



Working Papers

www.cesifo.org/wp

Public Pensions in a Multi-Period Mirrleesian Income Tax Model

Spencer Bastani
Sören Blomquist
Luca Micheletto

CESIFO WORKING PAPER NO. 6206

CATEGORY 1: PUBLIC FINANCE

NOVEMBER 2016

An electronic version of the paper may be downloaded

- *from the SSRN website:* www.SSRN.com
- *from the RePEc website:* www.RePEc.org
- *from the CESifo website:* www.CESifo-group.org/wp

ISSN 2364-1428

Public Pensions in a Multi-Period Mirrleesian Income Tax Model

Abstract

Using an OLG model with skill uncertainty and private savings, we investigate whether an optimally designed set of public pension transfers can usefully supplement a nonlinear labor income tax as a welfare-enhancing policy instrument. We consider a Mirrleesian setting where agents' skills are private information and highlight that, even though pensions, by crowding out private savings, adversely affect the achievement of the golden-rule, they can be used as a mimicking-detering device that makes it easier for the government to achieve the desired redistributive goals.

JEL-Codes: H210, H550, H630.

Keywords: public pensions, dynamic optimal income taxation, capital income taxation, tagging.

Spencer Bastani
Department of Economics and Statistics
Linnaeus University
Växjö / Sweden
spencer.bastani@lnu.se

Sören Blomquist
Department of Economics
Uppsala University
Uppsala / Sweden
soren.blomquist@nek.uu.se

Luca Micheletto
Department of Law
University of Milan
Italy – 20122 Milan
luca.micheletto@unimi.it

4 November 2016

We thank Pierre Pestieau for a helpful discussion and Robert Holzmann for insightful comments.

1 Introduction

A key feature of Mirrleesian models of optimal nonlinear income taxation is that, due to a problem of asymmetric information (individual skills are private information of the agents and are not directly observable by the government), the government must design the income tax subject to a set of self-selection (incentive-compatibility) constraints. This means that the income tax must be chosen in such a way that agents find in their own interest to choose the (pre-tax income, post-tax income)-bundles that the government wishes them to select rather than acting as “mimickers” by choosing a bundle intended for a different type.

In a static setting where agents supply labor in one single period and where there are N types of agents, differing only in terms of market ability, the government will aim at offering N different bundles in the (pre-tax income, post-tax income)-space, one for each skill type. More generally, in dynamic settings where agents supply labor in several periods, the government would like to differentiate the (pre-tax income, post-tax income)-bundles offered to agents not only on the basis of the specific skill type of an agent at a given period, but also on the basis of the agent’s past history of skill realization. This is due to the fact that in a dynamic setting one really has as many different types of agents as the number of different lifetime skill realization histories. However, the possibility to screen among agents with the same current skill but different skill-realization histories is severely limited under a *history-independent* nonlinear income tax, i.e. a nonlinear income tax which only depends on current earned income. The reason is that, when choosing the (pre-tax income, after-tax income)-bundles to be offered to agents, the government faces a very large set of incentive-compatibility constraints (hereafter, IC-constraints) since in every period agents can choose any of the bundles defined by the income tax schedule. Under a *history-dependent* income tax, on the other hand, i.e. a income tax which is allowed to be a general function of both the current income and the past earned incomes of an agent, screening becomes easier since the government has a more powerful tool to punish agents who were to adopt mimicking strategies.¹ This implies that the set of IC-constraints faced by the government is smaller under a history-dependent income tax than under a history-independent income tax.² Moreover, as has been documented by some recent contributions,³ the stronger screening power of a history-dependent

¹As an illustrative example, consider the case where there is an agent of skill type j who plans to earn at period t the same pre-tax income intended by the government for a different type (who can be different either because he is of skill type $i \neq j$ at period t or because he has a different history of past earned incomes) in order to get the same tax treatment. Under a history-independent income tax, if the two agents earn at period t the same pre-tax income, they will also pay the same income tax; under a history-dependent tax, instead, for the tax liability at period t to be the same it is not enough that the two agents earn the same income; it must also be the case that their past record of earned incomes is the same. This stricter requirement can be exploited by the government to make mimicking less appealing.

²The desirability of using history-dependent income taxes is at the heart of the so called *new dynamic public finance* literature. See, e.g., Golosov et al. (2007), Kocherlakota (2010) and Golosov and Tsyvinski (2015).

³See, e.g., Farhi and Werning (2013) and Weinzierl (2013).

income tax generates sizeable welfare gains over an ordinary, history-independent, nonlinear income tax.

However, as pointed out by Diamond and Saez (2011), the structure of an optimal fully history-dependent tax is likely to be too complex to be relevant for actual public policies. This observation has spurred the interest for exploring alternative, but less complex, instruments which might allow achieving a substantial share of the welfare gains attainable via fully history-dependent taxes. One such a device, which has received considerable attention in the optimal tax literature, is tagging by age, i.e. the possibility to make the tax schedule a function of both earned income and the age of the taxpayer.⁴

In this paper we focus on the welfare-enhancing role of a different policy instrument, namely public pensions, and we look at the mechanisms by which it can usefully supplement an optimal nonlinear labor income tax. In doing that, we will contribute to the relatively small normative literature that considers the interaction between nonlinear taxes and pension schemes in second-best settings characterized by asymmetric information between the government and the private agents.⁵ However, while most of the contributions in this area (see, e.g., Cremer et al., 2004, Choné and Laroque, 2014, Shourideh and Troshkin, 2015) emphasize how the individuals' extensive margin of choice (when to retire) should be optimally distorted, we will neglect the retirement decision, assuming that there is a fixed mandatory retirement age, and focus instead attention on how public pensions can allow improving upon the social optimum achieved by relying only on (annual) nonlinear labor income taxation and linear interest income taxation.

As a vehicle for our analysis, we consider a three-period OLG economy with heterogeneous agents in which individuals work for two periods and then retire, and we let young workers face uncertainty about their second-period skill realization.

We highlight that, even though public pensions, by crowding out private savings, adversely affect the achievement of the golden-rule, they can be used as a mimicking-detering device that makes it easier for the government to achieve the desired redistributive goals. As we will see, the welfare effects of public pensions depend, in general, on the assumptions that are made about the possibility for agents to borrow against their future pension entitlements. When agents are free to borrow against their future pension entitlements,

⁴Akerlof (1978) is the seminal contribution exploring the idea of using observable, and hard to manipulate, individual characteristics as “tags” to lower the efficiency costs of redistributive programs. Subsequent contributions on tagging and income taxation include Boskin and Sheshinski (1983), Immonen et al. (1998), Cremer et al. (2010), Mankiw and Weinzierl (2010), Bastani (2013), and Kanbur and Tuomala (2016). Blomquist and Micheletto (2008) provides the first analysis considering tagging by age in the context of a Mirrleesian setting. The idea of age-dependent taxation is described as one of the most promising avenue to improve the tax system in Banks and Diamond (2010). Recent contributions by Farhi and Werning (2013) and Weinzierl (2011) seem to suggest that age-dependent taxes can capture a substantial share of the total welfare gains achievable through more sophisticated fully history-dependent tax schemes.

⁵As stated by Diamond (2009), “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”.

public pensions can be used as an instrument to (almost) replicate the optimal allocation that would have been implemented had the labor income tax schedule been allowed to be history-dependent. When instead agents cannot borrow against their future pension entitlements, this might no longer be the case but, on the other hand, an optimally designed set of pension transfers can relax some IC-constraints that could not have been affected if agents were unrestricted in borrowing against their future pension entitlements. In particular, the pension system can be designed to relax those binding IC-constraints where mimickers plan to save less for retirement than the agents that they plan to mimic.

To shed light on the potential magnitude of the welfare gains descending from using public pensions as an additional policy tool, we perform some simulations based on U.S. data (assuming that agents cannot borrow against their future pension transfers). As a benchmark system to compute welfare gain measures we consider the solution to the government's problem under an optimal age-independent (hereafter, AI) nonlinear labor income tax with zero public debt, no pensions, and an optimally chosen interest income tax. Our main results can be summarized as follows.

For a utilitarian social welfare function, when the nonlinear labor income tax is AI and the government has no access to public debt policy, the possibility to introduce pension transfers generates a sizeable welfare gain, equivalent to about 2.48% of the aggregate GDP. If, on top of allowing the government to optimally choose the set of pension transfers, we also assume that the government faces no constraints in managing public debt policy, the welfare gains amount to 4.73%. Once the government is unrestricted in its use of debt policy and pensions transfers are optimally chosen, the welfare gains that could be achieved by moving from an optimal AI nonlinear labor income tax to an optimal age-dependent (hereafter, AD) income tax are virtually nil. In particular, an optimal AD nonlinear income tax, coupled with optimally chosen pensions and an optimal interest income tax, would deliver a welfare gain of 6.18% over the benchmark system, implying a 0.01% welfare gain in moving from an AI labor income tax to an AD labor income tax. However, when the government has no access to public debt policy, there remain large welfare gains (about 3%), even with pensions being used as an additional policy instrument, in moving from an optimal AI labor income tax to an optimal AD labor income tax.

