International Tax and Public Finance, 12, 7–28, 2005

Voting on Pensions with Endogenous Retirement Age*

GEORGES CASAMATTA georges.casamatta@univ-tlse1.fr
GREMAQ-CNRS, Université de Toulouse, 21 allée de Brienne, 31000 Toulouse, France

HELMUTH CREMER mut@cict.fr
GREMAQ and IDEI, Université de Toulouse, 21 allée de Brienne, 31000 Toulouse, France

PIERRE PESTIEAU P.Pestieau@ulg.ac.be
CREPP, Université de Liège, CORE, Delta and CEPR, 7 Boulevard du Rectorat, 4000 Liège, Belgium

Abstract

It is often argued that the observed trend towards early retirement is due mainly to the implicit tax imposed on continued activity of elderly workers. We study the relevance of such a distortion in a political economy model with endogenous age of retirement. The setting is a two-period overlapping generations model. Individuals differ in their productivity. In the first period they work a fixed amount of time; in the second, they choose when to retire and then receive a flat rate pension benefit. Pensions are financed by a payroll tax on earnings in the first and in the second period of life. Such a tax is non distortionary in the first period; it is distortionary in the second period. We allow for some rebating of the second period tax. Individuals vote on the level of the payroll tax given the rebate which can range from 0 (biased system) to 100% (neutral system). We provide sufficient conditions for the existence of a voting equilibrium and study its properties. Under these conditions, high tax rates are supported by all the old and by low productivity young individuals. We show that the pivotal voter is a young individual. The number of *young* individuals who have higher wage than the pivotal voter equals half the *total* population. We also show that the introduction of a bias increases the political support for the pension system. Finally, we study the simultaneous determination of the bias and the tax rate through a voting procedure and show that the equilibrium (if any) implies a bias which is always positive and may or not be larger than one.

Keywords: social security, retirement age, majority voting

JEL Code: H55, O41, O9

1. Introduction

Over the last forty years, labor force participation of the elderly has been dramatically decreasing in almost all industrialized countries. Participation rates for men aged 60 to 64

*This paper has been presented at CES/IFO workshop on pensions, Munich; Public economics workshop, GREQAM, Marseille; IIPF Conference, Linz; Recent advances in political economy, DELTA, Paris; CEPR workshop Dynamic Aspects of Policy Reforms, Vienna and in seminars at Toulouse and the Balearic Islands. We thank all the participants for their helpful comments. In particular, we are grateful to Mathias Kifmann, Eytan Sheshinski, Bertrand Wigniolle and especially to Vincenzo Galasso, for their very insightful discussions. Last but not least, we thank the referee and the editor, Jay Wilson, for their detailed and constructive remarks.

were above 70% in the early 60s; they have fallen to 57% in Sweden and to below 20% in Belgium, France, Italy and the Netherlands by the mid 90s.[1] At the same time, people are living longer and longer. In the European Union, life expectancy at age 65 has increased by more than one year per decade since 1950. This puts an enormous pressure on the financial viability of Pay-As-You-Go (PAYG) pension systems and the situation will become even more problematic when the "baby boomers" will come to retirement.

Gruber and Wise (1999) attribute this large decline in the labor force participation to the incentives created by social security systems. Continued work at later ages may be subject to two burdens: the traditional payroll tax which is nowhere age-dependent and forgone benefits when social security wealth decreases with the age of retirement. This double burden which Gruber and Wise call an implicit tax represents an allocative distortion that induces early retirement. As they show it is higher in France, Belgium and Italy than in Japan, Sweden or the US. Average retirement is also much earlier in these first countries than in the second.

Why do we have such an implicit tax on continued activity? On pure efficiency grounds, we would like to avoid any distortion in the labor retirement choice of aged workers and let them choose the age of retirement such that the marginal utility of retiring is equal to the worker productivity times marginal utility of consumption. This is the case in the *laissez-faire* equilibrium when there is no pension system and no taxes. More generally, the pension system could be designed to preserve this first-best trade-off and we can then think of it as "neutral." However, Cremer, Lozachmeur and Pestieau (2004) show that when the design of the pension system reflects some redistributive concern and when non-uniform lump-sum transfers are not available (because individual characteristics like productivity or health status are not publicly observable) a neutral system is no longer desirable. The optimal system is then "biased" as it implies an implicit tax on continued activity which creates downward distortions.

In this paper we consider a setting where the optimal second-best pension system also implies a bias. We do characterize this second-best solution as a benchmark, but our main focus is positive rather than normative. We study the implications of a bias on the political support for pension systems. Specifically, we characterize the equilibrium (majority voting) size of the pension system for a given bias in the benefit formula. Finally, we also study the endogenous determination of the bias through a specific voting procedure.

To shed light on these questions, we consider a model where each individual lives two periods. He works one unit of time in the first, pays a proportional tax to the pension system and saves. In the second period he works for some time and then retires. His consumption then is financed by disposable earnings, gross returns of savings and a flat pension benefit. Individuals differ according to age, they are young or old, and according to productivity. In the second period, only a fraction of the payroll tax is levied on elderly workers, the remaining being implicitly rebated. Given such a fraction which implies a distortion, a bias, on the retirement decision, young and old vote on their preferred payroll tax rate.

We provide sufficient conditions for the existence of a voting equilibrium and study its properties. Under these conditions, high tax rates are supported by all the old and by low productivity young individuals. We show that the pivotal voter is a young individual: the number of *young* individuals who have a higher wage than the pivotal voter equals half the *total* population. It also appears that the critical level of productivity below which the young favor a strictly positive payroll tax increases when a bias is introduced. *Consequently, the introduction of a bias increases the political support for the pension system.* Finally, we study the simultaneous determination of the bias and the tax rate through a voting procedure and show that the equilibrium (if any) implies a bias which is always positive and may or not be larger than one.

Before turning to these positive considerations, we briefly discuss the second-best solution where (linear) taxes and the bias are chosen to maximize a utilitarian social welfare function (Section 4). We show that the optimal system is also biased, but that the bias should be less than 100%, at least under some plausible conditions. Intuitively, the optimal implicit tax in the second period is determined very much like Sheshinski's optimal linear income tax; i.e. by balancing redistributive benefits against distortions. Redistributive benefits arise because the implicit tax is proportional to income while the pension is flat. The distortion is related to the decrease in retirement age. Like in Sheshinski's setting the optimal (implicit) tax is positive which implies that a full rebate of the second period tax is not called for. Put differently, there is a bias. However, under realistic assumptions on the range of wage inequalities, the distortion on retirement age leads to a lower tax in the second period than in the first one. Consequently, the optimal bias can be expected to be less then 100%.

The effect of the introduction of a pension system on the retirement decision has been studied by Sheshinski (1978) and Crawford and Lilien (1981). These authors argue that when there are no borrowing constraints an "actuarially fair" pension system (benefits equal contributions) does not affect the retirement decision. This is because private savings are simply replaced by public pensions. The introduction of a pension system which is not marginally fair leads to a decrease in the price of leisure with respect to consumption.[2] If the substitution effect dominates the income effect, it induces people to retire earlier.[3]

To our knowledge, only two papers deal with the retirement decision in a political economy environment. Lacomba and Lagos (1999) study the problem of a direct vote on the (mandatory) retirement age. More closely related to our study, Conde Ruiz and Galasso (2003, 2004) develop a model in which the vote takes place simultaneously on the payroll tax rate and on the decision on whether to introduce an early retirement provision. They show that the early retirement provision may be sustained at equilibrium by a coalition of the poor workers, who want to retire early, and old people with incomplete earnings history, who would receive no pension without this provision. This analysis and ours can be considered as complementary. Indeed, we do not investigate the issue of introducing an early retirement age.

