On the redistributive power of pensions*

Philippe Choné[†] and Guy Laroque[‡]

November 24, 2015

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 non-linear age-independent 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 feedback effects on pension incentives. Breaking down the virtual surplus associated with the government problem, we find that feedback effects tend to reinforce the redistributive role of taxes while pensions take on more of the provision of incentives to work.

*This work has benefited from the financial support of the European Research Council under grant WSCWTBDS. 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.

[†] CREST-ENSAE.

[‡] Department of Economics, UCL and Sciences-Po.

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, Rogerson, and Wallenius (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 one has to choose among the many possible formulations of the problem. Diamond and Mirrlees (1978) focus on the health shocks associated with aging and the social insurance features of pension schemes. Gorry and Oberfield (2012) and Michau (2014) introduce both an intensive and an extensive margins, but keep the fixed cost of going to work constant over the lifetime. Cremer, Lozachmeur, and Pestieau (2004) consider a model in which individuals differ in their productivity and their capacity to work long, assuming that the productivity and the intensity of labor disutility can each take two values. Lozachmeur (2006) and Cremer, Lozachmeur, and Pestieau (2008) study a two period of life overlapping generations model which allows them to discuss political economy aspects of the problem. Also in an overlapping generations with two period lives and intensive labor supply à la Mirrlees, Brett (2012) describes the comparative statics effects of a change of the trend in population growth on the steady state.

Here we consider a deterministic overlapping generations model in continuous time, where all agents have the same length of life. 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.[1] The government wants to redistribute lifetime

[1] Our setup is similar in a number of respects with that used by Rogerson (2011), Shourideh and Troshkin (2012) or Weinzierl (2011): overlapping generations, deterministic trajectories.

welfare across dynasties. The policy instruments are an age-independent nonlinear tax schedule and a pension scheme. The tax schedule depends on the current productivity of the workers. The pension scheme depends on a single aggregate statistic of the work life, e.g., the total time spent working over the life cycle, the before-tax lifetime earnings, the after-tax lifetime earnings.[2]

In this environment, we argue that pension systems constitute powerful redistributive devices. Redistribution requires extracting rent from agents with high productivities and/or low opportunity costs of work. We show that by letting any agent choose her pension plan the government can dramatically reduce informational rents. Rent extraction, however, is to be weighed against efficiency concerns. The rent-efficiency tradeoff is better understood by considering the virtual surplus associated with the government problem.

By appropriately breaking down the expression of the virtual surplus, we are able to rigorously disentangle the efficiency force and the redistribution force in the design of the optimal tax system. In a first step, we examine how labor supply behavior depends on taxes and pensions. When a pension system is in force, an agent' labor supply may be constrained because she needs to achieve a certain lifetime performance to receive her pension benefit. The pressure placed by a pension regime on the agents' labor supply depends on the shape of income taxes. As a result, income taxes affect labor supply not only directly, but also indirectly through "feedback effects" on pension incentives.

Our main contribution is to clarify, using the decomposition of the virtual surplus, how pension feedback effects enter the redistribution and efficiency forces. In a redistributive economy, we find that the feedback effects exacerbate the redistributive role of taxes and attenuate their incentive role: income tax tends to specialize into redistribution while pensions are used to provide the appropriate incentives to work. This qualitative property holds generally in redistributive economies. In a specialized framework –same marginal utility of consumption for all agents, monotonicity of productivity and pecuniary cost of work with age– we show that an optimal combination of a tax schedule with a pension policy eliminates any rent coming from observable productivity differences across agents. 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 article is organized as follows. Section 2 presents the economic environment, the government instruments, and the second-best problem. Section 3

Here labor supply is extensive rather than intensive à la Mirrlees while saving and/or borrowing are unrestricted.

[2]It should be noted that the framework of this paper is in the Ramsey tradition rather than in the Mirrlees one. We put *a priori* restrictions on the shape of the government instruments, and do not derive them from assumptions on asymmetric information, contrary to Diamond and Mirrlees (1978), Grochulski and Kocherlakota (2010) or Golosov, Troshkin, and Tsyvinski (2011).

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 indeed have their labor supply constrained by the pension system, and that the pension feedback effects increase the redistribution role and decrease the efficiency role of income taxes.

2 Model

We consider an economy in continuous time. All agents have the same life length, normalized to one. 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. She has a lifetime utility function of the form

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

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

The type of 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 the age a varies determines a curve in the (w,δ)-space, that we call a *trajectory*. We assume that the functions w, δ, and u are differentiable.