Finally, when the nonlinear income tax is AD, the welfare gains of optimally choosing pensions are about 4% (irrespective on the assumptions about the availability of debt policy).

Qualitatively similar results are obtained under a max-min social welfare function. Two are the main differences between the utilitarian and the max-min case. One is that, whereas under a utilitarian objective function, debt policy is always (both with and without pensions) a redundant instrument for the government

when the labor income tax is AD, this is no longer the case under a max-min objective function. Specifically, when public debt is constrained to be non-negative and pensions are optimally chosen, the economy does not reach the golden rule even if the nonlinear income tax is AD. However, even though this implies that, when debt policy is unconstrained, public pensions are more generous and most agents only rely on them to finance consumption during retirement, the welfare gain descending from the availability of debt policy is negligible, about 0.04%. The other difference between the utilitarian and the max-min case is that, whereas in the former case public pensions tend to mostly favor the agents who were of intermediate skill type when young, in the latter case pensions favor the agents who were of the lowest skill type when young.

The plan of the paper is the following. Section 2 presents the main ingredients of the model, while the government's problem is described in Section 3. Section 4 provides a discussion of the mechanisms by which public pensions can represent a welfare-enhancing policy tool. Section 5 provides a description of the calibration used in the numerical simulations. Section 6 presents the results of the numerical simulations. Finally, Section 7 offers some concluding remarks.

2 The Model

Consider a three-period OLG economy where agents, heterogeneous in terms of productivity (skills), work in the first two periods and retire in the third. Agents are indexed by their skill and age, and start their lives as one of m^y possible productivity types. Let a young agent's productivity be denoted by θ_i , with $i \in \{1, \dots, m^y\}$, and let π_i denote the proportion of young agents of type i , with $\sum_{i=1}^{m^y} \pi_i = 1$. With probability π_{ij} an agent who is of productivity i when young has productivity θ_{ij} , with $j \in \{1, \dots, m^o\}$, when middle-aged, where $\pi_{ij} \geq 0 \forall (i, j) \in \{1, \dots, m^y\} \times \{1, \dots, m^o\}$ and $\sum_{j=1}^{m^o} \pi_{ij} = 1$. Denoting the sets $\{1, \dots, m^y\}$ and $\{1, \dots, m^o\}$ by I and J respectively, we interpret θ_σ as referring to the productivity of a middle-aged agent if σ is a tuple $(i, j) \in I \times J$, and as referring to the productivity of a young agent if σ is a single index $i \in I$.

Labor supply and consumption of a young agent of type i born in period t is denoted by $h_{i,t}$ and $c_{i,t}$ respectively, whereas $h_{ij,t}$ and $c_{ij,t}$ denote the labor supply and consumption at time t of a j -type middle-aged agent (born in $t-1$) who was of type i when young. In the third period of lifetime labor supply is nil, and we denote by $c_{ij,t}^R$ the consumption in period t of a retiree who was of type i when young and of type j when middle-aged. Starting from an initial size N_0 , population grows at rate n , so that the size of the

cohort born at time t is $N_t = (1+n)^t N_0$. Labor in efficiency units and consumption are defined as

$$\tilde{L}_t = \sum_{i=1}^{m^y} \pi_i \theta_i \ell_{i,t} + (1+n)^{-1} \sum_{i=1}^{m^y} \pi_i \left(\sum_{j=1}^{m^o} \pi_{ij} \theta_{ij} \ell_{ij,t} \right) \quad (1)$$

and

$$\tilde{C}_t = \sum_{i=1}^{m^y} \pi_i c_{i,t} + (1+n)^{-1} \sum_{i=1}^{m^y} \pi_i \left[\sum_{j=1}^{m^o} \pi_{ij} \left(c_{ij,t} + (1+n)^{-1} c_{ij,t}^R \right) \right], \quad (2)$$

whereas total effective labor and consumption in the economy are given by $L_t = N_t \tilde{L}_t$ and $C_t = N_t \tilde{C}_t$.

Firms operate a CRS technology $F(K_t, L_t)$ admitting zero equilibrium profits. The capital depreciation rate, the interest rate, and the wage rate per efficiency unit of labor are denoted, respectively, by δ , r_t and ω_t . With perfectly competitive markets, factors earn their marginal products and we have $r_t = f'(k_t) - \delta$ and $\omega_t = f(k_t) - k_t f'(k_t)$, where $k_t \equiv K_t/L_t$ and $f(k_t) = F(K_t/L_t, 1)$. An agent's labor income is defined as the product of labor in efficiency units and the equilibrium wage rate: $y_{i,t} = \omega_t \theta_i \ell_{i,t} \equiv w_{i,t} \ell_{i,t}$ and $y_{ij,t} = \omega_t \theta_{ij} \ell_{ij,t} \equiv w_{ij,t} \ell_{ij,t}$.

In each period, individual preferences are represented by a twice differentiable, strictly quasi-concave utility function $u(c_t, \ell_t)$, where $u_c > 0$ and $u_\ell < 0$. Denoting by β a time-preference parameter, the expected lifetime utility of an agent with productivity i when young is given by:

$$EU_{i,t} = u(c_{i,t}, \ell_{i,t}) + \sum_{j=1}^{m^o} \pi_{ij} [\beta u(c_{ij,t+1}, \ell_{ij,t+1}) + \beta^2 u(c_{ij,t+2}^R)]. \quad (3)$$

We consider two alternative systems of labor income taxation. One is represented by an AI nonlinear labor income tax where the same tax schedule applies to both young and middle-aged workers. The other is a system of AD taxation where the government chooses two different nonlinear tax schedules, one for the young and one for the middle-aged. Thus, we consider tax systems of the form $\{T_1(y_i), T_2(y_{ij})\}$, where $T_1(\cdot)$ is the tax function that applies to young agents and $T_2(\cdot)$ the tax function for middle-aged agents, and where $T_1(\cdot)$ is constrained to be equal to $T_2(\cdot)$ under an AI tax. With respect to capital income taxation, we will maintain throughout our analysis the assumption that interest income can only be taxed at a proportional rate. However, we will also allow for the possibility that the interest tax rate can be made AD.⁶ This will

⁶Linearity of interest income taxation is usually justified based on the assumption that savings transactions can only be observed *anonymously*, not at an *individual* level, preventing the tax authority to tax the returns on savings nonlinearly. Notice that this informational assumption would also represent a hindrance for any attempt by the government to try to implement an AD interest income tax. Thus, allowing for the possibility of AD interest income taxation requires relaxing the assumption that savings can only be observed anonymously.

enable us, when performing numerical simulations, to explore whether there would be large welfare gains in moving from a system of AI interest income taxation to a system of AD interest income taxation.

Finally, as a further policy instrument at the government's disposal, we allow for the possibility of running a PAYG pension system where agents get in the third period of their life a pension transfer which is a (general) function of their lifetime labor earnings profile. When allowing for this additional policy instrument, we will make the assumption that, in each period, the pension transfers paid to the current retirees are entirely financed by a proportional, AI, social security contribution rate α on the labor income of the current workers.⁷

Denoting after-tax labor income when young and when middle-aged by, respectively, $b_i = y_i - T_1(y_i) - \alpha y_i$ and $b_{ij} = y_{ij} - T_2(y_{ij}) - \alpha y_{ij}$, and denoting the savings in the first and second period by respectively $s_{i,t}$ and $s_{ij,t+1}$, (3) can be rewritten as:

$$EU_{i,t} = u\left(b_{i,t} - s_{i,t}, \frac{y_{i,t}}{w_{i,t}}\right) + \beta \sum_{j=1}^{m^o} \pi_{ij} u\left(b_{ij,t+1} + (1 + r_{t+1}(1 - \tau_{1,t+1}))s_{i,t} - s_{ij,t+1}, \frac{y_{ij,t+1}}{w_{ij,t+1}}\right) + \\ + \beta^2 \sum_{j=1}^{m^o} \pi_{ij} u\left((1 + r_{t+2}(1 - \tau_{2,t+2}))s_{ij,t+1} + P(y_{i,t}, y_{ij,t+1})\right).$$

where $\tau_{1,t+1}$ denotes the interest income tax rate that applies to the return on savings made by young agents, $\tau_{2,t+2}$ denotes the interest income tax rate that applies to the return on savings made by middle-aged agents, and $P(y_{i,t}, y_{ij,t+1})$ represents the pension transfer received in the third period.

For given values of $b_{i,t}$, $y_{i,t}$, $b_{ij,t+1}$, $y_{ij,t+1}$, $\tau_{1,t+1}$, $\tau_{2,t+2}$ and $P_{ij,t+2}$, an individual chooses first- and second-period savings to maximize his expected lifetime utility. When a young agent makes plans for second-period savings, these plans are contingent on the realization of his productivity when middle-aged.