2. The Model

Individuals live for two periods and they are differentiated according to their wage level per unit of time (productivity). The distribution of productivities has support $[w_-, w_+]$, density function $f(\cdot)$, and cumulative distribution function $F(\cdot)$. We assume that the median productivity, w_m, is lower than the mean, $\bar{w}$. The intertemporal utility function is:

$$U(c, d) = u(c) + \beta u(d),$$

where c is the first period consumption and d is the second period consumption; β is a factor of time preference, which is, by assumption, equal to $1/(1+r)$ where r is the interest rate. The utility function $u(\cdot)$ is increasing and concave: $u'(\cdot) > 0$, $u''(\cdot) < 0$. Moreover, we assume that $\lim_{c\to 0} u'(c) = +\infty$ and that the coefficient of relative risk aversion is not greater than 1: $R_r(c) = -cu''(c)/u'(c) \leq 1$. Second period consumption, d, is to be distinguished from overall spending in the second period, x. Overall spending includes consumption plus the monetary disutility of $z \in [0, 1]$, which is the fraction of the period the individual continues to work. This variable is interpreted as the retirement age. We assume a quadratic specification for the disutility of work, so that $d = x - \gamma z^2/2$. The parameter γ specifies the intensity of this disutility.

First and second periods are of equal length, normalized to 1. Labor supply is assumed to be inelastic in the first period. In the second period, individuals decide which fraction of time, z, they spend working.[4] Observe that with our specific form of the labor disutility function, there are no income effects in labor supply decisions which thus depends only on the relative price of leisure and consumption.

First and second period consumption for an individual with productivity w are respectively given by:

$$c = w(1-\tau) - s$$
$$x = s(1+r) + wz(1-\theta\tau) + P,$$

where $\tau \in [0, 1]$ is the payroll tax rate[5] and $s \geq 0$ is the amount of savings; P corresponds to the *total* pension received and, by assumption, *does not depend on* z. The parameter $\theta \in [0, 1]$ measures the bias of the pension system. We define a *neutral system* as a system that does not distort individual decisions concerning retirement age. In other words, it does not modify the relative price of leisure and consumption, compared to the situation with no pension scheme. In a neutral system, the marginal benefit of working one more year is then w. This is the case in our setting when $\theta = 0$. When $\theta > 0$, the relative price of leisure and consumption becomes $w(1-\theta\tau)$. Consumption is therefore more expensive and individuals are induced to retire earlier.[6]

Note that P does not depend on w; the pension system is Beveridgean. Everyone contributes for an amount proportional to his labor income but the benefit received does not vary across individuals. This means that the pension system considered operates income redistribution across individuals of the same generation.[7]

3. Individual Saving and Retirement Decisions

In this section, we characterize the savings and retirement decisions of old and young individuals, *for given* τ, P *and* θ. We denote (z^y, s^y) the optimal decisions of young individuals, where z^y is the retirement age and s^y is the amount of savings. Decisions concerning savings have been made in the past for old people. Their only decision is to choose when to retire. The optimal retirement decision of an old individual is denoted z^o.

3.1. The Old

The program of old individuals is the following:

$$\max_{z} s(1+r) + wz(1-\theta\tau) + P - \gamma\frac{z^2}{2} \tag{1}$$

subject to

$$0 \leq z \leq 1.$$

The first-order condition for an interior value of z is:

$$w(1-\theta\tau) - \gamma z^o = 0.$$

This leads to

$$z^o = \frac{w(1-\theta\tau)}{\gamma}. \tag{2}$$

In order to ensure that $z^o \leq 1$ for everyone, we assume $\gamma \geq w_+$. All the individuals choose to work in the second period (except when $\theta\tau = 1$). The higher the productivity of an individual, the later he retires: consumption being cheaper for more productive individuals, they choose to work and consume more, provided of course that there is no income effect. On the other hand, increasing the bias of the system or the payroll tax rate increases the price of consumption with respect to leisure and consequently induces people to retire earlier. When $\theta\tau = 0$, there are no distortions so that z^o is equal to w/γ which corresponds to the first best level. Finally, a higher disutility of work yields lower retirement ages.

3.2. The Young

The program of young individuals is the following:

$$\max_{z,s} u[w(1-\tau) - s] + \beta u[s(1+r) + wz(1-\theta\tau) + P - \gamma z^2/2] \tag{3}$$

subject to

$$0 \leq z \leq 1 \quad \text{and} \quad 0 \leq s \leq w(1-\tau).$$

The young anticipate that they will choose their retirement age according to (2). Consequently, we have

$$d = s(1+r) + \frac{w^2(1-\theta\tau)^2}{2\gamma} + P.$$

Recalling that by assumption $\beta(1+r) = 1$, the first order condition for an interior solution of s is:

$$-u'(c) + u'(d) = 0.$$

Individuals equalize first and second period consumptions (net of the disutility of labor). For individuals choosing an interior solution, we obtain:

$$s^y = \frac{w(1-\tau) - \frac{w^2(1-\theta\tau)^2}{2\gamma} - P}{(2+r)}. \tag{4}$$

3.3. Budget Constraint

A feasible pension scheme must satisfy the government budget constraint:

$$\begin{aligned} N^o \int_{w_-}^{w_+} Pf(w)\,dw &= N^y\tau \int_{w_-}^{w_+} wf(w)\,dw + N^o\theta\tau \int_{w_-}^{w_+} wzf(w)\,dw \\ \Leftrightarrow P &= (1+n)\tau\bar{w} + \theta\tau\bar{y} \\ &= (1+n)\tau\bar{w} + \frac{\theta\tau(1-\theta\tau)}{\gamma}E(w^2), \end{aligned} \tag{5}$$

where N^y and N^o are respectively the numbers of young and old individuals, n is the rate of population growth, $y = wz$, $\bar{y} = \int_{w_-}^{w_+} yf(w)\,dw$ and $E(w^2) = \int_{w_-}^{w_+} w^2 f(w)\,dw$. We assume that

$$\frac{1+n}{1+r} \leq \sqrt{E(w^2)}/\bar{w}. \tag{6}$$

Because $\sqrt{E(w^2)} > \bar{w}$, condition (6) requires that n is not too large compared to r. It is always satisfied when $n \leq r$.

The total pension received by a given individual is the sum of (per capita) tax revenues on first and second period incomes. The tax base in the first period, $(1+n)\bar{w}$, is fixed whereas it depends on $\theta\tau$ in the second period. Put differently, taxation only gives rise to distortions on second period income. Differentiating (6) yields:

$$P'(\tau) = (1+n)\bar{w} + \frac{\theta - 2\theta^2\tau}{\gamma}E(w^2) \tag{7}$$

and

$$P''(\tau) = \frac{-2\theta^2}{\gamma}E(w^2) \leq 0. \tag{8}$$

The budget curve, represented on Figure 1, is concave, always above the line $\tau(1+n)\bar{w}$ and P is equal to $\tau(1+n)\bar{w}$ when $\theta\tau = 0$ or 1.

