Società italiana di economia pubblica

WORKING PAPERS
No. 118/2002

RETIREMENT AND SOCIAL SECURITY IN A PROBABILISTIC VOTING MODEL

.

Paola Profeta

:

JEL Classification: H53, H55, J26
Keywords: Political Economy, Multidimensional Voting, Merit Goods, Single-mindedness, Endogenous Ideological Homogeneity,

January, 2002

Retirement and Social Security in a Probabilistic Voting Model*

Paola Profeta
Università di Pavia and Università Bocconi

January 2002

Abstract

Why are social security transfers associated with retirement rules? This paper focuses on the political interactions between retirement and social security. Using a probabilistic voting approach, it analyzes why old people are induced to retire in order to receive pension transfers from the young. A crucial hypothesis is that leisure in old age represents a "merit good", which is positively valued by all agents in the society, young and old. Thus, the politicians choose to tax the labor income of the old, to induce them to retire. Retirement increases the level of ideological homogeneity of the old. In fact, once retired, the elderly are more "single-minded", since they only care about redistributive issues, such as pensions. This increase in their political power allows them to win the political game and to receive a positive transfer from the young (social security).

Keywords: Political Economy, Multidimensional Voting, Merit Goods, Single-mindedness, Endogenous Ideological Homogeneity, .

JEL Classification: H53, H55, J26

*Paola Profeta. Dipartimento di Economia Pubblica, Università di Pavia, paola.profeta@unipv.it. I thank Vincenzo Galasso for helpful discussions and suggestions, and exceptional support. I am grateful to Ignacio Conde Ruiz, Michael Reiter, Guido Tabellini, Heinrich Ursprung, participants to Universitat Pompeu Fabra seminars, to Università Bocconi Macroeconomic lunch seminars and to the 2001 annual meeting of the IIPF. I am especially grateful to Xavier Sala-i-Martin for his continuous advice and suggestions. All remaining errors are mine. Part of this paper was written while I was visiting Columbia University of New York.

1. Introduction

The relation between social security and retirement is a very crucial topic in the current debate on pension systems. The link from social security to retirement, i.e. the fact that the existence, design and size of the pensions programs have an impact on the labor participation decisions of elderly workers, is well understood. Social security provides income for the elderly people upon which they can decide to retire, and specific provisions included in the programs may differently affect the labor force participation rates. As recently summarized by Gruber and Wise (1999), in many countries around the world social security provisions include strong incentives to leave the labor force early. The link from retirement to social security has received much less attention. Many studies have analyzed the determinants of the size of the social security systems around the world focusing on the demographic factors, the financial "performance" of the system, the income distribution, and the income factors (see Galasso and Profeta (2001) for a review), while neglecting the role of retirement. However, a common feature of the current social security programs is that social security transfers are subject to retirement regulations. In spite of differences in many other relevant features, in the majority of countries around the world, the elderly are forced or induced to exit the labor market in order to collect their pensions. According to Mulligan and Sala-i-Martin (1999) in the 75% of the countries in their sample, social security benefit formulas induce retirement or explicitely require retirement[1]. Thus, should retirement be included among the determinants of social security? Is there a link running from retirement to social security? This is also a crucial question in the current debate on social security reforms, since many countries have introduced more flexibility in the mandatory retirement rule to cope with the sustainability of current programs under the aging process.

This paper examines retirement and social security programs in a multidimensional voting model, and provides a political economy explanation of the contemporaneous existence of pension transfers and retirement rules. Probabilistic voting is used to solve a bi-dimensional policy decision on retirement regulation and pension benefits.[2] In the political equilibrium of a probabilistic voting game, old people choose to retire and receive social security benefits from the young, and

[1] Different criteria are considered: retirement is required, benefits after the earliest retirement age decline with labor income, "unfair" credits are paid to those delaying retirement, current retirees are covered by previous law inducing retirement.

[2] Traditional political economy models of social security based on the median voter's theorem are not able to solve multidimensional policy decisions, since in a multidimensional space Nash equilibria of a majoritarian voting game generally fail to exist. The literature provides few types of solutions of this problem: the agenda-setting, the structure-induced equilibrium (Shepsle (1979)) and probabilistic voting.

the retirement regulation is necessary to obtain a positive level of social security.

The intuition for the result is the following. In a society composed of young and old individuals, both groups attach a positive value to old-age leisure. When making their optimal choice of leisure, the old do not consider that their consumption of leisure provides a positive externality to the young. As a consequence, votes maximizing politicians are induced by the young to impose a tax on the wage income of the old, so that the old increase their level of leisure. Therefore, retirement in old age arises as a solution to an externality problem. Additional political aspects are needed to explain the existence of social security. Since they retire, the old are a more ideologically homogeneous group than the young. This is because with no labor income they do not have to care about different issues related to their occupations or jobs, and focus exclusively on redistributive issues. Thus, the old can exert more political power than the young, win the political game, and obtain a positive transfer from the young, i.e. social security.