The aggregate savings of generation t is given by $N_t s_t = N_t \sum_{i=1}^{m^y} \pi_i s_{i,t}$ and the aggregate savings in period t of the generation born in period $t - 1$ is given by $N_{t-1} s_{t-1} = N_{t-1} \sum_{i=1}^{m^y} \pi_i \left(\sum_{j=1}^{m^o} \pi_{ij} s_{ij,t} \right)$. Agents invest either in government bonds D_t or in physical capital. The capital stock in period $t + 1$, is equal to that part of investment that goes into physical capital in period t :

$$K_{t+1} = N_t s_t + N_{t-1} s_{t-1} - D_t. \quad (4)$$

⁷Given that we assume that the social security contribution rate is proportional and AI and the nonlinear labor income tax is nonlinear, it does not matter whether one assumes that social security contributions are subtracted or not from the base of the nonlinear labor income tax.

Dividing (4) by $L_{t+1} = N_{t+1}\tilde{L}_{t+1}$, and defining $d_{t+1} \equiv D_t/N_{t+1}$, gives

$$\tilde{L}_{t+1} (1+n) k_{t+1} = s_t + \frac{s_{t-1}}{1+n} - d_{t+1} (1+n). \quad (5)$$

Denoting by S_t the quantity $N_t s_t + N_{t-1} s_{t-1}$, the resource constraint of the economy is:

$$(1-\delta) K_t + F(K_t, L_t) = C_t + S_t. \quad (6)$$

Dividing (6) by L_t gives

$$(1-\delta) k_t + f(k_t) = C_t/L_t + S_t/L_t. \quad (7)$$

Combining (5) and (7) with the private budget constraints relating before- to after-tax income, and using the definitions of C_t and L_t , one obtains the government's budget constraint:

$$\begin{aligned} & \sum_{i=1}^{m^y} \pi_i [T_{1,t}(w_{i,t}\ell_{i,t}) + \alpha_t w_{i,t}\ell_{i,t}] + (1+n)^{-1} \sum_{i=1}^{m^y} \pi_i \left[\sum_{j=1}^{m^o} \pi_{ij} (T_{2,t}(w_{ij,t}\ell_{ij,t}) + \alpha_t w_{ij,t}\ell_{ij,t}) \right] + \\ & (1+n)^{-1} r_t \left[\tau_{1,t} \sum_{i=1}^{m^y} \pi_i s_{i,t-1} + (1+n)^{-1} \tau_{2,t} \sum_{i=1}^{m^y} \pi_i \left(\sum_{j=1}^{m^o} \pi_{ij} s_{ij,t-1} \right) \right] + (1+n) d_{t+1} \\ = & (1+r_t) d_t + (1+n)^{-2} \sum_{i=1}^{m^y} \pi_i \sum_{j=1}^{m^o} \pi_{ij} P_{ij,t}, \end{aligned}$$

where $P_{ij,t} \equiv P_{ij,t}(w_{i,t-2}\ell_{i,t-2}, w_{ij,t-1}\ell_{ij,t-1})$ denotes the pension transfer paid in period t to a retiree who was of type i when young and of type j when middle-aged.

In the following section we describe the government's problem. To simplify the analysis, we will be focusing on stationary equilibria where prices are constant, tax and debt policy is constant, and all per capita quantities are constant. Formally, a stationary equilibrium is defined as:

Definition 1 *A stationary equilibrium is a competitive equilibrium with the property that, for all periods $t \geq \bar{t}$, $\{T_{1,t}(y_i), T_{2,t}(y_{ij}), \tau_{1,t}, \tau_{2,t}, P_{ij,t}\} = \{T_1(y_i), T_2(y_{ij}), \tau_1, \tau_2, P_{ij,t}\}$, $\{r_t, \omega_t\} = \{r, \omega\}$, $d_t = d$, $\{\tilde{C}_t, \tilde{L}_t\} = \{C_t, L_t\}$, and $k_t = k$.*

3 General structure of the government's problem

The government knows the skill distribution at each age and the Markov probabilities relating these distributions. Moreover, it can observe pre-tax incomes, $y = \omega\theta\ell$, whereas neither individual skills, θ , nor labor supplies, ℓ , are publicly observable. The government has at its disposal a nonlinear labor income tax schedule that can be conditioned on the age of the taxpayer (assumed to be observable). Agents maximize expected utility based on the link between pre-tax earnings and post-tax earnings implied by the tax schedules. Thus, the government's problem can equivalently be stated as choosing the allocations $\{b_i, y_i\}_{i \in I}$ and $\{b_{ij}, y_{ij}\}_{(i,j) \in I \times J}$ subject to a set of IC-constraints and a public budget constraint.

Instead of choosing a single income point as in a static optimal tax problem, in a dynamic setting agents choose a strategy. A strategy specifies which income point is chosen by an agent in each period of work and for each state of the world (i.e., for each skill realization in the second period of life). Agents' strategies are independent of each other and there is no aggregate uncertainty. Each strategy implies (unique) savings decisions consistent with the chosen income points and the agent's first order conditions for savings. The IC-constraints require that each agent (weakly) prefers to choose the strategy intended for his type rather than behave as a mimicker by choosing a deviating strategy. Formally, a strategy corresponding to an agent i is, under an AD tax, a plan $\sigma^i = (\sigma_1^i, \sigma_2^i)$ where $\sigma_1^i \in I$ is the reported type when young and σ_2^i is a functional $\sigma_2^i : J \rightarrow I \times J$ determining the income point chosen when middle-aged as a function of the second period skill realization $j \in J$. The set of all strategies available to agent i is denoted Γ_i . Truth-telling implies that a young agent of ability type i chooses the income point (b_i, y_i) in the first period and the income point (b_{ij}, y_{ij}) in the second period if j is his skill realization when middle-aged. We denote the truthful strategy for an agent of type i by $\tilde{\sigma}^i = (\tilde{\sigma}_1^i, \tilde{\sigma}_2^i)$ with $\tilde{\sigma}_1^i = i$ and $\tilde{\sigma}_2^i(j) = (i, j)$, $\forall j \in J$. Under an AI tax, a strategy corresponding to an agent i is instead a plan $\sigma^i = (\sigma_1^i, \sigma_2^i)$ where $\sigma_1^i \in (I \times J \cup I)$ is the reported type when young and σ_2^i is a functional $\sigma_2^i : J \rightarrow (I \times J \cup I)$ determining the income point chosen when middle-aged as a function of the second-period skill realization $j \in J$.

Assuming that the government maximizes a concave social welfare aggregator $W(\cdot)$ whose arguments are the steady-state expected lifetime utilities of the m^y types of agents forming a representative cohort,⁸ the government's problem can be described as:

$$\max_{d, \tau_1, \tau_2, \{b_i, y_i\}_{i \in I}, \{b_{ij}, y_{ij}\}_{(i,j) \in I \times J}, \{P_{ij}\}_{(i,j) \in I \times J}} W \left(V_1(\tilde{\sigma}^1), \dots, V_{m^y}(\tilde{\sigma}^{m^y}) \right)$$

⁸This is equivalent to maximizing the discounted sum of the expected lifetime utilities of all agents of all subsequent generations.

subject to the set of IC-constraints:

$$\forall i \in I : V_i(\tilde{\sigma}^i) \geq V_i(\sigma^i), \quad \forall \sigma^i \in \Gamma_i, \quad (8)$$

the government's (per-period) budget constraint:

$$\begin{aligned} \sum_{i=1}^{m^y} \pi_i (y_i - b_i) + (1+n)^{-1} \sum_{i=1}^{m^y} \pi_i \left(\sum_{j=1}^{m^o} \pi_{ij} (y_{ij} - b_{ij}) \right) + (1+n)^{-1} r \left[\tau_1 \sum_{i=1}^{m^y} \pi_i s_i + (1+n)^{-1} \tau_2 \sum_{i=1}^{m^y} \pi_i \left(\sum_{j=1}^{m^o} \pi_{ij} s_{ij} \right) \right] + \\ + (1+n)d = (1+r)d + (1+n)^{-2} \sum_{i=1}^{m^y} \pi_i \sum_{j=1}^{m^o} \pi_{ij} P_{ij}, \end{aligned}$$

and the capital market equilibrium condition:

$$(1+n)k\tilde{L} = s + (1+n)^{-1}s_{-1} - d(1+n), \quad (9)$$

where we have defined $\tilde{L} \equiv \sum_{i=1}^{m^y} \pi_i (y_i/\omega) + (1+n)^{-1} \sum_{i=1}^{m^y} \pi_i \left(\sum_{j=1}^{m^o} \pi_{ij} (y_{ij}/\omega) \right)$, and