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 we can remove the age argument in consumption and write simply $c(a;\theta) = c(\theta)$. Also we denote by $y(\theta)$ 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., any agent works whenever her opportunity cost of work is lower than or equal to her productivity, $\ell(a;\theta) = 1$ if and only if $\delta(a;\theta) < w(a;\theta)$.

Utilitarian optimum (First-Best) The utilitarian optimum is the allocation that maximizes $\int_\Theta u(c(\theta))\,\mathrm{d}F(\theta)$ subject to the feasibility constraint (1). It is the feasible efficient allocation such that marginal utilities are equal:

$$u_c(c;\theta) = \lambda,$$

for θ in Θ.

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.[3]

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 made of 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 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:

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

[3]Suppose that $\delta(a)$ is a disutility cost instead of a pecuniary one, i.e., agent θ, 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. Then agent θ works at age a under laissez-faire if and only if $u'(c)w(a) > \delta(a)$, where c is her constant, instantaneous consumption level. Hence, this specification entails an *income effect* in labor supply: participation decreases with c. Using Pareto-optimality conditions, it can be checked that laissez-faire is efficient. The pecuniary model adopted in this paper avoids these complications.

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 aggregate working time 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 kinds of existing legislations.[4] Note in particular that we do not allow the tax schedule to be age dependent, contrary to Weinzierl (2011), nor do we have the financial market imperfections that underlie some of the pension regimes in practice.[5] 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 $\ell(a)$, and

[4]The analysis up to section 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

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

where $n(a;\theta) = 1$ if $z(w(a;\theta))\ell^z(a;\theta)$ 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.

[5] In a previous version of the paper, we had studied a situation where the tax schedule differed before and after the retirement age (in fact it was fully confiscatory after the retirement age).

pension level Z so as to maximize her lifetime utility, i.e.,[6]

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

where $\ell(a)$ belongs to $\{0, 1\}$, and Z satisfies the pension requirement (2). Agent θ's lifetime net output is

$$y^z(\theta) = \int_{a=0}^1 [w(a;\theta) - \delta(a;\theta)]\,\ell^z(a;\theta)\,\mathrm{d}a.$$

The second best program consists in maximizing the sum of utilities

$$\int_\Theta u(c^z(\theta))\,\mathrm{d}F(\theta)$$

under the feasibility constraint

$$\int_\Theta \{y^z(\theta) - c^z(\theta)\}\;\mathrm{d}F(\theta) \geq 0, \tag{5}$$

with respect to the tax and pension schedules, when the agents choose their optimal consumption production and work levels.

3 Labor supply

We first explain how the agents' labor supply depend 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

$$\int_0^1 [R(w(a;\theta)) - \delta(a;\theta)]\ell(a;\theta)\,\mathrm{d}a \tag{6}$$

over $\ell(.)$, subject to the pension requirement (2). Lifetime consumption is $c(Z;\theta) = \gamma(Z;\theta) + P(Z)$. At the optimum, when agent θ picks her preferred retirement plan $Z(\theta)$, she enjoys lifetime consumption $c^z(\theta) = c(Z(\theta);\theta)$ given by (4).

Denoting by $\pi(Z;\theta)$ the Lagrange multiplier associated with the pension constraint (2) for agent θ, we rewrite lifetime earnings as

$$\begin{aligned} \gamma(Z;\theta) \quad &= \quad \max_{\ell^z(a;\theta)} \left\{ \int_0^1 [R(w(a;\theta)) - \delta(a;\theta)]\,\ell^z(a;\theta)\,\mathrm{d}a \right. \\ &\qquad\qquad \left. + \;\pi(Z;\theta)\left(\int_0^1 z(w(a;\theta))\ell^z(a;\theta)\,\mathrm{d}a - Z\right)\right\}. \end{aligned} \tag{7}$$

[6]We do not introduce explicitly the subsistence income s, a benefit often paid to the unemployed in extensive models. The consumption equation would then take the form

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

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

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

$$\ell(a;Z;\theta) = \mathbb{1}_{R(w(a;\theta)) + \pi(Z;\theta)z(w(a;\theta)) - \delta(a;\theta) \geq 0}. \tag{8}$$

