \frac{\hat{T}'}{\hat{T}^{F\prime}}. \tag{50}$$

The underlying logic mirrors the statutory $\mathcal{ABCD}$ formula derived above. The $\mathcal{B}$-term continues to capture the mechanical redistributive effect of longevity heterogeneity. The pension wedge correction term now compares the private effective marginal tax rate $\hat{T}'$, which governs individual incentives, with the fiscal effective marginal tax rate $\hat{T}^{F\prime}$, which governs the budgetary effects of earnings responses. Under actuarially fair annuity pricing, these effective rates coincide and the correction term equals one. Otherwise, the incentive and fiscal effects of the pension system must be distinguished.

3.3.2 Multi-period horizon

The two-period model facilitates the exposition. I next show that the same logic extends to a stylized multi-period environment with age-dependent wages, preferences, and survival rates, and period-by-period reporting.

In this setting, individuals live for at most N periods. They work in periods $t = 1, \ldots, R-1$ and are retired in periods $t = R, \ldots, N$. Individuals differ by a one-dimensional type $\theta \in \left(\underline{\theta}, \overline{\theta}\right) \subset \mathbb{R}_{++}$, which determines the wage rates $w_t(\theta)$ in working periods $t = 1, \ldots, R-1$ and the unconditional probabilities $p_t(\theta)$ of being alive in periods $t = 1, \ldots, N$. Individuals have access to a linear technology for intertemporal transfers. Let $q_{t+1}(\theta)$ denote the private price, in period t, of one unit of resources in period $t+1$. The corresponding period-t price of resources in period s is $q^{t,s}(\theta) := 1$ for $s = t$ and

$$q^{t,s}(\theta) := \prod_{\tau=t+1}^{s} q_{\tau}(\theta) \text{ for } s > t. \tag{51}$$

In working period $t < R$, an individual of type θ chooses labor supply $l_t(\theta)$ and earns $y_t(\theta) = w_t(\theta)l_t(\theta)$. The individual pays income taxes $T_t\left(y_t\left(\theta\right)\right)$ and pension contributions $S_t\left(y_t\left(\theta\right)\right)$. In retirement period $t \geq R$, the individual receives a stationary pension benefit $B\left(y_1\left(\theta\right), \ldots, y_{R-1}\left(\theta\right)\right)$, which is calculated based on the income history according to the additive benefit formula

$$B\left(y_1\left(\theta\right), \ldots, y_{R-1}\left(\theta\right)\right) = \sum_{\tau=1}^{R-1} b_\tau\left(y_\tau\left(\theta\right)\right). \tag{52}$$

Appendix A.4 provides the detailed setup of the individual and planner optimization problems in this multi-period setting. The planner problem can be stated as an optimal control problem with state vector $(x_1, \ldots, x_{R-1})$ and control vector $(l_1, \ldots, l_{R-1})$:

$$\max_{x,l} \int_{\underline{\theta}}^{\overline{\theta}} \tilde{V}\left(x_1\left(\theta\right), \ldots, x_{R-1}\left(\theta\right); \theta\right) dF\left(\theta\right) \tag{53}$$

subject to

$$\int_{\underline{\theta}}^{\overline{\theta}} \sum_{t=1}^{R-1} \beta^{t-1} p_t\left(\theta\right) \left[w_t\left(\theta\right) l_t\left(\theta\right) - x_t\left(\theta\right) - v_t\left(l_t\left(\theta\right)\right)\right.$$
$$\left. + \left[Q_t\left(\theta\right) - P_t^F\left(\theta\right)\right] b_t\left(w_t\left(\theta\right) l_t\left(\theta\right)\right)\right] dF\left(\theta\right) - E \geq 0,$$
$$\text{and} \quad \frac{dx_t}{d\theta} = l_t\left(\theta\right) \frac{w_t'\left(\theta\right)}{w_t\left(\theta\right)} v_t'\left(l_t\left(\theta\right)\right) + Q_t'\left(\theta\right) b_t\left(w_t\left(\theta\right) l_t\left(\theta\right)\right).$$

Here,

$$Q_t\left(\theta\right) := \sum_{\tau=R}^{N} q^{t,\tau}\left(\theta\right) \tag{54}$$

is the private value of a unit of annual pension benefits generated by income in period t, whereas

$$P_t^F\left(\theta\right) := \sum_{\tau=R}^{N} \beta^{\tau-t} \frac{p_\tau\left(\theta\right)}{p_t\left(\theta\right)} \tag{55}$$

is its expected fiscal cost, both expressed in period-t resources.

The optimality conditions for the tax system are straightforward extensions of the $\mathcal{ABCD}$ formula given in Proposition 3, as the next result shows.

Proposition 5 ($\mathcal{ABCD}$ formula for N periods) *Consider the model extension with N periods. If the tax functions $T_1, \ldots, T_{R-1}$ are optimal, taking the public pension system as given,*

then for each $t = 1, \ldots, R-1$ *they satisfy*

$$\frac{T_t'}{1-T_t'} = \frac{\varepsilon_{y_t,w_t}}{\varepsilon_{y_t,(1-T_t')}} \frac{\int_\theta^{\bar{\theta}} \left[1 - \frac{u_t'(\tilde{\theta})}{\lambda}\right] p_t\left(\tilde{\theta}\right) dF\left(\tilde{\theta}\right)}{(1-F)\, p_t} \frac{(1-F)\, w_t'}{f w_t} \frac{T_t'}{T_t' + \omega_t^F}, \tag{56}$$

where λ *is the multiplier of the intertemporal government budget constraint,* $\omega_t^F := S_t' - P_t^F b_t'$ *is the fiscal wedge in period* t*, and* $u_t' = u_t'\left(c_t - v_t\left(l_t\right)\right)$ *is the marginal utility of period-*t *consumption net of labor disutility.*

A comparison of Equations (42) and (56) shows that the multi-period formula preserves the structure of the two-period characterization. The mechanical value of redistribution is now adjusted for type-dependent survival to period t. That is, the tax revenue collected from inframarginal types above θ is weighted by their probability $p_t\left(\tilde{\theta}\right)$ of being alive in that period, relative to the survival probability $p_t(\theta)$ of the marginal type. Moreover, the distributional term is the inverse Pareto coefficient of the period-t wage distribution. Note that the wage in period t is distributed according to $\tilde{F}_t\left(w_t(\theta)\right) = F(\theta)$. Differentiation of this condition yields $\tilde{f}_t\left(w_t(\theta)\right) w_t'(\theta) = f(\theta)$. Hence, the inverse Pareto coefficient of the period-t wage distribution equals

$$\frac{1 - \tilde{F}_t\left(w_t(\theta)\right)}{\tilde{f}_t\left(w_t(\theta)\right) w_t(\theta)} = \frac{(1 - F(\theta))\, w_t'(\theta)}{f(\theta)\, w_t(\theta)}. \tag{57}$$

Thus, relative to the two-period formula, the $\mathcal{B}$-term incorporates survival-weighted mechanical redistribution, while the $\mathcal{C}$-term maps the type distribution into the relevant period-t wage distribution. The remaining terms retain their earlier interpretation: the $\mathcal{A}$-term captures the strength of behavioral responses, and the $\mathcal{D}$-term converts the fiscal effective wedge into the statutory income-tax wedge.

4 Quantitative evaluation

This section evaluates the quantitative implications of the $\mathcal{ABCD}$ formula.[10] I calibrate the model to the U.S. economy and simulate optimal tax systems under alternative assumptions about life expectancy and private retirement asset markets. The analysis first shows that the

[10]The computational approach follows Mankiw et al. (2009) and iterates the right-hand side of the optimality condition until a fixed point is found. To confirm the optimality of the fixed point tax policy, I verify that the resulting income profile $y(\theta)$ is increasing in skill θ.

fiscal pension wedge is quantitatively important even when life expectancy is homogeneous: ignoring this wedge leads to misspecified optimal tax schedules because earnings responses affect public revenues through both income taxes and the pension system. I then assess how accounting for life-expectancy inequality further changes optimal tax policy and its welfare consequences.

I consider two benchmark assumptions about the private intertemporal pricing of retirement resources. The first is actuarially fair annuity pricing, under which the private price of retirement resources equals their expected discounted payout, $q(\theta) = \beta p(\theta)$. The second is uniform intertemporal pricing, under which all individuals face a common price $q(\theta) = \bar{q}$, as would arise under pooled annuity pricing or a non-annuitized storage technology. In the uniform pricing case, I set $\bar{q} = \beta \bar{p}$, where $\bar{p}$ is the average survival probability.

4.1 Parameterization

The utility function is $u(c) = \frac{1}{1-\sigma} c^{1-\sigma}$, with a coefficient of relative risk-aversion given by $\sigma = 2$. Considering that individuals retire approximately 40 years after entering the labor force, I assume that each model period has a duration of 40 years and set the discount factor to $\beta = 0.98^{40}$, corresponding to an annual risk-free interest rate of two percent. Labor disutility is of the iso-elastic form $v(l) = \left(1 + \frac{1}{\phi}\right)^{-1} l^{1+\frac{1}{\phi}}$ with parameter $\phi = 0.33$ (Chetty, 2012). With this specification, the labor supply problem admits a closed-form solution, yielding

$$\theta = y^{\frac{1}{1+\phi}} \left[1 - T' - S' + qB'\right]^{-\frac{\phi}{1+\phi}}. \tag{58}$$

I use this relationship to infer the skill distribution from the income distribution.

Following Sachs et al. (2020) and Koehne and Sachs (2022), I model the baseline income distribution as a lognormal distribution with an appended Pareto tail. Incomes up to \$225,000 are drawn from a lognormal distribution with parameters $\tilde{\mu} = 10.741$ and $\tilde{\sigma} = 0.947$. The parameter $\tilde{\mu}$ is chosen to match the 2022 U.S. per capita income of \$77,800, while $\tilde{\sigma}$ is set to obtain a Pareto coefficient of 2.2 at the boundary of the lognormal segment. The Pareto tail begins at \$225,000, with a Pareto coefficient declining from $\Pi = 2.2$ to a constant level of $\Pi = 1.5$ for incomes above \$525,000.[11]

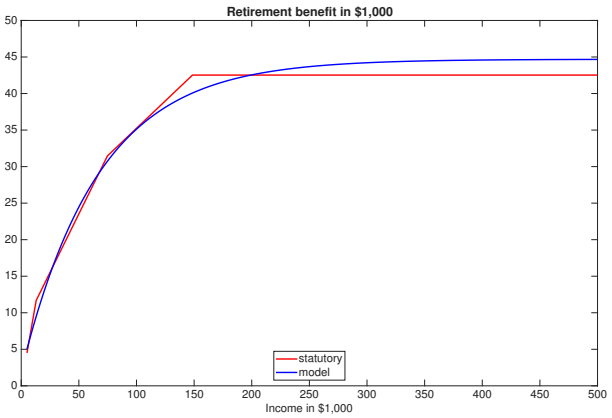
[11] The pattern of Pareto coefficients follows Diamond and Saez (2011), adjusted to 2022 dollars.

The baseline tax system has the standard parametric form $T(y) = y - \nu y^{1-\tau}$ used, for example, by Benabou (2002) and Heathcote et al. (2017). Because the Social Security system is modeled separately in the model, I choose a parameter of $\tau = 0.151$ that represents the progressivity of the U.S. tax and transfer system abstracting from intergenerational redistribution (Heathcote et al., 2014). The parameter ν controls the level of taxation and is chosen to match the average rate of the U.S. income tax. Recent statistics published by the IRS (2022) report an adjusted gross income of \$12.592 trillion and a total income tax of \$1.711 trillion, generating an average tax rate of 13.6 percent. To match this rate, I set $\nu = 5.121$.

