Soc Choice Welf (2018) 50:519–546
https://doi.org/10.1007/s00355-017-1094-0

ORIGINAL PAPER

On the redistributive power of pensions

Philippe Choné[1] · **Guy Laroque**[2,3]

Received: 4 December 2015 / Accepted: 27 October 2017 / Published online: 14 November 2017
© The Author(s) 2017. This article is an open access publication

Abstract We study the tradeoff between efficiency and redistribution in a model with overlapping generations, extensive labor supply, and perfect financial markets. The government instruments are a pension scheme and a age-independent nonlinear income tax schedule. At the second-best optimum, the pension system constrains the agents' labor supply behavior, forcing them to work to achieve a required lifetime performance. Income taxes affect labor supply directly, but also indirectly through pension incentives. The indirect effect of taxes counteracts the usual forces in the efficiency-redistribution tradeoff: *through the interplay with the pension system*, decreasing taxes induces redistribution and reduces productive efficiency.

Funding is acknowledged from the European Research Council (reference ERC-2010-AdG-269440-WSC WTBDS). Previous versions of the paper have been presented at conferences and seminars in Amsterdam, Koln, Nashville, London, Lugano, Marseille, Munich and Uppsala. We are most grateful for the comments of Felix Bierbrauer, Soren Blomquist, Monika Buetler Mikhail Golosov, Jean-Baptiste Michau, Nicola Pavoni, Richard Rogerson, Tom Sargent, Laurent Simula and Eytan Sheshinski.

✉ Guy Laroque
g.laroque@ucl.ac.uk

Philippe Choné
philippe.chone@ensae.fr

1 Centre de Recherche en Economie et Statistique (CREST), Universite Parris-Saclay, 5 av. Henry le Chatelier, 91120 Palaiseau, France

2 Department of Economics, University College London, London, UK

3 Department of Economics, Sciences-Po, Paris, France

1 Introduction

In the past ten years, following Prescott (2004)'s claim that the differences in work habits in the US and in Europe were largely due to the differences in the tax systems, a number or researchers have estimated labor supply elasticities both at the microeconomic and macroeconomic levels. It seems that an important, perhaps previously neglected, element is the extensive margin and its reaction to financial incentives at the beginning and at the end of the working life. This has led to a number of models with endogenous retirement dates in a life cycle setup, e.g. Prescott et al. (2009), Rogerson and Wallenius (2009) and Ljunqvist and Sargent (2014). However there is still little work on the interaction between nonlinear taxes and pension schemes. Indeed Diamond (2009) states

> Apart from some simulation studies, theoretical studies of optimal tax design typically contain neither a mandatory pension system nor the behavioral dimensions that lie behind justifications commonly offered for mandatory pensions. Conversely, optimizing models of pension design typically do not include annual taxation of labor and capital incomes. Recognizing the presence of two sets of policy institutions raises the issue of whether normative analysis should be done separately or as a single overarching optimization.

To make progress in this direction, the present study takes as a premise that two simple policy instruments are available to the (utilitarian) government: an age-independent nonlinear tax schedule and a pension scheme that depends on a single aggregate statistic of the work life. We adopt a Ramsey approach rather than a Mirrlees one. In this respect, we sharply differ from the dynamic optimal taxation literature that allows for very general tax instruments, possibly depending on the agents' age and on their entire labor income histories (Grochulski and Kocherlakota 2010; Weinzierl 2011; Shourideh and Troshkin 2012; Michau 2014). To study the redistributive power of pensions in this Ramsey environment, we regard pension benefits as transfers to the agents that are conditioned upon a verifiable variable. In practice, the government offers a menu of pension plans, and each agent chooses her preferred option depending on her characteristics.

Our main goal is to understand how the incentive constraints associated with pension choices affect the design of optimal taxes. To this aim, following the optimal taxation literature (Saez 2001, 2002; Jacquet et al. 2013), we proceed by evaluating welfare effects of small tax reforms around the optimum. The literature defines the mechanical or equity effect as the welfare change that would occur if there were no behavioral responses. Studies differ in the kind of behavioral responses they consider. Responses at the intensive margin give rise to substitution effects while responses at the extensive margin give rise to participation effects. The literature derives optimal tax formulae in these various contexts, sometimes allowing for income effects in labor supply. As noticed by Diamond (1998), however, the formulae are much simpler in the absence of income effects. Here, we assume away income effects, which are difficult to handle when a retirement system is in force.

We focus on the extensive margin, appropriate to discuss the retirement decision.[1] In a first step, we examine how labor supply behavior depends on taxes and pensions. Under a pension system, agents face constraints on labor supply because they need to achieve a certain lifetime performance to receive their pension benefits. These constraints tend to reduce the (positive) labor supply responses to tax cuts compared to the case without pension. There is, however, a distinct, subtler effect which turns out to go in the same direction. The strength of the labor supply constraints placed by the pension system is itself endogenous, a function of the tax schedule: a tax cut reduces the pressure placed by the pension regime on labor supply because it becomes easier for the agent to meet a given pension requirement. This second effect, which we call "*feedback effect of taxes on pension incentives*", further reduces the labor supply response to tax cuts.

Our main contribution is to analyze the welfare effects of tax reforms when pensions are optimally designed by the government. Redistribution requires extracting rent from agents with high productivities and/or low opportunity costs of work. Rent extraction, however, is to be weighed against efficiency concerns. In principal-agent models, the rent-efficiency tradeoff is best described with the notion of virtual surplus, first introduced by Myerson (1981) in the context of auctions.[2] Using this methodology, we distinguish the efficiency effect of taxes (effect of taxes on net production) from their redistributive effect (effect on lifetime consumption inequalities), while accounting for endogenous pension choices.

In a redistributive economy, we find that the presence of a pension scheme attenuates the usual forces that govern the welfare effects of tax reforms. While usually taxes are useful to redistribute wealth across agents and are detrimental to efficiency, the feedback effects of taxes on pension incentives introduce new forces that play in the opposite directions, making taxes both less redistributive and less detrimental to productive efficiency. In a specialized framework—same marginal utility of consumption for all agents, decreasing productivity and increasing pecuniary cost of work with age—we show that an optimal combination of tax and pension eliminates all rents coming from observable productivity differences. Furthermore it suppresses all upward labor supply distortions and reduces downward distortions. At the optimum the two instruments fully specialize: pensions provide the incentives to work, while taxes do all the redistribution.

The study which is closest to ours is that of Cremer et al. (2004). Our framework, however, differs from theirs in a number of important dimensions. First, Cremer et al. (2004) adopt a mechanism design approach and when it comes to implementation allow for a general, hybrid policy instrument that incorporates both pensions and taxes. Second, at the implementation stage, they restrict their analysis to a simple form of heterogeneity (two types). Third, their pension is based on retirement age.

[1] We check in an extension that most of the theoretical analysis carries over when the intensive margin is considered, see Appendix B.

[2] As explained by Laffont and Martimort (2002), the virtual surplus is a convenient way to account for incentive constraints in a variety of situations, e.g. nonlinear pricing, regulation, etc. In the present context, the virtual surplus embodies the constraint that the agent choices coincide with those expected by the government.

Here, the pension scheme lets entitlements depend on the agents' entire work history (through an aggregate statistic), and consequently the pension system exerts pressure on labor supply all over the agents' life cycles. As a consequence, the interaction between the fiscal and pension instruments is more complex than the implicit tax on prolonged activity that is well explained by e.g. Gruber and Wise (1999) and Cremer and Pestieau (2016).

Finally, it should be noted that the present study focuses on redistribution and ignores a number of important aspects of pension systems, such as uncertainty and insurance (Diamond and Mirrlees 1978; Golosov et al. 2016), political considerations (Cremer et al. 2008), and demographics (Brett 2012).

The article is organized as follows. Section 2 presents the economic environment, the government instruments, and the second-best problem. Section 3 explains how a pension system may constrain labor supply decisions and how this constraint depends on the shape of the nonlinear tax schedule, thus introducing the "feedback effects" of taxes on pension incentives. Section 4 shows that at the second-best optimum of a redistributive economy interior types have their labor supply constrained by the pension system, and that the pension feedback effects increase the redistributive role and decrease the efficiency role of income taxes. Section 4.6 provides a calibrated example with two types.

2 Model

We consider a deterministic overlapping generations model in continuous time, where all agents have the same length of life, normalized to one. At each date labor supply is extensive, either 0 or 1. Dynasties differ in their (deterministic) profiles of productivity and pecuniary cost of going to work, as well as by their instantaneous utility for consumption. The government wants to redistribute lifetime welfare across dynasties.

Formally, agents have different types, indexed by θ, where θ is distributed with the cdf F on the set Θ. At age a, $0 \leq a \leq 1$, an agent of type θ produces at most $w(a;\theta)$ units of a single homogeneous good but suffers a pecuniary cost $\delta(a;\theta)$, measured in units of good, if she works. Agent θ lifetime utility is

$$\int_0^1 u[c(a;\theta);\theta]\,\mathrm{d}a,$$

where $u(.;\theta)$ is an increasing concave utility index and $c(a;\theta)$ denotes consumption at age a.

An agent is thus characterized by a couple of exogenous, nonnegative functions $(w(\cdot;\theta),\delta(\cdot;\theta))$ and by the instantaneous utility index $u(\cdot;\theta)$. The pair $(w(a;\theta),\delta(a;\theta))$ as age a varies describes a curve in the (w,δ)-space, that we call a *trajectory*. We assume that the functions w, δ, and u are differentiable. To illustrate general qualitative properties, we use in some parts of the paper a specialized framework, which may be seen as describing the end of the agents' life, when productivity declines and cost of work rises. In practice, this is presumably during that period that most agents retire.