4. Optimal Solution: First- and Second-Best

Even though our approach is mainly positive, it is worth looking at the solution chosen by a utilitarian social planner. For simplicity the formal analysis in this section concentrates on the case where $r = n$. However, we shall also sketch the case where r and n differ. The

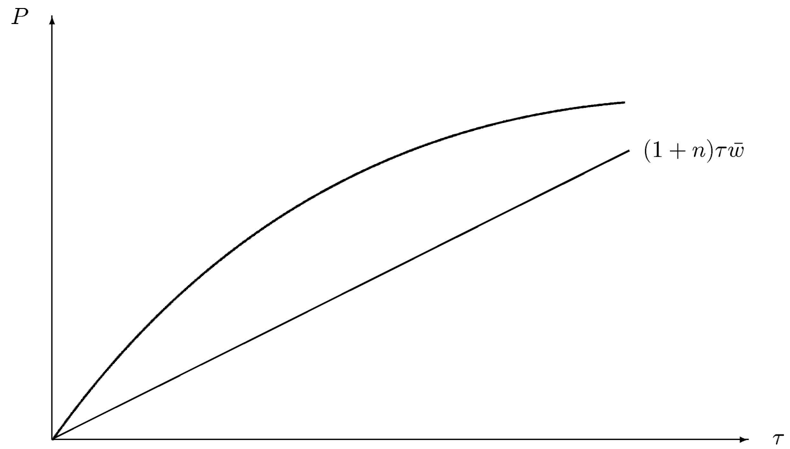

Figure 1. The budget curve.

results do not change when $n > r$. On the other hand, when $r > n$ some of the results change but the main conclusion is not affected. We successively consider the first-best allocation and then the second-best solution achieved with the instruments (and information) available in our setting. These solutions provide interesting benchmarks for the interpretation of the voting equilibria considered below. In particular they allow one to study the rationale (if any) of a distorting tax (biased system) from a normative optimal taxation perspective.

The first-best solution is obtained by solving:

$$\max \int_{w_-}^{w_+} [u(c(w)) + \beta u(x(w) - \gamma z(w)^2/2)] f(w)\, dw$$

subject to

$$\int_{w_-}^{w_+} \left[c(w) + \frac{x(w)}{1+n} - \frac{w}{1+n}(1+n+z(w)) \right] f(w)\, dw = 0. \tag{9}$$

The first order conditions imply:

$$u'(c(w)) = \beta(1+n)u'(d(w)) = \mu \tag{10}$$

and

$$z(w) = \frac{w}{\gamma}$$

where μ is the Lagrange multiplier associated with the resource constraint. These conditions are rather standard and intuitive. Consumption is equalized across individuals and even across periods if $\beta(1+n) = 1$.[8] Labor supply varies across individuals such that the marginal utility of retirement γ_z is equal to the marginal productivity of labor, w. [9] This first-best solution can be decentralized with lump-sum taxes and transfers and naturally with

$\theta = 0$ (τ would then be non-distortionary). It is plain that this solution cannot be achieved with the instruments considered in our setting.

We now consider the second-best problem. To make this a meaningful benchmark, policies are restricted in the same way as in the voting process considered below.[10] Instruments are thus limited to the parameters τ, θ and P. Without loss of generality we will introduce a new variable $\tilde{\tau} \equiv \theta\tau$ so that the social planner acts as if he determines a tax rate for the first period τ and another one, $\tilde{\tau}$, for the second period.

The optimality problem can be stated as the maximization of:

$$\pounds = \int_{w_-}^{w_+} \left[u(w(1-\tau) - s^y) + \beta u\left(s^y(1+r) + P + wz^y(1-\tilde{\tau}) - \frac{\gamma z^{y2}}{2} \right) \right] f(w)\, dw,$$

with respect to τ and $\tilde{\tau}$, subject to

$$P = (1+n)\tau\, \bar{w} + \tilde{\tau} \int_{w_-}^{w_+} w\, z(w) f(w)\, dw. \tag{11}$$

where s^y and z^y are respectively defined by (4) and (2). Substituting for (2) and (11), the utility in the second period is:

$$u\left(s^y(1+r) + \frac{w^2(1-\tilde{\tau})^2}{2\gamma} + (1+n)\tau\, \bar{w} + \frac{\tilde{\tau}(1-\tilde{\tau})}{\gamma} E(w^2) \right).$$

The first-order conditions are given by:

$$\begin{aligned} \frac{\partial \pounds}{\partial \tau} &= -\int_{w_-}^{w_+} [wu'(c) - \beta(1+n)\bar{w}u'(d)] f(w)\, dw \\ \frac{\partial \pounds}{\partial \tilde{\tau}} &= -\int_{w_-}^{w_+} u'(d) \left(\frac{w^2(1-\tilde{\tau})}{\gamma} - \frac{1-2\tilde{\tau}}{\gamma} E(w^2) \right) f(w)\, dw. \end{aligned} \tag{12}$$

To interpret (12), assume first that there is no liquidity constraint (savings can be negative). Then, $c(w) = d(w)$ for all w and we have $\tau = 1$. When there is a liquidity constraint, on the other hand, τ has to be less than 1 to ensure a positive consumption in the first period. However, we continue to have a positive tax. Recall that the tax on the first period income does not affect labor supply. The standard linear income tax problem then calls for a one hundred percent tax which is redistributed in a lump sum way. However, here the redistribution takes place in the second period (through the flat pension). When there is no liquidity constraint this is of no relevance and the traditional result goes through. With a liquidity constraint, a 100% tax is of course not desirable. In addition, a high tax then adversely affects individuals with a low w: they have to pay taxes on an already low first-period income while the lump sum refund (pension) only occurs in the second period. Note that when $r > n$, the tax rate may be less than one, even without a liquidity constraint. This is because the redistributive benefits of the tax (and the lump sum transfer it finances) has to be balanced against the inefficiency of the pay-as-you go system (which has then a lower return than private savings).

The condition for $\tilde{\tau}$ is equivalent to the standard formula for an optimal linear tax:

$$\tilde{\tau} = \frac{\int_{w_-}^{w_+} u'(d)[w^2 - E(w^2)]f(w)\,dw}{\int_{w_-}^{w_+} u'(d)[w^2 - 2\,E(w^2)]f(w)\,dw}$$

where the numerator is the covariance between the marginal utility of $d(w)$ and the square of the productivity levels. This covariance is negative. We thus have:

$$\tilde{\tau} = \frac{\operatorname{cov}(u'(d), w^2)}{\operatorname{cov}(u'(d), w^2) - E(w^2)\int_{w_-}^{w_+} u'(d)f(w)\,dw}, \tag{13}$$

namely $0 < \tilde{\tau} < 1$. Observe that if $\tilde{\tau}$ were levied on w rather than on wz, in other words if there was no distortion, $\tilde{\tau}$ would be equal to 1. This is useful for what follows. Further, observe that the design of the second period tax (and thus of the bias) is not directly affected by the relationship between r and n. With a utilitarian objective, it is the liquidity constraint for τ and the labor disincentive for $\tilde{\tau}$ which prevent the tax rates to be maximal. These two factors are also crucial in our political economy model.

To sum up, and recalling the definition $\tilde{\tau} = \theta\tau$ we have shown that the (utilitarian) second-best policy always calls for a positive bias. In the case where there is no liquidity constraint, we can also be sure that $\theta < 1$. Consequently, it is then always optimal to have some rebating of the second period tax.[11] When there is a liquidity constraint, on the other hand, the comparison between τ and $\tilde{\tau}$ is in principle ambiguous and the bias can be smaller as well as larger than one. However, for reasonable assumptions on the wage distribution and labor supply elasticities $\tau < \tilde{\tau}$ (i.e., $\theta > 1$) will not arise. Intuitively, $\tau < \tilde{\tau}$ could arise if the redistributive benefits of the second period tax (relative to the first period tax) were to outweigh the distortion. This requires a highly unequal wage distribution; recall that (with a liquidity constraint) the individuals with the lowest wage tend to be penalized by a high first-period tax.

5. Majority Voting Equilibrium Tax Rate

We now turn to the study of the voting equilibrium. For the time being we assume that the bias, θ, is exogenously given so that the policy choice involves a single dimension, namely τ (with P then being automatically determined by the budget constraint (6)).