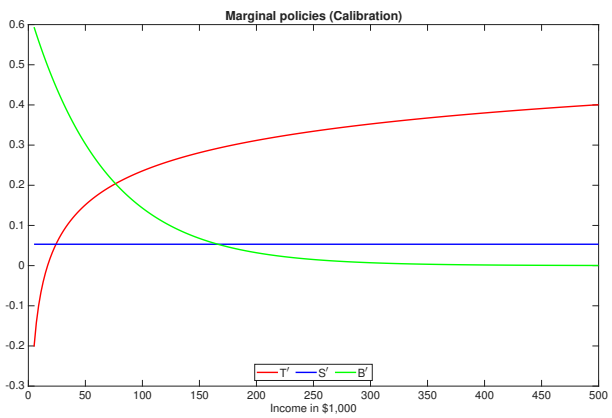
Retirement benefits follow the progressive formula of the U.S. Social Security system: marginal replacement rates of 90 percent for incomes below \$12,288, 32 percent between \$12,288 and \$74,064, and 15 percent between \$74,064 and the maximum creditable income of \$147,000 (SSA, 2022). I approximate this schedule using $B(y) = \delta e^{\varepsilon \cdot y} + \gamma$, with parameters $\delta = -42749.162$, $\varepsilon = -1.497 \cdot 10^{-5}$ and $\gamma = 44690.990$ obtained by least-squares estimation. This function provides a close fit of the statutory benefits (Figure 1a), yielding an R^2 of 0.994. Benefits are financed by a budget-balancing uniform contribution rate of 5.32 percent. Given the approximated tax and pension system (Figure 1b) and the lognormal-Pareto distribution of income, I use Equation (58) to calibrate the skill distribution following the approach of Saez (2001).

The probability $p(\theta)$ of reaching the retirement phase is calibrated based on estimates of male life expectancy by income percentile reported by Chetty et al. (2016). To approximate the remaining life expectancy of individuals retiring in 2022, I subtract 66 years and 4 months (the full retirement age for the cohort born in 1956) from the estimated life expectancy. I then fit a cubic spline to construct a smooth schedule across the income distribution. Finally, I convert life expectancies into survival probabilities assuming that the retirement period has a duration of 40 years. Following these steps, I obtain survival probabilities that range from 0.25 for bottom earners to 0.52 for top earners (Figure 1c), corresponding to expected retirement durations between 10 and 21 years. The homogeneous scenario sets the survival probability to the mean value $\bar{p}$, while the heterogeneous scenario uses the full type-dependent profile p.

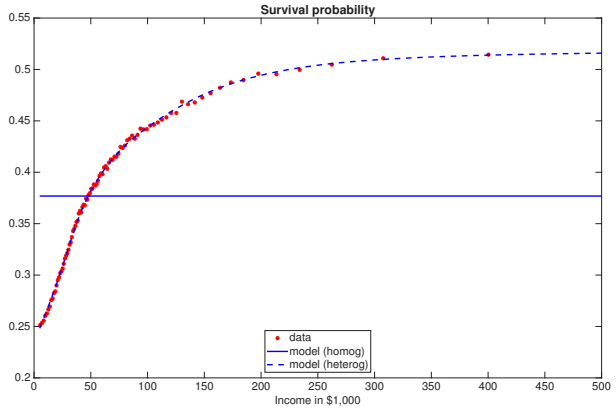
Given the baseline tax and pension system and survival probabilities, I infer government expenditures E from the government budget constraint (14), and I hold E fixed across policies, balancing the budget in each case through a uniform lump-sum transfer. Finally, I consider two



(a) Exponential approximation of retirement benefits



(b) Baseline marginal policy rates



(c) Survival probability p in model with homogeneous life expectancy (solid line) and heterogeneous life expectancy (dashed line).

Figure 1: Baseline policy and survival probability. The baseline income tax uses $T(y) = y - \nu y^{1-\tau}$ with $\tau = 0.151$ (progressivity) and $\nu = 5.121$ (level); retirement benefits use an exponential approximation of the Social Security schedule. Survival probabilities are calibrated from life expectancy by income percentile reported by Chetty et al. (2016).

benchmarks for the private intertemporal price of retirement resources: actuarially fair annuity pricing, with $q(\theta) = \beta p(\theta)$, and uniform intertemporal pricing, with $q(\theta) = \beta \bar{p}$, where $\bar{p}$ denotes the mean of $p(\theta)$ across types.

4.2 Optimal tax policy under homogeneous life expectancy

I begin by examining optimal taxes under the simplifying assumption of homogeneous survival probabilities, serving as a baseline for the subsequent analyses. Given the calibrated parameters, I solve for the optimal tax policy under actuarially fair annuity pricing. Because all individuals have identical survival rates, this case coincides with the uniform pricing benchmark.

The solid line in Figure 2 displays the optimal tax policy. The optimal marginal tax rates

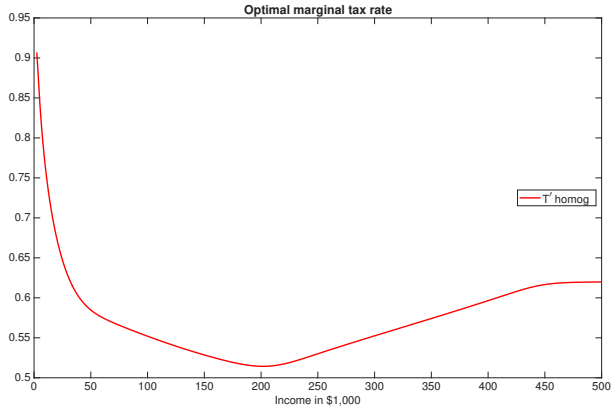


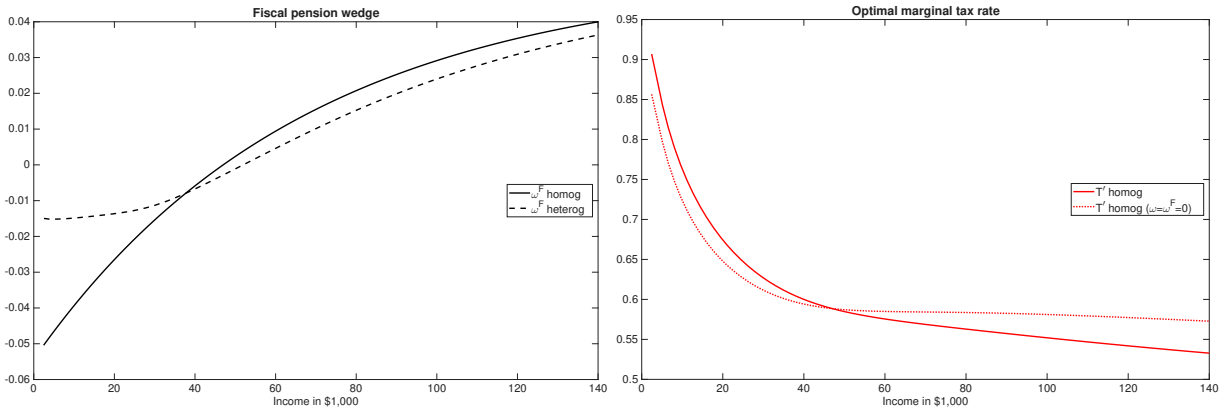
Figure 2: Optimal marginal tax rate under homogeneous life expectancy

are approximately U-shaped, consistent with typical simulations of Mirrleesian models (Saez, 2001; Mankiw et al., 2009; Sachs et al., 2020). The optimal marginal tax rates range from 91 percent for bottom earners to 51 percent for middle earners. For top earners in the Pareto tail, the optimal marginal tax rates converge to 62 percent.

To quantify the impact of the fiscal pension wedge on the optimal tax system under homogeneous life expectancy, I compare the baseline optimum to a counterfactual optimum in which marginal pension wedges are set to zero throughout the income distribution. In this counterfactual, labor supply incentives are governed by the statutory marginal tax rate alone, and earnings affect government revenue only through statutory taxation. Figure 3a shows that, in the baseline economy, the fiscal pension wedge rises with income, reflecting the progressivity of the benefit schedule. Ignoring this wedge overstates the fiscal cost of earnings responses at low incomes, where high marginal pension accruals partly offset the fiscal return to earnings, and understates the fiscal cost of earnings responses at higher incomes, where marginal pension contributions exceed the expected fiscal cost of marginal pension benefits. As shown by the dotted line in Figure 3b, the counterfactual optimum therefore features lower marginal tax rates for low earners and higher marginal tax rates for high earners. The resulting deviations of the marginal tax rates are economically significant, exceeding five percentage points.

4.3 Optimal tax policy under heterogeneous life expectancy

To assess the implications of life-expectancy inequality, I compare the baseline optimum to optimal tax schedules computed under heterogeneous survival probabilities. In this scenario, individual survival rates follow the heterogeneous profile shown in Figure 1c. I maintain the



(a) Fiscal pension wedge (b) Optimal marginal tax rate

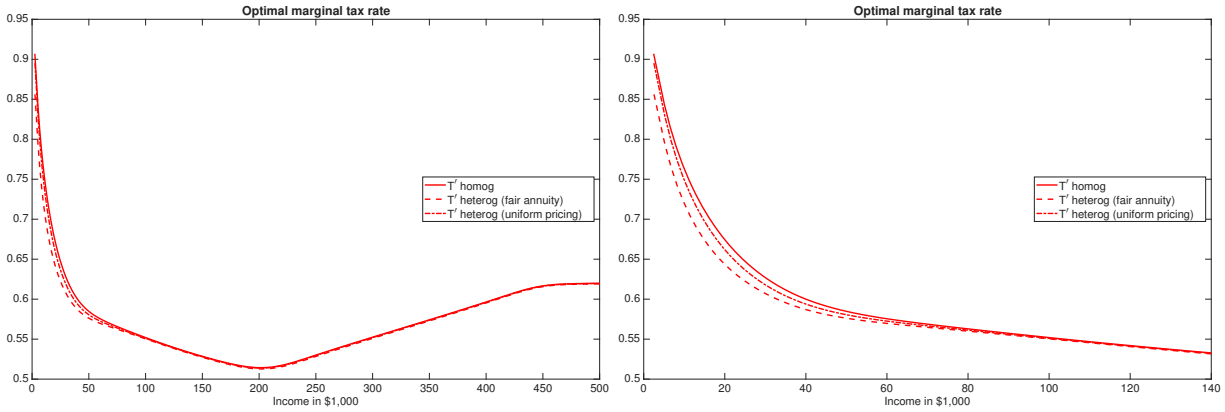
Figure 3: Impact of fiscal pension wedge on optimal policy. Left: fiscal pension wedge under homogeneous (solid) vs. heterogeneous (dashed) life expectancy; Right: optimal policy under homogeneous life expectancy (solid) vs. counterfactual with zero pension wedge (dotted).

skill distribution, the pension system, and the level of government expenditure from the baseline model.

As shown in the theoretical analysis, heterogeneous survival affects the optimal tax schedule through two distinct channels. The first is an efficiency channel operating through the fiscal pension wedge, $\omega^F = S' - \beta p B'$. When survival probabilities vary with income, the expected fiscal cost of marginal pension benefits changes across types. Figure 3a shows that this force compresses the fiscal pension-wedge profile relative to the homogeneous-survival benchmark: the wedge becomes less negative at the bottom and less positive for higher earners. Taken in isolation, this channel lowers optimal marginal tax rates at the bottom, where marginal earnings generate negative pension-related fiscal externalities, and raises them at middle and higher incomes, where the compression of positive fiscal externalities works in the opposite direction.