Assumption 2.1 (Specialized framework) *The productivities $w(a;\theta)$ decrease with age and the pecuniary costs of work $\delta(a;\theta)$ increase with age. Moreover, the agents utility index $u(c)$ does not depend on θ.*

At each date t, for each θ in Θ, the economy contains a continuum of agents of type θ of all ages a in [0, 1]; overtime the older agents die and are replaced by newborn of the same type. All cohorts are of the same size, with one agent of each type, and the economy is stationary. An *allocation* specifies the nonnegative consumption $c(a;\theta)$ and the labor supply $\ell(a;\theta)$ in {0, 1} of all types θ along their lives.

Furthermore we assume that there are perfect markets for transferring wealth across time, with a zero interest rate. The agents use these markets to smooth their consumption overtime, so that we can remove the age argument in consumption and write simply $c(a;\theta) = c(\theta)$. Also we note $y(\theta)$ the lifetime net output produced by agent θ, i.e., $y(\theta) = \int_0^1 [w(a;\theta) - \delta(a;\theta)]\,\ell(a;\theta)\,\mathrm{d}a$.

Feasibility An allocation is *feasible* if and only if total consumption does not exceed total output net of production cost:

$$\int_\Theta c(\theta)\,\mathrm{d}F(\theta) \leq \int_\Theta y(\theta)\,\mathrm{d}F(\theta). \tag{1}$$

Efficiency An allocation is *efficient* whenever output net of production costs is maximized, i.e., $\ell(a,\theta) = 0$ if $\delta(a;\theta) > w(a;\theta)$ and $\ell(a,\theta) = 1$ if $\delta(a;\theta) < w(a;\theta)$.

Utilitarian optimum (first-best) The utilitarian optimum is the allocation that maximizes $\int_\Theta u([c(\theta);\theta]\,\mathrm{d}F(\theta)$ subject to the feasibility constraint (1). It is the feasible efficient allocation such that marginal utilities are equal: for all θ in Θ

$$u_c[c(\theta),\theta] = \lambda.$$

Laissez-faire Laissez-faire induces the agents to maximize their lifetime consumption

$$c(\theta) = \int_0^1 \max(0,\, w(a;\theta) - \delta(a,\theta))\,\mathrm{d}a.$$

They work whenever their productivity is larger than their opportunity cost of work, so the laissez-faire equilibrium is efficient. In general, laissez-faire yields an allocation that differs from the utilitarian optimum.

In all the paper we suppose that the utilitarian government observes the employment status of the agents and, when they work, their productivity w. It never observes the pecuniary cost δ, which is private information.

The government instruments The government has access to two policy instruments, an income tax and a retirement scheme. The first policy instrument is a time invariant income tax schedule. The tax schedule is described by a nondecreasing function $R(w)$, the age-independent after-tax income of a worker with before tax wage w.

The second instrument is a pension scheme that relates a lifetime statistic Z, to a (possibly negative) government lifetime transfer $P(Z)$, which represents the present

value of all contributions and benefits associated with the retirement plan. An agent is entitled to receive $P(Z)$ provided that her lifetime performance is at least equal to Z:

$$\int_0^1 z(w(a;\theta))\,\ell(a;\theta)\,\mathrm{d}a \geq Z. \tag{2}$$

As the above equation shows, we assume in this study that the pension system relies on a single performance indicator that is linear in labor supply. The contribution of working at age a to the pension requirement, $z(w(a;\theta))$, depends positively—and possibly nonlinearly—on the observed productivity at that age. To illustrate, we consider three stripped down legislations:

- Regime L, $z(w) = 1$: all working years bring identical contributions to the pension requirement, which here coincides with total time worked over life;
- Regime W, $z(w) = w$: working years are weighted by the corresponding productivity, and the pension statistic is lifetime gross earnings;
- Regime N, $z(w) = R(w)$: the pension statistic is net lifetime earnings.

In practice, the pension transfers depend on individual labor histories through a number of channels. The regimes that we analyze are far from exhausting the existing legislations.[3] Note in particular that we do not allow the tax schedule to be age dependent, contrary to Weinzierl (2011), nor do we deal with the financial market imperfections that underlie some of the pension regimes in practice. Our pension regimes can be seen as a restricted way of introducing age-dependent transfers.

Second best program Facing the tax schedule $R(\cdot)$ and a pension regime associated with transfers $P(\cdot)$, the consumer chooses her labor supply $\ell(a) \in \{0, 1\}$ and pension plan Z so as to maximize her lifetime utility

$$c(\theta) = \max_{\ell, Z} \int_0^1 [R(w(a;\theta))) - \delta(a;\theta)]\ell(a)\,\mathrm{d}a + P(Z) \tag{4}$$

under the pension requirement (2). The second best program consists in designing the tax and pension schedules to maximize the sum of utilities

$$\int_\Theta u[c(\theta);\theta]\,\mathrm{d}F(\theta)$$

[3] The analysis up to Sect. 4.4 holds when the contribution of current work to the lifetime pension requirement is made dependent on age, i.e., z is a function of both productivity and age. It would be more difficult to deal with pension formulas that consider the N 'best' years rather than the whole life cycle as in (2). The performance indicator would then be

$$\int_0^1 z(w(a;\theta))\ell(a;\theta)n(a;\theta)\,\mathrm{d}a \tag{3}$$

where $n(a;\theta) = 1$ if $z(w(a;\theta))\ell(a;\theta)$ is above its $(1-N)$-th percentile and $n(a;\theta) = 0$ otherwise. If $n(a;\theta) = 1$, the work record at age a is retained for the computation of the lifetime pension requirement, otherwise it is discarded. Due to the definition of $n(a;\theta)$, the performance indicator (3) is nonlinear in labor supply. The pension decision involves both the choice of n and of Z and is more complicated than what we consider here.

under the feasibility constraint (1) when the agents choose their optimal consumption production and work levels according to program (4).

Notice that the above program does not involve a subsistence income s, a benefit often paid to the unemployed in extensive models. This is because the consumption equation would then take the form

$$c(\theta) = \max_{\ell, Z} \; s + \int_0^1 [\tilde{R}(w(a;\theta)) - \delta(a;\theta) - s]\ell(a)\,\mathrm{d}a + \tilde{P}(Z),$$

and the subsistence income appears to be superfluous by letting $R = \tilde{R} - s$ and $P = \tilde{P} + s$.

3 Labor supply

We first explain how labor supply depends on the tax schedule R and the pension requirement Z. It is useful to introduce the optimal lifetime earnings, the function $\gamma(Z;\theta)$ which is the maximum of

$$\int_0^1 [R(w(a;\theta)) - \delta(a;\theta)]\ell(a;\theta)\,\mathrm{d}a \tag{5}$$

over $\ell(.)$, subject to the pension requirement (2). Thus, lifetime consumption is $c(Z;\theta) = \gamma(Z;\theta) + P(Z)$. We also denote lifetime net output by $y(Z;\theta)$. Letting $\pi(Z;\theta)$ be the Lagrange multiplier associated with the pension constraint (2) for agent θ, we rewrite optimal lifetime earnings as

$$\gamma(Z;\theta) = \max_{\ell(a;\theta)} \left\{ \int_0^1 [R(w(a;\theta)) - \delta(a;\theta)]\,\ell(a;\theta)\,\mathrm{d}a + \pi(Z;\theta)\left(\int_0^1 z(w(a;\theta))\ell(a;\theta)\,\mathrm{d}a - Z \right) \right\}. \tag{6}$$

Agent θ's labor supply at age a if she picks plan Z is therefore given by

$$\ell(a;Z;\theta) = 1\!\!1_{R(w(a;\theta)) + \pi(Z;\theta)z(w(a;\theta)) - \delta(a;\theta) \geq 0}. \tag{7}$$

For our purpose, it is conceptually important to distinguish objects that are function of the pension requirement Z from objects that are evaluated at the agents' preferred plans. In this section, we take for granted that agent θ has chosen her preferred pension plan $Z(\theta)$ by solving the full program (4). Accordingly, we use the notations $\ell(a;\theta) = \ell(a;Z(\theta),\theta)$ for labor supply, $c(\theta) = c(Z(\theta);\theta)$ for lifetime consumption, $y(\theta) = y(Z(\theta);\theta)$ for lifetime net output, and $\pi(\theta) = \pi(Z(\theta);\theta)$ for the pension multiplier. Later in the analysis, we shall need to make the dependence in the pension requirement Z apparent.

Equation (7) clearly shows the absence of income effects in labor supply over the life cycle. In particular, labor supply is independent of the level of pension benefits, which

affects the agents' lifetime consumption.[4] To choose her labor supply agent θ must consider her adjusted tax schedule or financial incentive to work $R(w) + \pi(\theta)z(w)$. She works in regions where her trajectory is located below her incentive schedule $(w, R(w)+\pi(\theta)z(w))$, i.e., her opportunity cost of work δ is smaller than the financial incentive to work. The first component $R(w)$ represents the instantaneous after-tax income while the second term $\pi(\theta)z(w)$ represents the (deferred) pension benefit associated with before-tax earning w. The multiplier $\pi(\theta)$ can be thought of as an implicit conversion rate between after-tax earnings and pension benefits for agent θ.

A change in after-tax schedule R has two effects. First, at a given level of the pension multiplier $\pi(\theta)$, agent θ is subject to the incentive schedule $(w, R(w)+\pi(\theta)z(w))$ as productivity varies. Her work status changes at "switch points" where her trajectory crosses her incentive schedule. The static effect of a change in R outside switch points is zero. We show in the Appendix that a marginal tax rise around a switch point w, $\mathrm{d}R < 0$ on $[w, w+\mathrm{d}w]$, directly decreases agent θ's labor supply around w by $\eta|\mathrm{d}R|$, where $\eta > 0$ is the "static" elasticity.

Lemma 1 (Static labor supply elasticity) *The static labor of agent θ's supply elasticity around a switch point is given by*

$$\eta(w;\theta) = |\delta_a(a;\theta) - R'(w)w_a(a;\theta) - \pi(\theta)z_w w_a|^{-1}. \tag{8}$$

In the specialized framework of Assumption 2.1, the static elasticity is nondecreasing in the pension multiplier π.