5.1. Preferred Tax Rates

Define

$$V^y(\tau, \theta; w) = u[w(1-\tau) - s^y] + \beta u[s^y(1+r) + P(\tau, \theta) + w^2(1-\theta\tau)^2/2\gamma] \tag{14}$$

and

$$V^o(\tau, \theta; w) = u[s^o(1+r) + P(\tau, \theta) + w^2(1-\theta\tau)^2/2\gamma], \quad (15)$$

which represent the utility levels attained by type w individuals, young or old, for given τ and θ. Preferred tax rates for young and old individuals, denoted respectively τ^y and τ^o, are obtained by solving the following programs:

$$\max_{\tau \in [0,1]} V^i(\tau, \theta; w), \quad i = y, o.$$

The following proposition states the properties of the most preferred tax rates.

Proposition 1.

(i) *Preferred tax rates of young individuals are non-increasing with productivity.*
(ii) *No young individual chooses a corner solution at $\tau = 1$. The preferred tax rate of young individuals with productivity $w \leq \tilde{w}$ is positive, with $\tilde{w} = \bar{w}(1+n)/(1+r)$ when $\theta = 0$ while $\tilde{w} > \bar{w}(1+n)/(1+r)$ when $\theta > 0$. Consequently, the introduction of a bias increases the political support for the pension system.*
(iii) *Old individuals choose corner solutions with either $\tau^o = 1$ for everyone or alternatively $\tau^o = 1$ for the poor old and $\tau^o = 0$ for the rich old.*

The formal proof is given in Appendix 1. Here we provide a sketch of the main intuitions. The intuition for the first result is as follows. Consider for simplicity the case $\theta = 0$. An increase of w has three effects. First, when w increases, first period income also increases. Consequently, the individual also wants a higher second period consumption (which is a normal good) and thus a higher (first-period) tax.

Second, the relative price of first and second periods consumptions decreases. This is because with a PAYG system this relative price is given by $\bar{w}(1+n)/w$. Put differently, it costs less (in terms of first period consumption) to buy one unit of second period consumption for a low productivity individual than for a high one. By this substitution effect, high productivity individuals are induced to buy less second period consumption. For utility functions such that $R_r(\cdot) < 1$, this substitution effect dominates the first effect and low productivity individuals want tax rates larger than high productivity individuals. Note that when $R_r(\cdot) = 1$ (logarithmic utility function), income and substitution effects neutralize each other and preferred tax rates are constant with respect to productivity.

These two effects has already been pointed out in the related literature such as Casamatta, Cremer and Pestieau (2000) or Tabellini (2000); see Galasso and Profeta (2002) for a survey. The novelty here is a third effect arising from individuals also working in the second period. Because second period income increases with productivity, high productivity individuals raise their first period consumption (which is a normal good) by reducing the payroll tax rate. This effect reinforces the second one. As a consequence, preferred tax rates are decreasing with productivity even when $R_r(\cdot) = 1$.

The first part of point (ii) is obvious. When the tax rate equals one, marginal utility of consumption tends to infinity and all individuals prefer a smaller tax rate. To illustrate the second part, let us write the first-order derivative of a young individual life cycle utility at the point $\tau = 0$ (saving being optimally chosen):

$$\left.\frac{dV^y}{d\tau}\right|_{\tau=0} = -wu'(c) + \beta(1+n)\bar{w}u'(d) + \beta\frac{\theta}{\gamma}(E(w^2) - w^2)u'(d)$$
$$= (-w + \beta(1+n)\bar{w} + \beta\frac{\theta}{\gamma}(E(w^2) - w^2))u'(d),$$

where we have used the fact that saving is positive when $\tau = 0$ which implies that $u'(c) = u'(d)$. A first observation is that, in a neutral system ($\theta = 0$), individuals choose a positive tax rate *if and only if* $w \leq \beta\bar{w}(1+n) = \bar{w}(1+n)/(1+r)$. This is because with $\theta = 0$ there is no taxation of second period income. Individuals favoring a positive tax rate are those for whom the rate of return of the PAYG system, $(1+n)\bar{w}/w$ is higher than the rate of return of private savings, $1+r$. This is a standard result in the literature. Now, if one introduces a bias in the system which is a new feature of our analysis, second period incomes are redistributed from individuals with a productivity level higher than $\sqrt{E(w^2)}$ towards individuals with a lower productivity.[12] Therefore, individuals such that $w \leq \bar{w}(1+n)/(1+r)$ continue to favor a positive tax rate but some individuals with a higher productivity also do. Consequently, it appears that the introduction of a bias increases the political support for the pension system.

The last point of the proposition says that old individuals choose corner solutions for the tax rates. To understand this, differentiate the objective function of the old (15), with respect to the tax rate. This yields:

$$\frac{dV^o}{d\tau} = \left((1+n)\bar{w} + \frac{\theta(1-2\theta\tau)}{\gamma}E(w^2) - \frac{\theta(1-\theta\tau)}{\gamma}w^2\right)u'(d). \tag{16}$$

When $\theta = 0$, the RHS of (16) is necessarily positive. In a neutral system (and by continuity in a slightly biased system), the welfare of an old individual is an increasing function of the tax rate. Consequently, every old individual wants the tax rate to be as high as possible and choose $\tau = 1$. This is the usual case encountered in the literature.

When θ is increased, one can see, by evaluating the above expression at $\tau = 0$, that old people with productivity $w^2 < E(w^2) + (1+n)\bar{w}\gamma/\theta$ want a positive tax rate. We prove in the appendix that they in fact most prefer $\tau = 1$. On the other hand, old individuals with a higher productivity dislike a marginal increase in the tax rate (starting from $\tau = 0$). This does not mean that their optimal tax rate is 0. Instead, we show that *their objective function may be convex*. To see this, note that (16) is positive at $\tau = 1$ when $\theta = 1$. When $\theta = 1$ and τ approaches 1, everyone stops working and the old rich do not suffer anymore from the redistribution towards the poor. On the other hand, their pension increases with the tax rate. They thus favor a marginal increase in the tax rate.

To sum up, old people with productivity $w^2 < E(w^2) + (1+n)\bar{w}\gamma/\theta$ favor the maximal tax rate $\tau = 1$. Those with a higher productivity may prefer 0 or 1 but no one has an interior optimal tax rate (between 0 and 1). The productivity level above which the optimal tax rate becomes 1 is derived in the appendix.

5.2. Voting Equilibrium

We turn to the determination of the equilibrium payroll tax rate under majority voting. We have proved in Appendix 1 that, for utility functions such that $R_r(\cdot) \leq 1$, preferences of the young over tax rates satisfy the single-crossing condition established by Gans and Smart (1996). This means that we can order young individuals and alternatives such that when an individual prefers the higher of two alternatives, all the individuals with a lower productivity display the same preference. Similarly, one can show that preferences of the old also satisfy this single-crossing property. However, these properties are not sufficient to guarantee the existence of a Condorcet winner. A sufficient condition for that would be that the single crossing property holds when the entire population (young and old jointly) is considered. However, single crossing within each subgroup does not imply single crossing for the entire population. To ensure existence of a voting equilibrium we thus need additional assumptions. Specifically, we restrict our attention to cases where the utility of the old is monotonically increasing with the tax rate. This is true when the marginal utility of the richest old individual is positive at $\tau = 0$ which, from (16), is the case when $\theta/\gamma \leq (1+n)\bar{w}/(w_+^2 - E(w^2))$. This condition is satisfied when the bias parameter, θ, is small enough, or when γ is large enough. With the old preferring a maximum tax rate and the preferred tax of the young decreasing with productivity (Proposition 1, (i)) the construction of the equilibrium is straightforward. Specifically, and recalling that $n > 0$, the pivotal voter is a young individual so that the number of *young* individuals who have a higher wage (and thus want a lower tax than the pivotal voter) equals half the *total* population. This is stated formally in the following proposition.