The second channel is an equity channel operating through the mapping from disposable income into social marginal welfare. Under uniform intertemporal pricing, this channel is driven by survival probabilities alone. Longer-lived individuals attach greater expected value to retirement consumption, shift resources toward retirement, and have lower working-life consumption. This raises their marginal welfare value of a transfer and weakens the usual redistributive motive when life expectancy rises with income. Under actuarially fair annuity pricing, the survival-probability effect is accompanied by income and substitution effects induced by survival-dependent annu-

ity prices. With $\sigma = 2$, the intertemporal elasticity of substitution is relatively low, so the substitution response in working-life consumption is limited. In the calibration, the composite annuity-price effect therefore reinforces the survival-probability effect.



(a) Optimal marginal tax rate (b) Optimal marginal tax rate for low incomes

Figure 4: Optimal policies under homogeneous (solid) and heterogeneous (dashed, dash-dotted) life expectancy. Dashed: actuarially fair annuity pricing; dash-dotted: uniform pricing.

The optimal marginal tax schedules reflect the net effect of these efficiency and equity forces and are shown in Figure 4. Under both pricing benchmarks, the equity channel dominates the efficiency-side force that would otherwise raise optimal marginal tax rates for non-bottom earners. As a result, the optimal marginal tax rates under heterogeneous survival lie below those of the homogeneous-survival baseline at all income levels. The reductions are concentrated at the bottom (where efficiency and equity channels work in the same direction) and are larger under actuarially fair annuity pricing. Relative to the homogeneous baseline, optimal marginal tax rates fall by up to 4.9 percentage points under fair annuity pricing (dashed line), compared with at most 1.4 percentage points under uniform pricing (dash-dotted line).

Figure 5 translates these policy changes into welfare effects, measured as consumption-equivalent variations relative to the optimal homogeneous tax policy, while evaluating welfare in the corresponding heterogeneous-survival environment. Under actuarially fair annuity pricing (dashed line), the lowest-income types lose more than 2 percent of consumption. Although their own marginal tax rates fall, the lower rates throughout the distribution reduce government revenue and thus the uniform lump-sum grant that balances the budget, and this loss dominates for the types who depend on the grant most. Most other individuals gain, with benefits peaking at roughly 1 percent for middle earners. Under uniform pricing (dash-dotted line), the same

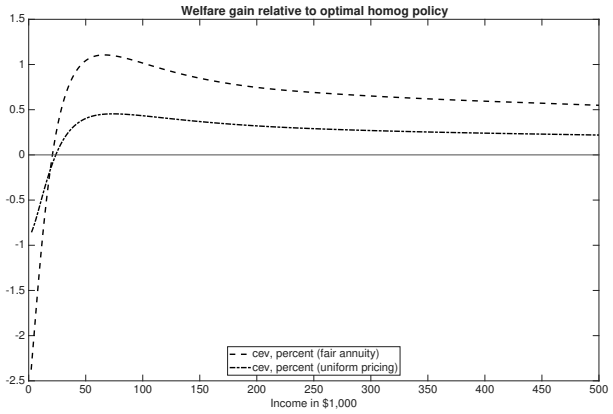


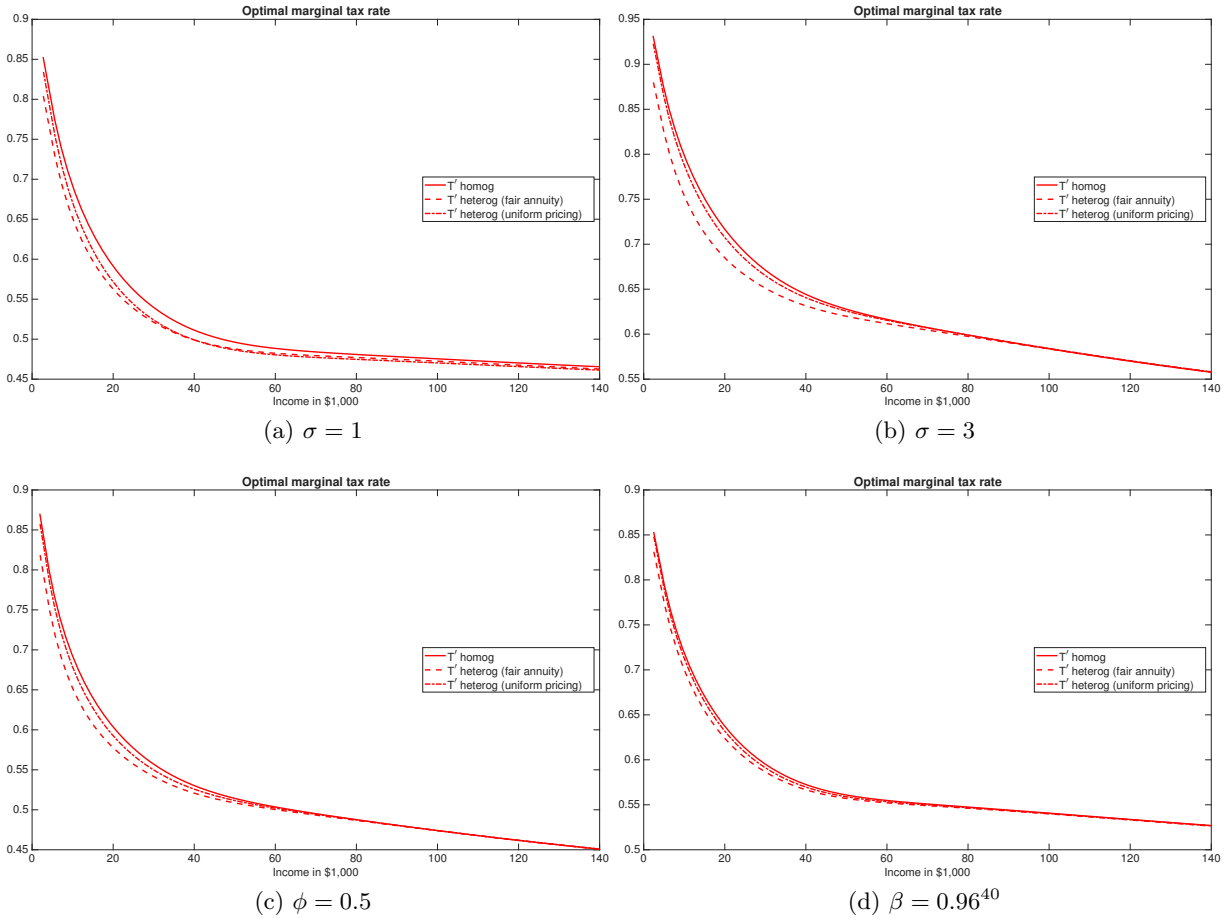
Figure 5: Consumption-equivalent welfare effects relative to optimal homogeneous policy. Welfare effects are evaluated in the heterogeneous life expectancy model. Dashed: actuarially fair annuity pricing; dash-dotted: uniform pricing.

qualitative pattern emerges, but the effects are attenuated. The lowest-income types again lose, with losses below 1 percent of consumption, while gains for the rest of the distribution are smaller and peak at about 0.5 percent.

Overall, the policy experiments show that life-expectancy inequality affects optimal taxation through the interaction of efficiency and equity channels. The compression of the fiscal pension wedge lowers optimal marginal tax rates at the bottom, while the adjusted equity channel weakens the mechanical redistributive gain from high marginal tax rates throughout the income distribution. Together, these forces reduce marginal tax rates most strongly at the bottom. Because lower bottom tax rates reduce inframarginal tax payments by most taxpayers, they limit the income support provided by the budget-balancing grant. This force generates welfare losses for the lowest-income types, while most other taxpayers gain from the lower overall redistribution.

Figure 6 documents the robustness of these findings across alternative levels of risk aversion, labor supply elasticity, and intertemporal discount factor. In all cases, life-expectancy inequality reduces optimal marginal tax rates most strongly at the bottom of the income distribution, generating a redistributive pattern that shifts resources away from the lowest-income types and toward middle and higher earners. The gap between the optimal tax rates derived under fair annuity pricing and uniform pricing is larger when risk aversion is higher. The pattern is consistent with the theory because higher risk aversion implies a lower intertemporal elasticity of substitution, limiting the substitution response to annuity-price changes and making the

income effect on working-life consumption more important for marginal welfare.



(a) $\sigma = 1$ (b) $\sigma = 3$

(c) $\phi = 0.5$ (d) $\beta = 0.96^{40}$

Figure 6: Optimal policies for alternative parameterizations

5 Concluding remarks

This paper shows that accounting for inequality in life expectancy changes the efficiency and equity rationales for redistribution. In the quantitative application to the U.S. economy, these forces reduce optimal marginal tax rates most strongly at the bottom of the income distribution, with the magnitude of the adjustment depending on the pricing of private retirement resources. Overall, the results imply that, in the calibrated economy, incorporating life-expectancy inequality shifts optimal redistribution away from the lowest-income types.

Future work could extend the framework in several directions. A natural next step would be to introduce stochastic relationships between skill and longevity, thereby modeling the joint provision of redistribution and insurance in a two-dimensional type space. A further avenue

concerns endogenous health and longevity, although progress in this direction remains limited by the still incomplete empirical understanding of the income-health gradient. In light of these difficulties, the exogenous treatment of life expectancy offers a tractable benchmark for subsequent extensions. Finally, recent survey evidence indicates that socioeconomic disparities in life expectancy are perceived as unfair (Jessen et al., 2024), suggesting a possible complementary extension with longevity-dependent social welfare weights, although such an extension would require a separate normative argument about how concerns over unequal lifespans should enter the welfare criterion.

A Appendix

A.1 Proofs

Proof of Proposition 1. To simplify the notation, suppress type dependence and let $z := y - T - S$ denote net income before pension benefits. The intertemporal first-order condition can be written as

$$G\left(c_2^M, p, q\right) := qu'\left(z + qB - qc_2^M\right) - \beta p u'\left(c_2^M\right) = 0.$$

Since $c_1^M = z + qB - qc_2^M$, I obtain

$$G_{c_2} = -q^2 u''\left(c_1^M\right) - \beta p u''\left(c_2^M\right) > 0.$$

For the effect of the survival probability, I obtain

$$G_p = -\beta u'\left(c_2^M\right) < 0.$$

The implicit-function theorem implies

$$\frac{\partial c_2^M}{\partial p} = -\frac{G_p}{G_{c_2}} = \frac{\beta u'\left(c_2^M\right)}{-q^2 u''\left(c_1^M\right) - \beta p u''\left(c_2^M\right)} > 0.$$

Since $c_1^M = z + qB - qc_2^M$, I obtain

$$\frac{\partial c_1^M}{\partial p} = -q\frac{\partial c_2^M}{\partial p} < 0.$$

Using $u'' < 0$ gives

$$\frac{\partial u'\left(c_1^M\right)}{\partial p} = u''\left(c_1^M\right)\frac{\partial c_1^M}{\partial p} > 0.$$

For the effect of the private intertemporal price, I obtain

$$G_q = u'\left(c_1^M\right) + q\left(B - c_2^M\right)u''\left(c_1^M\right).$$

Hence, using the implicit-function theorem, I obtain

$$\frac{\partial c_2^M}{\partial q} = -\frac{u'\left(c_1^M\right) + q\left(B - c_2^M\right)u''\left(c_1^M\right)}{-q^2u''\left(c_1^M\right) - \beta pu''\left(c_2^M\right)}.$$

If $c_2^M > B$, the numerator is positive and I obtain

$$\frac{\partial c_2^M}{\partial q} < 0.$$

Finally, I obtain

$$\frac{\partial c_1^M}{\partial q} = B - c_2^M - q\frac{\partial c_2^M}{\partial q}.$$

Substituting the expression for $\partial c_2^M/\partial q$ and using the first-order condition

$$qu'\left(c_1^M\right) = \beta pu'\left(c_2^M\right)$$

yields

$$\frac{\partial c_1^M}{\partial q} = \frac{\beta p\left[u'\left(c_2^M\right) + \left(c_2^M - B\right)u''\left(c_2^M\right)\right]}{-q^2u''\left(c_1^M\right) - \beta pu''\left(c_2^M\right)}.$$