In the specialized framework of Assumption 2.1, the static elasticity is weakly smaller when the pension requirement is binding ($\pi > 0$) than in the absence of retirement scheme ($\pi = 0$). The agents respond less vigorously to tax rises because they understand that they need to meet a pension requirement to receive the corresponding benefits. Specifically, the static elasticity under regime L is in fact independent from the pension multiplier because the derivative z_w is identically zero in that case. In the other regimes, the static elasticity decreases with the pension multiplier in the particular case where productivity declines and labor cost rises as time passes, i.e., where the agent's lifetime trajectory is decreasing in the (w, δ)-space.

Second, when the pension constraint (2) is binding, a change in the after-tax schedule affects agent θ's pension multiplier $\pi(\theta)$, and thus modifies indirectly the incentive schedule $R(w)+\pi(\theta)z(w)$. We refer to $\partial\pi/\partial R$ as the *feedback effect of taxes on pension incentives*. Increasing R around a switch point causes $\pi(\theta)$ to decrease, i.e.,

[4] Letting labor supply decisions depend on the level of pension benefits is possible but technically challenging. It can be done by assuming that $\delta(a)$ is a disutility cost instead of a pecuniary one, i.e., agent θ,
Footnote 4 continued
when working, produces $w(a)$ and has instantaneous utility $u(c(a)) - \delta(a)$, while she has instantaneous utility $u(c(a))$ when not working. The consumer maximizes $\int_0^1 [u(c(a)) - \delta(a)\ell(a)]\,\mathrm{d}a$ subject to the budget constraint $\int_0^1 [c(a) - w(a)\ell(a)]\,\mathrm{d}a = 0$. Switching from no work into work on the interval $[a, a+\mathrm{d}a]$ increases the (constant) consumption level by $\mathrm{d}c = w(a)\,\mathrm{d}a$, while it increases intertemporal utility by $\mathrm{d}U = u'(c)\,\mathrm{d}c - \delta(a)\,\mathrm{d}a$, Then agent θ works at age a under laissez-faire if and only if $u'(c)w(a) > \delta(a)$. Hence, this specification entails an *income effect* in labor supply: participation decreases with c. The analysis of optimal pension and tax schemes in this context is considerably more difficult, and is left for further research.

translates into less pressure placed by the pension scheme on the agent's labor supply. In this sense, the effects of pensions and taxes on labor supply are substitutes.

Proposition 1 (Feedback effect of taxes on pension incentives) *Tax cuts reduce the pressure exerted by the pension system on labour supply, thereby contributing to* decrease *labor supply.*

When the pension requirement is binding, a tax cut leads to lower pension multipliers. In other words, it weakens the labor supply incentives created by the pension system. The feedback effect affects labor supply in a nonlocal way, see (7). By "nonlocal", we mean that a change of R around a particular switch point alters labor supply around *all* switch points of the agent. We show in the Appendix that the *static effect* of taxes on labor supply locally (i.e., around the productivity level where taxes are changed) dominates the *feedback effect*. More precisely, labor supply around a switch point weakly increases following a local increase in after-tax income (a tax cut) around that switch point—while it decreases around the other switch points.

In the specialized framework of Assumption 2.1, trajectories are decreasing and therefore cannot intersect the incentive schedule $R(w) + \pi z(w)$ more than once. Hence they have at most one switch point. Suppose that an agent's labor supply is constrained by the pension system ($\pi(\theta) > 0$). Then the static effect and the feedback effect of income taxes cancel out exactly for that agent. This is because when an agent's trajectory has a single switch point, labor supply is entirely determined by the pension requirement (2), so changing R around the switch point has no effect on labor supply: the total labor supply elasticity in this particular case is zero.

4 Designing taxes and pensions

Section 4.1 exhibits conditions under which the single crossing property holds and attention can thus be restricted to local incentive constraints. In Sect. 4.2, we link the pattern of active binding incentive constraints to the shape of social weights. In Sect. 4.3, we look for the optimal pension requirements Z and derive a property of the agents' labor supply at the optimum. Section 4.4 presents necessary conditions for the optimal tax schedule. Section 4.5 shows that the two pension schemes L and N are equivalent when the agent's trajectories are decreasing and spell out how pensions allow to improve upon the second best optimum from a situation where the only available instrument is income tax. Finally Sect. 4.6 presents a calibrated example with two types.

4.1 Incentive constraints

To make sure that the pension level $Z(\theta)$ be chosen by agent θ, it is standard to replace the maximization with respect to Z with a set of incentive constraints and to use the consumption levels c rather than the transfers P as unknowns. Agent θ prefers her allocation to that of agent θ' if and only if

$$c(\theta) \geq \gamma(Z(\theta'); \theta) + P(Z(\theta')) = c(\theta') - \gamma(Z(\theta'); \theta') + \gamma(Z(\theta'); \theta). \quad (9)$$

Under a given pension regime, the government selects the retirement scheme P (or equivalently consumptions c) and the tax schedule R that maximize its utilitarian objective $\int_\Theta u(c(\theta);\theta)\,\mathrm{d}F(\theta)$ subject to the feasibility constraint (1) and the family of incentive constraints (9).

Assumption 4.1 (Type ordering) *The productivities and pecuniary costs are ordered: for all ages a, $w(a;\theta)$ is nondecreasing in θ and $\delta(a;\theta)$ decreases with θ.*

Throughout the remainder of the paper, we maintain the above assumption. Under that assumption, lifetime earnings (5) increase with agent type and the pension constraint (2), at given Z, is milder as θ rises. It follows that the optimal lifetime earnings $\gamma(Z;\theta)$ increase with θ, formally $\gamma_\theta(Z;\theta) > 0$. The next result shows that the slope $\partial\gamma(Z;\theta)/\partial Z$ also increases with θ. As $c(Z,\theta) = \gamma(Z;\theta) + P(Z)$, the latter implies single-crossing, $\partial c(Z;\theta)/\partial Z$ also increases with θ. All these properties hold for any given nondecreasing after-tax schedule $R(w)$ and any given nondecreasing pension requirement $z(w)$.

Lemma 2 *Under Assumption 4.1, the single-crossing property holds:*

$$\frac{\partial^2\gamma(Z;\theta)}{\partial\theta\partial Z} \geq 0. \tag{10}$$

It follows that an allocation $(P(\theta), Z(\theta))$, θ in Θ, is incentive compatible if and only if the envelope condition

$$\frac{\mathrm{d}c}{\mathrm{d}\theta} = \gamma_\theta(Z(\theta);\theta) \tag{11}$$

is satisfied together with the monotonicity requirement that $Z(\theta)$ is nondecreasing in θ.

The pension multiplier $\pi = -\partial\gamma/\partial Z$ is a measure of the pressure placed by the pension system on labor supply behavior. The single-crossing property expresses the fact that a given pension requirement Z places less pressure on more productive agents.

Virtual surplus The single-crossing property simplifies the government problem. Denoting by λ the multiplier associated with the feasibility constraint (1) and $\mu(\theta)$ that associated with the local incentive constraints (11), the Lagrangian can be written as

$$\mathcal{L} = \int_\Theta \left[u(c(\theta);\theta) - \lambda c(\theta) + \lambda y(Z;\theta)\right]\mathrm{d}F(\theta) + \int_\Theta \mu(\theta)\left[\frac{\mathrm{d}c}{\mathrm{d}\theta} - \gamma_\theta(Z(\theta);\theta)\right]\mathrm{d}\theta. \tag{12}$$

We now introduce the virtual surplus

$$S^v(Z;\theta) = \lambda y(Z;\theta) - \frac{\mu(\theta)}{f(\theta)}\gamma_\theta(Z;\theta), \tag{13}$$

which expresses the tradeoff between productive efficiency and rent extraction: on the one hand, the government wants the net output y produced by each agent to be as large

as possible; on the other, it wants to minimize the inequality in lifetime consumption, i.e., to avoid that high-type agents enjoy much higher lifetime consumption than low-type agents.

We can therefore rewrite the Lagrangian as

$$\mathcal{L} = \int_{\Theta} [u(c(\theta); \theta) - \lambda c(\theta)] \, \mathrm{d}F(\theta) + \int_{\Theta} \mu(\theta) \frac{\mathrm{d}c}{\mathrm{d}\theta} \, \mathrm{d}\theta + S^{v,\mathrm{Tot}}, \tag{14}$$

where $S^{v,\mathrm{Tot}}$ is the expected virtual surplus

$$S^{v,\mathrm{Tot}} = \int_{\Theta} S^{v}(Z(\theta); \theta) \, \mathrm{d}F(\theta). \tag{15}$$

The first two terms on the right-hand side of (14), which depend on the lifetime consumptions $c(\theta)$ and are controlled by the pension transfers $P(Z(\theta))$, are studied in Sect. 4.2. The expected virtual surplus embodies the tradeoff between productive efficiency and incentive constraints. It depends on the pension requirements $Z(\theta)$ and on the tax schedule; these two components are successively examined in Sects. 4.3 and 4.4.

4.2 Taxonomy of economies

We now introduce a simple taxonomy of economies which is associated with different patterns of binding incentive constraints. We say that an economy

- is *redistributive* if $u_c(c(\theta); \theta)$ decreases with θ, or more generally if the average weight of the agents less productive than any interior type χ exceeds the cost of public funds, $\int_0^{\chi} [u_c(c(\theta); \theta) - \lambda] \, \mathrm{d}F(\theta) > 0$;
- is *anti-redistributive* if $u_c(c(\theta); \theta)$ increases with θ, or more generally if for any interior type χ, $\int_0^{\chi} [u_c(c(\theta); \theta) - \lambda] \, \mathrm{d}F(\theta) < 0$;
- favors *middle classes* if the social weights of intermediate agents are above average, while those of low and high types are below average, so that $\int_0^{\chi} [u_c(c(\theta); \theta) - \lambda] \, \mathrm{d}F(\theta)$ is negative (positive) for small (large) χ.