When making her labor supply decision, agent θ takes into consideration the adjusted tax schedule or financial incentive to work $R(w) + \pi^z(\theta)z(w)$.[7] She works in regions where her trajectory is located below her incentive schedule $(w, R(w) + \pi^z(\theta)z(w))$, 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 $\pi^z(\theta)z(w)$ represents the (deferred) pension benefit associated with before-tax earning w. The multiplier $\pi^z(\theta)$ can therefore 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 $\pi^z(\theta)$, agent θ is subject to the incentive schedule $(w, R(w) + \pi^z(\theta)z(w))$ 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 increase $\mathrm{d}R$ of after-tax income $R(w)$ on $[w, w + \mathrm{d}w]$ around a switch point w directly increases agent θ's labor supply around w by $\eta^z \, \mathrm{d}R$, where the "static" elasticity η^z is given by

$$\eta^z(w;\theta) = |\delta_a(a;\theta) - R'(w)w_a(a;\theta) - \pi^z(\theta)z_w w_a|^{-1}. \tag{9}$$

The static elasticity under regime L does not depend on 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. This is because for a higher pension multiplier $\pi^z(\theta)$ the adjusted incentive schedule is steeper and hence the agent spends less time around any given switch point.

Second, when the pension constraint (2) is binding, a change in the after-tax schedule affects agent θ's pension multiplier $\pi^z(\theta)$, and thus modifies indirectly the incentive schedule $R(w) + \pi^z(\theta)z(w)$. Increasing R around a switch point causes $\pi^z(\theta)$ to decrease, i.e., translates into less pressure placed by the pension scheme on the agent's labor supply. The Frechet-derivatives of the multiplier with respect to after-tax income in regimes L, W and N are given by equations (A.3) and (A.4) in the appendix. Here, we informally write the negative *feedback effect* of taxes on the pension multipliers as:

$$\frac{\partial \pi(Z;\theta)}{\partial R} \leq 0. \tag{10}$$

[7] To simplify notations, when the agent picks her preferred retirement plan $Z(\theta)$, we denote labor supply as $\ell^z(a;\theta) = \ell(a;Z(\theta),\theta)$, lifetime net output $y^z(\theta) = y(Z(\theta);\theta)$, the pension multiplier $\pi^z(\theta) = \pi(Z(\theta);\theta)$ and lifetime earnings $\gamma^z(\theta) = \gamma(Z(\theta);\theta)$.

When the pension requirement is binding, a tax cut leads to lower pension multipliers, i.e. weakens the labor supply incentives placed by the pension system. The feedback effect affects labor supply in a nonlocal way through (8). By "nonlocal", we mean that a change of R around a particular switch point alters labor supply around *all* switch points of agent θ.

Lemma 1 (Labor supply)**.** *Income taxes positively affect the labor supply incentives exerted by the pension system, thus counteracting the negative static effect given by* (9)*.*

The life cycle effect of taxes is nonlocal in the sense that a tax rise at some point in the productivity distribution may affect labor supply at other points of that distribution.

We show in the appendix that the *static effect* of taxes on labor supply locally dominates the *feedback effect*. More precisely, labor supply around a switch point weakly increases following a local increase in after-tax income around that switch point –while it decreases around the other switch points. If there are no other switch points and the agent labor supply is constrained by the pension system ($\pi^z(\theta) > 0$), the static effect and the feedback effect of income taxes cancel out exactly. This is because when an agent trajectory has only one 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.

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 Section 4.2, we link the pattern of active binding incentive constraints to the shape of social weights. In Section 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.

4.1 Incentive constraints

To make sure that the level of pension $Z(\theta)$ is 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 does not strictly prefer agent θ''s allocation if and only if

$$c^z(\theta) \geq \gamma(Z(\theta');\theta) + P(Z(\theta')) = c^z(\theta') - \gamma^z(Z(\theta');\theta') + \gamma(Z(\theta');\theta). \quad (11)$$

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 (11).