There are two crucial elements in this political economy model which drive the results. First, leisure when old is a "merit good". The young value the consumption of this good by the elderly, and support policies which induce retirement to favor old age leisure. This policy takes the form of a tax on old age labor income, which increases their leisure. This explains the existence of retirement regulations which are supported by elderly, as well as by young people. In fact, in western countries, unions strongly support retirement regulations, a fact that cannot be explained by the "political power" of the old[3]. Although merit goods have never been adopted in the context of political economy theories of social security, they have an old tradition.[4]. The fundamental idea is that these goods represent a value for the all community, which is considered by the individual preferences (Musgrave, 1988). In the public finance literature, they have been examined by Harberger (1984). More closely related to this paper, in the context of redistribution, Mulligan and Philipson (1999) have argued that merit goods imply in-kind redistribution of the goods which the donor considers meritorious for the donee. They suggest that many programs targeted to the poor, such as government health insurance, compulsory schooling, public housing, can be considered of this type. The rich seem to value consumption of these goods by the poor, and are willing to redistribute income to them, but only for these specific purposes. Mandatory retirement related to social security can also be interpreted as a program of this type. A government program that redistributes from the young to the old represents a mechanism (a "merit-good contract") that the young use to

[3]As in Mulligan and Sala-i-Martin (1999a).

[4]Philosophical, psychological, and moral interpretations of merit preferences are due to Brennan and Lomansky (1983), Brennan and Walsh (1977), Head (1966), Musgrave (1957, 1986, 1988), Olson (1980).

"help" the old to enjoy old-age leisure, which they feel to be meritorious. These merit goods motives for redistribution are supported by evidence from the private non-profit sector, in which several private organizations (churches or others) complement, or even substitute for the mandatory government programs in gathering help from rich to poor and elderly individuals. Similar reasoning, in the spirit of merit goods, may justify redistribution in the public sector.

The second crucial idea is that the degree of ideological homogeneity of the members in each group, which ultimately determines the political power of the group, depends positively on their level of leisure. To my knowledge this is the first attempt to formally endogenize the degree of ideological homogeneity in a probabilistic voting model. This idea resembles Mulligan and Sala-i-Martin's (1999a) notion of "single-mindedness". While workers care about a large variety of issues, related to their different occupations and industries, non-workers (e.g. retirees) are more united in their political action. The old turn out to be more politically successful because they focus on a single issue, such as pensions. Mulligan and Sala-i-Martin (1999a) provide empirical support for this "single-mindedness" hypothesis using cross-country government finance data and cross-country political participation surveys.

The model is developed as follows. Society is divided in two groups of voters, young and old. The two groups may have different wage and size. Every individual in a group has a specific political preference, i.e., an ideology, which contributes to her voting decision. The degree of ideological homogeneity of members of the same group is captured by a density function. This function is endogenous. It depends positively on leisure, reflecting the fact that when the individuals in a group enjoy more leisure, they are on average more united in their ideological preferences, since they all care mainly about redistributive policies. In the limit, when they don't work, they focus on a single issue. This density function determines the political power of each group, since more ideologically homogenous groups are more politically successful. Two candidates are involved in the electoral competition. They act simultaneously and do not cooperate. Before the election takes place, they make binding commitments to policy platforms; rational voters select their most preferred policy platform. These policy platforms are multidimensional. They include two instruments of intragenerational redistribution, i.e., the group-specific tax rates on wage income (with taxes rebated lump-sum to members of each group), and an instrument of intergenerational redistribution, the lump-sum transfer. The first policy is affected by the "merit goods" motive. Since there exists an externality, a corrective positive tax on old age wage income is introduced, to increase the leisure of the old and induce them to retire. On the contrary, the wage income tax rate for the young is zero, since the tax is distortionary. The outcome of the intergenerational policy depends on the political power of each group. Since they are induced to retire, the endogeneity of the density function implies that the old will have a higher level of ideological homogeneity. This in turn increases their political power and allows them to receive a positive transfer from the other group (social security). Thus, the contemporaneous existence of social security transfers and retirement rules emerges as political equilibrium of a multidimensional probabilistic voting game.

The paper proceeds as follows: Section 2 introduces the model, its general features and the individual's and group's problem. Section 3 solves the model for the optimal choice of the multidimensional policy platform. Section 4 concludes.

2. The Model

2.1. General Features

I consider a two-periods overlapping generations model. Society is composed of two groups of voters, young and old, denoted by $i = y, o$. In each group there is a continuum of voters with unit mass. The two groups have different size: $n^o \neq n^y$, where n^i is the size of group i. Individual's preferences are identical within groups and depend on consumption and leisure, according to a quasi-linear utility function[5]. Agents are endowed with one unit of time in youth and in old age.

The preferences of the old depend on their consumption (c^o) and leisure (l^o):

$$u(c^o, l^o) = c^o + \psi^o H(l^o) \tag{2.1}$$

where ψ^o represents the intrinsic preference of the old for leisure and H is increasing and concave in leisure: $H' > 0, H'' < 0$.

The old consume all their income:

$$c^o = w^o(1 - \tau^o)(1 - l^o) + A + b^o + T^o \tag{2.2}$$

where c^i is the private consumption of group i , l^i is the leisure of group i, w^i is the unitary wage per hour worked, τ^i is the tax rate on wage income, A is saving from the previous period (asset income), b^i and T^i are transfers to the individual in group i.