Lemma 3 *In* (anti-)redistributive *economies, all the (upward) downward incentive constraints are binding. In* middle class *societies, upward (downward) incentive constraints are binding in the low (high) end of the population.*

Proof As a preliminary observation, we notice that increasing all pension benefits $P(\theta)$ by the same small amount in (14) yields the first-order condition:

$$\int_{\Theta} u_c(c(\theta), \theta) \, \mathrm{d}F(\theta) = \lambda. \tag{16}$$

We increase $\mathrm{d}c/\mathrm{d}\theta$ by $\delta c'$ between θ and $\theta + \delta\theta$, leaving c unchanged below θ. This increases c by $\delta c' \delta\theta$ above θ and therefore changes the Lagrangian by

$$\delta\mathcal{L} = (\delta c').(\delta\theta) \left[\int_{\theta}^{\bar{\theta}} [u_c(c(t); t) - \lambda] \, \mathrm{d}F(t) + \mu(\theta) \right].$$

Using (16) yields

$$\mu(\theta) = \int_{\underline{\theta}}^{\theta} [u_c(c(t); t) - \lambda] \, \mathrm{d}F(t). \tag{17}$$

From the above taxonomy of economies, we find that the Lagrange multiplier of the envelope condition, μ, is positive in redistributive economies and negative in anti-redistributive economies. In middle class societies, it is negative (positive) for low (high) types. □

Hereafter, we restrict attention to redistributive economies, for which we have seen that μ is positive for interior types and zero at extremal types. A particular class of redistributive economies obtains by assuming that the agents have the same utility function $u(c)$. Indeed, under this circumstance, $u_c(c(\theta); \theta)$ boils down to $u'(c(\theta))$ and we know from (11) that $c(\theta)$ increases with θ, hence $u'(c(\theta))$ decreases with θ.

4.3 The optimal choice of the pension requirements

We now study the optimal choice of the pension requirements $Z(\theta)$, keeping the income tax schedule fixed. (The determination of R is examined in the next section.) We also keep the lifetime consumption levels $c(\theta)$ fixed as they are controlled by the pension transfers $P(\theta)$. Pension requirements thus affect the Lagrangian (14) only through the expected virtual surplus (15), which incorporates the incentive constraints associated with pension choices.

Proposition 2 *The pension requirement $Z(\theta)$ that should be imposed on agent θ equalizes the tradeoff between productive efficiency and rent extraction:*

$$\lambda \frac{\partial y(Z; \theta)}{\partial Z} = \frac{\mu(\theta)}{f(\theta)} \frac{\partial^2 \gamma(Z; \theta)}{\partial Z \partial \theta} \geq 0. \tag{18}$$

When trajectories have only one switch point, labor supply is distorted downwards or undistorted.

Equation (18) is obtained by maximizing the virtual surplus (13) with respect to Z. Positivity follows from the single-crossing property (10). To compute the derivative of agent θ's lifetime net output $y(Z; \theta)$ with respect to the pension requirement Z, we observe that an increase in Z increases the pension multiplier π, which increases labor supply. To get her pension transfer P the agent has to work longer. For instance, in regime L, the adjusted after-tax schedule is shifted upwards, see Fig. 1. Assuming that the pension constraint (2) is binding, we show in the Appendix that

$$\frac{\partial y(Z; \theta)}{\partial Z} = \frac{\sum_{\sigma \in \mathcal{S}_\theta} (w_\sigma - \delta_\sigma) z(w_\sigma) \eta(w_\sigma)}{\sum_{\sigma \in \mathcal{S}_\theta} z^2(w_\sigma) \eta(w_\sigma)}, \tag{19}$$

where $\mathcal{S}_\theta$ is the set of agent θ's switch points. The term $(w_\sigma - \delta_\sigma)/z(w_\sigma)$ is a measure of the local distortion of labor supply around the concerned switch point. At any efficient

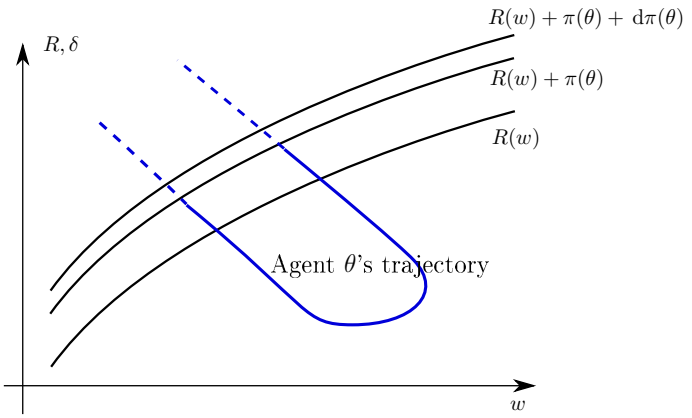

Fig. 1 In regime L, a small increase in Z raises the adjusted after tax-schedule $R + \pi$ by $\mathrm{d}\pi > 0$

allocation, this term is zero at all switch points. Labor supply is locally distorted downwards (upwards) when this term is positive (negative).

When an agent's trajectory has a single switch point,[5] the formula simplifies into

$$\frac{\partial y(Z;\theta)}{\partial Z} = \frac{w_\sigma - \delta_\sigma}{z(w(\sigma))},$$

which is nonnegative because of (18). It follows that the agent's labor supply is distorted downwards or undistorted: the agent never retires too late. This last result strikingly contrasts with the situation that prevails in the absence of a pension system. Choné and Laroque (2017) show that when the government's sole instrument is an age-independent nonlinear income tax schedule, it may be optimal to distort the labor supply of low type agents upwards. In particular, when the agents differ essentially in the productivity dimension, the occupations with the lowest productivity levels are occupied by low-type agents, and subsidizing such occupations allows to redistribute towards those agents. The availability of a pension scheme renders this utilization of income taxes suboptimal.

4.4 Optimal income tax

We now turn to the design of the optimal income tax schedule $R(w)$, keeping the pension requirements fixed, i.e., holding the equality $Z = Z(\theta)$ for each θ. As done above, we keep the lifetime consumption levels $c(\theta)$ fixed as they are controlled by the pension transfers $P(\theta)$. The income tax schedule thus affects the Lagrangian (14) only through the expected virtual surplus (15), which incorporates the incentive constraints associated with pension choices.

[5] This is the case in the specialized framework of Assumption 2.1 that we study in Sect. 4.5.

To identify the effects of the tax schedule on efficiency and redistribution, we rewrite the expected virtual surplus as the sum $S^{v,\mathrm{Tot}} = S^{v,\mathrm{Red}} + S^{v,\mathrm{Eff}}$, where

$$S^{v,\mathrm{Eff}} = \lambda \int_\Theta y(Z(\theta);\theta)\,\mathrm{d}F(\theta) \quad \text{and} \quad S^{v,\mathrm{Red}} = -\int_\Theta \mu(\theta)\gamma_\theta(Z(\theta);\theta)\,\mathrm{d}\theta \tag{20}$$

represent its efficiency part and its redistributive part respectively. The efficiency part is nothing else than the expected total net lifetime output. The redistributive part is a measure of the inequalities in lifetime consumption across agents (recall that, by (11), $\gamma_\theta(Z(\theta),\theta)$ equals $\mathrm{d}c/\mathrm{d}\theta$), weighted by the Lagrange multipliers $\mu(\theta)$ associated with the incentive constraints.

We call *redistribution force* and *efficiency force* the effect of the income tax schedule on respectively the redistributive part and the efficiency part of the virtual surplus. Following the optimal taxation literature (e.g., Diamond 1998; Saez 2001, 2002; Jacquet et al. 2013), we consider small tax reforms, i.e., small variations of the tax schedule, and compute their effect on each component of the virtual surplus.

Redistribution From the type ordering assumption (Assumption 4.1) and the envelope property (11), we know that lifetime consumption increases with the agent type: $\mathrm{d}c/\mathrm{d}\theta = \gamma_\theta(Z(\theta);\theta) > 0$. For the redistributive motive, the government would like to reduce locally the inequalities in consumption, that is the size of $\mathrm{d}c/\mathrm{d}\theta$, all the more that the incentive multiplier $\mu(\theta)$, computed in (17), is largest.

Before examining the redistributive force under a pension scheme, we start by recalling how the optimization works in the absence of such a scheme. Then the Lagrangian takes the simple form

$$\int_\Theta [u(c(\theta);\theta) - \lambda c(\theta)]\,\mathrm{d}F(\theta) + \int_\Theta y(\theta)\,\mathrm{d}F(\theta). \tag{21}$$

The optimal taxation studies mentioned above typically consider small tax cuts for part of the population. Their *mechanical* effect is obtained when behavioral responses are neutralized. Suppose that the government decreases taxes or equivalently increases the after-tax income $R(w)$ below some productivity level w. This would raise lifetime consumption $\int_0^1 [R(w(a;\theta)) - \delta(a;\theta)]\ell(a;\theta)\,\mathrm{d}a$ for agents who work below that productivity. More precisely, for agent θ, the increase in lifetime earnings depends on the time she spends working below that productivity

$$T(w;\theta) = \int_0^1 \mathbb{1}_{w(a;\theta)\le w}\,\ell(a;\theta)\,\mathrm{d}a.$$

Similarly, a tax cut in a small productivity interval $[w, w + \mathrm{d}w]$ increases lifetime consumption for agents who work in that productivity interval. For agent θ, the increase in lifetime earnings depends on the time she spends working in that interval, which is simply the derivative of $T(w;\theta)$ with respect to w, denoted hereafter $T_w(w;\theta)$. The *mechanical* effect (net of budgetary cost) of a tax cut around productivity w is the

change of the first integral in (21) and is thus given by

$$\text{Net mechanical effect} = \int_{\Theta} [u_c(c(\theta);\theta) - \lambda]\, T_w(w;\theta)\, \mathrm{d}F(\theta). \tag{22}$$

As the expectation of the marginal utilities of income coincides with the cost of public funds λ, the above term is equal to the covariance of the marginal utilities with the time spent working at the considered productivity level.