$$\begin{aligned} V_i(\tilde{\sigma}^i) &\equiv u \left(b_i - s_i, \frac{y_i}{w_i} \right) + \beta \sum_{j=1}^{m^o} \pi_{ij} u \left(b_{ij} + (1+r(1-\tau_1))s_i - s_{ij}, \frac{y_{ij}}{w_{ij}} \right) + \\ &+ \beta^2 \sum_{j=1}^{m^o} \pi_{ij} u \left((1+r(1-\tau_2))s_{ij} + P(y_i, y_{ij}) \right), \\ V_i(\sigma^i) &\equiv u \left(b_{\sigma_1^i} - s_{\sigma_1^i}, \frac{y_{\sigma_1^i}}{w_i} \right) + \beta \sum_{j=1}^{m^o} \pi_{ij} u \left(b_{\sigma_2^i(j)} + (1+r(1-\tau_1))s_{\sigma_1^i} - s_{\sigma_2^i(j)}, \frac{y_{\sigma_2^i(j)}}{w_{ij}} \right) + \\ &+ \beta^2 \sum_{j=1}^{m^o} \pi_{ij} u \left((1+r(1-\tau_2))s_{\sigma_2^i(j)} + P(y_{\sigma_1^i}, y_{\sigma_2^i(j)}) \right), \end{aligned}$$

with $s_{\sigma_1^i}$ denoting the savings chosen in the first period by a type i who adopts the strategy σ^i , and $s_{\sigma_2^i(j)}$ denotes the savings prescribed by strategy σ^i in the second period conditional on the agent's skill realization when middle-aged being j .

Having described the structure of the government's problem, it is worth summarizing how, in the absence of public pensions, a nonlinear AD labor income tax enables the policy maker to Pareto-improve upon the optimum achieved through an AI tax. As highlighted in Bastani et al. (2013), the potential gains of an AD nonlinear labor income tax as compared to an AI one are twofold. First, by shutting down some mimicking

strategies, it reduces the number of IC-constraints faced by the government.⁹ Second, in a general equilibrium environment with endogenous capital accumulation and constraints on the management of public debt policy, only an AD tax can always achieve the golden-rule level of capital without interfering with the redistributive goals of the government.¹⁰

We are now ready to analyze how public pensions can be used as an additional welfare-enhancing policy instrument. For simplicity, the discussion provided in Section 4 will be based on the assumption that, conditional on the skill type of an agent when young, two are the possible skill realizations when middle-aged: either an agent experiences an unfavorable skill shock or a favorable one. Notice that we will use this terminology even though we do not assume that, if an agent is hit by an unfavorable skill shock, his wage rate in the second period is necessarily lower than his wage rate in the first period.

4 Public pensions as a welfare-enhancing device

Introducing public pensions as an additional policy instrument requires the government to optimally choose how they are related to the labor income earned by agents during their working years. Formally, the government has to optimally select the set of transfers $P(y_i, y_{ij})$, where y_i refers to labor income earned in the first period and y_{ij} to labor income earned in the second period, and where no restriction is *ex ante* placed on the shape of the $P(y_i, y_{ij})$ function, except that $P(y_i, y_{ij}) \geq 0$.

As mentioned above, the potential gains of an AD nonlinear labor income tax, as compared to an AI one, come from two different sources: mimicking-detering effects (an AD tax reduces the number of IC-

⁹Under an AD tax an individual cannot choose an income point intended for an agent belonging to a different age group.

¹⁰This implies that under an optimal AD nonlinear tax the economy reaches a steady state where the golden-rule capital-labor ratio is achieved even if debt policy is unavailable. Notice that, under an AD nonlinear tax one can always marginally change the after-tax labor incomes of the young, and at the same time adjust the after-tax labor incomes of the middle-aged, in such a way that the present-value lifetime tax payment of all agents is left unaffected, all young agents change their savings by the same amount, the public budget is kept balanced, and all the IC-constraints continue to be satisfied if they were initially satisfied (see Appendix A.2 in Bastani et al. (2013)). For instance, increasing marginally all the after-tax labor incomes assigned to the various young workers, b_i , and decreasing by $1 + r(1 - \tau_1)$ all the after-tax labor incomes assigned to the various middle-aged workers, b_{ij} , would not affect the IC-constraints faced by the government (since middle-aged workers are prevented, under an AD tax, from choosing a bundle on the income tax schedule for the young) but it would imply that all young workers respond by marginally raising their savings in the first period in order to keep consuming the same amount of material goods and leisure in all periods and possible states of the world. In this sense one can claim that the absolute level of private savings does not matter for self-selection purposes; savings can be controlled (by a proper choice of the labor income taxes when young and when middle-aged) with the sole purpose of achieving the golden-rule capital stock and without interfering with the redistributive goals of the government. Instead, under an optimal AI tax the economy does not necessarily reach a steady state where the golden-rule is satisfied. Whether it does or not depends crucially on the assumptions about debt policy. If debt policy is fully unrestricted, it can be used by the government to achieve the golden-rule while the tax policy can be chosen with the sole purpose of fulfilling the intra-generational redistributive objectives. However, if there are restrictions on the use of debt policy, the tax instruments are burdened with the pursuit of two objectives, achieving the golden-rule and redistributing among agents of a given cohort, which are not necessarily coherent. For instance, there is no guarantee that trying to boost the private savings by marginally increasing all the after-tax labor incomes assigned to the various young workers, b_i , and decreasing by $1 + r(1 - \tau_1)$ all the after-tax labor incomes assigned to the various middle-aged workers, b_{ij} , would not violate the IC-constraints faced by the government (since middle-aged workers can, under an AI tax, choose a bundle on the single income tax schedule).

constraints faced by the government) and capital-accumulation effects (in the presence of restrictions on debt policy, only an AD tax can always achieve the golden-rule level of capital without interfering with the redistributive goals of the government).

Let's consider how pensions can play a role as a welfare-enhancing device when used as a policy tool that supplements an optimal nonlinear labor income tax (with or without interest income taxation). Given that the calibrated wage process that we consider in our numerical simulations is such that, at the solution to the government's tax problem in the absence of public pensions, the steady-state optimum is dynamically efficient both under an AD tax and under an AI tax, it is clear that pensions cannot generate welfare gains that are related to capital-accumulation effects. By providing agents with a transfer in their retirement period, pensions will lessen the need for agents to save for retirement purposes. Thus, in terms of capital-accumulation effects, pensions will not be desirable since they represent a hindrance for the achievement of the golden-rule.¹¹ This also means that, if pensions can have a role as a welfare-enhancing device, such a role must be traced back to mimicking-detering effects. With respect to this kind of effects, two cases need to be distinguished, depending on whether one assumes that agents are free to borrow against their future pension entitlements or whether they aren't. In discussing these two cases we will for simplicity assume that the government faces no restrictions on its debt policy, so that it can control an independent policy instrument that allows the economy to reach a steady state where the golden rule is achieved.¹²

In the former case, the pension transfers can be combined with a reform in the labor income tax schedule such that one makes non-binding all the IC-constraints associated with deviating strategies where a mimicker does not perfectly match the lifetime labor income profile of a non-deviating agent.

Intuitively, suppose that, at an optimum without pensions, one of the binding IC-constraints is associated with the following deviating strategy: a young type 2 agent chooses the income point (y_2, b_2) intended for him

¹¹This is true both when the underlying nonlinear labor income tax is AI and when it is AD. In particular, when pension transfers are provided to retired agents, it is no longer true that an AD tax can always achieve the golden-rule level of capital accumulation without interfering with the redistributive goals of the government, at least if one assumes that all the elements in the set $\{b_{ij}\}$ are constrained to be non-negative (which is equivalent to assume that the average tax rate on the labor income earned by a middle-aged agent cannot exceed 100%). If pension transfers are sufficiently generous, the savings by middle-aged workers will eventually become nil. At that point, the only source of private savings will be savings made by the young for the purpose of supplementing their after-tax labor income when middle-aged. By reallocating the income tax burden from the young to the middle-aged workers (lowering the after-tax labor income of the middle-aged and raising it for the young), the government can provide a boost to private savings, but this effect might not be strong enough to allow the economy to reach the golden-rule in steady-state. Savings by the young can be boosted, without affecting the IC-constraints faced by the government, by increasing marginally all the after-tax labor incomes assigned to the various young workers, b_i , and decreasing by $1 + r(1 - \tau_1)$ all the after-tax labor incomes assigned to the various middle-aged workers, b_{ij} . But such a saving-increasing intertemporal reshuffling of the income tax burden can be replicated, without affecting the IC-constraints faced by the government, only until a point is reached when one of the elements in the set $\{b_{ij}\}$ becomes equal to zero. Once this happens, and under a constraint that all the elements in the set $\{b_{ij}\}$ must be non-negative, the element that has reached the lower bound cannot be further downward adjusted, and then there is no guarantee that a further reshuffling of the tax burden from the young to the middle-aged will not violate one of the IC-constraints faced by the government.