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.*

Under the above assumption, the lifetime earnings (6) increase with agent type and the 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 goes one step further, showing that the slopes $\partial\gamma(Z;\theta)/\partial Z$ also increase with θ. As $c(Z,\theta) = \gamma(Z;\theta) + P(Z)$, the latter property implies that the single-crossing property holds for the iso-consumption curves. 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:*

$$\frac{\partial\gamma(Z;\theta)}{\partial Z} \text{ increases with } \theta \text{ for all } Z.$$

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

$$\frac{\mathrm{d}c^z}{\mathrm{d}\theta} = \gamma_\theta(Z(\theta);\theta) \tag{12}$$

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

Proof. We already know that $\pi(\theta) = -\partial\gamma(Z;\theta)/\partial Z$, together with $\ell(a)$, are solutions of the system made of (2) and (8). Moreover an increase in π raises $\ell(a)$ by (8), 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 $\delta(a;\theta)$. It follows that the left-hand side of (2) increases with the type θ. Therefore $\pi(\theta)$ decreases with θ under Assumption 4.1, which yields the single-crossing property. The second part of the Lemma is standard. □

The pension multiplier $\pi = -\partial\gamma/\partial Z$ reflects 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 the 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 (12), the Lagrangian can be written as

$$\mathcal{L}^z = \int_\Theta [u(c^z(\theta);\theta) - \lambda c^z(\theta) + \lambda y^z(\theta)] \, \mathrm{d}F(\theta) + \int_\Theta \mu(\theta) \left[\frac{\mathrm{d}c^z}{\mathrm{d}\theta} - \gamma_\theta(Z(\theta);\theta) \right] \mathrm{d}\theta. \tag{13}$$

We now introduce the virtual surplus

$$S^v(Z;\theta) = \lambda y(Z;\theta) - \frac{\mu(\theta)}{f(\theta)} \gamma_\theta(Z;\theta), \tag{14}$$

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

$$\mathcal{L}^z = \int_\Theta [u(c^z(\theta);\theta) - \lambda c^z(\theta)] \, \mathrm{d}F(\theta) + \int_\Theta \mu(\theta) \frac{\mathrm{d}c^z}{\mathrm{d}\theta} \, \mathrm{d}\theta + \int_\Theta S^v(Z(\theta);\theta) \, \mathrm{d}F(\theta). \tag{15}$$

The first two terms, which depend on the lifetime consumptions $c^z(\theta)$ and are controlled by the pension transfers $P(Z(\theta))$, are studied in section 4.2. The last one, namely 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 Sections 4.3 and 4.4.

4.2 Taxonomy of economies

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

- is *redistributive* if $u_c(c^z(\theta);\theta)$ 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, $\int_0^\chi [u_c(c^z(\theta);\theta) - \lambda] \, \mathrm{d}F(\theta) > 0$;
- is *anti-redistributive* if $u_c(c^z(\theta);\theta)$ increases with θ, or more generally if for any interior type χ, $\int_0^\chi [u_c(c^z(\theta);\theta) - \lambda] \, \mathrm{d}F(\theta) < 0$;
- 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 $\int_0^\chi [u_c(c^z(\theta);\theta) - \lambda] \, \mathrm{d}F(\theta)$ 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 benefit $P(\theta)$ by the same small amount in (15) yields the first-order condition:

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

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

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

Using (16) yields

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

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, $u_c(c^z(\theta); \theta)$ boils down to $u'(c^z(\theta))$ and we know from (12) that $c^z(\theta)$ increases with θ, hence $u'(c^z(\theta))$ decreases with θ.

4.3 The optimal choice of the pension requirements

Keeping c and R fixed and assuming no bunching, we maximize the virtual surplus (14) with respect to Z for each θ separately, and find the first-order condition:

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

Proposition 1. *Suppose that Assumption 4.1 holds. Then, at the second-best allocation of a redistributive economy, any agent with an interior type has her labor supply constrained by the pension system, i.e., $\pi(\theta) > 0$ for all $\theta \in (\underline{\theta}, \bar{\theta})$.*