The denominator is positive. Therefore,

$$\frac{\partial c_1^M}{\partial q} < 0$$

if and only if

$$u'\left(c_2^M\right) + \left(c_2^M - B\right) u''\left(c_2^M\right) < 0.$$

Equivalently,

$$1 - \frac{B}{c_2^M} > -\frac{u'\left(c_2^M\right)}{c_2^M u''\left(c_2^M\right)} = \psi\left(c_2^M\right).$$

This proves the result. ■

Proof of Proposition 2. Under actuarially fair annuity pricing, the intertemporal first-order condition becomes

$$\beta p(\theta) u'\left(c_1^M(\theta)\right) = \beta p(\theta) u'\left(c_2^M(\theta)\right).$$

Since $\beta p(\theta) > 0$ and u is strictly concave, this implies

$$c_1^M(\theta) = c_2^M(\theta).$$

Denote the time-independent consumption level by $c^M(\theta)$. The lifetime budget constraint is

$$c^M(\theta) + \beta p(\theta) c^M(\theta) = z(\theta) + \beta p(\theta) B(y(\theta)).$$

Solving for $c^M(\theta)$ gives

$$c^M(\theta) = \frac{z(\theta) + \beta p(\theta) B(y(\theta))}{1 + \beta p(\theta)}.$$

Differentiating with respect to $p(\theta)$, holding $z(\theta)$ and $B(y(\theta))$ fixed, yields

$$\frac{\partial c^M(\theta)}{\partial p(\theta)} = \frac{\beta B(y(\theta)) \left[1 + \beta p(\theta)\right] - \beta \left[z(\theta) + \beta p(\theta) B(y(\theta))\right]}{(1 + \beta p(\theta))^2}.$$

Hence,

$$\frac{\partial c^M(\theta)}{\partial p(\theta)} = \frac{\beta \left[B(y(\theta)) - z(\theta)\right]}{(1 + \beta p(\theta))^2}.$$

Using

$$c^M(\theta) - B(y(\theta)) = \frac{z(\theta) - B(y(\theta))}{1 + \beta p(\theta)},$$

this result can be rewritten as

$$\frac{\partial c^M(\theta)}{\partial p(\theta)} = -\frac{\beta}{1 + \beta p(\theta)} \left[c^M(\theta) - B(y(\theta))\right].$$

Since $c_1^M(\theta) = c^M(\theta)$, the expression also gives the effect on working-life consumption. If $c^M(\theta) > B(y(\theta))$, then

$$\frac{\partial c_1^M(\theta)}{\partial p(\theta)} < 0.$$

Finally, because $u'' < 0$, I obtain

$$\frac{\partial u'\left(c_1^M(\theta)\right)}{\partial p(\theta)} = u''\left(c_1^M(\theta)\right)\frac{\partial c_1^M(\theta)}{\partial p(\theta)} > 0.$$

This proves the result. ■

Proof of Lemma 1. The individual labor supply problem (9) generates the first-order condition

$$\theta\left(1 - T'(y) - S'(y) + q(\theta)B'(y)\right) = v'\left(\frac{y}{\theta}\right). \tag{59}$$

Assuming a linearized tax and pension system, differentiation of (59) with respect to θ yields:

$$\left(1 - T' - S' + q(\theta)B'\right) + \theta q'(\theta)B' = v''\left(\frac{y}{\theta}\right)\frac{\frac{dy}{d\theta}\theta - y}{\theta^2}.$$

Equivalently,

$$\frac{dy}{d\theta} = \frac{\theta\left(1 - T' - S' + q(\theta)B'\right) + \theta^2 q'(\theta)B'}{v''\left(\frac{y}{\theta}\right)} + \frac{y}{\theta}.$$

Equivalently,

$$\frac{dy}{d\theta} = \frac{v'\left(\frac{y}{\theta}\right) + \theta^2 q'(\theta)B'}{v''\left(\frac{y}{\theta}\right)} + \frac{y}{\theta}.$$

Hence, the elasticity of income with respect to skill is

$$\varepsilon_{y,\theta} = \frac{dy}{d\theta}\frac{\theta}{y} = 1 + \frac{v'\left(\frac{y}{\theta}\right)}{\frac{y}{\theta}v''\left(\frac{y}{\theta}\right)} + \frac{\theta^2 q'(\theta)B'}{\frac{y}{\theta}v''\left(\frac{y}{\theta}\right)} = 1 + \frac{v'(l)}{lv''(l)} + \frac{\theta^2 q'(\theta)B'}{lv''(l)}.$$

Assuming once more a linearized tax and pension system, differentiation of (59) with respect to the net-of-tax rate $1 - T'$ yields:

$$\theta = v''\left(\frac{y}{\theta}\right)\frac{1}{\theta}\frac{dy}{d(1 - T')}.$$

Equivalently,

$$\frac{dy}{d(1 - T')} = \frac{\theta^2}{v''\left(\frac{y}{\theta}\right)} = \frac{\theta^2}{v''(l)}.$$

Hence, the elasticity of income with respect to the net-of-tax rate is

$$\varepsilon_{y,(1-T')} = \frac{dy}{d\left(1-T'\right)}\frac{\left(1-T'\right)}{y} = \frac{\theta\left(1-T'\right)}{lv''\left(l\right)}.$$

Using (59), I obtain

$$\theta\left(1-T'\right) = v'\left(l\right) + \theta\left(S' - q\left(\theta\right)B'\right).$$

Hence,

$$\varepsilon_{y,(1-T')} = \frac{v'\left(l\right)}{lv''\left(l\right)} + \frac{\theta\left(S' - q\left(\theta\right)B'\right)}{lv''\left(l\right)}$$

Similarly, I can derive the elasticity with respect to the effective net-of-tax rate $1-\hat{T}'$. Assuming a linearized tax and pension system, differentiation of the first-order condition

$$\theta\left(1-\hat{T}'\right) = v'\left(\frac{y}{\theta}\right)$$

yields

$$\theta = v''\left(\frac{y}{\theta}\right)\frac{1}{\theta}\frac{dy}{d\left(1-\hat{T}'\right)}.$$

Hence,

$$\varepsilon_{y,\left(1-\hat{T}'\right)} = \frac{dy}{d\left(1-\hat{T}'\right)}\frac{\left(1-\hat{T}'\right)}{y} = \frac{\theta^2}{v''\left(\frac{y}{\theta}\right)}\frac{\frac{v'\left(\frac{y}{\theta}\right)}{\theta}}{y} = \frac{v'\left(l\right)}{lv''\left(l\right)}.$$

For completeness, I also derive the elasticities $\varepsilon_{y,(1-S')}$ and $\varepsilon_{y,B'}$. Differentiation of the first-order condition (59) with respect to $1-S'$, assuming a linearized tax and pension system, yields:

$$\theta = v''\left(\frac{y}{\theta}\right)\frac{1}{\theta}\frac{dy}{d\left(1-S'\right)}.$$

Hence,

$$\varepsilon_{y,(1-S')} = \frac{dy}{d\left(1-S'\right)}\frac{\left(1-S'\right)}{y} = \frac{v'\left(l\right)}{lv''\left(l\right)} + \frac{\theta\left(T' - q\left(\theta\right)B'\right)}{lv''\left(l\right)}.$$

Differentiation of (59) with respect to B', assuming a linearized tax and pension system, yields

$$\theta q\left(\theta\right) = v''\left(\frac{y}{\theta}\right)\frac{1}{\theta}\frac{dy}{dB'}.$$

Hence,

$$\varepsilon_{y,B'} = \frac{dy}{dB'}\frac{B'}{y} = \frac{\theta^2 q(\theta)}{v''\left(\frac{y}{\theta}\right)}\frac{B'}{y} = \frac{v'(l)}{lv''(l)} - \frac{\theta\left(1 - T'(y) - S'(y)\right)}{lv''(l)}.$$

■

Proof of Proposition 3. Problem (24) is a constrained optimal control problem with state variable x and control variable l. Denote the co-state variable by $\mu(\theta)$ and the multiplier of the feasibility constraint by λ. The Hamiltonian of problem (24) is:

$$\begin{aligned} H =& \tilde{V}(x(\theta);\theta) f(\theta) + \lambda f(\theta)\left[\theta l(\theta) - x(\theta) - v(l(\theta)) - \left[\beta p(\theta) - q(\theta)\right] B(\theta l(\theta))\right] \\ & - \lambda E + \mu(\theta)\left[\frac{l(\theta)}{\theta} v'(l(\theta)) + q'(\theta) B(\theta l(\theta))\right]. \end{aligned}$$

The maximum principle generates the following necessary conditions:

$$\begin{aligned} 0 = \frac{\partial H}{\partial l} =& \mu(\theta)\left[\frac{v'(l(\theta))}{\theta} + \frac{l(\theta)}{\theta} v''(l(\theta)) + \theta q'(\theta) B'(\theta l(\theta))\right] \\ & + \lambda f(\theta)\left[\theta - v'(l(\theta)) - \theta\left[\beta p(\theta) - q(\theta)\right] B'(\theta l(\theta))\right], \\ \dot{\mu} = -\frac{\partial H}{\partial x} =& -\left[u'(x(\theta) - q(\theta) c_2(\theta)) - \lambda\right] f(\theta). \end{aligned}$$

where $c_2(\theta)$ is the solution of the individual's intertemporal problem (22).

Using the transversality condition, $\mu\left(\overline{\theta}\right) = 0$, I obtain:

$$\mu(\theta) = \mu(\theta) - \mu\left(\overline{\theta}\right) = -\int_{\theta}^{\overline{\theta}} \dot{\mu}(\theta)\, d\theta = \int_{\theta}^{\overline{\theta}} \left[u'\left(x\left(\tilde{\theta}\right) - q\left(\tilde{\theta}\right) c_2\left(\tilde{\theta}\right)\right) - \lambda\right] dF\left(\tilde{\theta}\right). \quad (60)$$

The state x is free at the lower endpoint as well, so the maximum principle delivers the second transversality condition $\mu(\underline{\theta}) = 0$. Evaluating Equation (60) at $\underline{\theta}$ therefore yields

$$\lambda = \int_{\underline{\theta}}^{\overline{\theta}} u'\left(x\left(\tilde{\theta}\right) - q\left(\tilde{\theta}\right) c_2\left(\tilde{\theta}\right)\right) dF\left(\tilde{\theta}\right),$$

so that the multiplier on the resource constraint equals the population average of the marginal welfare weights. This normalization is what allows the $\mathcal{B}$-term to be read as a comparison of the average weight u'/λ among types above θ to a population average of one.