Proposition 2. *If* $\theta/\gamma \leq (1+n)\bar{w}/(w_+^2 - E(w^2))$, *a voting equilibrium on* τ *exists. It is given by* $\tau^{mv}(\theta) = \tau^y(\theta; w_{\text{piv}})$, *where* w_{piv} *is the productivity of the pivotal individual which is implicitly determined by the following condition*:

$$N^y(1 - F(w_{\text{piv}})) = \frac{N^o + N^y}{2}$$
$$\Leftrightarrow F(w_{\text{piv}}) = \frac{n}{2(1+n)}. \tag{17}$$

Observe that condition (17) also implies that $N^o + N^y F(w_{\text{piv}}) = (N^o + N^y)/2$ so that the young with a lower wage plus the old (i.e., the individuals who want a higher tax) represent also half of the population.

Using Proposition 1 (ii) we establish the following proposition.

Proposition 3. *The tax rate* τ^{mv} *defined by Proposition* 2 *is positive if*

$$w_{\text{piv}} < \bar{w}(1+n)/(1+r). \tag{18}$$

Interestingly, condition (18) is necessarily satisfied when $n = r$. To see this observe that $n/(2(1+n)) < 1/2$. Consequently, the pivotal voter has a productivity level below the median level w_m which in turn is below the mean level $\bar{w}$. This argument also makes it clear that condition (18) continues to hold when n is smaller than r, but not by "too much". However, as the differential between n and r increases the condition becomes more and more difficult to satisfy. Intuitively, the pivotal individual benefits from the redistribution implied by the PAYG system and thus favors a positive tax as long as the return of PAYG, namely n, is sufficiently close to the interest rate. The results in Propositions 2 and 3 have a familiar flavor. Similar properties are obtained for example in Casamatta, Cremer and Pestieau (2000) and Tabellini (2000). Note however that in this earlier work condition (18) is both necessary and sufficient whereas it is only sufficient here. As emphasized earlier, people with a productivity level above $\bar{w}(1+n)/(1+r)$ may sustain the PAYG system when $\theta > 0$. In other words, the political support for the PAYG system is increased when it is biased.

To conclude, it is important to stress that the condition for existence imposed in Proposition 2 is only a *sufficient* condition. A less stringent (though more complex) condition would be to require that the utility of the richest old be higher at the point $\tau^y(\theta; w_{\text{piv}})$ than at $\tau = 0$, that is:

$$V^o[\tau^y(\theta; w_{\text{piv}}), \theta; w_+] \geq V^o[0, \theta; w_+]. \tag{19}$$

If this condition is satisfied for the richest old, it holds effectively for the entire old generation. Any tax rate lower than $\tau^y(\theta; w_{\text{piv}})$ is thus rejected by the coalition of all the old and the young to the left of w_{piv}. A tax rate higher than $\tau^y(\theta; w_{\text{piv}})$, on the other hand, is (by definition) rejected by the young to the right of w_{piv}. Consequently, as long as condition (19) holds, $\tau^y(\theta; w_{\text{piv}})$ continues to be a Condorcet winner even when the preferences of the old are not monotonically increasing.

6. Endogenous Bias

In the previous section, we have assumed that θ was exogenously given. We now turn to the political determination of θ. The natural approach would be to have the pair (θ, τ) chosen jointly in a majority vote. However, it is well known that a Condorcet winner is unlikely to exist when the issue space is multidimensional and this model is not an exception. To overcome this difficulty we consider a more restricted voting procedure, introduced by Shepsle (1979), and assume that the parameters θ and τ are chosen simultaneously and independently. The equilibrium (if any) of this procedure is then a pair (θ^S, τ^S) such that the level of each variable is a Condorcet winner given the level of the other variable. In other words, it is given by the intersection of the two "reaction functions" $\tau^{mv}(\theta)$ and $\theta^{mv}(\tau)$ specifying the majority equilibrium level of a variable given the level of the other variable. The equilibrium tax for a given bias has been studied in Section 5.2 where $\tau^{mv}(\theta)$ is defined in Proposition 2. We now turn to the determination of $\theta^{mv}(\tau)$, the equilibrium bias for a given tax rate.

6.1. Voting Equilibrium Level of θ for a Given Level of τ

With τ given, the variable left to be determined is θ, P being determined by the budget constraint. Differentiating (1) and (3) shows that the slope of an indifference curve in the

(θ, P) plane is given by

$$\frac{dP}{d\theta} = \frac{\tau w^2(1-\theta\tau)}{\gamma} > 0,$$

for the young as well as for the old individuals. This expression is increasing with w. Consequently, preferences are single-crossing so that a majority voting equilibrium on θ exists.

An individual's most preferred level of θ is obtained by solving:

$$\max_{\theta\in[0,1/\tau]} V^i(\tau,\theta;w), \quad i = y, o,$$

This yields

$$\theta^i = \frac{E(w^2) - w^2}{\tau(2E(w^2) - w^2)}, \quad i = y, o,$$

and

$$\frac{d\theta^i}{dw} = -\frac{2\tau w E(w^2)}{[\tau(2E(w^2) - w^2)]^2} < 0, \quad i = y, o,$$

so that *the most preferred level of θ decreases with productivity*. Notice that these expressions apply for old as well as for young individuals. Consequently, individual's most preferred levels of θ do not depend on age. A young and an old individual with the same w prefer the same θ. Putting these elements together, it follows that the Condorcet winner is then the level of θ most preferred by the individuals with median productivity and we have:

$$\theta^{mv}(\tau) = \frac{E(w^2) - w_m^2}{\tau\left(2E(w^2) - w_m^2\right)} > 0. \tag{20}$$

This hyperbolic expression is represented on Figure 2 for $\tau \in [0, 1]$. It thus appears that the equilibrium bias is *positive and decreasing in the tax rate*.

Intuitively, (20) defines the second period tax rate $\tilde{\tau} = \theta\tau$, which is optimal from the median voter's perspective. This rate is given by

$$\tilde{\tau} = \tau\theta^{mv}(\tau) = \frac{E(w^2) - w_m^2}{\left(2E(w^2) - w_m^2\right)}.$$

The crucial property of this expression is that the level of this second period tax does not depend on the first period tax, τ.[13] In other words, when τ changes, the median voter's most preferred level of $\tilde{\tau}$ remains constant and is achieved simply by adjusting θ.