The analysis seems at first glance very different when a pension scheme is in force. In particular, the income tax schedule now affects the Lagrangian (14) only through the expected virtual surplus (15), which does not contain the integral term in (21) that generates the mechanical effect. Yet, as stated in the next Proposition, that very same mechanical effect is still present under an optimal pension scheme. The novelty is that the effect of taxes on lifetime consumption inequalities now involves a second term, which comes from their *feedback effect on pension incentives*.

Proposition 3 *The redistribution force describes how taxes affect consumption inequalities. It has two components:*

- *the standard mechanical effect net of budgetary cost, which is the covariance of working times and net social marginal utility of incomes, see* (22)*;*
- *a contribution from the feedback effect of taxes on pension incentives. In redistributive economies, the feedback effect of tax cuts at any productivity level* reduces *consumption inequalities, see* (24).

In the Appendix, we show that the redistributive component of the expected virtual surplus can be rewritten as

$$S^{v,\text{Red}} = \int_{\underline{\theta}}^{\bar{\theta}} [u'(c(\theta);\theta) - \lambda]\, \gamma(Z(\theta);\theta)\ \mathrm{d}F(\theta) - \int_{\Theta} \mu(\theta)\pi(\theta)Z'(\theta)\, \mathrm{d}\theta. \tag{23}$$

Applying the analysis to the case without pensions presented above, we find that the effect of a tax cut on the first term is nothing else than the standard mechanical effect.[6] Under an optimal pension scheme, however, the equity force does not boil down to that mechanical effect. Tax reforms change the pressure placed by the pension system on the agent labor supply, which indirectly affects lifetime consumption inequalities. Formally the contribution of the feedback effect to the redistribution force is

$$\text{Feedback redistributive effect} = -\int_{\Theta} \mu(\theta)\frac{\partial \pi(\theta)}{\partial R} Z'(\theta)\, \mathrm{d}\theta \geq 0. \tag{24}$$

As stated in Proposition 1, tax cuts reduce the pressure exerted by the pension system on the agents' labor supplies: i.e. $\partial\pi/\partial R \leq 0$. Moreover, $\mu(\theta)$ is positive in a redistributive economy; w_θ is positive under Assumption 4.1; and Z' is positive in the

[6] Replace $\gamma(Z(\theta);\theta)$ by its value $\int_0^1 [R(w(a;\theta)) - \delta(a;\theta)]\ell(a;\theta)\, \mathrm{d}a$.

three considered pension regimes. In sum, tax cuts, by reducing pension multipliers, contribute *positively* to redistribution.

Efficiency We now turn to the efficiency part of the expected virtual surplus, namely the expected net output y multiplied by the cost of public funds. A tax cut (an increase in $R(w)$) around a productivity level where no agent switches work status does not change the total net output. A tax cut around a switch point of agent θ has two effects. First, given the level of the pension multiplier $\pi(\theta)$, there is the following *static* effect

$$\mathrm{d}y(Z(\theta);\theta)_{\pi=\mathrm{cst}} = [w-\delta]\eta(w;\theta)\,\mathrm{d}R, \tag{25}$$

where the static labor supply elasticity η is given by (8). Recall that the static elasticity accounts for the fact that the agent has to meet the pension requirement $Z(\theta)$ to receive the pension benefit $P(\theta)$; it is based on the *adjusted* incentives, $R(w)+\pi(\theta)z(w)$. The above effect, however, does not account for the change in the pension multiplier: it ignores the endogeneity of the pension multiplier $\pi(\theta)$.

As mentioned in Sect. 3, a tax cut lowers the pension multiplier $\pi(\theta)$. We observe furthermore that $\partial y/\partial\pi$ has the same sign as $\partial y/\partial Z$. Indeed, a rise in Z has the same effect (up to a positive factor) on output as a rise in π.[7] We know from (18) that $\partial y/\partial Z$ is positive at the second-best optimum of a redistributive economy. A decrease in π or Z therefore contributes to reduce output y.

Proposition 4 *The efficiency force describes how taxes affect the expected net output. This force has two components:*

- *the direct effect on net output, the sign of which depends on the direction of the labor supply distortion, see* (25);
- *an indirect contribution from the feedback effect on pension incentives. In redistributive economies, the feedback effect of tax cuts* reduces *the expected net output.*

To summarize, Propositions 3 and 4 suggest that the presence of an optimal pension scheme attenuates both the redistributive role and the efficiency role of taxes. In the absence of a pension scheme, taxes are used to redistribute and may be detrimental to productive efficiency. The new feedback effect of taxes on pension incentives counteracts these usual forces: through the feedback effect, decreasing taxes *induces* redistribution and *reduces* productive efficiency.

4.5 A specialized framework

In this section, we work in the specialized framework of Assumption 2.1. While restrictive, this assumption is consistent with many different patterns: the trajectories may very well cross, possibly many times, meaning that the same characteristics (productivity, cost) are reached by different agents at different ages. The assumption yields a sharper description of second best optimal retirement and of the interactions between the two instruments, income tax and pensions.

[7] This is because the partial derivative $\partial\pi/\partial Z$ is positive. Its expression is given by equation (A.2) in the Appendix.

When the trajectories are decreasing, the quantity $R(w(a)) + \pi z(w(a)) - \delta(a)$ decreases with age, implying that the agents work up to a retirement age $\bar{a}$ where $R(w(\bar{a}))+\pi z(w(\bar{a}))-\delta(\bar{a}) = 0$ and do not work afterwards. The problem is therefore much simplified, and we restate it below. Given a menu of pension plans $(P(\theta), Z(\theta))$ and the tax schedule $R(w)$, agent θ consumption is

$$c(\theta) = \max_{\theta'}[P(\theta') + \gamma(Z(\theta'); \theta)]$$

where

$$\gamma(Z;\theta) = \max_{\bar{a}} \int_0^{\bar{a}} [R(w(a;\theta)) - \delta(a;\theta)]\,\mathrm{d}a \quad \text{subject to} \quad \int_0^{\bar{a}} z(w(a;\theta))\,\mathrm{d}a \geq Z.$$

The Lagrangian of the agent problem given Z and $R(.)$ is

$$\int_0^{\bar{a}} [R(w(a;\theta)) - \delta(a;\theta)]\,\mathrm{d}a + \pi(\theta)\left[\int_0^{\bar{a}} z(w(a;\theta))\,\mathrm{d}a - Z\right]$$

and the above term is equal to $\gamma(Z;\theta)$ whether or not the pension requirement condition is binding. The retirement age $\bar{a}(\theta)$ satisfies the equation

$$R(w(\bar{a}(\theta);\theta)) - \delta(\bar{a}(\theta);\theta) + \pi(\theta)z((w(\bar{a};\theta)) = 0.$$

If the pension requirement is binding ($\pi > 0$), under Assumption 2.1, the pension system forces the agent to work longer than she would do facing the same tax schedule in the absence of pensions:

$$\pi(\theta) = -\frac{R(w(\bar{a}(\theta);\theta)) - \delta(\bar{a}(\theta);\theta)}{z(w(\bar{a}(\theta);\theta))} > 0. \tag{26}$$

The envelope theorem yields

$$\gamma_\theta(Z(\theta);\theta) = \int_0^{\bar{a}(\theta)} \left\{\left[R'(w(a;\theta)) + \pi(\theta)z'(w(a;\theta))\right] w_\theta(a;\theta) - \delta_\theta(a;\theta)\right\}\,\mathrm{d}a. \tag{27}$$

The virtual surplus defined by (13) takes the form

$$\begin{aligned} f(\theta)\,S^v(Z;\theta) = \int_0^{\bar{a}(\theta)} &\{\lambda\,[w(a;\theta) - \delta(a;\theta)]\,f(\theta) - \mu(\theta) \\ &\times \left(\left[R'(w(a;\theta)) + \pi(\theta)z'(w(a;\theta))\right] w_\theta(a;\theta) - \delta_\theta(a;\theta)\right)\}\,\mathrm{d}a. \end{aligned} \tag{28}$$

As $\mu(\theta)$ and w_θ are positive, it is desirable to set the nonnegative square-bracketed term in (28) closest possible to zero. Indeed it can be made equal to zero in regimes L

and N. Then labor supplies should be set so as to maximize

$$\int_0^{\bar{a}(\theta)} \{\lambda\ [w(a;\theta) - \delta(a;\theta)]\ f(\theta) + \mu(\theta)\,\delta_\theta(a;\theta)\}\ \mathrm{d}a, \tag{29}$$

which, assuming concavity, yields the optimal retirement age

$$\lambda\ [w(\bar{a}(\theta);\theta) - \delta(\bar{a}(\theta);\theta)]\ f(\theta) = -\mu(\theta)\,\delta_\theta(\bar{a}(\theta);\theta) > 0. \tag{30}$$

The above first-order condition applies (18) to our specialized framework. If it has a unique solution $\bar{a}(\theta)$ that increases with θ (no bunching), this configuration is implementable under regimes L and N by making the after-tax schedule entirely flat, i.e., by choosing $R' = 0$. Indeed, under such a schedule, the square-bracketed term in (28) is zero in these two regimes.

To implement this allocation, the government has to choose the constant value of R low enough so that the pension system effectively constrains the labor supply of all agents, i.e., so that $\pi(\theta)$ is positive as indicated in (26). The pension requirement $Z(\theta)$ is $L(\theta) = \bar{a}(\theta)$ in regime L and $N(\theta) = \bar{a}(\theta)R$ in regime N. Labor supplies are necessarily distorted downwards as the right-hand side of (30) is positive.