Combining Equation (60) with the first-order condition for l yields:

$$-\lambda\left[\theta - v' - \theta\left[\beta p - q\right]B'\right]f = \frac{v' + lv'' + \theta^2 q' B'}{\theta}\int_{\theta}^{\overline{\theta}}\left[u' - \lambda\right]dF,$$

where I have dropped the arguments of all functions for brevity. Equivalently,

$$\frac{\theta - v' - \theta\left[\beta p - q\right]B'}{v'} = \frac{v' + lv'' + \theta^2 q' B'}{v'}\frac{1}{\theta f}\int_{\theta}^{\overline{\theta}}\left[1 - \frac{u'}{\lambda}\right]dF.$$

Equivalently,

$$\frac{\theta - v' - \theta\left[\beta p - q\right]B'}{v'} = \frac{v' + lv'' + \theta^2 q' B'}{v'}\frac{1-F}{\theta f}\frac{\int_{\theta}^{\overline{\theta}}\left[1 - \frac{u'}{\lambda}\right]dF}{1-F}.$$

Equivalently,

$$\frac{\theta - v' - \theta\left[\beta p - q\right]B'}{v'} = \frac{v' + lv'' + \theta^2 q' B'}{v'}\frac{1-F}{\theta f}\left(1 - \frac{\int_{\theta}^{\overline{\theta}} u'\,dF}{\lambda\left(1-F\right)}\right).$$

Next, I connect this optimality condition to the marginal tax rates using individual first-order conditions. The first-order condition of problem (9) is

$$1 - \hat{T}' = \frac{v'}{\theta}.$$

Hence,

$$\frac{\theta - v' - \theta\left[\beta p - q\right]B'}{v'} = \frac{1 - \frac{v'}{\theta} - \left[\beta p - q\right]B'}{\frac{v'}{\theta}} = \frac{\hat{T}' - \left[\beta p - q\right]B'}{1 - \hat{T}'}.$$

Using

$$\hat{T}' - \left[\beta p - q\right]B' = T' + S' - \beta p B' = T' + \omega^F,$$

this condition can be rewritten as

$$\frac{\theta - v' - \theta\left[\beta p - q\right]B'}{v'} = \frac{T' + \omega^F}{1 - \hat{T}'}$$

Hence,

$$\frac{T' + \omega^F}{1 - \hat{T}'} = \frac{v' + lv'' + \theta^2 q' B'}{v'}\frac{1-F}{\theta f}\left(1 - \frac{\int_{\theta}^{\overline{\theta}} u'\,dF}{\lambda\left(1-F\right)}\right).$$

Equivalently,

$$\frac{T' + \omega^F}{1 - T'} = \frac{v' + lv'' + \theta^2 q' B'}{v'} \frac{1 - F}{\theta f} \left(1 - \frac{\int_\theta^{\overline{\theta}} u' \, dF}{\lambda (1 - F)} \right) \frac{1 - \hat{T}'}{1 - T'}.$$

The private effective marginal tax rate and the statutory marginal tax rate are related as follows:

$$\hat{T}' = T' + \left(S' - qB' \right).$$

Using this relationship and the first-order condition

$$1 - \hat{T}' = \frac{v'}{\theta},$$

I obtain

$$\frac{T' + \omega^F}{1 - T'} = \frac{v' + lv'' + \theta^2 q' B'}{v'} \frac{1 - F}{\theta f} \left(1 - \frac{\int_\theta^{\overline{\theta}} u' \, dF}{\lambda (1 - F)} \right) \frac{\frac{v'}{\theta}}{\frac{v'}{\theta} + S' - qB'}.$$

Equivalently,

$$\frac{T' + \omega^F}{1 - T'} = \frac{v' + lv'' + \theta^2 q' B'}{v' + \theta \left(S' - qB' \right)} \frac{1 - F}{\theta f} \left(1 - \frac{\int_\theta^{\overline{\theta}} u' \, dF}{\lambda (1 - F)} \right).$$

Next, I rewrite the optimality condition with the help of the elasticities obtained in Lemma 1. Using Equations (37) and (38), I obtain

$$\frac{v' + lv'' + \theta^2 q' B'}{v' + \theta \left(S' - qB' \right)} = \frac{1 + \frac{v' + \theta^2 q' B'}{lv''}}{\frac{v' + \theta (S' - qB')}{lv''}} = \frac{\varepsilon_{y,\theta}}{\varepsilon_{y,(1-T')}}.$$

The optimality condition for the tax system is hence

$$\frac{T' + \omega^F}{1 - T'} = \frac{\varepsilon_{y,\theta}}{\varepsilon_{y,(1-T')}} \left(1 - \frac{\int_\theta^{\overline{\theta}} u' \, dF}{\lambda (1 - F)} \right) \frac{1 - F}{\theta f},$$

which establishes Equation (41).

Equivalently,

$$\frac{T'}{1 - T'} = \frac{\varepsilon_{y,\theta}}{\varepsilon_{y,(1-T')}} \left(1 - \frac{\int_\theta^{\overline{\theta}} u' \, dF}{\lambda (1 - F)} \right) \frac{1 - F}{\theta f} \frac{T'}{T' + \omega^F},$$

which establishes Equation (42).

Finally, I state the optimality condition in terms of the income distribution. Let H be the

cumulative distribution function of income y and $h = H'$ be its density function. Assuming a linearized tax and pension system, the individual first-order condition for labor supply (73) implies

$$dy = \left(\frac{v'\left(\frac{y}{\theta}\right) + \theta^2 q'(\theta) B'}{v''\left(\frac{y}{\theta}\right)} + \frac{y}{\theta} \right) d\theta.$$

Using $h(y)\, dy = f(\theta)\, d\theta$, I obtain

$$\frac{1}{y \cdot h(y)} = \frac{dy}{y \cdot f(\theta)\, d\theta} = \frac{1}{\theta \cdot f(\theta)} \left(1 + \frac{v'(l)}{l v''(l)} + \frac{\theta^2 q'(\theta) B'}{l v''(l)} \right) = \frac{\varepsilon_{y,\theta}}{\theta \cdot f(\theta)}.$$

Hence, using $H(y) = F(\theta)$, I obtain

$$\frac{1 - H(y)}{y \cdot h(y)} = \frac{1 - F(\theta)}{\theta \cdot f(\theta)} \varepsilon_{y,\theta}$$

and can transform Equation (42) into Equation (43). ■

Proof of Proposition 4. Using

$$p(\theta) = p_0(\theta) + \eta \left(p_0(\theta) - \bar{p} \right),$$

I obtain

$$\frac{\partial p(\theta)}{\partial \eta} = p_0(\theta) - \bar{p}.$$

Differentiation of

$$\mathcal{D} = \frac{T'}{T' + \omega^F} = \frac{T'}{T' + S' - \beta p(\theta) B'}$$

with respect to η, holding T', S', and B' fixed, thus yields

$$\frac{\partial \mathcal{D}}{\partial \eta} = \frac{\beta T' B'}{\left(T' + S' - \beta p(\theta) B'\right)^2} \left(p_0(\theta) - \bar{p} \right).$$

Assuming $T', B' > 0$, the prefactor is positive. In that case, I obtain

$$\frac{\partial \mathcal{D}}{\partial \eta} \gtreqless 0 \quad \Longleftrightarrow \quad p_0(\theta) \gtreqless \bar{p}.$$

Finally, if $B' = 0$, the derivative is zero. ■

Proof of Proposition 5. Denote the co-state variables of Problem (53) by $\mu_1, \ldots, \mu_{R-1}$ and

the multiplier of the feasibility constraint by λ. The Hamiltonian of problem (53) is:

$$\begin{aligned}H =&\tilde{V}\left(x_1\left(\theta\right),\ldots,x_{R-1}\left(\theta\right);\theta\right)f\left(\theta\right)\\&+\lambda f\left(\theta\right)\sum_{t=1}^{R-1}\beta^{t-1}p_t\left(\theta\right)\left[w_t\left(\theta\right)l_t\left(\theta\right)-x_t\left(\theta\right)-v_t\left(l_t\left(\theta\right)\right)+\left[Q_t\left(\theta\right)-P_t^F\left(\theta\right)\right]b_t\left(w_t\left(\theta\right)l_t\left(\theta\right)\right)\right]\\&-\lambda E+\sum_{t=1}^{R-1}\mu_t\left(\theta\right)\left[l_t\left(\theta\right)\frac{w_t'\left(\theta\right)}{w_t\left(\theta\right)}v_t'\left(l_t\left(\theta\right)\right)+Q_t'\left(\theta\right)b_t\left(w_t\left(\theta\right)l_t\left(\theta\right)\right)\right].\end{aligned}$$

For each $t=1,\ldots,R-1$, the maximum principle implies:

$$\begin{aligned}0=\frac{\partial H}{\partial l_t}&=\mu_t\left[\frac{w_t'}{w_t}v_t'+l_t\frac{w_t'}{w_t}v_t''+w_tQ_t'b_t'\right]+\lambda f\beta^{t-1}p_t\left[w_t-v_t'+w_t\left[Q_t-P_t^F\right]b_t'\right],\\\dot{\mu_t}=-\frac{\partial H}{\partial x_t}&=-\left[\beta^{t-1}p_tu_t'-\lambda\beta^{t-1}p_t\right]f,\end{aligned}$$

where I have dropped the arguments of all functions for brevity and used the envelope theorem to obtain $\frac{\partial\tilde{V}}{\partial x_t}=\beta^{t-1}p_tu_t'$.

Using the transversality condition $\mu_t\left(\overline{\theta}\right)=0$, I obtain:

$$\mu_t\left(\theta\right)=\mu_t\left(\theta\right)-\mu_t\left(\overline{\theta}\right)=-\int_\theta^{\overline{\theta}}\dot{\mu_t}\left(\tilde{\theta}\right)d\tilde{\theta}=\int_\theta^{\overline{\theta}}\left[\beta^{t-1}p_t\left(\tilde{\theta}\right)u_t'\left(\tilde{\theta}\right)-\lambda\beta^{t-1}p_t\left(\tilde{\theta}\right)\right]f\left(\tilde{\theta}\right)d\tilde{\theta}.$$

Combining this result with the first-order condition for l_t yields

$$-\lambda f\beta^{t-1}p_t\left[w_t-v_t'+w_t\left[Q_t-P_t^F\right]b_t'\right]=\frac{w_t'v_t'+l_tw_t'v_t''+\left(w_t\right)^2Q_t'b_t'}{w_t}\int_\theta^{\overline{\theta}}\left[\beta^{t-1}p_tu_t'-\lambda\beta^{t-1}p_t\right]dF.$$

Equivalently,

$$\frac{w_t-v_t'-w_t\left[P_t^F-Q_t\right]b_t'}{v_t'}=\frac{w_t'v_t'+l_tw_t'v_t''+\left(w_t\right)^2Q_t'b_t'}{v_t'}\frac{1}{p_tw_tf}\int_\theta^{\overline{\theta}}\left[1-\frac{u_t'}{\lambda}\right]p_tdF.$$

Equivalently,

$$\frac{w_t-v_t'-w_t\left[P_t^F-Q_t\right]b_t'}{v_t'}=\frac{v_t'+l_tv_t''+\frac{\left(w_t\right)^2}{w_t'}Q_t'b_t'}{v_t'}\frac{w_t'\left(1-F\right)}{w_tf}\frac{\int_\theta^{\overline{\theta}}\left[1-\frac{u_t'}{\lambda}\right]p_tdF}{p_t\left(1-F\right)}.$$

Next, I connect this optimality condition to the marginal tax rates using individual first-order conditions. The first-order condition of the individual labor supply problem in period t

is

$$1-\hat{T}_t'=1-T_t'-S_t'+Q_tb_t'=\frac{v_t'}{w_t}, \tag{61}$$

where $\hat{T}_t' := T_t' + S_t' - Q_t b_t'$ denotes the private effective marginal tax rate in period t. Hence,

$$\frac{w_t-v_t'-w_t\left[P_t^F-Q_t\right]b_t'}{v_t'}=\frac{1-\frac{v_t'}{w_t}-\left[P_t^F-Q_t\right]b_t'}{\frac{v_t'}{w_t}}=\frac{\hat{T}_t'-\left[P_t^F-Q_t\right]b_t'}{1-\hat{T}_t'}.$$

Using

$$\hat{T}_t'-\left[P_t^F-Q_t\right]b_t'=T_t'+S_t'-P_t^Fb_t'=T_t'+\omega_t^F,$$

where $\omega_t^F = S_t' - P_t^F b_t'$ is the fiscal wedge in period t, this condition can be rewritten as

$$\frac{w_t-v_t'-w_t\left[P_t^F-Q_t\right]b_t'}{v_t'}=\frac{T_t'+\omega_t^F}{1-\hat{T}_t'}.$$

Hence,

$$\frac{T_t'+\omega_t^F}{1-\hat{T}_t'}=\frac{v_t'+l_tv_t''+\frac{(w_t)^2}{w_t'}Q_t'b_t'}{v_t'}\frac{w_t'\left(1-F\right)}{w_tf}\frac{\int_\theta^{\overline{\theta}}\left[1-\frac{u_t'}{\lambda}\right]p_tdF}{p_t\left(1-F\right)}.$$