¹²We do this in order to focus attention on the effects of pensions on the IC-constraints faced by the government.

in the first period and then plan to select either the point (y_{11}, b_{11}) or the point (y_{12}, b_{12}) in the second period depending on his realized productivity shock (i.e. he plans to choose in the second period the income points intended for a middle-aged worker who was of type 1 when young).¹³ To relax this IC-constraint, one could supplement (labor- and interest-) income taxation with a set of pension transfers, and implement the following reform: keep y_{11} and y_{12} fixed at the pre-reform values and lower b_{11} and b_{12} by, respectively, $db_{11} < 0$ and $db_{12} < 0$, while at the same time introducing pension transfers $P(y_1, y_{11}) = -[1 + r(1 - \tau_2)]db_{11}$ and $P(y_1, y_{12}) = -[1 + r(1 - \tau_2)]db_{12}$ (while no pension is paid out to agents who earn y_{11} or y_{12} but having earned y_2 in the first period, i.e. $P(y_2, y_{11}) = 0$ and $P(y_2, y_{12}) = 0$). If middle-aged agents are unrestricted in borrowing against future pension entitlements, one can keep replicating the type of reform described above, raising $P(y_1, y_{11})$ and $P(y_1, y_{12})$ while at the same time lowering b_{11} and b_{12} , in a way that is both welfare-neutral for agents of type 1 and welfare-reducing for a type 2 agent contemplating to adopt the assumed deviating strategy.¹⁴

An optimally designed set of pension transfers can then be used, both under an AI labor income tax and under an AD labor income tax, as an instrument to (almost) replicate the allocation that, absent a pension system, would have been implemented by a history-dependent labor income tax system.¹⁵ As a consequence, if the government has the possibility to optimally select the set of transfers $P(y_i, y_{ij})$, moving from an optimally designed AI labor income tax to an optimally designed AD tax might be of second-order importance in terms of welfare gains due to mimicking-detering effects.¹⁶ It is worth noticing, however, that with agents being unrestricted in borrowing against their future pension entitlements, the pension system cannot be used as an instrument to relax binding IC-constraints associated with deviating strategies where a mimicker perfectly replicate the lifetime labor income profile of a non-deviating agent (for instance, the case of a type 2 agent choosing y_1 in first period and either y_{11} or y_{12} in the second period).

¹³In accordance with the notation introduced in Section 3, the income point (y_{11}, b_{11}) denotes the bundle intended for a middle-aged agent who was of type 1 when young and who experienced an unfavorable skill shock in the second period. The income point (y_{12}, b_{12}) denotes the bundle intended for a middle-aged agent who was of type 1 when young and who experienced a favorable skill shock in the second period.

¹⁴The mimicking-detering effect would obviously be even stronger if one were to remove the constraint that $P(y_i, y_{ij}) \geq 0$, so that $P(y_2, y_{11})$ and $P(y_2, y_{12})$ could take large negative values.

¹⁵With unrestricted debt policy at disposal, and if one didn't impose the restriction that $P(y_i, y_{ij}) \geq 0$, an optimally designed set of pension transfers would enable the government, both under an AI labor income tax and under an AD labor income tax, to perfectly replicate the allocation that, absent a pension system, would have been implemented by a history-dependent labor income tax system. The reason is that, without the restriction $P(y_i, y_{ij}) \geq 0$, there would be no limit to the punishment that could be inflicted on an agent who plans to adopt a deviating strategy that does not perfectly match the lifetime labor income profile of a non-deviating agent. With $P(y_i, y_{ij}) \geq 0$, there is instead an implicit upper bound on the penalty. In this case, the allocation implemented by supplementing the nonlinear labor income tax with public pensions might not perfectly replicate the one that could be achieved through an unconstrained history-dependent tax function.

¹⁶The welfare gain of moving from an optimally designed AI labor income tax to an optimally designed AD tax would be nil if one were to remove the restriction that $P(y_i, y_{ij}) \geq 0$. With the constraint $P(y_i, y_{ij}) \geq 0$ in place, there might still be (some small) welfare gains in moving from an AI labor income tax to an AD tax.

Let's now consider the case when agents have limited or no ability to borrow against future pension transfers. In this case the pension system might be less effective in slackening the IC-constraints associated with deviating strategies where a mimicker does not replicate the lifetime labor income profile of a non-deviating agent. The reason is that there is now a limit to the extent that the government can exploit the mimicking-deterring reforms outlined above (reforms consisting in raising the pension transfer and lowering the after-tax labor income when middle-aged) without lowering the utility of a non-deviating agent. In fact, when a non-deviating agent becomes borrowing-constrained in the second period of life, further increasing his pension transfer, while reducing in a present-value-neutral way his after-tax labor income in the second period, makes him worse off.¹⁷ The extent to which this represents a serious hindrance for the government will in general vary according to the specific parameterization of the model considered (distribution of skills in the first and second period, transition probabilities, proportion of different type of agents, etc.).¹⁸

On the other hand, when agents have limited or no ability to borrow against future pension transfers, an optimally designed set of pension transfers may relax IC-constraints that could not be affected if agents were unrestricted in borrowing against their future pension entitlements. This is the case for the IC-constraints associated with lifetime labor income profiles that perfectly mimic those of non-deviating agents. Consider for instance the case of a binding IC-constraint associated with a deviating strategy where a type 2 agent fully mimics a type 1 agent by choosing the income point (y_1, b_1) in first period and either the point (y_{11}, b_{11}) or the point (y_{12}, b_{12}) in the second period. If both a type 1 agent and a type 2 mimicker had the same probabilities of getting an unfavorable skill shock or a favorable one, then their saving behavior in both the first and second period would be identical and it would not be possible to rely on the pension transfers to alleviate this IC-constraint. However, if they face different probabilities of getting a favorable or unfavorable skill shock, then their saving behavior would be different in both the first and the second period. In this case, if in the absence of pensions a type 2 mimicker would save less when middle-aged (for retirement) than a type 1 agent, the government could rely on a proper choice of the pension transfers to alleviate this IC-constraint.¹⁹ The reason is that by providing pensions the government can push a mimicker to a corner

¹⁷Reconsidering the example that we have sketched above, once a type 1 non-deviating agent becomes borrowing constrained in the second period, one can no longer keep raising $P(y_1, y_{11})$ and $P(y_1, y_{12})$ while at the same time lowering b_{11} and b_{12} in a way that is welfare-neutral for him.

¹⁸Moreover, if one didn't impose the restriction that $P(y_i, y_{ij}) \geq 0$, the government would have another instrument to impose an indefinitely large penalty on the mimicker without affecting the utility of non-deviating agents. Thus, without the restriction that $P(y_i, y_{ij}) \geq 0$, it would still be true, even when agents have limited or no ability to borrow against future pension entitlements, that the government could implement the allocation that, absent a pension system, would have been achieved by a history-dependent labor income tax system.

¹⁹If instead, at the optimum without pensions, a type 2 mimicker would like to save more when middle-aged (for retirement) than a type 1 agent, pension transfers could not be used as an instrument to alleviate this IC-constraint.

solution where he is forced to under-consume when middle-aged and over-consume during retirement.²⁰

To summarize, when agents have limited or no ability to borrow against future pension entitlements, an optimally designed set of pension transfers may relax IC-constraints that could not have been affected otherwise. However, this requires that, at an optimum in the absence of pensions, middle-aged mimickers would like to save less than the agent they mimic. In turn, this condition is more likely to be met when the probability to get a favorable skill shock in the second period is higher for the potential mimicker than for the agent being mimicked.

As a final remark, one should always take into account that, when the government faces constraints on the management of debt policy, the mimicking-deterring effects that were highlighted above will in the end need to be traded off against the adverse capital accumulation effects generated by the PAYG system.