The preferences of the young are given by:

$$u(c^y, l^y) = c^y + \psi^y H(l^y) + \varphi^y H(l^o) + \beta u(c^{o\prime}, l^{o\prime}) \tag{2.3}$$

[5] Quasi-linearity simplifies the model since the income effects only show up in the linear component, i.e. consumption. It is a common assumption in this kind of redistribution models. See Persson and Tabellini (2000) for a review of these models.

where β is the individual discount factor , ψ^y is the intrinsic preference of young for leisure and $'$ refers to the next period variables: young knows that they will be old in the next period.

The intrinsic value of leisure for the old is assumed not to be lower than the intrinsic value of leisure for the young: $\psi^o \geq \psi^y$. Individuals desire to supply less labor when old, because old-age leisure has an higher value for them than leisure in youth (for example, because effective time endowments in old age is reduced due to health considerations).

The utility function of the young includes the leisure of the old, weighted by the parameter φ^y. Following Mulligan and Philipson (1999), the leisure of the old is assumed to be a "merit good", which therefore provides positive utility to both old and young. Young enjoy if the old work less and spend more time in leisure. Notice that this merit good assumption implies that the utility function of the young differs from the one of the old; however young people know that when they become old they will have the utility function of the old.

The young can consume or accumulate their income. Their budget constraint is:

$$c^y + A = w^y(1-\tau^y)(1-l^y) + b^y + T^y \tag{2.4}$$

where R is the discount rate.

There are two types of transfers: intergenerational, i.e., across cohorts, and intragenerational, i.e., within cohorts. They satisfy the following constraints:

$$\begin{aligned} T^o &= \tau^o w^o(1-l^o) \\ T^y &= \tau^y w^y(1-l^y) \end{aligned} \tag{2.5}$$

$$\begin{aligned} n^o b^o + n^y b^y + \alpha \left| n^o b^o \right| \left| n^y b^y \right| &= 0, \alpha > 0 \\ b^o b^y &< 0 \end{aligned} \tag{2.6}$$

The constraints make explicit that the tax τ is an instrument of intragenerational redistribution and the transfer b is a policy for intergenerational redistribution. The first and the second equations indicate that labor income tax revenues are rebated lump-sum to the members of the group. The old and the young set a distortionary tax on income and redistribute revenues lump sum to the members of their group. Notice that this policy is in principle inefficient, and that the optimal level of tax rate should be zero for both groups. However, as it will be shown below, a positive tax rate of the old may be induced by the externality of the old age leisure in the young utility.

The third equation represents the balanced budget constraint of the intergenerational program. The total transfers between groups and the amount of resources needed to carry on this process have to sum up to zero. The intergenerational transfer involves an efficiency loss. Resources can be shifted from one group to another, but this procedure entails a cost. The cost is represented by the term $\alpha |n^o b^o| |n^y b^y|$, i.e., it depends in a quadratic way on the size of the total transfer. This term may represent bureaucracy's costs, or rents to the politicians. Thus, social security is a system which redistributes resources from the young to the old group or viceversa, at a cost that depends on its own size.[6]The last constraint rules out the case of both transfers being negative, which would represent a system in which bureaucracy extracts resources from both groups, and there is no redistribution.

In this setting, a tax on labor income and a social security transfer are contemporaneously introduced[7]. The use of those instruments allows me to take into account the fact that the majority of social security programs around the world combine transfers from young to old with strong incentives to retire (Mulligan and Sala-i-Martin (1999b)), and that social security benefits are only weakly related to contributions.

To rule out the case in which a difference in wage levels is solely responsible for the existence of retirement[8], I assume that current old and young have the same wage and that an individual receives the same wage in youth and in old age ($w^o = w^y = w^{o\prime} = w$). In other words, the model will explain retirement independently from any "ex-ante" economic difference between the two groups.[9]

The public policy vector q is defined by a quadruple $q = (\tau^o, \tau^y, b^o, b^y)$, composed of the two tax rates and the intergenerational transfers.

---

[6]Notice that the results derived in the paper hold true also if α is assumed to be zero. However, in this case, because of the quasi-linear specification, there would be a corner solution, with old people extracting all resources from the young. The model would thus provide the same predictions, namely the existence of retirement and social security transfers, with the maximum value of transfers to the old.

[7]To treat separately the two sources of transfers, I introduce different budget constraints. The results would not change with a unique budget constraint in which total revenues from taxes (from the young and the old group) were used to finance the intergenerational transfers.

[8]Retirement will be defined as the positive difference between the level of leisure chosen in old age and in youth.

[9]Clearly, the result holds true if the old have lower unitary wages than the young. Specifically, I can assume that when an individual becomes old, her unitary wage is not larger than the wage received in her youth, and not larger than the wage received by the current young: $w^o \leq w^y$ and $w^{o\prime} \leq w^y$. The lower labor productivity of the old can be justified by the depreciation of human capital. In Mulligan (1998), this hypothesis is supported by cross-sectional age-average hourly earnings. Mulligan and Sala-i-Martin (1999a) argue that the labor productivity of the old is lower than what age-earnings profile often suggest, due to Lazear-type (1979) long-term employment contracts, which imply that earnings are not just payment for labor services rendered at the time, but also a return on past investment. Kotlikoff and Gokhale (1992) provide estimations which support this hypothesis.