Equivalently,

$$\frac{T_t'+\omega_t^F}{1-T_t'}=\frac{v_t'+l_tv_t''+\frac{(w_t)^2}{w_t'}Q_t'b_t'}{v_t'}\frac{w_t'\left(1-F\right)}{w_tf}\frac{\int_\theta^{\overline{\theta}}\left[1-\frac{u_t'}{\lambda}\right]p_tdF}{p_t\left(1-F\right)}\frac{1-\hat{T}_t'}{1-T_t'}.$$

The private effective marginal tax rate and the statutory marginal tax rate in period t are related as follows:

$$\hat{T}_t'=T_t'+S_t'-Q_tb_t'.$$

Using this relationship and the first-order condition

$$1-\hat{T}_t'=\frac{v_t'}{w_t},$$

I obtain

$$\frac{T_t'+\omega_t^F}{1-T_t'}=\frac{v_t'+l_tv_t''+\frac{(w_t)^2}{w_t'}Q_t'b_t'}{v_t'}\frac{w_t'\left(1-F\right)}{w_tf}\frac{\int_\theta^{\overline{\theta}}\left[1-\frac{u_t'}{\lambda}\right]p_tdF}{p_t\left(1-F\right)}\frac{\frac{v_t'}{w_t}}{\frac{v_t'}{w_t}+S_t'-Q_tb_t'}.$$

Equivalently,

$$\frac{T_t' + \omega_t^F}{1 - T_t'} = \frac{v_t' + l_t v_t'' + \frac{(w_t)^2}{w_t'} Q_t' b_t'}{v_t' + w_t \left(S_t' - Q_t b_t'\right)} \frac{w_t' \left(1 - F\right)}{w_t f} \frac{\int_\theta^{\overline{\theta}} \left[1 - \frac{u_t'}{\lambda}\right] p_t dF}{p_t \left(1 - F\right)}. \tag{62}$$

Next, I rewrite this condition with the help of elasticities. The first-order condition (61) can be rearranged as

$$w_t \left(1 - T_t' - S_t' + Q_t b_t'\right) = v_t' \left(\frac{y_t}{w_t}\right). \tag{63}$$

Assuming a linearized tax and pension system, differentiation of (63) with respect to w_t yields:

$$\left(1 - T_t' - S_t' + Q_t b_t'\right) + w_t Q_t' \frac{1}{w_t'} b_t' = v_t'' \left(\frac{y_t}{w_t}\right) \frac{\frac{dy_t}{dw_t} w_t - y_t}{\left(w_t\right)^2}.$$

Equivalently,

$$\frac{dy_t}{dw_t} = \frac{w_t \left(1 - T_t' - S_t' + Q_t b_t'\right) + \frac{(w_t)^2}{w_t'} Q_t' b_t'}{v_t'' \left(\frac{y_t}{w_t}\right)} + \frac{y_t}{w_t}.$$

Equivalently,

$$\frac{dy_t}{dw_t} = \frac{v_t' + \frac{(w_t)^2}{w_t'} Q_t' b_t'}{v_t''} + \frac{y_t}{w_t}.$$

Hence, the elasticity of income with respect to the wage is

$$\varepsilon_{y_t, w_t} = \frac{dy_t}{dw_t} \frac{w_t}{y_t} = 1 + \frac{v_t'}{l_t v_t''} + \frac{\left(w_t\right)^2 Q_t' b_t'}{w_t' l_t v_t''}.$$

Assuming once more a linearized tax and pension system, differentiation of (63) with respect to the net-of-tax rate $1 - T_t'$ yields:

$$w_t = v_t'' \frac{1}{w_t} \frac{dy_t}{d\left(1 - T_t'\right)}.$$

Equivalently,

$$\frac{dy_t}{d\left(1 - T_t'\right)} = \frac{\left(w_t\right)^2}{v_t''}.$$

Hence, the elasticity of income with respect to the net-of-tax rate is

$$\varepsilon_{y_t, (1 - T_t')} = \frac{dy_t}{d\left(1 - T_t'\right)} \frac{\left(1 - T_t'\right)}{y_t} = \frac{w_t \left(1 - T_t'\right)}{l_t v_t''}.$$

Using (63), I obtain

$$w_t\left(1-T_t'\right)=v_t'+w_t\left(S_t'-Q_tb_t'\right).$$

Hence,

$$\varepsilon_{y_t,(1-T_t')}=\frac{v_t'}{l_tv_t''}+\frac{w_t\left(S_t'-Q_tb_t'\right)}{l_tv_t''}.$$

Hence, the elasticity ratio $\varepsilon_{y_t,w_t}/\varepsilon_{y_t,(1-T_t')}$ equals

$$\frac{\varepsilon_{y_t,w_t}}{\varepsilon_{y_t,(1-T_t')}}=\frac{1+\frac{v_t'}{l_tv_t''}+\frac{(w_t)^2Q_t'b_t'}{w_t'l_tv_t''}}{\frac{v_t'}{l_tv_t''}+\frac{w_t(S_t'-Q_tb_t')}{l_tv_t''}}=\frac{v_t'+l_tv_t''+\frac{(w_t)^2}{w_t'}Q_t'b_t'}{v_t'+w_t\left(S_t'-Q_tb_t'\right)}.$$

Using this result, (62) can be stated as

$$\frac{T_t'+\omega_t^F}{1-T_t'}=\frac{\varepsilon_{y_t,w_t}}{\varepsilon_{y_t,(1-T_t')}}\frac{\int_\theta^{\overline{\theta}}\left[1-\frac{u_t'}{\lambda}\right]p_tdF}{p_t\left(1-F\right)}\frac{w_t'\left(1-F\right)}{w_tf}.$$

Equivalently,

$$\frac{T_t'}{1-T_t'}=\frac{\varepsilon_{y_t,w_t}}{\varepsilon_{y_t,(1-T_t')}}\frac{\int_\theta^{\overline{\theta}}\left[1-\frac{u_t'}{\lambda}\right]p_tdF}{p_t\left(1-F\right)}\frac{w_t'\left(1-F\right)}{w_tf}\frac{T_t'}{T_t'+\omega_t^F}.$$

which establishes Equation (56). ■

A.2 Tax perturbation approach

Using the tax perturbation approach of Saez (2001), I derive an optimal tax condition in terms of the endogenous income distribution rather than the exogenous skill distribution.

Denote the cumulative distribution function of income y by H and its density function by $h=H'$. Consider an increase of the statutory marginal tax rate T' by $d\tau$ in a small band $(y,y+dy)$. Mechanically, this reform raises an extra tax revenue of

$$dM=dy\,d\tau\left(1-H\left(y\right)\right). \tag{64}$$

Since the extra revenue is financed by all individuals with incomes above y, there is a mechanical welfare cost (in terms of public funds) of

$$dW=-dy\,d\tau\left(1-H\left(y\right)\right)\frac{\int_y^{\overline{y}}u'dH}{\lambda\left(1-H\left(y\right)\right)}. \tag{65}$$

Moreover, individuals in the band $(y, y+dy)$ change their income in response to the increased marginal tax rate. There are $h(y)\,dy$ individuals in the band and they change their income by $-\varepsilon_{y,(1-T')} y\, d\tau/(1-T')$. To convert this behavioral response into public funds, the income change must be valued using the government's fiscal marginal return from income. The relevant fiscal effective marginal tax rate is

$$\hat{T}^{F\prime}(y;\theta) = T'(y) + S'(y) - \beta p(\theta) B'(y). \tag{66}$$

Thus, the behavioral effect on public funds is

$$dB = -dy\, d\tau\, h(y)\, \varepsilon_{y,(1-T')} y \frac{\hat{T}^{F\prime}}{1-T'}. \tag{67}$$

If the tax system is optimal, $dM + dW + dB = 0$ and, hence,

$$\frac{\hat{T}^{F\prime}}{1-T'} = \frac{1}{\varepsilon_{y,(1-T')}} \left(1 - \frac{\int_y^{\overline{y}} u' dH}{\lambda (1-H(y))}\right) \frac{(1-H(y))}{y \cdot h(y)}. \tag{68}$$

Equivalently,

$$\frac{T'}{1-T'} = \frac{1}{\varepsilon_{y,(1-T')}} \left(1 - \frac{\int_y^{\overline{y}} u' dH}{\lambda (1-H(y))}\right) \frac{(1-H(y))}{y \cdot h(y)} \frac{T'}{\hat{T}^{F\prime}}. \tag{69}$$

A.3 Discussion of taxable income elasticity and inverse Pareto coefficient

The elasticity term, given by $\mathcal{A} = \varepsilon_{y,\theta}/\varepsilon_{y,(1-T')}$ in the skill-based optimality condition and by $1/\varepsilon_{y,(1-T')}$ in the income-based formulation, captures the strength of behavioral responses induced by the local marginal tax rate. When the elasticity of taxable income is high, increasing the marginal tax rate induces a large behavioral response and therefore generates a large efficiency cost.

Without a pension system, the elasticity term in Equation (42) can be simplified to the form $\mathcal{A} = 1 + 1/\varepsilon_{y,(1-T')}$ known from Diamond (1998), as the next result shows.

Lemma 2 (Simplification of $\mathcal{A}$) *If $q'B' = 0$, the elasticity ratio $\varepsilon_{y,\theta}/\varepsilon_{y,(1-T')}$ satisfies*

$$\frac{\varepsilon_{y,\theta}}{\varepsilon_{y,(1-T')}} = \frac{1 + \varepsilon_{y,(1-T')}\big|_{S'-qB'=0}}{\varepsilon_{y,(1-T')}}. \tag{70}$$

If $q'B' \neq 0$, then the ratio $\varepsilon_{y,\theta}/\varepsilon_{y,(1-T')}$ cannot be expressed solely in terms of the taxable-income elasticities $\varepsilon_{y,(1-T')}$, $\varepsilon_{y,(1-S')}$, $\varepsilon_{y,B'}$.

Proof of Lemma 2. By Equations (37) and (40), I obtain

$$\begin{aligned} \varepsilon_{y,\theta} &= 1 + \frac{v'(l)}{lv''(l)} + \frac{\theta^2 q'(\theta) B'}{lv''(l)}, \\ 1 + \varepsilon_{y,(1-T')}\big|_{S'-qB'=0} &= 1 + \frac{v'(l)}{lv''(l)}, \end{aligned}$$

where $\varepsilon_{y,(1-T')}\big|_{S'-qB'=0}$ represents the elasticity of taxable income with respect to the net-of-tax rate, assuming a privately valued pension wedge of zero.

If $q'(\theta) B' = 0$, I obtain

$$\varepsilon_{y,\theta} = 1 + \frac{v'(l)}{lv''(l)} = 1 + \varepsilon_{y,(1-T')}\big|_{S'-qB'=0},$$

which establishes the first claim.