Having discussed the channels by which a PAYG system can affect the government's problem, we are now ready to present some results based on numerical simulations aimed to illustrate the potential welfare gains that could be obtained by using an optimally designed set of pension transfers as an additional policy instrument. In performing the simulations we will make the assumption that agents cannot borrow against their future pension transfers.²¹

5 Calibration

5.1 Parameterization

We use the same parameterization and skill process that was used in Bastani et al., 2013. Each period corresponds to 20 years. Based on an annual depreciation rate of 8% and a population growth rate of 1.1%, we calculate the 20 year analogues which yields $n = (1.011)^{20} - 1$ and $\delta = 1 - 0.92^{20}$. The individual rate of time preference β is set to $\beta = 0.988^{20}$. Production is Cobb-Douglas, i.e. $F(K_t, L_t) = AK_t^\alpha L_t^{1-\alpha}$, and the share of capital in production is $\alpha = 1/3$. The production scale factor A is chosen so that the equilibrium

²⁰Starting from an initial equilibrium without pensions and where the nonlinear income tax is optimally chosen, this can be done by implementing the following reform. Keep y_{11} and y_{12} fixed at the pre-reform value and lower b_{11} and b_{12} by, respectively, $-c_{11}^R/[1+r(1-\tau_2)] < 0$ and $-c_{12}^R/[1+r(1-\tau_2)] < 0$, where c_{11}^R and c_{12}^R denote the consumption during retirement, at the income tax optimum without pensions, of, respectively, a type 1 agent who was hit by an unfavorable skill shock when middle-aged, and a type 1 agent who was hit by a favorable skill shock when middle-aged. At the same time, introduce pension transfers $P(y_1, y_{11}) = c_{11}^R$ and $P(y_1, y_{12}) = c_{12}^R$. Clearly, the expected well-being of a young type agent who does not plan to enact a deviating strategy would be unaffected by the reform (he would no longer save when middle-aged, but he would continue to work and consume in all periods and state of the world as in the pre-reform equilibrium). But suppose that at the pre-reform equilibrium with no pensions a type 2 mimicker, either when choosing in the second period the point (y_{11}, b_{11}) or the point (y_{12}, b_{12}) , was saving less for retirement than the corresponding non-deviating agent choosing the same income point. Under the assumption that agents are not free to borrow against their future pension entitlements, the reform would make the mimicker worse-off by preventing him from achieving his preferred intertemporal consumption profile.

²¹This seems the most realistic case to consider. Moreover, we will also maintain the assumption that $P(y_i, y_{ij}) \geq 0$.

rental price for one efficiency unit of labor is equal to one when the golden rule is satisfied. Agents maximize their expected lifetime utility given an instantaneous utility function defined as $u(c, \ell) = \frac{c^{1-\gamma}}{1-\gamma} - \frac{\ell^\kappa}{\kappa}$ with $\kappa = 3$ (implying a Frisch elasticity of 0.5) and $\gamma = 2$.

5.2 Skill process

The skill process consists of three parts, the skill distributions in the first- and second periods, and the transition probabilities linking these distributions.

We infer the distribution of skills from the distribution of market wage rates. Since we focus on a zero-productivity-growth stationary state, we calibrate the skill distribution for young and middle-aged agents using wage distributions obtained from the same year t . In particular, we use the 2003 wave of the Current Population Survey (CPS) to obtain the skill distribution for both young and middle-aged.

To capture how agents' skills evolve over the life-cycle it is necessary to adopt a procedure which is based on wage distributions taken from two separate periods in time. Because our period length is 20 years and we are interested in how individuals' wages change over time, we need a panel data set spanning a long period of time and for this reason we use the National Longitudinal Study of Youth (NLSY79) to retrieve the transition probabilities. Since our model abstracts from productivity growth, we are likely to overestimate the increase in skills which occurs over time by computing transition probabilities based on cohort-specific wage distributions in year t and $t + 20$. We deal with this problem by looking at how agents change over time their *relative position* in their cohort-specific wage distribution.²² We focus on agents who were 25 years old during the period 1982-1988 and determine how they move in the wage distribution by calculating their wages 20 years later during the period 2002-2008 (when they are 45). Since the number of IC-constraints severely limits the number of types we can handle in our numerical simulations, we divide the wage distribution for the young into three groups and the wage distribution for the middle-aged into four groups. We also restrict some of the transition probabilities by assuming $\pi_{ij} > 0$ for $0 \leq j - i \leq 1$.

The wage rates and the transition matrix are given in Table 1.

²²That is, in both the first- and second period each individual's percentile position in the wage distribution is calculated. We then see how individuals move between percentiles in the wage distribution over time.

Table 1: Transition matrix and hourly wages.¹

Young	Middle-aged				Wages
	1	2	3	4	
1	0.41	0.59	0	0	8.19
2	0	0.43	0.57	0	12.38
3	0	0	0.59	0.41	19.58
Wages	9.67	15.42	21.40	34.75	

¹ Wages expressed in 2003 US Dollars.

The numbers in the matrix show the proportion of agents belonging to a given group in the first period that move to a particular group in the second period. However, these numbers are consistent with different assumptions about how individuals perceive their wage process. The assumption that we make is that individuals perceive these proportions as probabilities. Thus, if a proportion p_{ij} moves from group i in first period to group j in second period, the assumption would be that individuals who belong to group i in first period regard π_{ij} as the probability that they will end up in group j in the second period.

6 Results

The results that we are going to present are obtained assuming that the government maximizes either an *ex-ante* utilitarian social welfare function or an *ex-ante* max-min social welfare function.²³ To obtain a revenue-based measure of the welfare gains attainable under different assumptions about the set of policy instruments available to the government, we consider an equivalent-variation-type of welfare gain measure, taking as a benchmark the solution to the government’s problem under an AI nonlinear labor income tax with zero public debt, no pensions, and an optimally chosen (AI) interest income tax rate. More precisely, we proceed as follows. We first calculate the minimum amount of extra revenue that should be injected into the government’s budget, in the benchmark optimal AI tax problem, in order to achieve the same social welfare level as under a different policy regime. Once we have found this minimum amount of extra revenue, we divide it by the aggregate GDP at the AI tax benchmark to get a revenue-based measure of the welfare gains. All results are displayed in Tables 2-5.

Let’s start with the results obtained for a utilitarian SWF (see fifth column of Table 2 and Table 3).²⁴

²³In the first case the government maximizes the sum of the lifetime expected utilities of the three different types of young agents; in the second case the government maximizes the expected lifetime utility of the least skilled young agents.

²⁴In Tables 3 and 4, column “Pension transfers”, P_{i1} (with $i = 1, 2, 3$) denotes the pension paid to an agent who was of skill type i when young and who experienced an unfavorable skill shock when middle-aged; P_{i2} (with $i = 1, 2, 3$) denotes instead the pension paid to an agent who was of skill type i when young and who experienced a favorable skill shock when middle-aged. The implicit rates of return are calculated based on the equilibrium value of the contribution rate α , on the social security contributions paid

Table 2: Overview of welfare gains

Labor Tax	Debt	Capital Tax	Pensions	WG Utilitarian	WG Max-min
AI	Zero	AI	No	Benchmark	Benchmark
AI	Zero	AI	Yes	2.48%	4.94%
AI	Free	AI	No	1.44%	2.24%
AI	Free	AI	Yes	6.17%	8.26%
AD	Zero	AI	No	2.19%	3.35%
AD	Zero	AI	Yes	6.18%	8.27%
AD	Free	AI	Yes	6.18%	8.31%

Table 3: Utilitarian SWF with pensions

Instruments		Pension transfers	Implicit rate of return
AI $T, d = 0, AI \tau$	$r=79.28\%$	P11=0.10	-61%
	$\tau=93.81\%$	P12=2.19	190%
	$\alpha=4.62\%$	P21=0.41	-22%
		P22=0.48	-20%
		P31=0.03	-93%
		P32=2.24	60%
AI T, d free, AI τ	$r=24.46\%$	P11=0.11	-83%
	$\tau=16.22\%$	P12=2.27	33%
	$\alpha=12.17\%$	P21=1.69	-3%
		P22=6.06	101%
		P31=2.55	-7%
		P32=4.91	18%
AD $T, d = 0, AI \tau$	$r=24.46\%$	P11=0.01	-97%
	$\tau=21.32\%$	P12=1.90	30%
	$\alpha=10.52\%$	P21=2.99	58%
		P22=6.83	140%
		P31=0.12	-91%
		P32=3.63	5%

First of all, we can see that, when the nonlinear labor income tax is AI, public debt is fixed at zero, and the interest income tax is optimally chosen, allowing the government to optimally choose the set of pension transfers generates a sizeable welfare gain, equivalent to about 2.48% of the aggregate GDP under the benchmark tax system. It is worth noticing that such a large welfare gain is obtained with quite small pension transfers: the equilibrium contribution rate is just 4.62%.²⁵ The reason why the optimal pension transfers are in this case so low is the fact that, without access to an unrestricted debt policy (in particular, by the different types of agents. For a given labor earning trajectory and corresponding pension entitlement, the implicit rate of return IRR has been calculated by solving the following type of equation: $\alpha [y_i (1 + IRR)^2 + y_{ij} (1 + IRR)] = P(y_i, y_{ij})$.

²⁵The share of consumption during retirement that is financed through public pensions ranges from a minimum of 0.5% (for type 3 agents who experienced an unfavorable skill shock when middle-aged) to a maximum of 45.9% (for type 1 agents who experienced a favorable skill shock when middle-aged).

a negative public debt), making the pension transfers more generous entails a trade-off between relaxing IC-constraints and moving the economy even further away from the golden rule (in the absence of pensions).²⁶ Finally, we can also see that, based on the computed equilibrium contribution rate, the implicit rate of returns on social security contributions vary a lot depending on the realized lifetime labor earning trajectory of an agent, and tend to be higher for agents having experienced a favorable skill shock when middle-aged.