To handle political equilibria, since the space is inherently multi-dimensional, I use a model with probabilistic voting (as in Persson, Roland and Tabellini (1998a), Lindbeck and Weibull (1987), which in turn build on probabilistic voting models by Hinich et al.(1972), Coughlin and Nitzan (1981a, 1981b), Coughlin (1992)).

Consider two parties, or candidates, labeled A and B. Before the election takes place, the parties commit to a policy platform, q^A and q^B. They act simultaneously and do not cooperate. Each party chooses the platform which maximizes its expected number of votes[10].

Platforms are chosen when the election outcome is still uncertain. The two parties differ along some other dimension relevant to the voters than the announced policies, and which may reflect ideological elements. Moreover, voters are heterogenous with respect to their ideological preferences. Voter j in group i votes for party A if

$$V^i(q^A) + \delta + \sigma^j > V^i(q^B) \tag{2.7}$$

where $V^i(q^A)$ is the indirect utility of voters in group i under government policy q^A and the term $(\delta + \sigma^j) \gtrless 0$ reflects voter j's ideological preferences for party A. This term includes two components, δ, which is common to all voters, and σ^j, which is idiosyncratic.

The first component, δ, reflects the general popularity of party A. This is a random variable, which I assume to be uniformly distributed on $(-1/2d, 1/2d)$. Its expected value is zero and the density is d. This component represents the source of electoral uncertainty, since it is assumed that δ is realized between the announcement of the party platforms and the election.

The second component, σ^j, reflects the individual ideology of voter j. Voters are distributed within each group according to a uniform distribution on $(-1/2s^i, 1/2s^i)$. The density is s^i and the mean is zero.[11]

Furthermore, I assume that the density is a positive function of the level of leisure:

$$s^i = s\left(l^i\right)$$

with $s' > 0$.

This is a crucial assumption, that can be motivated as follows: s^i can be seen as representing the level of "political single-mindedness" of the group i. I think

[10]This approach is extensively used in literature (see Couglin (1992) for a review). Alternatively, and without changing the results, the objective of the party can be to maximize the probability of winning, which in turn depends on the electoral rule, as in Persson, Roland and Tabellini (1998a).

[11]In general, both δ and σ^j may have expected values that differ from zero. Suppose σ^j is a random variable with uniform distribution on $(-1/2s^i + \overline{\sigma}^i, 1/2s^i + \overline{\sigma}^i)$. The density is s^i and the means is $\overline{\sigma}^i$. The specific means $\overline{\sigma}^i$ reflect the across groups difference in average ideology. Here, I assume that voters of both groups are on average ideologically neutral: $\overline{\sigma}^o = \overline{\sigma}^y = 0$.

of every citizen as having a fixed amount of political resources which she may allocate among different issues. Higher s^i means that individuals in group i are more homogeneous in their political action, focusing on a single "issue". Here, higher levels of leisure induce higher ideological homogeneity. This generalizes the "single-mindedness" hypothesis, introduced by Mulligan and Sala-i-Martin (1999a). While workers care about several opposite issues (members of different industries and different occupations tend to focus on programs that subsidize their own industry or occupation), nonworkers tend to have fewer special interests and are more united in their political action. They mainly care about redistributive programs. Moreover, workers may care about several opposite issues, depending on their labor income: low income agents rely more on redistributive programs, while high income agents care about several issues (labor income, property taxes, etc..). On the contrary, nonworkers, who do not have other sources of income than the transfers, are more homogeneous in their economic interests.[12]

Each group has neutral voters, called "swing voters", who are indifferent between party A and B. The identity of the swing voters is crucial when a party considers whether to deviate from a common policy announcement, $q^A = q^B$, or not. Suppose party A decides a unitary increase in the transfer to group o financed by a budget-balanced decrease in the transfer to group y. Party A expects a gain of votes from group o equal to the number of swing voters in group o, and a loss of votes from group y equal to the number of swing voters in group y. If group o has a higher number of swing voters, this will lead to a net gain of votes. As a consequence, each party tries to attract the more mobile voters. Figure 1 illustrates these features[13]. Formally, the swing voter[14] in group i is identified by $\sigma^{s.v.}$ where

$$\sigma^{s.v.} = V^i(q^B) - V^i(q^A) - \delta \tag{2.8}$$

Voters with σ^j lower than $\sigma^{s.v.}$ vote for B and voters with σ^j higher than $\sigma^{s.v.}$ vote for A.

[12] A microeconomic foundation of the function $s^i = s\left(l^i\right)$with $s' > 0$ can be derived by considering the optimization problem of individuals who differ in their sensitivity with respect to leisure. In this environment, agents who care more about leisure are more affected by redistributive policies than by alternative policies, such as occupational related subsidies or provision of public goods. In this sense, they are more ideologically homogenous and their votes are easier to be captured by redistributive policies.

[13] Notice that, when the two parties choose the same platform, the swing voter has a type $\sigma^{s.v.} = -\delta$. Party A's increase of the transfer to group o shifts the swing voter in group o to the left, and the swing voter in group y to the right, since there are less young voting for A.

[14] Notice that the existence of a swing voter depends on the supports of the distribution of δ and σ.

Therefore, the vote share of party A in group i can be expressed by

$$\pi^{A,i} = s^i \left[V^i(q^A) + \delta - V^i(q^B)\right] + \frac{1}{2} \tag{2.9}$$

To summarize, the timing of the game is the following: (i) the candidates, who know the distributions of δ and σ, choose their platforms, q^A and q^B; (ii) δ gets realized; and (iii) individuals in the two groups cast their vote.