Proposition 5 *Consider a redistributive economy that satisfies the assumptions of the specialized framework and operates under regime L or N:*

1. *At the optimum, the two instruments, pensions and income tax, specialize. Pensions manage labor supply and efficiency, while taxes redistribute.*
2. *The optimal marginal tax rate is 100%.*
3. *Labor supplies are distorted downwards, i.e., the agents retire earlier than in the first best.*

Under the system W, the bracketed term in (28) cannot be set to zero. In this regime, we have $z' = 1$, and the pension multiplier $\pi(\theta)$ in (27) contributes to make the allocation more unequal, i.e., to increase the inequality between lifetime consumptions $c(\theta)$ with the agent type θ.

4.6 A calibrated example

In this final section, following the suggestion of a referee, we work out the shapes of the optimal tax and pension systems in a simple toy economy. There are two types of agents in equal numbers, whose activities we describe from the age of 16 until 85. The time path of their productivity reproduces the French monthly labor cost in francs per month as seen in the 1999 cross section, for the skilled (H for high) and unskilled (L for low) productivities, see the wage estimates in Laroque and Salanié (2002). At age a, the productivities of the typical unskilled and skilled men are

$$w_H(a) = \exp[0.06a - 0.0006a^2 + 7] + 500, \quad w_L(a) = w_H(a) - 1500.$$

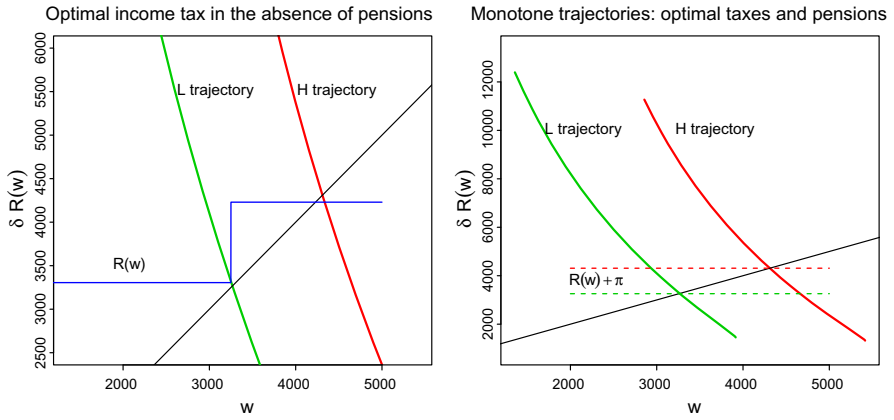

Fig. 2 The distortions induced by taxes

Productivity increases from age 16 up to age 50, decreases afterwards until age 85. While there is data on the cost of labor to the employer, much less is known on the cost supported by the worker. We adopted similar inverted time profiles

$$\delta_H(a) = \exp[-0.075a + 0.0008a^2 + 9.5] - 1000 \\ + 4(x - 50)^2(x \geq 50), \delta_L(a) = 1.1\delta_H(a).$$

This pattern with two types satisfies the ordering property stated in Assumption 4.1. The government is utilitarian with the utility function being $u(c) = \ln(c)$.

We first study the simpler case where both the wage and disutility of work vary monotically with age: wage decreases past 50, disutility of work increases. This amounts to focus the attention on the second half of life, from 50 onwards. The trajectories are represented on Fig. 2 in the (productivity, cost of work) space, with that of the skilled H type dominating that of the unskilled L type. The point reached at 50 is at the bottom right, and the trajectory is described from bottom to top and from right to left, from 50 to 85.

We start with the situation where there are no pensions: the government only redistributive tool is a time independent income tax. The optimal tax $R(w)$ is represented on the left panel of Fig. 2. The agents work when their work disutility is smaller than their after tax wages, i.e. their trajectory is below the after tax income schedule. At the laissez-faire, switching points between leisure and work are at the intersection of the trajectories with the 45-degree line. The computation, retraced on Fig. 2, shows that the skilled stop working (approximately 3 months) earlier than what they would do in the first best, their labor supply is distorted downwards, while the unskilled work more, their labor supply is (barely, 2 and half months) distorted upwards, as mentioned in Choné and Laroque (2017). The lifetime consumptions of the two agents (in monthly terms) are 1062 and 1234, instead of a common 1149 at the first best. The corresponding welfare values, sum of the logarithms of these consumption levels, are

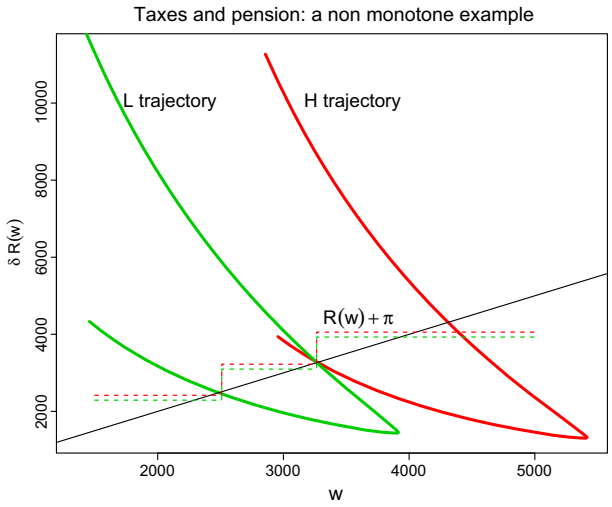

Fig. 3 Non monotonic trajectories: the optimal taxes and pensions

respectively 14.0863 and 14.0924. Thus, in this two types example, the tax schedule alone is a very powerful instrument that brings the economy close to the first best.

Still in the monotone trajectory example, the right panel of Fig. 2 represents the optimum when the government is able to complement the tax with a pension scheme. Compared with the tax only case on the left panel, the after tax schedule $R(w)$ is modified into $R(w) + \pi$, where π is the Lagrange multiplier associated with the pension constraint, a number larger for type H than for type L. The dashed lines on the figure represent $R(w)+\pi_H$ and $R(w)+\pi_L$ for type H and type L. Upwards labor supply distortions disappear, and agent H retires at the efficient time, consistently with Proposition 4.5. The lifetime consumptions of the two agents (in monthly terms) are 1107 and 1190 per month, hence a substantial reduction in consumption inequality. Welfare is equal to 14.0911. Adding the pension system reduces the welfare gap with the first best from $14.0924 - 14.0863 = 61 \times 10^{-5}$ to 13×10^{-5}.

Regarding taxes, without loss of generality, we have used piecewise constant functions to represent tax schedules in the above numerical analysis. In the absence of a pension scheme and with monotonic trajectories, the optimal after tax schedule is an horizontal line with a step in the middle (left panel of Fig. 2), consistent with Choné and Laroque (2017). With a pension system the tax schedule is flat.

The latter property (100% marginal tax rate everywhere when there are pensions and trajectories are monotone) does not extend to economies with more complex trajectories. Figure 3 illustrates the outcome when trajectories are non-monotonic and the government can use both taxes and a pension scheme. Here the trajectories start at the top of the lower branches at age 16, go down at the bottom at 50, and come back to the top on the upper branches at age 85. Using a pension based on the length of the active life, we find that the skilled types start work earlier (5 months) at the second best than at the first best at the beginning of life, but also retire earlier (5 months), so that on average their life time labor supply seems undistorted, consistently with the

theoretical analysis of Sect. 4.3 (see in particular the comments after Proposition 2). The unskilled types work less than at the first best, 3 months at entry, with no retirement distortion, and overall they work less than at the first best.

5 Conclusion

In this article, we have uncovered a fundamental channel through which pension schemes may help redistribute welfare across agents with different intertemporal profiles of productivity and opportunity cost of work. The main intuition is that by letting the agents choose their preferred option among a menu of pension plans, the government can indirectly control their labor supply decisions. In practice, agents take into account the deferred pension benefit associated with their current earnings when deciding to work. The implicit conversion rate between after-tax earnings and pension benefits reflects the pressure placed by the pension system on an agent's labor supply. This pressure itself depends on the shape of the income tax schedule, a phenomenon we refer to as the "feedback effect" of taxes on pension incentives.

The subtle interplay between taxes and pensions is better understood by studying the virtual surplus associated with the government problem. In particular, an adequate decomposition of that surplus allows to separate the impact of tax reforms on redistribution and on efficiency. The main insight from our analysis is that the presence of the feedback effect counteracts the usual forces in the equity-redistribution tradeoff. *Through the interplay with the pension system*, decreasing taxes has positive redistributive effects and increasing taxes favors efficiency. The overall impact of tax reforms, however, depends on fine details of the economy. In a particular, quite extreme example, we find that labor supply can be fully controlled by the pension system and a strong degree of taxation is therefore optimal.

An important policy question is the choice of the lifetime statistic(s) that should underlie the pension scheme. Each statistic is associated with a particular shape of the virtual surplus, but also with a distinct monotonicity condition (recall Lemma 2). As explained at the end of Sect. 4.5, basing the pension scheme on total working time allows the government to reduce the inequalities in lifetime consumption that come from productivity differences across agents.[8] However, we have been unable to obtain a formal optimality result, and leave the design of optimal pension schemes for future research.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

[8] With $z(w) = 1$, the nonnegative term $\pi(\theta)z'(w(a;\theta))$ in (27) is set to zero.

Appendix A: Proofs

Proof of Lemma 1 We say that there is *indifference* for agent θ at w if she has productivity w at the age a_θ, $w = w(a_\theta)$, and is indifferent between working and not working at this age, i.e., $R(w)+\pi(\theta)z(w)-\delta(a_\theta)=0$. A *switch point* is an indifference point such that the work status of the indifferent agent changes in a neighborhood of w, i.e., the trajectory of agent θ crosses the incentive schedule at w. When the slopes of the adjusted tax schedule and of the trajectory are different, $R'(w)+\pi(\theta)z_w \neq \delta_a/w_a$, the inverse of $|\delta_a(a;\theta)-R'(w)w_a(a;\theta)-\pi(\theta)z_w w_a(a;\theta)|$ is positive and finite.