6.2. Simultaneous Voting

Returning to the original problem, namely the determination of the equilibrium pair (θ^S, τ^S), we face the difficulty that $\tau^{mv}(\theta)$, may not be defined for all levels of θ; see Section 5. The

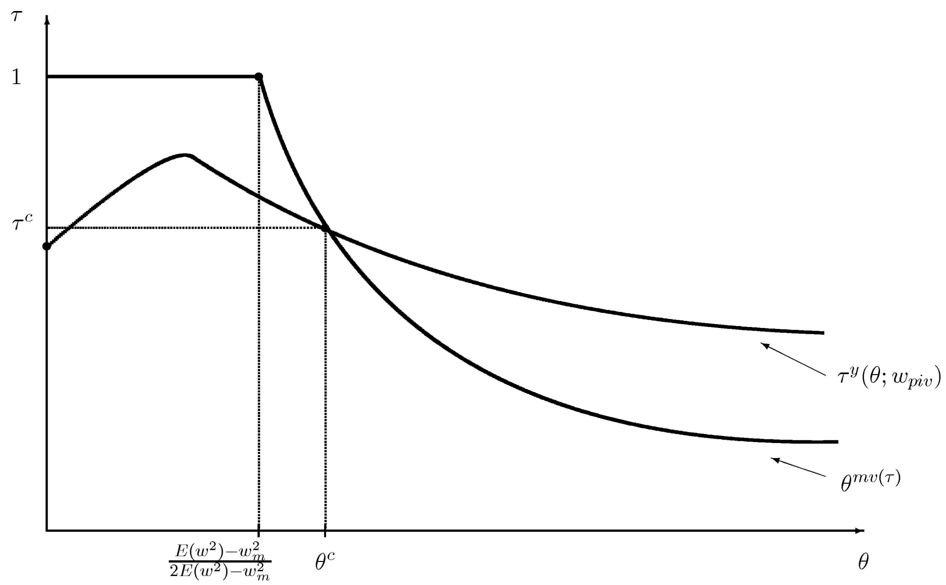

Figure 2. Determination of the simultaneous voting equilibrium.

simplest way to address this problem is to consider a candidate equilibrium (θ^c, τ^c), given by the intersection of $\theta^{mv}(\tau)$ with the curve $\tau^y(\theta; w_{\text{piv}})$, which gives the preferred tax of a young individual with productivity w_{piv} (defined by (17)). Put differently, the candidate equilibrium, (θ^c, τ^c), is defined by the conditions $\theta^c = \theta^{mv}(\tau^c)$, $\tau^c = \tau^y(\theta^c; w_{\text{piv}})$; see Figure 2 for an illustration. We show in Appendix B that $\theta^{mv}(\tau)$ and $\tau^y(\theta; w_{\text{piv}})$ intersect at least once so that the candidate equilibrium (θ^c, τ^c) exists. This does not imply that a simultaneous, issue by issue, voting equilibrium, (θ^S, τ^S), exists. However, (θ^c, τ^c) is an equilibrium when (sufficient) condition (19) is satisfied, that is when

$$\frac{w_+^2[1-(1-\theta^c\tau^c)^2]}{2\gamma} \leq (1+n)\tau^c\bar{w} + \frac{\theta^c\tau^c(1-\theta^c\tau^c)}{\gamma}E(w^2),$$

where

$$\theta^c\tau^c = \frac{E(w^2)-w_m^2}{\left(2E(w^2)-w_m^2\right)}. \tag{21}$$

For simplicity we assume in the remainder of this section that this is effectively the case. Then it is plain that this procedure yields a positive bias $\theta^S > 0$ and an interior solution for the tax rate $0 < \tau^S < 1$. We also have an interior solution for the second period tax $\tilde{\tau} = \theta^S\tau^S$. However, there is no obvious way to determine whether the bias is larger or smaller than one; inspection of expressions only allows one to conclude that both cases appear to be possible. Similarly, the comparison between the second period tax implied by (21) and the socially optimal level defined by (13) appears to be ambiguous. Not surprisingly, it is then also not

Table 1. Numercial illustrations: Simultaneous voting (Shepsle) equilibrium and utilitarian optimum.

	Shepsle			Util. Opt.		
	θ	τ	$\theta\tau$	θ	τ	$\theta\tau$
Skewed dist.						
$\varepsilon = 0.2$	0.44	0.82	0.36	0.25	0.30	0.08
$\varepsilon = 0.5$	0.80	0.45	0.36	0.56	0.29	0.16
$\varepsilon = 1$	1.17	0.31	0.36	0.96	0.26	0.25
Unif. dist.						
$\varepsilon = 0.2$	0.22	0.89	0.19	0.17	0.28	0.05
$\varepsilon = 0.5$	0.40	0.49	0.19	0.42	0.26	0.11
$\varepsilon = 1$	0.64	0.31	0.19	0.97	0.21	0.21

possible to compare the equilibrium bias θ^S to the optimal (second-best) bias studied in Section 4. To illustrate the results and to show that the comparisons between equilibrium and optimum are effectively ambiguous we now present some numerical examples.

6.3. Numerical Examples

Consider the following specifications for the various components of our model. Productivities are distributed on [1, 100] and we set $\gamma = 100$ and $r = n = 1$. We consider two possible distributions: a distribution skewed to the right with $w_m = 17.59 < \bar{w} = 22.33 < \sqrt{E(w^2)} = 26.82$ and a uniform distribution function with $w_m = \bar{w} = 50.5 < \sqrt{E(w^2)} = 58.03$. The utility function is isoelastic: $u(x) = x^{1-\varepsilon}/(1-\varepsilon)$, where ε is the coefficient of relative risk aversion. We consider three levels of ε: 0.2, 0.5 and 1.

Table 1 presents the simultaneous voting equilibrium (θ^S, τ^S) and the utilitarian (second-best) optimum (θ^*, τ^*) for each combination of parameter values. To facilitate interpretation, we also report $\theta\tau = \tilde{\tau}$, i.e., the second period implicit tax rate.

These results are very interesting and confirm a number of points suggested by the analytical expressions. In particular, they show that a bias which is larger than one can effectively arise at the voting equilibrium. Furthermore, we find cases where the equilibrium bias exceeds the optimal one, and cases where the opposite is true. Similarly, for the second period tax, $\theta^S\tau^S > \theta^*\tau^*$ as well as $\theta^S\tau^S < \theta^*\tau^*$ can effectively occur.

7. Conclusion

This paper has examined two related questions. First, we have studied the determination of the *size* of the pension system through the political process when retirement age is endogenous. Specifically, each individual chooses his retirement age by taking into account the (given) *benefit formula* which may or may not be biased (i.e., impose an implicit tax on continued activity of the elderly). We have shown how the endogenous retirement age decision affects voters preferences over tax rates and how this relationship is affected by the benefit formula. Overall, it has turned out that the general properties of the voting equilibrium (if any) are not qualitatively different from the case with exogenous retirement age, see Casamatta, Cremer and Pestieau (2000).[14] In particular, the feature that high tax rates are supported by all the old and by low productivity young individuals has reemerged within our current setting.

The second question has involved a more significant departure from the existing literature. It concerns the determination of the benefit formula and, more specifically, the bias of the pension system. The challenge we were addressing was to seek explanations for the empirically observed bias in retirement systems. This is a more ambitious problem and we can only claim to have provided some preliminary and partial solutions. We have shown that a biased system is optimal from a utilitarian perspective and that it also tends to emerge from a specific political process (simultaneous and independent vote). From that perspective, political economy considerations are not necessary to explain the emergence of a biased system. On the other hand, an extreme bias ($\theta > 1$) though often observed in reality is not consistent with welfare maximization.[15] It can, however, emerge from the political process. Specifically, we have shown that a very simple majority procedure is sufficient to obtain such a drastic departure from optimality.

Our setting is rather restrictive and our findings have to be qualified accordingly. Our main endeavor it to point to possible explanations for the high implicit tax on continued activity. We do not claim that the effects we discuss are always relevant, nor that they provide the unique or even main driving forces. To address these issues additional studies featuring (for instance) more general benefit formulas and alternative specifications of the political process are required. We leave these questions open for future research.

Appendix A: Proof of Proposition 1

(i) To prove that individually optimal tax rates of the young are decreasing with productivity, we show that the slope of indifference curves in the (τ, P) plane is increasing with productivity, that is $d^2P/d\tau dw > 0$.