2.2. The Party's Problem

Each party maximizes the expected total number of votes from the current young and old. Given the definition of $\pi^{A,i}$, the objective function of party A can be expressed as follows:

$$\max E(\sum_{i=y,o} n^i \pi^{A,i}) \tag{2.10}$$

Substituting the expression for $\pi^{A,i}$ and given the previous assumptions about the distribution functions, party A will choose q^A such as to maximize the following objective function:

$$\sum_{i=y,o} n^i s^i \left[V^i(q^A) - V^i(q^B)\right] \tag{2.11}$$

Clearly, if the number of swing voters is the same, the two groups get equal weight in the candidate's decision, which turns out to be maximizing the average voter's utility. However, if the two groups differ in how easily their votes can be swayed, the more ideologically homogeneous group has more swing voters, it is more responsive to policy, and gets a higher weight in the party's objective. In other terms, parties seek to please the more mobile voters.

2.3. The Individual's Problem

Every individual, young and old, solves her optimization problem. The old choose their level of leisure, whereas the young have to decide their level of leisure and their savings. To simplify the analysis, I will assume a logarithmic utility for the leisure: $H(l^i) = \log l^i$.

Consider first the problem of the old. Each individual in group *old* solves the following problem:

$$\begin{aligned} \max_{\{c^o, l^o\}} u(c^o, l^o) &= c^o + \psi^o \log l^o \\ s.t.\ c^o &= w(1-\tau^o)(1-l^o) + RA + b^o + T^o \end{aligned} \tag{2.12}$$

with w, τ^o, b^o, T^o, A given. It is easy to see that (from the first order condition and the budget constraint) the optimal levels of leisure and consumption are:

$l^o = \frac{\psi^o}{w(1-\tau^o)}$
$c^o = w(1-\tau^o) - \psi^o + A + b^o + T^o$
The corresponding indirect utility function is:

$$V^o = w(1-\tau^o) - \psi^o + A + b^o + T^o + \psi^o \log \psi^o - \psi^o \log w - \psi^o \log(1-\tau^o) \quad (2.13)$$

Each individual in group *young* solves the following problem:

$$\begin{aligned} \max_{\{c^y, l^y\}} u(c^y, l^y) &= c^y + \psi^y \log l^y + \varphi^y \log l^o + \beta(c^{o'} + \psi^{o\prime} \log l^{o'}) \quad (2.14) \\ s.t.\ c^y + \frac{c^{o\prime}}{R} &= w(1-\tau^y)(1-l^y) + b^y + T^y + \frac{w(1-\tau^{o\prime})(1-l^{o\prime}) + b^{o\prime} + T^{o\prime}}{R} \end{aligned}$$

for w, τ^y, b^y, T^y, $\tau^{o\prime}, b^{o\prime}$given. Let $\psi^{o\prime} = \psi^o$ and $w' = w$. Young people expect that $l^{o\prime} = \frac{\psi^o}{w(1-\tau^{o\prime})}$.

Notice that since the utility function is linear in consumption, the consumption profile only depends on the relation between the rate of time preference, β, and the exogenous discount rate, R.[15]. For the purpose of this analysis, I restrict savings to be non-negative $A \geq 0$ and impose that $\beta = 1/R$.

Again, it is easy to find the optimal decision in youth:
$l^y = \frac{\psi^y}{w(1-\tau^y)}$
$c^y + \beta c^{o\prime} = w(1-\tau^y) - \psi^y + b^y + T^y + 1/R(w(1-\tau^{o'}) - \psi^o + b^{o'} + T^{o'})$
The indirect utility function for the young is the following:

$$\begin{aligned} V^y &= w(1-\tau^y) - \psi^y + b^y + T^y + \psi^y \log \psi^y - \psi^y \log w - \psi^y \log(1-\tau^y) \quad (2.15) \\ &+\varphi^y \log \psi^o - \varphi^y \log w - \varphi^y \log(1-\tau^o) + \\ &1/R(w(1-\tau^{o'}) - \psi^o + b^{o'} + T^{o'} + \psi^o \log \psi^o - \psi^o \log w - \psi^o \log(1-\tau^{o'})) \end{aligned}$$

3. Solving the Model

I can now consider the optimization problem of party A (and symmetrically of party B):

$$\max_{\{q^A\}} \sum_{i=y,o} n^i s^i \left[V^i(q^A) - V^i(q^B)\right] \quad (3.1)$$

where $V^y(q^i)$ and $V^o(q^i)$ are defined respectively in equations 2.13 and 2.15. Moreover, the expressions for $T^o, T^{o\prime}$ and T^y in equilibrium are:
$T^o = \tau^o w(1-l^o) = \tau^o w - \frac{\tau^o \psi^o}{1-\tau^o}$
$T^y = \tau^y w(1-l^y) = \tau^y w - \frac{\tau^y \psi^y}{1-\tau^y}$

[15] Clearly, if $\beta > 1/R$ all consumption takes place in youth, and viceversa, if $\beta < 1/R$, in old age. For $\beta = 1/R$, agents are indifferent on how to split consumption.