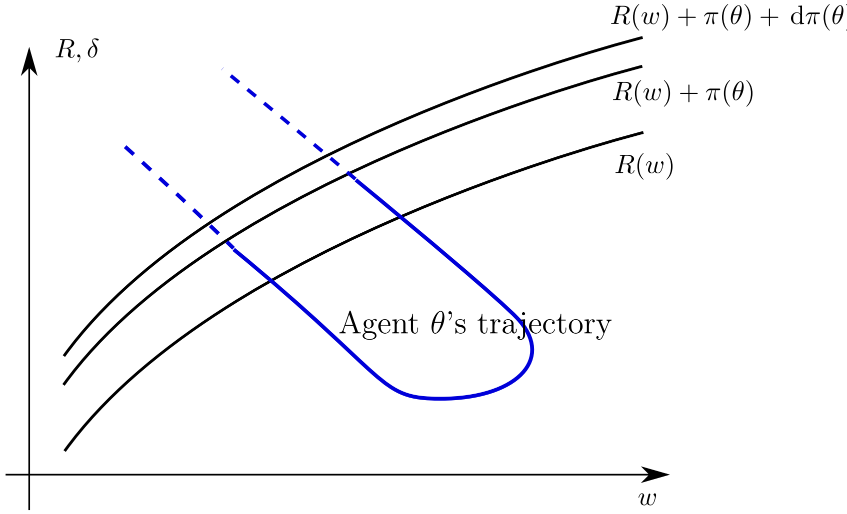

Figure 1: In regime L, a small increase in Z raises the adjusted after tax-schedule $R + \pi$ by $\mathrm{d}\pi > 0$

Proof. As the cross-derivative $\partial^2\gamma/\partial\theta\partial Z$ is positive according to Lemma 2 and by definition of a redistributive economy $\mu(\theta) > 0$ for interior types, the first-order conditions (18) imply that $\partial y^z/\partial Z > 0$ for those types. Consequently, the pension system must bite for those types, otherwise an increase in Z would have no effect on their labor supply or their net output. □

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 Figure 1. Assuming that the pension constraint (2) is binding, we show in the appendix that

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

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

It follows from Proposition 1 that for interior types, a weighted average of adjusted local distortions $(w_\sigma - \delta_\sigma)/z_\sigma$ over their switch points is positive. When the trajectory has a single switch point, the agent's labor supply is unambiguously distorted downwards.[8] By contrast, 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, see Choné and Laroque (2015) for details. The availability of a pension scheme renders this utilization of income taxes suboptimal.

4.4 Optimal income tax

We now maximize the virtual surplus (14) with respect to the tax schedule $R(w)$, keeping the pension requirements fixed, i.e., holding the equality $Z = Z(\theta)$ for each θ. We call *efficiency force* the effect of R through labor supply and net output. We call *redistribution force* the effect of R when labor supply and net output are kept fixed.

Redistribution force The redistribution force pushes the government to minimize the term

$$\int_\Theta \gamma_\theta(Z(\theta);\theta)\,\mathrm{d}F(\theta)$$

in the virtual surplus. This term reflects the inequality in consumption across types: A higher type enjoys a higher lifetime consumption; $\gamma_\theta(Z(\theta);\theta)$ measures the extent to which lifetime consumption increases with θ. For redistribution purposes, the government would like this term to be as small as possible.

Evaluating $\gamma(Z;\theta)$ at $Z = Z(\theta)$, we get from (7)

$$\gamma(Z(\theta);\theta) = \int_0^1 [R(w(a;\theta)) - \delta(a;\theta)]\,\ell^z(a;\theta)\,\mathrm{d}a \tag{20}$$

$$+ \pi(Z(\theta);\theta)\left(\int_0^1 z(w(a;\theta))\ell^z(a;\theta)\,\mathrm{d}a - Z(\theta)\right). \tag{21}$$

To compute $\gamma_\theta(Z(\theta);\theta)$, we need to differentiate the above expression with respect to θ *while holding* $Z(\theta)$ *fixed*. We first observe that the derivative of the pension multiplier π with respect to θ vanishes in the differentiation because the product $\pi_\theta(Z(\theta);\theta)\left[\int z(w)T^z_w(w;\theta)\,\mathrm{d}w - Z(\theta)\right]$ is identically zero.[9] As pension requirements and labor supply are kept fixed, the only occurrence of θ on line (21)

[8] For the extreme types $\theta = \underline{\theta}$ and $\theta = \bar{\theta}$, we know that $\mu = 0$ and therefore $\partial y^z/\partial Z = 0$. Either the agent is unconstrained by the pension system ($\pi = 0$); or she is ($\pi > 0$) and the corresponding weighted sum of local distortions (the right-hand side of (19)) is zero.