The equation of an indifference curve is derived by solving:

$$u[w(1-\tau) - s^y] + \beta u[s^y(1+r) + P + w^2(1-\theta\tau)^2/2\gamma] = c,$$

where c is a constant. Differentiating this expression, the slope of an indifference curve is

$$\frac{dP}{d\tau} = \frac{wu'(c) + \beta\frac{w^2}{\gamma}\theta(1-\theta\tau)u'(d)}{\beta u'(d)} > 0. \tag{22}$$

If savings are positive, $u'(c) = u'(d)$. This leads to

$$\frac{dP}{d\tau} = \frac{w + \beta \frac{w^2}{\gamma}\theta(1-\theta\tau)}{\beta} \geq 0.$$

Indifference curves are increasing. Moreover, $d^2P/d\tau dw > 0$: the slope of indifference curves is increasing with productivity.

If the individual does not want to save, the differentiation of (22) with respect to w leads to

$$\begin{aligned}\frac{d^2P}{d\tau dw} &= \frac{(u'(c) + w(1-\tau)u''(c))\beta u'(d) - \beta w u'(c)\frac{w}{\gamma}(1-\theta\tau)^2 u''(d)}{(\beta u'(d))^2} \\ &\quad + \frac{2w}{\gamma}\theta(1-\theta\tau) \\ &= \frac{(u'(c)(1 - R_r(c)))\beta u'(d) - \beta w u'(c)\frac{w}{\gamma}(1-\theta\tau)^2 u''(d)}{(\beta u'(d))^2} \\ &\quad + \frac{2w}{\gamma}\theta(1-\theta\tau).\end{aligned}$$

If $R_r(\cdot) \leq 1$, $d^2P/d\tau dw$ is positive. It follows that the slope of indifference curves is increasing with productivity. This leads to our conclusion that preferred tax rates are decreasing with productivity.

(ii) To prove that no young individual chooses a corner solution at $\tau = 1$ and that the preferred tax rate of the young with productivity $w < \bar{w}(1+n)/(1+r)$ is positive, we evaluate the impact of a marginal increase in τ on the utility of the young, namely $dV^y/d\tau$, respectively at $\tau = 1$ and at $\tau = 0$.

Differentiating (14) with respect to τ and using (7), we have

$$\begin{aligned}\frac{dV^y}{d\tau} &= -wu'(c) + \beta\left(P'(\tau) - \frac{w^2\theta(1-\theta\tau)}{\gamma}\right)u'(d) \\ &= -wu'(c) + \beta\left((1+n)\bar{w} + \frac{\theta - 2\theta^2\tau}{\gamma}E(w^2) - \frac{w^2\theta(1-\theta\tau)}{\gamma}\right)u'(d).\end{aligned}$$

At $\tau = 1$,

$$\begin{aligned}\left.\frac{dV^y}{d\tau}\right|_{\tau=1} &= -wu'(0) + \beta\left((1+n)\bar{w} + \frac{\theta - 2\theta^2}{\gamma}E(w^2) - \frac{w^2\theta(1-\theta)}{\gamma}\right) \\ &\quad \times u'\left((1+n)\bar{w} + \frac{E(w^2)\theta(1-\theta)}{\gamma}\right).\end{aligned}$$

It is clear that, if $\lim_{x\to 0} u'(x) = +\infty$, $dV^y/d\tau|_{\tau=1} < 0$.

At $\tau = 0$,

$$\left.\frac{dV^y}{d\tau}\right|_{\tau=0} = -wu'(c) + \beta\left((1+n)\bar{w} + \frac{\theta}{\gamma}E(w^2) - \frac{w^2\theta}{\gamma}\right)u'(d).$$

We argue now that $s^y > 0$ when $\tau = 0$. From (4),

$$s^y|_{\tau=0} = \frac{w - \frac{w^2}{2\gamma} - P(0)}{(2+r)} = \frac{w - \frac{w^2}{2\gamma}}{(2+r)} > 0 \Leftrightarrow w < 2\gamma.$$

Because $\gamma \geq w_+$, the condition $w < 2\gamma$ is satisfied for any w. It follows that $u'(c) = u'(d)$ and

$$\left.\frac{dV^y}{d\tau}\right|_{\tau=0} = u'(c)\left(-w + \beta\left((1+n)\bar{w} + \frac{\theta}{\gamma}E(w^2) - \frac{w^2\theta}{\gamma}\right)\right) > 0$$
$$\Leftrightarrow -w + \beta\left((1+n)\bar{w} + \frac{\theta}{\gamma}E(w^2) - \frac{w^2\theta}{\gamma}\right) > 0$$
$$\Leftrightarrow -w + \frac{1+n}{1+r}\bar{w} + \frac{1}{1+r}\left(\frac{\theta}{\gamma}E(w^2) - \frac{w^2\theta}{\gamma}\right) > 0. \tag{23}$$

If $w < \bar{w}(1+n)/(1+r)$ then, by assumption, $w < \sqrt{E(w^2)}$ and condition (23) is satisfied. When $\theta = 0$, only people with productivity $w < \bar{w}(1+n)/(1+r)$ want a positive tax rate. When $\theta > 0$, this is also the case for some individuals with higher productivity. This follows from continuity because inequality (23) is strict for $w = \bar{w}(1+n)/(1+r)$.

(iii) To prove that old individuals choose only corner solutions, we compare the slope of their indifference curves with the slope of the government budget constraint.

Indifference curves for old individuals are derived by solving

$$u\left(s(1+r) + P + \frac{w^2(1-\theta\tau)^2}{2\gamma}\right) = c.$$

Differentiating, we have

$$\frac{dP}{d\tau} = \frac{w^2}{\gamma}\theta(1-\theta\tau) > 0$$

and

$$\frac{d^2P}{d\tau^2} = -\frac{w^2\theta^2}{\gamma} < 0.$$

Comparing with (8), we find that the slope of indifference curves decreases more quickly than the slope of the budget curve if and only if $w^2 > 2E(w^2)$. At $\tau = 0$, the productivity level such that the slope of the indifference curve equals the slope of the budget curve is given by:

$$\frac{w^2}{\gamma}\theta = (1+n)\bar{w} + \frac{\theta}{\gamma}E(w^2)$$
$$\Leftrightarrow w_s^2 = \frac{\gamma}{\theta}(1+n)\bar{w} + E(w^2).$$

Noting that

$$E(w^2) = \int_{w_-}^{w_+} w^2 f(w)\,dw < w_+ \int_{w_-}^{w_+} wf(w)\,dw = w_+\bar{w}$$

and recalling that $\gamma \geq w_+$,

$$w_s^2 > 2E(w^2).$$

Following the discussion above, this indifference curve is always below the budget curve. Therefore, individuals with productivity w_s have a preferred tax rate equal to 1. Observing that the slope of indifference curves is increasing with productivity, all individuals with a productivity less than w_s want also a tax rate equal to 1. Individuals with a higher productivity do not choose an interior solution for τ. Indeed, their indifference curves being "more concave" than the budget curve, a point of tangency between an indifference curve and the budget curve corresponds to a minimizing tax rate. The individuals indifferent between $\tau = 0$ and $\tau = 1$ are such that

$$\begin{aligned} V^o(0; w^o) = V^o(1; w^o) &\Leftrightarrow u\left(s(1+r) + P(0) + \frac{w^{o^2}}{2\gamma}\right) \\ &= u\left(s(1+r) + P(1) + \frac{w^{o^2}(1-\theta)^2}{2\gamma}\right) \\ &\Leftrightarrow \frac{w^{o^2}}{2\gamma} = (1+n)\bar{w} + \frac{\theta(1-\theta)}{\gamma}E(w^2) + \frac{w^{o^2}(1-\theta)^2}{2\gamma} \\ &\Leftrightarrow w^{o^2} = \frac{2\gamma(1+n)\bar{w}}{\theta(2-\theta)} + 2\frac{1-\theta}{2-\theta}E(w^2), \quad \text{if} \quad \theta \neq 0. \end{aligned} \tag{24}$$

Note that this equation does not always have a solution. In particular, when $\theta = 0$, every old individual most prefer a tax rate equal to 1. Observe also that the indifferent old individual has a productivity level higher than $\bar{w}$. Finally, $dw^{o^2}/d\theta < 0$.