$T^{o\prime} = \tau^{o\prime} w(1 - l^{o\prime}) = \tau^{o\prime} w - \frac{\tau^{o\prime}\psi^o}{1-\tau^{o\prime}}$

The political equilibrium depends on the agents' expectations about future policy decisions. Here, the equilibrium concept is based on the assumption of no commitment for intragenerational and intergenerational policies. Each party decides the tax rates and the benefit levels for the two current groups, taking as given the tax rate and the benefit level of the next period. Under this assumption, only current generations are relevant for the choice of the politician.[16]. In a stationary equilibrium, the young expect their tax rate and their old age benefit to be equal to the level of the tax rate and the benefit for the current old.

Parties act simultaneously, taking the choice of the other party as given, and do not cooperate. Thus, taking q^B as given, party A solves the problem at equation 3.1, where $V^y(q^A)$ and $V^o(q^A)$ are defined at equations 2.13 and 2.15, and subject to the following constraints:

$n^o b^o + n^y b^y + \alpha \left| n^o b^o \right| \left| n^y b^y \right| = 0, \alpha > 0$

$b^o b^y < 0$

$\tau^{o\prime} given$

$b^{o\prime} given$

Since the density function s is endogenous, i.e., $s^o = s(l^o)$ and $s^y = s(l^y)$, the first order conditions are the following:

$$FOC\left\{\tau^o\right\}: n^o \frac{ds^o}{d\tau^o}(V^{oA} - V^{oB}) + n^o s^o \frac{dV^o}{d\tau^o} + n^y s^y \frac{dV^y}{d\tau^o} = 0 \tag{3.2}$$

$$FOC\left\{\tau^y\right\}: n^y \frac{ds^y}{d\tau^y}(V^{yA} - V^{yB}) + n^y s^y \frac{dV^y}{d\tau^y} = 0 \tag{3.3}$$

$$FOC\left\{b^o\right\}: n^o s^o = \lambda(n^o - n^o n^y \alpha b^y) \tag{3.4}$$

$$FOC\left\{b^y\right\}: n^y s^y = \lambda(n^y - n^o n^y \alpha b^o) \tag{3.5}$$

3.1. Tax Rates and Retirement

Since party B solves a symmetric problem, it can be showed that in equilibrium the two parties choose the same platform[17], $q^A = q^B$, and thus the utility levels reached by the individuals are the same, $V^{iA} = V^{iB}$ $(i = o, y)$.

[16]Notice that this is a very conservative assumption. In fact, I consider the more difficult environment for explaining the existence of social security transfers, since current young will not expect to receive a benefit when they become old. The result holds true if, as in Mulligan and Sala-i-Martin (1999a), I assume that the transfer that young people expect to receive in the next period is a percentage ρ $(0 \leq \rho \leq 1)$ of the transfer that current old people are receiving. If $\rho = 0$ there is no commitment, and if $\rho = 1$ there is full commitment, i.e., young people know that they will receive the same transfer as the current old.

[17]See Coughlin (1992).

Therefore, after substituting for the expressions of the derivatives of the indirect utilities with respect to the old group tax rate, the first order condition with respect to τ^o becomes:

$$n^o s^o \left(-\frac{\psi^o \tau^o}{(1-\tau^o)^2} \right) + n^y s^y \frac{\varphi^y}{1-\tau^o} = 0 \tag{3.6}$$

The first order condition for τ^y is instead:
$n^y s^y \frac{dV^y}{d\tau^y} = 0$, which can be written as

$$n^y s^y \left(-\frac{\psi^y \tau^y}{(1-\tau^y)^2} \right) = 0 \tag{3.7}$$

Proposition 3.1. *The old group sets a positive tax rate. The young group sets a zero tax rate.*

Proof. See Appendix. ■

The intuition is the following. For both groups, the tax has an economic cost, due to the decrease in consumption which cannot be compensated by the increase in leisure ($-\frac{1}{(1-\tau^i)^2} + \frac{1}{1-\tau^i} < 0$ as long as $\tau^i > 0$). Since lump-sum transfers are available, taxation will always be inefficient. For the young, this is the only effect of tax, and thus the tax rate is set equal to zero. For the old, there exists an additional effect, due to the value of the tax in the welfare of the young ($n^y s^y \frac{\varphi^y}{1-\tau^o}$). In other words, the "merit good" hypothesis implies that young people care about the leisure of the old. This represents a positive externality, since the old do not take into account that their choice of leisure has a positive value for the young. Thus, a positive tax on old wage income is introduced, which induces them to enjoy more leisure.

Definition 3.2. *Retirement is defined as* $R^o_{t+1} \equiv l^o_{t+1} - l^y_t = l^{o\prime} - l^y$ *if* $R^o_{t+1} > 0$.

This definition implies that leisure is interpreted as extensive margin, i.e. the number of years worked in each period. In a two-periods OLG model, leisure different from 1 for the old, but larger than leisure of the young would mean that the old work for few years (e.g. from 55 to 60).

Proposition 3.3. *Retirement exists* ($l^{o\prime} > l^y$).

Proof. See Appendix. ■

Corollary 3.4. *The old are more ideologically homogeneous ("more single-minded") than the young* ($s^o > s^y$).