[9] In regions where the pension requirement is binding, the bracketed term is zero. In regions where the pension requirement is not binding, the multiplier, and hence its derivative, are zero.

that needs to be differentiated appears via $z(w(a;\theta))$, which gives rise to the term shown on line (23):

$$\gamma_\theta(Z(\theta);\theta) \;=\; \frac{\partial}{\partial\theta}\int_0^1 \left[R(w(a;\theta)) - \delta(a;\theta)\right]\ell^z(a;\theta)\,\mathrm{d}a \tag{22}$$

$$+\;\; \pi^z(\theta)\int_0^1 z'(w(a;\theta))w_\theta(a;\theta)\ell^z(a;\theta)\,\mathrm{d}a. \tag{23}$$

Multiplying (22) and (23) by $-\mu(\theta)$ and integrating over the set of agent types, we write the redistributive component of the expected virtual surplus as the sum of two terms

$$-\int_\Theta \mu(\theta)\gamma_\theta(Z(\theta);\theta)\,\mathrm{d}\theta = D_1 + D_2.$$

Using (17), $\mu(\underline{\theta}) = \mu(\bar{\theta}) = 0$, and integrating by parts, we rewrite the term D_1 that comes from (22) as

$$D_1 = \int_\Theta \left[u_c(c^z(\theta);\theta) - \lambda\right]\left\{\int_0^1 \left[R(w(a;\theta)) - \delta(a;\theta)\right]\ell^z(a;\theta)\,\mathrm{d}a\right\}\mathrm{d}F(\theta).$$

We now compute the Frechet-derivatives of that term with respect to the after-tax schedule –labor supply being kept fixed. A tax cut (i.e., an increase in the after-tax income) below some productivity level w raises lifetime earnings of agents who work below productivity w. For agent θ, the increase in lifetime earnings depends on the time she spends working below that productivity

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

Similarly, a tax cut in a small productivity interval $[w, w+\mathrm{d}w]$ increases lifetime earnings of 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^z with respect to w, denoted hereafter by $T^z_w(w;\theta)$. To evaluate the welfare contribution of the agents' increased lifetime earnings, the utilitarian government weights those increases with their net social marginal utility of income, $u_c(c^z(\theta);\theta) - \lambda$. It follows that the direct redistributive effect of a tax cut around productivity w is given by

$$\text{Direct redist. effect} = \int_\Theta \left[u_c(c^z(\theta);\theta) - \lambda\right] T^z_w(w;\theta)\,\mathrm{d}F(\theta). \tag{24}$$

As from (16) 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 second redistributive term, D_2, coming from (23), is given by

$$D_2 = -\int_\Theta \mu(\theta)\left[\pi^z(\theta)\int_0^1 z'(w(a;\theta))w_\theta(a;\theta)\ell^z(a;\theta)\,\mathrm{d}a\right]\mathrm{d}\theta.$$

As labor supply is maintained constant, this term involves the after-tax schedule only through its effect on the pension multiplier. Hence the contribution of the feedback effect to the redistribution force:

$$\text{Feedback redist. effect} = -\int_{\Theta} \mu(\theta)\frac{\partial \pi^z(\theta)}{\partial R}\left[\int_0^1 z'(w(a;\theta))w_\theta(a;\theta)\ell(a;\theta)\,\mathrm{d}a\right]\mathrm{d}\theta. \tag{25}$$

As explained in Section 3, tax cuts lowers pensions multipliers, recall (10). Moreover, $\mu(\theta)$ is positive in a redistributive economy; w_θ is positive under Assumption 4.1; z' is positive in the three considered regimes. It follows that the feedback effect of income taxes on pension multipliers contributes *positively* to the redistribution force.

Proposition 2. *The redistribution force expresses the effect of the tax schedule on inequality in lifetime consumptions,* $\mathrm{d}c^z/\,\mathrm{d}\theta$. *This force has two components:*

- *the direct effect on lifetime earnings, which is the covariance of working times and net social marginal utility of incomes, see* (24)*;*
- *an indirect contribution from the feedback effect on pension multipliers, which is positive in redistributive economies, see* (25).

Efficiency force A change in labor supply leaves lifetime earnings $\gamma(Z;\theta)$ unchanged for all θ and all Z because the agent is indifferent between working and not working at any switch point. Such a change therefore does not affect the derivative γ_θ. It only affects net output $y^z(\theta)$ produced by each agent, which therefore is the efficiency component of the virtual surplus.

The "static" effect of a tax cut around a switch point of agent θ, at given level of the pension multiplier $\pi^z(\theta)$, is given by

$$\mathrm{d}y^z_{\pi=\text{cst}} = [w-\delta]\eta^z(w;\theta)\,\mathrm{d}R, \tag{26}$$

where the labor supply elasticity η^z has been computed in (9).