Consider a switch point w and replace R with $R+\mathrm{d}R$ on the interval $[w, w+\mathrm{d}w]$, with $\mathrm{d}R=(\delta_a/w_a - R' - \pi z_w)\,\mathrm{d}w$. The perturbation changes the status of the agent on the interval from working to non working. The time spent in the interval is

$$\mathrm{d}a = \mathrm{d}w/|w_a| = \eta\,\mathrm{d}R,$$

hence η is the absolute value of the derivative of labor supply with respect to the adjusted tax schedule R.

If agent θ's trajectory crosses the adjusted tax schedule more than once at w, Eq. (8) must be modified by adding up the contributions of all the ages at which the agent switches. □

Proof of Proposition 1 We now prove that the Frechet-derivative of pension multipliers with respect to R is a negative measure:

$$\frac{\partial\pi(Z;\theta)}{\partial R} \leq 0. \tag{A.1}$$

This property yields the result stated in the proposition because tax cuts are equivalent to increases in the after-tax income R. The multiplier $\pi(Z;R)$ is defined jointly by (2) and (7), which can be rewritten as $K(\pi)=Z$ with

$$K(\pi)=\int_0^1 z(w(a))\,\mathbb{1}_{R(w(a))+\pi z(w(a))-\delta(a)\geq 0}\,\mathrm{d}a.$$

We obtain the derivative of K with respect to π by using the function $\Theta(a,\pi)=R(w(a))+\pi z(w(a))-\delta(a)$:

$$\frac{\partial K}{\partial\pi}=\sum_{\sigma\in\mathcal{S}_\theta} z^2(w_\sigma)\eta(w_\sigma;R,\pi).$$

The inverse function theorem yields the following expression for the impact on π of a marginal increase $\mathrm{d}Z$ of the pension requirement

$$\mathrm{d}\pi=\frac{1}{\sum_{\sigma\in\mathcal{S}_\theta} z^2(w_\sigma)\eta(w_\sigma)}\,\mathrm{d}Z. \tag{A.2}$$

Similarly the Frechet-derivative of K with respect to R is given by

$$\frac{\partial K}{\partial R} = \sum_{\sigma \in \mathcal{S}_\theta} z(w_\sigma)\eta(w_\sigma; R, \pi)\zeta(w_\sigma)$$

in regime L and W and

$$\frac{\partial K}{\partial R} = T_w(w; \ell, \theta) + \sum_{\sigma \in \mathcal{S}_\theta} R(w_\sigma)\eta(w_\sigma; R, \pi)\zeta(w_\sigma)$$

in regime N, where $\zeta(w_\sigma)$ denotes the mass point at w_σ. Applying the implicit function theorem yields the Frechet-derivative of π with respect to R:

$$\frac{\partial \pi}{\partial R} = -\frac{\sum_{\sigma \in \mathcal{S}_\theta} z(w_\sigma)\eta(w_\sigma; R, \pi)\zeta(w_\sigma)}{\sum_{\sigma \in \mathcal{S}_\theta} z^2(w_\sigma)\eta(w_\sigma; R, \pi)} < 0 \tag{A.3}$$

in regime L and W, and by

$$\frac{\partial \pi}{\partial R} = -\frac{T_w(w; \ell, \theta) + \sum_{\sigma \in \mathcal{S}_\theta} R(w_\sigma)\eta(w_\sigma; R, \pi)\zeta(w_\sigma)}{\sum_{\sigma \in \mathcal{S}_\theta} R^2(w_\sigma)\eta(w_\sigma; R, \pi)} < 0 \tag{A.4}$$

in regime N. In this last regime, increasing R outside a switch point increases the after-tax income collected during the lifetime, modifies (2), hence the new term T at the numerator in the expression of π.

If tax rise $\mathrm{d}R < 0$ around switch point w, then π increases, and labor supply around switch point w' *increases* by

$$-\frac{\eta(w)\eta(w')z(w)z(w')}{\sum_{\sigma \in \mathcal{S}_\theta} z^2(w_\sigma)\eta(w_\sigma)}\mathrm{d}R > 0. \tag{A.5}$$

Around the switch point w itself, the net effect of the static effect and the feedback effect is negative on labor supply as

$$\eta(w) - \frac{z^2(w)\eta^2(w)}{\sum_{\sigma \in \mathcal{S}_\theta} z^2(w_\sigma)\eta(w_\sigma)} > 0.$$

where $\mathcal{S}_\theta$ is the set of agent θ's switch points. □

Proof of Lemma 2 We already know that $\pi(Z, \theta) = -\partial\gamma(Z; \theta)/\partial Z$, together with $\ell(a)$, are solutions of the system made of (2) and (7). Moreover an increase in π raises $\ell(a)$ by (7), and therefore the left-hand side of (2). Now under Assumption 4.1, note that $\ell(a; \theta)$, since by construction $R(w(a; \theta)) + \pi z(w(a; \theta))$ is nondecreasing in $w(a; \theta)$, is nondecreasing in $w(a; \theta)$. Hence the left-hand side of (2) is also nondecreasing in $w(a; \theta)$. Furthermore, $\ell(a; \theta)$ and therefore the left-hand side of (2) decreases with $\delta(a; \theta)$. Consequently, labor supply $\ell(a; \theta)$ is nondecreasing in θ. It follows that the

left-hand side of (2) increases with the type θ. Therefore $\pi(Z,\theta)$ decreases with θ under Assumption 4.1, which yields the single-crossing property.

The second part of the Lemma follows from Proposition 1 of Rochet (1987). □

Proof of Proposition 2 To compute the derivative of an agent's net output y with respect to the pension requirement Z, we use the chain rule

$$\frac{\partial y}{\partial Z} = \frac{\partial y}{\partial \pi}\frac{\partial \pi}{\partial Z},$$

where the second term, $\partial\pi/\partial Z$, is given by (A.2). Computing the first term, $\partial y/\partial\pi$ with the same method as above yields (19).

When w is not a switch point of agent θ and the pension constraint is binding, a marginal increase $\mathrm{d}R$ on $[w, w+\mathrm{d}w]$ has no effect on her labor supply in regimes L and W and has a second-order, indirect effect on labor supply around all switch points w'

$$-\frac{\eta(w')R(w')}{\sum_{\sigma\in\mathcal{S}_\theta} R^2(w_\sigma)\eta(w_\sigma)} T_w(w;\ell_\theta)\,\mathrm{d}R \tag{A.6}$$

in regime N.

Finally, when the pension constraint is binding, a marginal increase $\mathrm{d}Z$ of the pension requirement increases the time agent θ spends working in the neighborhood of all switch points w' by

$$\frac{\eta(w')z(w')}{\sum_{\sigma\in\mathcal{S}_\theta} z^2(w_\sigma)\eta(w_\sigma)}\,\mathrm{d}Z. \tag{A.7}$$

□

Proof of Proposition 3 We now establish Eq. (23). We first observe that we may assume, with no loss of generality, that the inequality (2) holds as an equality for all agents. If this is not the case, we just replace $Z(\theta)$ with $\bar{Z}(\theta) = \int_0^1 z(w(a;\theta))l(a;\theta)\,\mathrm{d}a$. As the new pension requirement is stronger, no other agent is attracted by the plan $(\bar{Z}(\theta), P(\theta))$, and therefore the incentive constraints are still satisfied. Hereafter, we therefore assume $\int_0^1 z(w(a;\theta))\ell(a;\theta)\,\mathrm{d}a = Z(\theta)$ for all θ.

Differentiating (6) with respect to θ and noticing that, from the above remark, the contribution of the derivative of the pension multiplier is zero, we get

$$\frac{\partial}{\partial\theta}\gamma(Z;\theta) = \int_0^1 \left[R'(w(a;\theta))w_\theta - \delta_\theta + \pi(Z;\theta)z'(w(a;\theta))w_\theta\right]\ell(a;Z,\theta)\,\mathrm{d}a.$$

Evaluating the above at $Z = Z(\theta)$ yields

$$\frac{\partial}{\partial\theta}\gamma(Z(\theta);\theta) = \int_0^1 \left[R'(w(a;\theta))w_\theta - \delta_\theta + \pi(\theta)z'(w(a;\theta))w_\theta\right]\ell^z(a;\theta)\,\mathrm{d}a,$$

where, as already mentioned, we use the notations $\pi(\theta)$ for $\pi(Z(\theta);\theta)$ and $\ell(a;\theta)$ for $\ell(a;Z,\theta)$. We therefore get

$$\frac{\partial}{\partial\theta}\gamma(Z(\theta);\theta)=\frac{\mathrm{d}}{\mathrm{d}\theta}\int_0^1[R(w)-\delta(a;\theta)+\pi(\theta)z(w)]\,\ell\,\mathrm{d}a-Z(\theta)\frac{\mathrm{d}\pi(\theta)}{\mathrm{d}\theta}.$$

We obtain the redistributive part of the expected virtual surplus, $S^{v,\mathrm{Red}}$, by multiplying the above term by $-\mu(\theta)$ and integrating over θ. We therefore write $S^{v,\mathrm{Red}}=S_1+S_2$ with

$$S_1=-\int_{\underline{\theta}}^{\bar{\theta}}\mu(\theta)\frac{\mathrm{d}}{\mathrm{d}\theta}\int_0^1[R(w)-\delta(a;\theta)+\pi(\theta)z(w)]\,\ell\,\mathrm{d}a\,\mathrm{d}\theta$$

and

$$S_2=\int_{\underline{\theta}}^{\bar{\theta}}\mu(\theta)\frac{\mathrm{d}\pi(\theta)}{\mathrm{d}\theta}Z(\theta)\,\mathrm{d}\theta.$$

Using (17), $\mu(\underline{\theta})=\mu(\bar{\theta})=0$, and integrating by parts, we rewrite S_1 and S_2 as

$$S_1=\int_\Theta(u'(c(\theta);\theta)-\lambda)\int_0^1[R(w)-\delta+\pi(\theta)z(w)]\ell\,\mathrm{d}a\,\mathrm{d}F(\theta)$$

and

$$S_2=-\int_{\underline{\theta}}^{\bar{\theta}}\pi(\theta)[(u'(c(\theta);\theta)-\lambda)Z(\theta)+\mu(\theta)Z'(\theta)]\,\mathrm{d}\theta.$$

Using again the equality $\int_0^1 z(w(a;\theta))\ell(a;\theta)\,\mathrm{d}a=Z(\theta)$ yields (23). □

Appendix B: Intensive labor supply

We have assumed in the main text that labor supply decisions involve only the extensive margin, representing labor supply by a dichotomous variable $\ell\in\{0,1\}$. In this section, we show that our method of analysis applies when labor supply decisions involve both the intensive margin and the extensive one. To this end, we allow ℓ to be any nonnegative number and the cost of work $\delta(\ell;a,\theta)$ to be a *function* of ℓ, which also depends on the agent's age and type.