For the second claim, note that the proof of Lemma 1 establishes the following elasticities:

$$\varepsilon_{y,(1-S')} = \frac{v'(l)}{lv''(l)} + \frac{\theta\left(T' - q(\theta) B'\right)}{lv''(l)} \quad , \tag{71}$$

$$\varepsilon_{y,B'} = \frac{v'(l)}{lv''(l)} - \frac{\theta\left(1 - T'(y) - S'(y)\right)}{lv''(l)} \quad , \tag{72}$$

Together with $\varepsilon_{y,(1-T')}$ stated in Equation (38), these elasticities depend on the privately valued pension wedge and on the marginal policy rates, but not on the derivative $q'(\theta)$. If $q'(\theta) B' \neq 0$, however, the elasticity $\varepsilon_{y,\theta}$ involves the term $q'(\theta)$. Hence, in this case the elasticity ratio cannot be reduced to a function of those taxable-income elasticities alone. ■

To see how the pension system prevents the simplification of the elasticity term, note that the first-order condition for labor supply is

$$\theta\left(1 - T' - S' + q(\theta) B'\right) = v', \tag{73}$$

which states that the marginal return to an hour of work (measured in terms of disposable lifetime income) equals the marginal disutility of an hour of work. The comparative statics of labor supply with respect to skill are therefore driven by two effects. First, a higher skill

raises the gross return to work (as indicated by the factor θ that multiplies the left-hand side). Second, a higher skill has an additional influence on the net return to work if the private value of marginal pension entitlements $q(\theta) B'$ depends on skill, i.e., if $q'B'$ is nonzero. This additional term does not exist in the static case and is not captured by conventional elasticity concepts with respect to marginal tax or transfer rates. Therefore, the elasticity $\varepsilon_{y,\theta}$ does not vanish in the optimality condition (42) if $q'B'$ is nonzero. However, when the condition is stated in terms of the income distribution as in Equation (43), the transformation between income and skills drops out and the elasticity term takes the form that is familiar from static sufficient statistics results. In this sense, the trade-offs of the tax designer (as viewed with the tax perturbation approach) do not change in any deeper way, even though the elasticity term is somewhat more elaborate than in the static case.

The term $\mathcal{C} = (1 - F) / (\theta f)$ has its usual form and captures the inverse Pareto coefficient of the skill distribution. Intuitively, this term relates equity gains and efficiency losses induced by the local marginal tax rate. The term increases with the proportion of taxpayers with skills above θ (whose tax level increases with the local marginal tax rate) and decreases with the productivity and the proportion of taxpayers with skill θ (whose choices are distorted by the local marginal tax rate).

A.4 Optimal control problem with multi-period horizon

In the multi-period setting of Section 3.3.2, the individual optimization problem for type θ is to choose consumption c_t for $t = 1, \ldots, N$, intertemporal transfers a_{t+1} for $t = 1, \ldots, N-1$, and income y_t for $t = 1, \ldots, R-1$ to maximize expected lifetime utility:

$$\max_{c_t, a_{t+1}, y_t} \sum_{t=1}^{R-1} \beta^{t-1} p_t(\theta)\, u_t\left(c_t - v_t\left(\frac{y_t}{w_t(\theta)}\right)\right) + \sum_{t=R}^{N} \beta^{t-1} p_t(\theta)\, u_t(c_t) \tag{74}$$

subject to the period budget constraints

$$\begin{aligned} c_t + q_{t+1}(\theta)\, a_{t+1} &= y_t + a_t - T_t(y_t) - S_t(y_t) \quad \text{for } t < R, \\ c_t + q_{t+1}(\theta)\, a_{t+1} &= a_t + B(y_1, \ldots, y_{R-1}) \quad \text{for } R \leq t < N, \\ c_t &= a_t + B(y_1, \ldots, y_{R-1}) \quad \text{for } t = N, \end{aligned} \tag{75}$$

where $a_1 = 0$.

To keep the multi-period extension close to the two-period optimal-control structure, I use a separable period-by-period reporting formulation of the multi-period planner problem. Each working period is associated with its own reporting margin, and the resulting net incomes are then aggregated through the intertemporal value function $\tilde{V}$ defined below. Specifically, if an individual of type θ reports a type of $\hat{\theta}$, then in working period t her net income (including the private value of newly acquired pension entitlements) minus labor disutility, valued in private period-t resources, is given by:

$$\hat{x}_t\left(\theta,\hat{\theta}\right) := y_t\left(\hat{\theta}\right) - T_t\left(y_t\left(\hat{\theta}\right)\right) - S_t\left(y_t\left(\hat{\theta}\right)\right) + \sum_{\tau=R}^{N} q^{t,\tau}(\theta)\, b_t\left(y_t\left(\hat{\theta}\right)\right) - v_t\left(\frac{y_t\left(\hat{\theta}\right)}{w_t(\theta)}\right). \quad (76)$$

Setting $x_t(\theta) := \hat{x}_t(\theta,\theta)$ and

$$Q_t(\theta) := \sum_{\tau=R}^{N} q^{t,\tau}(\theta), \quad (77)$$

I obtain

$$x_t(\theta) = y_t(\theta) - T_t(y_t(\theta)) - S_t(y_t(\theta)) + Q_t(\theta)\, b_t(y_t(\theta)) - v_t\left(\frac{y_t(\theta)}{w_t(\theta)}\right). \quad (78)$$

The factor $Q_t(\theta)$ is the private period-t value of one additional unit of annual pension benefits generated by income in period t. Since the benefit component $b_t(y_t)$ is paid in every retirement period $\tau = R,\ldots,N$, the individual values this entitlement by summing the private prices $q^{t,\tau}(\theta)$ of resources in those retirement periods. The envelope condition for period $t = 1,\ldots,R-1$ is given by:

$$\frac{dx_t}{d\theta} = \frac{\partial \hat{x}_t(\theta,\theta)}{\partial\theta} = l_t(\theta)\frac{w_t'(\theta)}{w_t(\theta)} v_t'(l_t(\theta)) + Q_t'(\theta)\, b_t(w_t(\theta)\, l_t(\theta)). \quad (79)$$

The government can save and borrow at the interest rate $r = 1/\beta - 1$. Its intertemporal budget balance condition is

$$\int_{\underline{\theta}}^{\overline{\theta}} \left[\sum_{t=1}^{R-1} \beta^{t-1} p_t(\theta)\left[T_t(y_t(\theta)) + S_t(y_t(\theta))\right] - \sum_{\tau=R}^{N} \beta^{\tau-1} p_\tau(\theta) \sum_{t=1}^{R-1} b_t(y_t(\theta))\right] dF(\theta) - E \geq 0. \quad (80)$$

Equivalently, the condition can be written as

$$\int_{\underline{\theta}}^{\overline{\theta}} \sum_{t=1}^{R-1} \beta^{t-1} p_t(\theta) \left[T_t(y_t(\theta)) + S_t(y_t(\theta)) - P_t^F(\theta) b_t(y_t(\theta)) \right] dF(\theta) - E \geq 0. \tag{81}$$

where the fiscal period-t value of one unit of pension benefits associated with b_t is

$$P_t^F(\theta) := \sum_{\tau=R}^{N} \beta^{\tau-t} \frac{p_\tau(\theta)}{p_t(\theta)}. \tag{82}$$

Thus, $Q_t(\theta)$ is the private value of a unit of annual pension benefits generated by income in period t, whereas $P_t^F(\theta)$ is its expected fiscal cost, both expressed in period-t resources.

Using

$$x_t(\theta) = w_t(\theta) l_t(\theta) - T_t(w_t(\theta) l_t(\theta)) - S_t(w_t(\theta) l_t(\theta)) + Q_t(\theta) b_t(w_t(\theta) l_t(\theta)) - v_t(l_t(\theta)), \tag{83}$$

the government's budget balance condition can be rewritten as

$$\begin{aligned} \int_{\underline{\theta}}^{\overline{\theta}} \sum_{t=1}^{R-1} \beta^{t-1} p_t(\theta) & \left[w_t(\theta) l_t(\theta) - x_t(\theta) - v_t(l_t(\theta)) \right. \\ & \left. + \left[Q_t(\theta) - P_t^F(\theta) \right] b_t(w_t(\theta) l_t(\theta)) \right] dF(\theta) - E \geq 0 \end{aligned} \tag{84}$$

For a given profile $(x_1, \ldots, x_{R-1})$ of net incomes minus labor disutilities, the intertemporal optimization problem of an individual of type θ is

$$\begin{aligned} \tilde{V}(x_1, \ldots, x_{R-1}; \theta) := \max_{A_2, \ldots, A_N} & \sum_{t=1}^{R-1} \beta^{t-1} p_t(\theta) u_t(x_t + A_t - q_{t+1}(\theta) A_{t+1}) \\ & + \sum_{t=R}^{N} \beta^{t-1} p_t(\theta) u_t(A_t - q_{t+1}(\theta) A_{t+1}), \end{aligned} \tag{85}$$

where $A_1 = A_{N+1} = 0$ and $q_{N+1} A_{N+1} = 0$. The variables A_t denote transformed intertemporal asset positions after the pension entitlements have been capitalized into the working-period resources x_t. Thus, pension benefits do not appear again in retirement-period consumption in this transformed intertemporal problem.

The intertemporal value function of an individual of type θ is defined as

$$V(\theta) := \tilde{V}(x_1(\theta), \ldots, x_{R-1}(\theta); \theta). \tag{86}$$

With these preparations, the government problem can be stated as the following optimal control problem with state vector $(x_1, \ldots, x_{R-1})$ and control vector $(l_1, \ldots, l_{R-1})$:

$$\max_{x,l} \int_{\underline{\theta}}^{\overline{\theta}} \tilde{V}(x_1(\theta), \ldots, x_{R-1}(\theta); \theta) \, dF(\theta) \tag{87}$$

subject to

$$\int_{\underline{\theta}}^{\overline{\theta}} \sum_{t=1}^{R-1} \beta^{t-1} p_t(\theta) \left[w_t(\theta) l_t(\theta) - x_t(\theta) - v_t(l_t(\theta)) \right.$$
$$\left. + \left[Q_t(\theta) - P_t^F(\theta) \right] b_t(w_t(\theta) l_t(\theta)) \right] dF(\theta) - E \geq 0,$$

and $$\frac{dx_t}{d\theta} = l_t(\theta) \frac{w_t'(\theta)}{w_t(\theta)} v_t'(l_t(\theta)) + Q_t'(\theta) b_t(w_t(\theta) l_t(\theta)).$$

A.5 Solution of parameterized model

A.5.1 Individual problem

Using $v(l) = \left(1 + \frac{1}{\phi}\right)^{-1} l^{1+\frac{1}{\phi}}$, the first-order condition for labor supply implies

$$l = \left[\theta \left(1 - T' - S' + qB'\right)\right]^{\phi}.$$

Income is given by

$$y = \theta l = \theta^{1+\phi} \left[1 - T' - S' + qB'\right]^{\phi}.$$

Inverting this relationship, skill can be expressed as

$$\theta = y^{\frac{1}{1+\phi}} \left[1 - T' - S' + qB'\right]^{-\frac{\phi}{1+\phi}}.$$

To find the marginal utility of consumption, I define the net lifetime income

$$y^n := y - T(y) - S(y) + q(\theta) B(y)$$

and solve the intertemporal problem:

$$\max_{c_2} u\left(y^n - v\left(l\right) - q\left(\theta\right)c_2\right) + \beta p\left(\theta\right)u\left(c_2\right).$$

The first-order condition

$$q\left(\theta\right)\left(y^n - v\left(l\right) - q\left(\theta\right)c_2\right)^{-\sigma} = \beta p\left(\theta\right)c_2^{-\sigma}$$

implies

$$c_2 = \frac{y^n - v\left(l\right)}{\left(\frac{q(\theta)}{\beta p(\theta)}\right)^{\frac{1}{\sigma}} + q\left(\theta\right)}.$$

Hence, the marginal utility of consumption (generated by a dollar received in the first period) is

$$u' = \left(y^n - v\left(l\right) - q\left(\theta\right)\frac{y^n - v\left(l\right)}{\left(\frac{q(\theta)}{\beta p(\theta)}\right)^{\frac{1}{\sigma}} + q\left(\theta\right)}\right)^{-\sigma}.$$

Equivalently,

$$u' = \left(1 + q\left(\theta\right)^{1-\frac{1}{\sigma}}\left(\beta p\left(\theta\right)\right)^{\frac{1}{\sigma}}\right)^{\sigma}\left[y^n - v\left(l\right)\right]^{-\sigma}.$$