Proof. See Appendix. ■

These propositions and the corollary show that retirement derives as the equilibrium outcome of a democratic voting process when leisure of the old is a "merit good" ($\frac{dV^y}{dl^o} > 0$). Young people's preferences induce the politicians to set a positive tax rate on the labor income of the old, to induce them to increase their level of leisure. This tax corrects the externality: the old do not take fully into account the effects of their choice of leisure on the economic welfare, since they do not consider that their leisure has a positive utility for the young. The tax has the function of inducing the old to increase their optimal level of leisure, in other words to retire. Retirement induces a higher ideological homogeneity among the old, under the assumptions on the s function.

3.1.1. The Tax Rate for a Constant Elasticity Density Function

In this section I consider a density function with a constant positive elasticity equal to ε:

$$s^i = \left(l^i\right)^{\varepsilon}, \varepsilon > 0$$

Under this assumption, equation 3.6, that defines the equilibrium tax rate for the old, can be written as follows:

$$\frac{\tau^o}{(1-\tau^o)^{\varepsilon+1}} = \frac{n^y}{n^o}\varphi^y \frac{(\psi^y)^{\varepsilon}}{(\psi^o)^{\varepsilon+1}}$$

It is immediate to show that the following results hold:

1. $\frac{\partial \tau^o}{\partial \psi^o} < 0$ and $\frac{\partial \tau^o}{\partial \psi^y} > 0$
2. $\frac{\partial \tau^o}{\partial \varphi^y} > 0$
3. $\frac{\partial \tau^o}{\partial \left(\frac{n^o}{n^y}\right)} < 0$

1. When the old intrinsic preference for leisure increases, the tax rate decreases. There are two effects. On one hand, the marginal disutility from the tax for the old increases; on the other hand, their political relevance increases, due to an increase of s^o, because the old are more ideologically homogeneous when they enjoy more leisure. The two effects imply that the optimal tax rate has to decrease. On the other side, when the young intrinsic preference for leisure increases, there is an increase of s^y, which implies that the optimal tax rate has to rise.

2. When the young preference for old leisure increases, the external effect of old age leisure on young increases, which implies a higher correcting tax.

3. When the relative size of the old increases, the negative effect of the tax increases (because old are negatively affected by the tax) and the positive effect of the tax decreases (because fewer young are enjoying the tax inducing retirement for the old). Thus a lower tax rate is required. This result implies that the aging process will decrease retirement, since there will be relatively fewer young that enjoy retirement of the old.[18]

3.2. Optimal Benefit Rate

In this probabilistic voting environment, the higher ideological homogeneity of the old implies that the old group has more "swing voters", and therefore more political power, which allows them to obtain intergenerational transfers from the young.

Proposition 3.5. *There exist social security transfers from young to old:* $b^o > 0$ $b^y < 0$

Proof. See Appendix. ■

The previous proposition shows that in equilibrium the old receive a positive transfer from the young. This result replicates the current PAYG social security systems, where current young finance the pensions of current retirees. The model derives this result from the higher homogeneity of the old group. Because of their higher leisure, the old are more ideologically homogeneous, which gives them more political power to obtain resources from the young group.

Proposition 3.6. *The equilibrium levels of the transfers between young and old are the following:*

$$b^y = \frac{1 - \sqrt{\frac{s^o}{s^y}}}{\alpha n^y}$$

$$b^o = \frac{1 - \sqrt{\frac{s^y}{s^o}}}{\alpha n^o}$$

[18] Alternative frameworks may assume that the merit goods externality depends on the number of old people through a general function $f(n^o)$, in which case: $u(c^y, l^y) = c^y + \psi^y H(l^y) + f(n^o)\varphi^y \log(l^o) + \beta u(c^{o'}, l^{o\prime})$. In a scenario where $f(n^o)$ is a positive function of the number of old (young care more about the leisure of the old when they are more numerous, since they think that more of them will be affected), and $f(n^o)/n^o$ is still a positive function of the number of old, the equilibrium tax rate will increase with the aging of the population. On the other side, in a scenario where $f(n^o)$ is a decreasing function of the number of old (young become used to see many old people and they are less interested in their leisure), or $f(n^o)/n^o$is a negative function of the number of old, the aging of the population implies lower tax rate, i.e. lower tax inducing retirement. Lacking a clear motivation for choosing one of the two scenarios, I prefer to adopt a more neutral formulation, which does not consider size effects in the merit good externality.

Proof. See Appendix. ■

The equilibrium level of transfers decreases with the level of α, which reflects the amount of resources appropriated by the bureaucrats and therefore not redistributed between the two groups.

The transfer to a group increases with the density of this group, and decreases with the density of the opposite group (higher s^y implies larger b^y and smaller b^o, while higher s^o implies larger b^o and smaller b^y). This result[19] derives from the relation between the higher density and the larger number of swing voters. The group which is relatively more ideologically homogeneous (i.e. has the higher density s) contains more swing voters. Therefore, shifting resources from the opposite group towards this group represents a net gain of votes for the candidate, and therefore an optimal policy. Notice that due to the presence of bureaucracy costs, which increase with the overall size of these transfers, it is not optimal to fully expropriate the group with lower density, and to redistribute all the resources to the group with higher density.