If, on top of allowing the government to optimally choose the set of pension transfers, we also assume that the government faces no constraints in managing public debt policy, we can see that the welfare gain, compared to the benchmark system, amounts to 4.73% (calculated from Table 2, column “WG Utilitarian”, as the difference between 6.17% and 1.44%). In this case, pensions can be used for the sole purpose of slackening binding IC-constraints, without the need to worry about adverse consequences on capital accumulation (the golden rule condition is satisfied, with the reduction in private savings being more than compensated by the increase in public saving, i.e. negative public debt).²⁷ The equilibrium contribution rate increases to 12.17%;²⁸ once again, the implicit rate of returns on social security contributions, even though they increase compared with the case when public debt was fixed at zero, vary a lot and are always larger for agents having experienced a favorable skill shock when middle-aged.

One can also notice that, once the government is unrestricted in its use of debt policy and pensions transfers are optimally chosen, the welfare gains that could be achieved by moving from an optimal AI nonlinear labor income tax to an optimal AD income tax would be virtually nil. An optimal AD nonlinear income tax, coupled with optimally chosen pensions and an optimal AI interest income tax, would deliver a welfare gain of 6.18% over the benchmark system, implying a 0.01% welfare gain in moving from an AI labor income tax to an AD labor income tax. It is also worth noticing that for a utilitarian SWF, when the nonlinear labor income tax is AD, it does not matter whether or not the government has access to debt policy.²⁹ This implies that, at least for the case when the government has no access to public debt policy, there remain large welfare gains (3.7%, calculated as the difference between 6.18% and 2.48% in the “WG Utilitarian” column of Table 2), even with pensions being used as an additional policy instrument,

²⁶Under the golden rule condition, the equilibrium interest rate is equal to the population growth rate, which in our setting is about 24%. In the benchmark AI system without pensions and with no public debt, the equilibrium interest rate is 63%. With an optimally chosen set of pension transfers, the equilibrium interest rate is raised to 79% and the capital/labor ratio falls from 30% to 25%.

²⁷Under golden-rule, the capital/labor ratio is 47.48%.

²⁸The maximum share of consumption during retirement financed through public pensions increases to 92% (for type 2 agents who experienced a favorable skill shock when middle-aged).

²⁹As discussed in Section 4, this result, that was shown to generally hold in a setting without pensions by Bastani et al. (2013), does not necessarily generalize to the case when pensions are used as an additional policy tool. However, at least for the case of a utilitarian SWF, our simulations show that, even when optimally choosing pensions and having no access to public debt policy, a nonlinear AD labor income tax enables the government to move the economy to the golden rule.

in moving from an optimal AI labor income tax to an optimal AD labor income tax. Regarding the set of optimal pension transfers under a nonlinear AD income tax, it resembles fairly close the one obtained under a nonlinear AI income tax with unrestricted debt policy.³⁰ Finally, when the nonlinear income tax is AD, the welfare gains of optimally choosing pensions are quite large, being close to 4% (calculated as the difference between 6.18% and 2.19% in the “WG Utilitarian” column of Table 2). This number is larger than the corresponding gain (2.48%) calculated under a nonlinear AI income tax with debt restricted to zero. The reason is that, as emphasized above, under a utilitarian SWF debt policy is redundant under a nonlinear AD tax both when pensions are used and when they are not (in both cases, the economy reaches the golden-rule). On the contrary, under a nonlinear AI tax, a zero public debt constraint represents a hindrance for the government due to the trade-off that arises between the pursuit of redistributive goals and the achievement of the golden-rule. In this case, if pensions provide a source of welfare gain due to the mimicking-deterring effects that they exert, they also move the economy further away from the golden rule.

Let’s now briefly consider how results change under a max-min SWF. The main difference pertains to the case of a nonlinear AD income tax. It is now no longer true that the availability of debt policy is irrelevant when the nonlinear labor income tax is AD and pension transfers are provided by the government. Specifically, when public debt is fixed at zero and pensions are optimally chosen, the economy does not reach the golden rule even if the nonlinear income tax is AD.³¹ However, even though this implies that, when debt policy is unconstrained, pensions are more generous and most agents only rely on pensions to finance consumption during retirement, the welfare gain descending from the availability of debt policy is negligible, about 0.04% (calculated as the difference between 8.31% and 8.27% in the last column of Table 2). Compared with the case when pensions are not provided, under a nonlinear AD income tax the welfare gain of using pensions amounts to about 5%.

Once again, with the government being unrestricted in its use of debt policy and pensions transfers being optimally chosen, the welfare gains that could be achieved by moving from an optimal AI nonlinear labor income tax to an optimal AD income tax would be virtually nil (0.05%, calculated as the difference between 8.31% and 8.26% in the last column of Table 2). If instead debt were to be restricted to zero, there would still be significant welfare gains, 3.3% (calculated as the difference between 8.27% and 4.94% in the last column of Table 2), in moving from an AI income tax (with pensions) to an AD income tax (with pensions).

³⁰The equilibrium contribution rate is 10.52%, and the maximum share of consumption during retirement financed through public pensions increases to 100% (for type 2 agents who experienced a favorable skill shock). The implicit rate of returns on social security contributions are always larger for agents having experienced a favorable skill shock when middle-aged.

³¹The equilibrium interest rate is 27.1%, slightly higher than the one prevailing under golden-rule (24.46%).

Comparing Table 3 and Table 4 we can also see that another difference between the utilitarian and the max-min case is that, whereas in the former case pensions tend to mostly favor those agents who were of the intermediate skill type when young (at least when the solution to the government’s problem is such that the economy reaches the golden-rule), in the latter case pensions definitely favor those agents who were of the lowest skill type when young.

Table 4: Max-min SWF with pensions

Instruments		Pension transfers	Implicit rate of return
AI T , $d = 0$, AI τ	$r=84.36\%$	P11=0.41	5%
	$\tau=89.87\%$	P12=0.92	36%
	$\alpha=6.72\%$	P21=2.87	126%
		P22=3.23	114%
		P31=0.23	-78%
		P32=1.02	-47%
AI T , d free, AI τ	$r=24.46\%$	P11=1.20	74%
	$\tau=14.64\%$	P12=2.33	112%
	$\alpha=9.23\%$	P21=1.25	5%
		P22=4.22	107%
		P31=0.45	-71%
		P32=3.21	2%
AD T , $d = 0$, AI τ	$r=27.09\%$	P11=0	-100%
	$\tau=46.65\%$	P12=0.92	-100%
	$\alpha=14.91\%$	P21=5.01	96%
		P22=6.04	91%
		P31=1.68	-42%
		P32=8.86	51%
AD T , d free, AI τ	$r=24.46\%$	P11=4.64	126%
	$\tau=34.37\%$	P12=5.20	94%
	$\alpha=23.39\%$	P21=1.55	-37%
		P22=6.09	38%
		P31=6.52	8%
		P32=7.91	0%

Before concluding, we would like to briefly discuss two further issues, one related to the welfare effects of interest income taxation and the other to the relative merits of taxing private savings according to a TtE scheme or a EEt scheme.³² Regarding the first issue, we have considered the solution to the government’s problem, with and without pensions, under various assumptions regarding the tax treatment of interest income: no taxation of interest income, AI proportional taxation of interest income, AD proportional taxation of interest income. Our results indicate that the only cases where there would be some non-negligible welfare

³²A TtE scheme involves saving out of taxed income, taxing on an accrual basis the fund’s investment return and tax free withdrawal of pension benefits. Under an EEt scheme, savings are deductible against the labor income tax and are then taxed, together with the earned interest, in the period when the funds are used for consumption purposes.

gains in moving from an AI interest income tax to an AD interest income tax are when the nonlinear labor income tax is AI and debt policy is unavailable (see Table 5). In these cases we find that, under a utilitarian SWF, the welfare gains of moving from an AI interest income tax to an AD interest income tax are about 0.4% when pensions are not provided and about 0.7% when pensions are optimally chosen; under a max-min SWF, the corresponding figures are 0.1% and 0.3%. In each of these cases, an optimal AD interest income tax would require a higher tax rate on the savings of middle-aged workers.