The consumption levels in the two periods can be expressed as

$$c_2 = \frac{y^n - v\left(l\right)}{q\left(\theta\right)\left(1 + q\left(\theta\right)^{\frac{1}{\sigma}-1}\left(\beta p\left(\theta\right)\right)^{-\frac{1}{\sigma}}\right)}$$

and

$$c_1 = y^n - q\left(\theta\right)c_2.$$

A.5.2 Optimal tax

Using the transversality condition (60), I obtain:

$$\mu\left(\theta\right) = \int_{\theta}^{\overline{\theta}}\left[u'\left(x\left(\tilde{\theta}\right) - q\left(\tilde{\theta}\right)c_2\left(\tilde{\theta}\right)\right) - \lambda\right]dF\left(\tilde{\theta}\right).$$

Due to the free initial state, I obtain the transversality condition $\mu(\underline{\theta}) = 0$ and hence

$$\lambda = \int_{\underline{\theta}}^{\overline{\theta}} u'\left(x\left(\tilde{\theta}\right) - q\left(\tilde{\theta}\right) c_2\left(\tilde{\theta}\right)\right) dF\left(\tilde{\theta}\right).$$

Using Equation (41), the optimality condition can be expressed in terms of the statutory marginal tax rates and the fiscal pension wedge as

$$\frac{T' + \omega^F}{1 - T'} = \frac{\varepsilon_{y,\theta}}{\varepsilon_{y,(1-T')}} \left(1 - \frac{\int_{\theta}^{\overline{\theta}} u'\, dF}{\lambda(1 - F)}\right) \frac{1 - F}{\theta f}.$$

The right-hand side can be decomposed into the following factors:

$$\begin{aligned}
\mathcal{A}(\theta) &= \frac{\varepsilon_{y,\theta}}{\varepsilon_{y,(1-T')}} = \frac{1 + \phi + \frac{\theta^2 q' B'}{l v''}}{\phi + \frac{\theta(S' - qB')}{l v''}}, \\
\mathcal{B}(\theta) &= 1 - \frac{\int_{\theta}^{\overline{\theta}} u'\left(\tilde{\theta}\right) f\left(\tilde{\theta}\right) d\tilde{\theta}}{\lambda(1 - F(\theta))}, \\
\mathcal{C}(\theta) &= \frac{1 - F(\theta)}{\theta f(\theta)}.
\end{aligned}$$

Solving the optimality condition for the statutory marginal tax rate yields

$$T' = \frac{\mathcal{A}(\theta)\,\mathcal{B}(\theta)\,\mathcal{C}(\theta) - \omega^F}{1 + \mathcal{A}(\theta)\,\mathcal{B}(\theta)\,\mathcal{C}(\theta)}.$$

Alternatively, defining the factor

$$\mathcal{D}(\theta) = \frac{T'(\theta)}{T'(\theta) + \omega^F(\theta)},$$

the same condition can be rearranged as

$$\frac{T'}{1 - T'} = \mathcal{A}(\theta)\,\mathcal{B}(\theta)\,\mathcal{C}(\theta)\,\mathcal{D}(\theta).$$

Equivalently, conditional on the value of $\mathcal{D}(\theta)$, I obtain

$$T' = \frac{\mathcal{A}(\theta)\,\mathcal{B}(\theta)\,\mathcal{C}(\theta)\,\mathcal{D}(\theta)}{1 + \mathcal{A}(\theta)\,\mathcal{B}(\theta)\,\mathcal{C}(\theta)\,\mathcal{D}(\theta)}.$$

References

Abraham, A., P. Brendler, and E. Carceles (2023): "Optimal Income Redistribution," Mimeo. 1

Adams, P., M. D. Hurd, D. McFadden, A. Merrill, and T. Ribeiro (2003): "Healthy, wealthy, and wise? Tests for direct causal paths between health and socioeconomic status," *Journal of Econometrics*, 112, 3–56. 2.1

Bagchi, S. (2019): "Differential mortality and the progressivity of social security," *Journal of Public Economics*, 177, 104044. 1

Banerjee, A., E. Duflo, G. Postel-Vinay, and T. Watts (2010): "Long-Run Health Impacts of Income Shocks: Wine and Phylloxera in Nineteenth-Century France," *Review of Economics and Statistics*, 92, 714–728. 2.1

Benabou, R. (2002): "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?" *Econometrica*, 70, 481–517. 4.1

Bommier, A., M.-L. Leroux, and J.-M. Lozachmeur (2011): "On the public economics of annuities with differential mortality," *Journal of Public Economics*, 95, 612–623. 1

Cesarini, D., E. Lindqvist, R. Östling, and B. Wallace (2016): "Wealth, Health, and Child Development: Evidence from Administrative Data on Swedish Lottery Players," *Quarterly Journal of Economics*, 131, 687–738. 2.1

Chetty, R. (2009): "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods," *Annual Review of Economics*, 1, 451–488. 3.2

——— (2012): "Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply," *Econometrica*, 80, 969–1018. 4.1

Chetty, R., M. Stepner, S. Abraham, S. Lin, B. Scuderi, N. Turner, A. Bergeron, and D. Cutler (2016): "The Association Between Income and Life Expectancy in the United States, 2001-2014," *Journal of the American Medical Association*, 315, 1750–1766. 1, 4.1, 1

Cremer, H., J.-M. Lozachmeur, and P. Pestieau (2004): "Social security, retirement age and optimal income taxation," *Journal of Public Economics*, 88, 2259–2281. 1, 1

Cutler, D., A. Deaton, and A. Lleras-Muney (2006): "The determinants of mortality," *Journal of Economic Perspectives*, 20, 97–120. 2.1

Cutler, D. M., A. Lleras-Muney, and T. Vogl (2011): "Socioeconomic status and health: dimensions and mechanisms," in *The Oxford Handbook of Health Economics*, ed. by S. Glied and P. C. Smith, Oxford University Press. 1, 2.1

Diamond, P. and E. Saez (2011): "The Case for a Progressive Tax: From Basic Research to Policy Recommendations," *Journal of Economic Perspectives*, 25, 165–90. 1, 11

Diamond, P. and J. Spinnewijn (2011): "Capital Income Taxes with Heterogeneous Discount Rates," *American Economic Journal: Economic Policy*, 3, 52–76. 3

Diamond, P. A. (1998): "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates," *American Economic Review*, 88, 83–95. 1, 2.4, 3, 3.2, 3.2, A.3

Erixson, O. (2017): "Health responses to a wealth shock: evidence from a Swedish tax reform," *Journal of Population Economics*, 30, 1281–1336. 2.1

European Commission (2013): *Health inequalities in the EU – Final report of a consortium - Consortium lead: Sir Michael Marmot*, European Commission Directorate-General for Health and Consumers. 1

Farhi, E. and I. Werning (2013): "Insurance and taxation over the life cycle," *Review of Economic Studies*, 80, 596–635. 2

Ferey, A. (2022): "Redistribution and Unemployment Insurance," Mimeo. 1, 1

Frijters, P., J. P. Haisken-DeNew, and M. A. Shields (2005): "The causal effect of income on health: Evidence from German reunification," *Journal of Health Economics*, 24, 997–1017. 2.1

Gerdtham, U.-G., P. Lundborg, C. H. Lyttkens, and P. Nystedt (2016): "Do Education and Income Really Explain Inequalities in Health? Applying a Twin Design," *The Scandinavian Journal of Economics*, 118, 25–48. 2.1

Goldman, N. (2001): "Social inequalities in health: disentangling the underlying mechanisms," *Annals of the New York Academy of Sciences*, 954, 118–139. 1

Golosov, M., M. Troshkin, and A. Tsyvinski (2016): "Redistribution and social insurance," *American Economic Review*, 106, 359–386. 2

Grochulski, B. and N. Kocherlakota (2010): "Nonseparable preferences and optimal social security systems," *Journal of Economic Theory*, 145, 2055 – 2077. 2

Heathcote, J., K. Storesletten, and G. L. Violante (2014): "Optimal Tax Progressivity: An Analytical Framework," IFS Working Papers W14/27, Institute for Fiscal Studies. 4.1

——— (2017): "Optimal Tax Progressivity: An Analytical Framework," *Quarterly Journal of Economics*, 132, 1693–1754. 4.1

Hosseini, R. and A. Shourideh (2019): "Retirement financing: An optimal reform approach," *Econometrica*, 87, 1205–1265. 1, 1, 6

IRS (2022): *Statistics of Income—2020 Individual Income Tax Returns*, 1304 (Rev. 11-2022), Internal Revenue Service. 4.1

Jang, Y., S. Pashchenko, and P. Porapakkarm (2023): "Mortality Regressivity and Pension Design," Available at SSRN 4509697. 1, 1

Jessen, L. J., S. Koehne, P. Nüss, and J. Ruhose (2024): "Socioeconomic Inequality in Life Expectancy: Perception and Policy Demand," Mimeo. 5

Jones, J. B. and Y. Li (2023): "Social Security reform with heterogeneous mortality," *Review of Economic Dynamics*, 48, 320–344. 1

Koehne, S. (2018): "On the Taxation of Durable Goods," *International Economic Review*, 59, 825–857. 2

Koehne, S. and M. Kuhn (2015): "Optimal taxation in a habit formation economy," *Journal of Public Economics*, 122, 31–39. 2

Koehne, S. and D. Sachs (2022): "Pareto-improving reforms of tax deductions," *European Economic Review*, 148. 4.1

Lampert, T., J. Hoebel, and L. E. Kroll (2019): "Soziale Unterschiede in der Mortalität und Lebenserwartung in Deutschland - Aktuelle Situation und Trends," *Journal of Health Monitoring*, 4, 3–15. 1

Leroux, M.-L. and G. Ponthiere (2013): "Utilitarianism and unequal longevities: A remedy?" *Economic Modelling*, 30, 888–899. 1

Mackenbach, J. P., I. Stirbu, A.-J. R. Roskam, M. M. Schaap, G. Menvielle, M. Leinsalu, and A. E. Kunst (2008): "Socioeconomic inequalities in health in 22 European countries," *New England Journal of Medicine*, 358, 2468–2481. 1

Mankiw, N. G., M. Weinzierl, and D. Yagan (2009): "Optimal Taxation in Theory and Practice," *Journal of Economic Perspectives*, 23, 147–74. 1, 10, 4.2

Meer, J., D. L. Miller, and H. S. Rosen (2003): "Exploring the health–wealth nexus," *Journal of Health Economics*, 22, 713–730. 2.1

Michau, J.-B. (2014): "Optimal redistribution: A life-cycle perspective," *Journal of Public Economics*, 111, 1–16. 1

Mirrlees, J. A. (1971): "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38, 175–208. 1, 2.4

——— (1976): "Optimal tax theory: A synthesis," *Journal of Public Economics*, 6, 327–358. 2.4

Piketty, T. and E. Saez (2013): "Chapter 7 - Optimal Labor Income Taxation," in *Handbook of Public Economics*, ed. by A. J. Auerbach, R. Chetty, M. Feldstein, and E. Saez, Elsevier, vol. 5 of *Handbook of Public Economics*, 391–474. 1

Sachs, D., A. Tsyvinski, and N. Werquin (2020): "Nonlinear tax incidence and optimal taxation in general equilibrium," *Econometrica*, 88, 469–493. 4.1, 4.2

Saez, E. (2001): "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 68, 205–29. 2.4, 3.2, 9, 4.1, 4.2, A.2

Sheshinski, E. (2008): *The Economic Theory of Annuities*, Princeton University Press. 1

Shourideh, A. and M. Troshkin (2017): "Incentives and efficiency of pension systems," Mimeo. 1, 1

SSA (2022): *Social Security Bulletin: Annual statistical supplement*, 13-11700, Social Security Administration. 4.1

Stantcheva, S. (2020): "Dynamic Taxation," *Annual Review of Economics*, 12, 801–831. 2