Appendix B: Proof That the Curves $\tau^y(\theta; w_{\text{piv}})$ and $\theta^{mv}(\tau)$ Intersect at Least Once

We already know that when $\theta = 0$, $\theta^{mv}(\tau)$ is above $\tau^{mv}(\theta)$. We now proceed to show that the converse also occurs for some value of θ. This, along with the continuity of the curves proves the claim.

To find a point where $\theta^{mv}(\tau)$ is below $\tau^{mv}(\theta)$ (and only as part of our strategy in the proof) we characterize the sequential (τ followed by θ) voting equilibrium (θ^{sv}, τ^{sv}) and use the property that by definition $\theta^{sv} = \theta^{mv}(\tau^{sv})$, i.e., ($\theta^{sv}$, τ^{sv}) is on the curve $\theta = \theta^{mv}(\tau)$ [because θ is determined in the second stage and is thus a majority equilibrium given τ]. Let define (θ^{sv}, τ^{sv}) as the solution of

$$\max_{\tau} u[w_{\text{piv}}(1-\tau) - s^y] + \beta u\left[s^y(1+r) + P(\tau, \theta) + w_{\text{piv}}^2(1-\theta\tau)^2/2\gamma\right]$$

subject to

$$\theta\tau = \frac{E(w^2) - w_m^2}{2E(w^2) - w_m^2}.$$

Substituting the constraint in the objective function, we obtain the following FOC on τ:

$$w_{\text{piv}} u'(c) = \beta(1+n)\bar{w}u'(d).$$

Note that for τ^{sv} to be positive, condition (18) must be satisfied.

Using this condition, we obtain

$$\frac{\partial V^y}{\partial \tau}(\tau^{sv}, \theta^{sv}; w_{\text{piv}}) = \frac{\beta\theta}{\gamma}\big[(1-2\theta^{sv}\tau^{sv})E(w^2) - (1-\theta^{sv}\tau^{sv})w_{\text{piv}}^2\big]u'(d)$$

$$> 0 \Leftrightarrow w_{\text{piv}}^2 < \frac{1-2\theta^{sv}\tau^{sv}}{1-\theta^{sv}\tau^{sv}}E(w^2).$$

Observing that $\theta^{sv}\tau^{sv} = w_m^2/E(w^2)$, this condition becomes $w_{\text{piv}} < w_m$, which is always true. We can conclude that at the point (θ^{sv}, τ^{sv}) which is on the curve $\theta^{mv}(\tau)$, w_{piv} wants to increase the tax rate. Therefore, $\tau^y(\theta, w_{\text{piv}})$ is above $\theta^{mv}(\tau)$ at that point.

Notes

1. A notable exception is Japan.
2. Departure from actuarial fairness can be of three types. First, the pension system may redistribute across individuals. Second, the aggregate level of benefits may outweigh the aggregate level of contributions, which is typically the case in a non mature PAYG system. Third the system may not be marginally fair which is the case when there is an implicit tax as defined above.
3. This takes an extreme form in our model where income effects are assumed away.
4. The assumption of fixed labor supply is made for the sake of simplicity. Because of this assumption, we have to restrict our tax instruments to avoid trivial solutions. With no tax distortions, if demogrant in the first period were available, it would call for a 100% tax on earnings in that period.
5. We implicitly assume that the tax rate chosen today remains the same in the next period: $\tau_{t+1} = \tau_t$. In other terms, the individuals believe that the decision they take today will apply for the whole future (or at least for their retirement period). One can show that this assumption of once-and-for-all voting could be endogenized by considering a repeated elections model in which implicit contracts among successive generations arise, such as in Hammond (1975) or Boldrin and Rustichini (2000).
6. A pension system might also induce people to retire earlier when the amount of pension benefit foregone if working one more year is not compensated by a corresponding increase in the pension level. This effect would be taken into account in our model if P were decreasing in z.
7. The results would not change if a contributory element were introduced in P. In that case we would have $P = P_C + P_F$ where P_C is proportional to contributions $\tau w[1+n+z\theta]$, while P_F is flat i.e., independent of w and z.
8. With our assumption that $\beta(1+r) = 1$ this holds when $r = n$.
9. When $r > n$, there are no pay as you go pensions in the first best and $(1+r)$ replaces $(1+n)$ in (9) and (10). This case can arise in our setting because r is exogenous (for instance because we have a small open economy). In a more sophisticated model with capital accumulation and endogenous r, the first-best necessarily implies $r = n$, namely the golden rule (at least as long as we restrict ourself to the steady state).
10. In other words, restrictions on the instruments are imposed to keep the voting problem tractable. The normative analysis could be pursued in a much more general setting; see Cremer, Lozachmeur and Pestieau (2004).
11. In the non linear case (benefit formula and tax), this result remains valid, except for the "top" individual; see Cremer, Lozachmeur and Pestieau (2004).
12. Second period income (i.e., wz) is less than the average if and only if $w < \sqrt{E(w^2)}$.
13. This property arises because z does not depend on the first period tax. This is because with a monetary disutility, z is not subject to income effets.

14. Except that existence may be more problematic.
15. Under plausible assumptions on wage distribution and labor supply (retirement age) elasticity.

References

Boldrin, M. and A. Rustichini. (2000). "Political Equilibria with Social Security," *Review of Economic Dynamics* 3, 41–78.

Casamatta, G., H. Cremer and P. Pestieau. (2000). "The Political Economy of Social Security," *Scandinavian Journal of Economics* 102, 503–522.

Conde Ruiz, J. I. and V. Galasso. (2003). "Early Retirement," *Review of Economic Dynamics* 6, 12–36.

Conde Ruiz, J. I. and V. Galasso. (2004). "The Macroeconomics of Early Retirement," *Journal of Public Economics*, forthcoming.

Crawford, V. and D. M. Lilien. (1981). "Social Security and the Retirement Decision," *Quarterly Journal of Economics* 95, 505–529.

Cremer, H., J.-M. Lozachmeur and P. Pestieau. (2004). "Social Security and Variable Retirement Schemes: An Optimal Taxation Approach," *Journal of Public Economics*, 88, 2259–2281.

Galasso, V. and P. Profeta. (2002). "Political Economy Models of Social Security: A Survey," *European Journal of Political Economy* 18, 1–29.

Gans, J. S. and M. Smart. (1996). "Majority Voting with Single-Crossing Preferences," *Journal of Public Economics* 59, 219–237.

Gruber, J. and D. Wise. (1999). "Intorduction and Summary." In J. Gruber and D. Wise (eds.), *Social Security and Retirement Around the World.* Chicago: University of Chicago Press.

Hammond, P. (1975). "Charity: Altruism or Cooperative Egoism." In E. Phelps (ed.), *Altruism, Morality and Economic Theory*. New York: Russel Sage Foundation.

Lacomba, J. A. and F. M. Lagos. (1999). "Social Security and Political Election on Retirement Age," *Mimeo*.

Shepsle, K. A. (1979). "Institutional Arrangements and Equilibrium in Multidimensional Voting Models," *American Journal of Political Science* 23, 27–59.

Sheshinski, E. (1978). "A Model of Social Security and Retirement Decisions," *Journal of Public Economics* 10, 337–360.

Tabellini, G. (2000). "A Positive Theory of Social Security," *Scandinavian Journal of Economics* 102, 523–545.