Table 5: Optimal interest income tax rates and welfare gains

T	d	τ	Pensions	Utilitarian		Max-min	
				τ	Welfare gain	τ	Welfare gain
AI	Zero	Zero	No	$\tau_1 = \tau_2 = 0$	-2.91%	$\tau_1 = \tau_2 = 0$	-1.20%
AI	Zero	AI	No	$\tau_1 = \tau_2 = 73\%$	–	$\tau_1 = \tau_2 = 84\%$	–
AI	Zero	AD	No	$\tau_1 = 51\%, \tau_2 = 82\%$	0.4%	$\tau_1 = 56\%, \tau_2 = 96\%$	0.1%
AI	Zero	Zero	Yes	$\tau_1 = \tau_2 = 0$	1.11%	$\tau_1 = \tau_2 = 0$	2.9%
AI	Zero	AI	Yes	$\tau_1 = \tau_2 = 81\%$	2.48%	$\tau_1 = \tau_2 = 89\%$	3.64%
AI	Zero	AD	Yes	$\tau_1 = 5\%, \tau_2 = 91\%$	2.55%	$\tau_1 = 11\%, \tau_2 = 97\%$	3.67%

Regarding the second issue, i.e. the relative merits of taxing private savings according to a TtE scheme or a EEt scheme, the first thing to notice is that throughout our analysis we have maintained the informational assumption that savings transactions could only be observed anonymously, not at an individual level, thus preventing the tax authority from taxing the returns on savings nonlinearly.³³ Consistently with this informational assumption, we have only considered tax schemes of the TtE -type. However, if one were to assume that private savings are publicly observable at an individual level, another possibility would be to tax them according to a consumption-type EEt scheme.³⁴ The question then arises whether, being informationally more demanding, an EEt scheme welfare dominates a TtE scheme. To answer this question we have also computed the solution to the government’s problem assuming that private savings are publicly observable and subject to an EEt scheme where they are taxed at the consumption stage, together with the earned interests, according to an optimally chosen proportional tax rate. The results that we have obtained show that, in the absence of public pensions, the EEt scheme is better than the TtE scheme either when the nonlinear labor income tax is AI and the social welfare function is utilitarian, or when the nonlinear labor income tax is AD. Instead, a TtE scheme welfare dominates an EEt scheme when either the government uses public pensions as a supplementary policy tool or when, in the absence of public pensions, the nonlinear

³³This assumption was relaxed when considering the possibility for the government to implement an AD interest income tax.

³⁴One should also notice that, when private savings are publicly observable at an individual level, the government could do better than using the available information to implement an EEt scheme. It could implement a nonlinear AD tax function that depends, in a non-separable way, both on labor income and on interest income.

labor income tax is AI and the social welfare function is max-min.

7 Concluding remarks

Using an OLG model with skill uncertainty and private savings, we have investigated whether an optimally designed set of public pension transfers can usefully supplement a nonlinear labor income tax as a welfare-enhancing policy instrument in a Mirrleesian setting where agents' skills are private information. We have highlighted that, even though pensions, by crowding out private savings, adversely affect the achievement of the golden-rule, they can be used as a mimicking-detering device that makes it easier for the government to achieve the desired redistributive goals. We have also emphasized that the welfare effects of pensions depend, in general, on the assumptions that are made about the possibility for agents to borrow against their future pension entitlements. When agents are free to borrow against their future pension entitlements, public pensions can be used as an instrument to (almost) replicate the optimal allocation that would have been implemented had the labor income tax schedule been allowed to be history-dependent. When instead agents cannot borrow against their future pension entitlements, this might no longer be the case but, on the other hand, an optimally designed set of pension transfers can relax some IC-constraints that could not have been affected if agents were unrestricted in borrowing against their future pension entitlements. In particular, the pension system can be designed to relax those binding IC-constraints where mimickers plan to save less for retirement than the agents that they plan to mimic.

To shed light on the potential magnitude of the welfare gains descending from using public pensions as an additional policy tool, and assuming that agents cannot borrow against their future pension transfers, we have conducted some numerical simulations based on U.S. data.

The results confirm that, if the government faces no constraints in the management of debt policy, using pensions as a policy instrument implies that the welfare gains that could be achieved by moving from an optimal age-independent nonlinear labor income tax to an optimal age-dependent income tax are virtually nil. If instead the government has no access to debt policy, there would still remain significant welfare gains, even with pensions being used as an additional policy instrument, in moving from an age-independent labor income tax to an age-dependent one. Depending on the social welfare function being maximized, the welfare gains descending from supplementing a nonlinear labor income tax with an optimal set of pension transfers range between 2.5% and 6%, in the case of a nonlinear age-independent labor income tax, and between 4% and 5% in the case of a nonlinear age-dependent labor income tax.

To conclude, we would like to emphasize some of the many limitations of our analysis. Apart from considering only a very limited number of types in our numerical simulations, we have restricted our attention to steady-state equilibria and we have neglected the possibility for agents to decide when to retire based on the incentives created by the pension system.³⁵ We have assumed that the government can commit to its future tax and pension policy. Moreover, we have also assumed that all agents are fully rational. For example, myopia is regarded as a standard justification for public pension systems and might also influence the effectiveness of public pensions as a mimicking-deterring device, especially if the degree of myopia is heterogeneous across agents.³⁶ Finally, we have implicitly assumed that agents cannot misreport their true earned income to the tax authority; if misreporting were possible, the scope for using pension transfers as a mimicking-deterring device would be weakened.

References

- [1] Akerlof, G. A. (1978) The economics of tagging as applied to the optimal income tax, welfare programs, and manpower planning. *American Economic Review* **68**, 8-19.
- [2] Andersen, T. and Bhattacharya, J. (2011) On myopia as rationale for social security. *Economic Theory* **47**, 135-158.
- [3] Banks, J. and Diamond, P. (2010) The base for direct taxation. In *Dimensions of Tax Design, The Mirrlees Review*, Oxford: Oxford University Press.
- [4] Bastani, S. (2013) Gender-Based and Couple-Based Taxation, *International Tax and Public Finance* **20**, Issue 4, 653-686
- [5] Bastani, S., Blomquist, S. and Micheletto, L. (2013) The welfare gains of age-related optimal income taxation. *International Economic Review* **54**, 1219-1249.
- [6] Blomquist, S. and Micheletto, L. (2008) Age-related optimal income taxation. *Scandinavian Journal of Economics* **110**, 45-71.
- [7] Boskin, M. J. and Sheshinski, E. (1983) Optimal tax treatment of the family: married couples. *Journal of Public Economics* **20**, 281-297.
- [8] Choné, P. and Laroque, G. (2014) Income taxes and retirement schemes. *CREST Working Paper* 2014-07.

³⁵For an optimal tax model where the age of retirement is a choice variable for agents (but where for each agent the labor supply is constant across periods during the working years), see Cremer et al. (2004). See also the discussion contained in Cremer and Pestieau (2016).

³⁶Myopia as a rationale for the existence of social security systems has been widely analyzed in the literature. See, among many contributions, Feldstein (1985), Tenhunen and Tuomala (2010), Andersen and Bhattacharya (2011), and Cremer and Pestieau (2011).

- [9] Cremer, H., Gahvari, F. and Lozachmeur, J.M. (2010) Tagging and income taxation: theory and an application. *American Economic Journal: Economic Policy* **2**, 31-50.
- [10] Cremer, H., Lozachmeur, J.M. and Pestieau, P. (2004) Social security, retirement age and optimal income taxation. *Journal of Public Economics* **88**, 2259-2281.
- [11] Cremer, H. and Pestieau, P. (2011) Myopia, redistribution and pensions. *European Economic Review* **55**, 165-175.
- [12] Cremer, H. and Pestieau, P. (2016) Taxing pensions. *CESifo Working Paper No. 5930* .
- [13] Diamond, P. and Saez, E. (2011) The case for a progressive tax: from basic research to policy recommendations. *Journal of Economic Perspectives* **25**, 165-190.
- [14] Diamond, P. (2009) Taxes and pensions. *Southern Economic Journal* **76** (1), 2-15.
- [15] Farhi, E. and Werning, I. (2013) Insurance and taxation over the life cycle. *Review of Economic Studies* **80**, 596-635.
- [16] Feldstein, M. (1985) The optimal level of social security benefits. *Quarterly Journal of Economics* **100**, 303-320.
- [17] Golosov, M. and Tsyvinski, A. (2015) Policy implications of dynamic public finance. *Annual Review of Economics* **7**, 147-171.
- [18] Golosov, M., Tsyvinski, A. and Werning, I. (2006) New dynamic public finance: a user's guide. In *NBER Macroeconomics Annual 2006*, MIT Press.
- [19] Immonen, R., Kanbur, R., Keen, M. and Tuomala, M. (1998) Tagging and taxing: the use of categorical and income information in designing tax/transfer schemes. *Economica* **65**, 179-192.
- [20] Kanbur, R. and Tuomala, M. (2016) Groupings and the gains from tagging. *Research in Economics* **70**, 53-63.
- [21] Kocherlakota, N. (2010) *The new dynamic public finance*. Princeton University Press.
- [22] Mankiw, N. G. and Weinzierl, M. (2010) The optimal taxation of height: a case study of utilitarian income redistribution. *American Economic Journal: Economic Policy* **2**, 155-176.
- [23] Shourideh, A. and Troshkin, M. (2015) Incentives and efficiency of pension systems. Mimeo.
- [24] Tenhunen, S. and Tuomala, M. (2010) On optimal lifetime redistribution policy. *Journal of Public Economic Theory* **12**, 171-198.
- [25] Weinzierl M. (2011) The surprising power of age-dependent taxes. *Review of Economic Studies* **78**, 1490-1518.