As mentioned in Section 3, a tax cut lowers the pension multiplier $\pi^z(\theta)$. We observe furthermore that $\partial y^z/\partial \pi$ has the same sign as $\partial y^z/\partial Z$. Indeed, a rise in Z has the same effect (up to a positive factor) on output as a rise in π.[10] We know from (18) that the sign of $\partial y^z/\partial Z$ is positive at the second-best optimum of a redistributive economy. It follows that the feedback effect of the tax schedule on the pension multipliers contributes negatively to the efficiency force.

Proposition 3. *The efficiency force expresses the effect of the tax schedule on net outputs* $y^z(\theta)$. *This force has two components:*

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

- *the direct effect on net output, the sign of which depends on the direction of the labor supply distortion, see (26);*
- *an indirect contribution from the feedback effect on pension multipliers, which is always negative in redistributive economies.*

The first order conditions that determine the optimal tax schedule put to zero the sum of the redistribution and efficiency forces. Thus propositions 2 and 3 lead to one of our major results:

Proposition 4. *In a redistributive economy, the feedback effects of pensions increase the redistribution role and decrease the efficiency role of income taxes.*

The presence of the pension instrument makes it optimal for the government to place more emphasis on redistribution and less on efficiency when setting taxes. This qualitative property holds generally in all redistributive economies. In the specialized framework of the next section, it becomes unrealistically sharp and leads to a full specialization of the policy instruments.

4.5 A specialized framework

Assumption 4.2 (Decreasing trajectories)**.** *The productivities $w(a;\theta)$ decrease with age and the pecuniary costs of work $\delta(a;\theta)$ increase with age. Moreover, the agents have the same utility function $u(c)$.*

Assumptions 4.1 and 4.2, while restrictive, are 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. What matters for income taxation and labor supply distortions is whether the trajectories intersect the tax schedule at low or at high productivity levels. This depends on fine properties of the (observable) productivity and (unobservable) cost of work, as the proposition below shows.

In this section, we work in a specialized framework under Assumptions 4.1 and 4.2. This yields a sharper description of second best optimal retirement and of the interactions between the two instruments, income tax and pensions.

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

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

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 $\gamma^z(Z;\theta)$ 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 4.2, the pension system forces the agent to work longer than she would do facing the same tax schedule in the absence of pensions:

$$\pi(\theta) = -\frac{R(w(\bar{a}(\theta);\theta)) - \delta(\bar{a}(\theta);\theta)}{z(w(\bar{a}(\theta);\theta))} > 0. \tag{27}$$

The envelope theorem yields

$$\gamma_\theta(Z(\theta);\theta) = \int_0^{\bar{a}(\theta)} \left\{ [R'(w(a;\theta)) + \pi(\theta) z'(w(a;\theta))] \, w_\theta(a;\theta) - \delta_\theta(a;\theta) \right\} \, \mathrm{d}a. \tag{28}$$

The virtual surplus defined by (14) takes the form

$$\begin{aligned} f(\theta) \, S^v(Z;\theta) \;\; = \;\; & \int_0^{\bar{a}(\theta)} \{ \lambda \, [w(a;\theta) - \delta(a;\theta)] \, f(\theta) \\ & - \mu(\theta) \left([R'(w(a;\theta)) + \pi(\theta) z'(w(a;\theta))] \, w_\theta(a;\theta) - \delta_\theta(a;\theta) \right) \} \, \mathrm{d}a. \end{aligned} \tag{29}$$

As $\mu(\theta)$ and w_θ are positive, it is desirable to set the nonnegative square-bracketed term in (29) 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

$$\int_0^{\bar{a}(\theta)} \{ \lambda \; [w(a;\theta) - \delta(a;\theta)] \, f(\theta) + \mu(\theta) \, \delta_\theta(a;\theta) \} \, \mathrm{d}a, \tag{30}$$

which, assuming concavity, yields the optimal retirement age of any agent

$$\lambda \; [w(\bar{a}(\theta);\theta) - \delta(\bar{a}(\theta);\theta)] \, f(\theta) = -\mu(\theta) \, \delta_\theta(\bar{a}(\theta);\theta) > 0. \tag{31}$$

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 (29) is zero in these two regimes.

To implement this allocation, the government only 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 (27) –consistently with the general result stated in Proposition 1. 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 (31) 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 (29) cannot be set to zero. In this regime, we have $z' = 1$, and the pension multiplier $\pi(\theta)$ in (28) contributes to make the allocation more unequal, i.e., to make lifetime consumptions $c(\theta)$ more increasing with the agent type θ.