The type of an agent is thus characterized by a couple of exogenous, nonnegative functions $(w(a;\theta),\delta(l;a,\theta))$ and by the instantaneous utility index $u(c;\theta)$. We assume that the functions w, δ, and u are differentiable. We allow, however, the cost of work δ to be discontinuous at $\ell=0$ to reflect an indivisible cost of work (extensive margin).

An *allocation* specifies the nonnegative consumption $c(a;\theta)$ and the labor supply $\ell(a;\theta) \geq 0$ of all types θ along their lives. We assume that there are perfect markets for transferring wealth across time, with a zero interest rate. The agents use these markets to smooth their consumption overtime, so we can remove the age argument in consumption and write simply $c(a;\theta) = c(\theta)$. We denote by $y(\theta)$ the lifetime net output produced by agent θ, i.e., $y(\theta) = \int_0^1 [w(a;\theta)\ell(a;\theta) - \delta(\ell(a;\theta);a,\theta)]\,\mathrm{d}a$.

In this context, the tax schedule R is a function of the product $w\ell$, and we write the pension requirement as

$$\int_0^1 z(w(a;\theta),\ell(a;\theta))\,\mathrm{d}a \geq Z. \tag{B.1}$$

The pension systems L, W and N are unchanged with respectively $z(w,\ell) = \ell, z(w,\ell) = w\ell$, and $z(w,\ell) = R(w\ell)$. Facing the tax schedule $R(\cdot)$ and a pension regime associated with transfers $P(\cdot)$, the consumer chooses her labor supply $\ell(a)$, and pension level Z so as to maximize her lifetime utility

$$c(\theta) = \max_{\ell,Z} \int_0^1 [R(w(a;\theta)\ell(a)) - \delta(\ell(a);a,\theta)]\,\mathrm{d}a + P(Z)$$

under the constraint (B.1). We rewrite consumption as the sum of lifetime earnings and pension benefit: $c = P + \gamma$. Denoting by $\pi(Z;\theta)$ the Lagrange multiplier associated with the pension constraint (2) for agent θ, we rewrite lifetime earnings as

$$\begin{aligned}\gamma(Z;\theta) = \max_{\ell(a;\theta)} \Bigg\{ &\int_0^1 [R(w(a;\theta)\ell(a;\theta)) - \delta(\ell(a;\theta);a,\theta)]\,\mathrm{d}a \\ &+ \pi(Z;\theta)\left(\int_0^1 z(w(a;\theta),\ell(a;\theta))\,\mathrm{d}a - Z\right)\Bigg\}.\end{aligned} \tag{B.2}$$

For agent θ, labor supply incentives are driven by the adjusted tax schedule $R(w\ell) + \pi(\theta)z(w,\ell)$. The labor supply of agent θ at age a maximizes

$$R(w\ell) + \pi(\theta)z(w,\ell) - \delta(\ell,a;\theta).$$

The pressure placed by the pension system on labor supply now takes two forms: as already seen, the agents perceive a fall in the absolute level of taxes (R is replaced with $R + \pi z > R$); in addition, the agents now perceive a lower *marginal* tax rate. Labor supply, if positive, satisfies:

$$wR'(w\ell) + \pi(\theta)z_\ell(w,\ell) - \delta_\ell(\ell,a;\theta) = 0.$$

In the absence of a pension system ($\pi = 0$), the marginal tax rate is $T' = 1 - R'$. Under system L, W and N, the perceived marginal tax rate decreases to $T' - \pi/w$, $T' - \pi$, and $T' - \pi(1 - T')$ respectively. Under mild regularity conditions, labor supply increases with π because a lower absolute level of taxes increases labor supply at the

extensive margin and a lower marginal tax rate does the same at the intensive margin. Proposition 1 is still valid in this context provided that "tax cuts" are understood in a more general sense, including both decreases in the absolute level of taxes and decreases in marginal tax rates.[9]

We reinforce Assumption 4.1 by assuming that both the absolute level and the *marginal* cost of work, i.e., $\delta(\ell; a, \theta)$ and its derivative $\delta_\ell(\ell; a, \theta)$, decrease with θ. Under these assumptions, the single-crossing condition stated in Lemma 2 remains valid. This is because labor supply increases with θ: higher types have higher productivity and lower fixed and marginal cost of work. It is therefore easier for them to meet any given pension requirement Z, and $\pi(Z, \theta)$ decreases with θ.

The taxonomy of economies described in Lemma 3 also remains valid as it does not involve labor supply behaviors but only the levels of lifetime consumptions $c(\theta)$. The virtual surplus is still defined by (13). We find the same expression for the multiplier $\mu(\theta)$ associated with incentive constraints, namely equation (17).

The first part of Proposition 2, which follows from virtual surplus maximization is still valid. The second part, however, does not necessarily hold because due to the intensive margin in labor supply the total lifetime output $y(Z(\theta); \theta)$ is not enough to completely characterize labor supply at each age.

Finally the efficiency and redistributive components of the virtual surplus are still defined by (20). Equation (23) still holds (proof unchanged). The results stated in Proposition 3 and Proposition 4 carry over to this context: providing more incentives through taxes reduces the incentives provided by pensions, thereby counteracting the usual forces in the equity-redistribution tradeoff.

References

Brett C (2012) The effects of population aging on optimal redistributive taxes in an overlapping genrerations model. Int Tax Public Financ 19:777–799

Choné P, Laroque G (2017) Optimal income tax in an extensive labor supply life-cycle model. CESifo Econ Stud 63(1):78–96

Cremer H, Lozachmeur J-M, Pestieau P (2004) Social security, retirement age and optimal income taxation. J Public Econ 88:2259–2281

Cremer H, Lozachmeur J-M, Pestieau P (2008) Social security and retirement decision: a positive and normative approach. J Econ Surv 22:213–233

Cremer H, Pestieau P (2016) Taxing pensions. Discussion Paper 9821, IZA

Diamond P (2009) Taxes and pensions. South Econ J 76(1):2–15

Diamond P, Mirrlees J (1978) A model of social insurance with variable retirement. J Public Econ 10:295–336

Diamond PA (1998) Optimal income taxation: an example with a U-shaped pattern of optimal marginal tax rates. Am Econ Rev 88(1):83–95

Golosov M, Troshkin M, Tsyvinski A (2016) Redistribution and social insurance. Am Econ Rev 106(2):359–386

Grochulski B, Kocherlakota N (2010) Nonseparable preferences and optimal social security systems. J Econ Theory 145:2055–2077

Gruber J, Wise DA (eds) (1999) Social security and retirement around the world. University of Chicago Press, Chicago

[9] An example of tax cut consists in lowering T' to $T' - \eta < T'$ in $(w_0, w_0 + \varepsilon)$, the schedule R being unchanged below w_0 and increased by $\varepsilon\eta > 0$ above.

Jacquet L, Lehmann E, Van der Linden B (2013) Optimal redistributive taxation with both extensive and intensive responses. J Econ Theory 148:1770–1805
Laffont J-J, Martimort D (2002) The theory of incentives. Princeton University Press, Princeton
Laroque G, Salanié B (2002) Labour market institutions and employment in France. J Appl Econom 17(1):25–48
Ljunqvist L, Sargent T (2014) Career length: effects of curvature of earning profiles, earning shocks, and social security. Rev Econ Dyn 17:1–20
Michau J-B (2014) Optimal redistribution: a life-cycle perspective. J Public Econ 111:1–16
Myerson RB (1981) Optimal auction design. Math Oper Res 6(1):58–73
Prescott E (2004) Why do Americans work so much more than Europeans? Fed Reserv Bank Minneap Q Rev 28:2–13
Prescott E, Rogerson R, Wallenius J (2009) Lifetime aggregate labor supply with endogenous workweek length. R Econ Dyn 12:23–36
Rochet J-C (1987) A necessary and sufficient condition for rationalizability in a quasi-linear context. J Math Econ 16:191–200
Rogerson R, Wallenius J (2009) Micro and macro elasticities in a life cycle model with taxes. J Econ Theory 144:2277–2292
Saez E (2001) Using elasticities to derive optimal income tax rates. Rev Econ Stud 68(1):205–229
Saez E (2002) Optimal income transfer programs: intensive versus extensive labor supply responses. Q J Econ 117(3):1039–1073
Shourideh A, Troshkin M (2012) Providing efficient incentives to work: retirement ages and the pension system, Discussion paper, Yale University
Weinzierl M (2011) The surprising power of age-dependent taxes. Rev Econ Stud 78:1490–1518