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 taxes on redistribution and on efficiency. The main insight from our analysis is that the presence of the feedback effect tends to increase the role of taxes for redistribution and to downplay their role for efficiency.

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 Section 4.5, basing the pension scheme on total working time has the advantage of minimizing a source of inequality in lifetime consumption across agents.[11] However, we have been unable to obtain a formal optimality result, and leave the design of optimal pension schemes for future research.

[11] With $z(w) = 1$, the nonnegative term $\pi(\theta)z'(w(a;\theta))$ in (28) is minimized.

References

BRETT, C. (2012): "The Effects of Population Aging on Optimal Redistributive Taxes in an Overlapping Genrerations Model," *International Tax and Public Finance*, 19, 777–799.

CHONÉ, P., AND G. LAROQUE (2015): "Optimal income tax in an extensive labor supply life-cycle model," *CESifo Economic Studies*, Forthcoming.

CREMER, H., J.-M. LOZACHMEUR, AND P. PESTIEAU (2004): "Social security, retirement age and optimal income taxation," *Journal of Public Economics*, 88, 2259–2281.

——— (2008): "Social security and retirement decision: a positive and normative approach," *Journal of Economic Surveys*, 22, 213–233.

DIAMOND, P. (2009): "Taxes and Pensions," *Southern Economic Journal*, 76(1), 2–15.

DIAMOND, P., AND J. MIRRLEES (1978): "A model of social insurance with variable retirement," *Journal of Public Economics*, 10, 295–336.

GOLOSOV, M., M. TROSHKIN, AND A. TSYVINSKI (2011): "Redistribution and Social Insurance," Discussion Paper 17642, National Bureau of Economic Research.

GORRY, A., AND E. OBERFIELD (2012): "Optimal Taxation over the Life Cycle," *Review of Economic Dynamics*, 16(1), 100–119.

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

LJUNQVIST, L., AND T. SARGENT (2014): "Career Length: Effects of Curvature of Earning Profiles, Earning Shocks, and Social Security," *Review of Economic Dynamics*, 17, 1–20.

LOZACHMEUR, J.-M. (2006): "Optimal Age-Specific Income Taxation," *Journal of Public Economic Theory*, 8, 697–711.

MICHAU, J.-B. (2014): "Optimal Redistribution: A Life-Cycle Perspective," *Journal of Public Economics*, 111, 1–16.

PRESCOTT, E. (2004): "Why do Americans work so much more than Europeans?," *Federal Reserve Bank of Minneapolis Quarterly Review*, 28, 2–13.

Prescott, E., R. Rogerson, and J. Wallenius (2009): "Lifetime aggregate labor supply with endogenous workweek length," *Review of Economic Dynamics*, 12, 23–36.

Rogerson, R. (2011): "Individual and Aggregate Labor Supply With Coordinated Working Times," *Journal of Money, Credit and Banking*, 43, 7–37.

Rogerson, R., and J. Wallenius (2009): "Micro and macro elasticities in a life cycle model with taxes," *Journal of Economic Theory*, 144, 2277–2292.

Shourideh, A., and M. Troshkin (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," *Review of Economic Studies*, 78, 1490–1518.

A Appendix

Labor supply elasticity 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., $R(w) + \pi(\theta)z(w) - \delta(a_\theta) = s$. 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 $|\delta_a(a;\theta) - R'(w)w_a(a;\theta) - \pi(\theta)z_w|$ 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

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

hence η^z 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, equation (9) must be modified by adding up the contributions of all the ages at which the agent switches.

$$-\frac{\eta^z(w)\eta^z(w')z(w)z(w')}{\sum_{\sigma\in\mathcal{S}_\theta} z^2(w_\sigma)\eta^z(w_\sigma)}\,\mathrm{d}R, \tag{A.1}$$

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

Effect on taxes on pension multiplier The multiplier $\pi(Z;R)$ is defined jointly by (2) and (8), 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 $\Theta(a,\pi) = R(w(a)) + \pi z(w(a)) - \delta(a)$:

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

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

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

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

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

in regime L and W and

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

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:

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

in regime L and W, and by

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

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.

Net output 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).

Finally, 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'

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

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

$$\frac{\eta^z(w')z(w')}{\sum_{\sigma \in \mathcal{S}_\theta} z^2(w_\sigma)\eta^z(w_\sigma)}\,\mathrm{d}Z. \tag{A.6}$$