Finally, the equilibrium level of the transfer for each group depends on the intrinsic preference for leisure of the two groups, on the preference of the young for the leisure of the old, and on the relative size of the two groups, through the effects that these variables have on the density of each group. In fact, I have shown in the previous section how these variables affect the optimal level of the tax rate, and thus the level of leisure and the density. In particular, $\frac{\partial \tau^o}{\partial\left(\frac{n^o}{n^y}\right)} < 0$ implies that $\frac{\partial s^o}{\partial\left(\frac{n^o}{n^y}\right)} < 0$, since, when the number of old increases, the cost of the tax increases, and therefore the optimal tax is reduced, thus inducing lower retirement and lower ideological concentration of the old, who thus obtain a lower transfer. On the contrary, when there are more young people, the optimal tax rate increases, and therefore retirement increases and the old are more ideologically concentrated and more politically successful.

4. Concluding Remarks

The model developed in this paper represents the first attempt to solve the puzzle of the association between retirement rules and social security in a political economy voting model. The explanation is based on two crucial features: old age leisure is a "merit good", and the political power of a group relies on its ideological homogeneity, which depends positively on the level of leisure. Under these circumstances, a democratic voting-maximizer policy-maker would set a positive tax on the wages of the old people, with revenues redistributed lump-sum within

[19] Notice that the sign of the transfers does not depend on the size of the groups, but rather on the relative density between the two groups.

the generation. This tax induces the old to retire, and thus to become more ideologically homogeneous. This makes them politically more powerful, and allows them to receive a positive intergenerational transfer from the young.

The model delivers additional interesting predictions. The endogenous political power of each group, and the results on the equilibrium level of the tax rate on old age wage income, may help to explain why retirement regulations have been changing so often in many countries, and what are the driving elements of these changes. Along this line, the model predicts why the aging process is a crucial element. The aging of the population is expected to decrease the retirement regulations, as it is happening in the U.S., where retirement used to be mandatory, and is becoming more flexible. And the reduced retirement would then decrease the per capita, as well as the aggregate social security size.

Moreover, Lindert (1996) finds evidence that a larger number of young people would imply a larger social security size. To my knowledge, this evidence is a feature unexplained by previous contributions in the political economy literature of social security. In the model here developed, a larger number of young people would lead to a higher tax inducing retirement of the old, thus increasing social security. This result suggests that reforms of retirement regulations implied by the aging process may allow to solve the demographic problem of social security.

The formulation here adopted is quite general, and could be extended in several directions. I will discuss two of them. First, while the model assumes the "merit goods" hypothesis to explain the existence of mandatory retirement, the interactions between merit goods motives and politics may explain many redistributive policies involving mandatory programs (public education, public health). This is a new promising area of research.

Second, it would be interesting to estimate empirically the concentration of different groups of voters (the number of swing voters for each group) by age, income and other characteristics. It is well known that highly educated, rich, white, elderly, media exposed males participate more in the voting process (Delli Caprini and Keeter, 1996). This evidence seems to support the idea developed in the model (elderly vote more). However, in my framework what is relevant is not the number of voters in a group (voting turnout), but rather the number of swing voters, i.e., of people who vote mainly according to the proposed policy platforms and are ready to change party as the policy changes (voter choice). This represents both a limit and a possible extension of the analysis of this paper. On one side, the assumption that, when voting on redistributive policies, the within-group concentration depends on the level of leisure, has to be supported with empirical estimates; on the other side, a general measure to estimate the number of swing voters related to a specific redistributive program has to be developed. There are still very few contributions in this direction. Stromberg (2000) studies the impact of mass media on government spending: mass media users have more swing voters and therefore more political power. He analyzes a major New Deal relief program implemented in the middle of the expansion period of the radio and shows that counties with many radio listeners received more relief funds. Though in a different framework, this analysis provides evidence that groups that have more swing voters enjoy more political power and obtain larger transfers.

5. Technical Appendix

5.1. Proof of Proposition 3.1

Proposition 3.1: The old group sets a positive tax rate. The young group sets a zero tax rate. Therefore, the old group is taxed heavier than the young group.

Proof: The old group sets a positive tax rate: From equation (3.6) it is: $\frac{\tau^o}{1-\tau^o} = \frac{n^y}{n^o}\frac{s^y}{s^o}\frac{\varphi^y}{\psi^o}$. Since the right hand side is positive (all terms are positive), the left hand side has to be positive, which implies $\tau^o > 0$.

The young group sets a zero tax rate: Directly from equation (3.7) which is always negative when $\tau^y > 0$.

Therefore the old group is taxed heavier than the young group: $\tau^o > \tau^y = 0$. Q.E.D.

5.2. Proof of Proposition 3.3

Proposition 3.3: Define retirement: $R^o_{t+1} \equiv l^o_{t+1} - l^y_t = l^{o\prime} - l^y$ if $R^o_{t+1} > 0$. At the steady state retirement exists ($l^{o\prime} > l^y$).

Proof: Since from the previous proposition, $\tau^o > \tau^y$, in steady state $\tau^{o\prime} = \tau^o > \tau^y$. Since $\psi^o \geq \psi^y$ by hypothesis, then $l^{o\prime} = \frac{\psi^o}{w(1-\tau^{o\prime})} > l^y = \frac{\psi^y}{w(1-\tau^y)}$. Q.E.D.

5.3. Proof of Corollary 3.4

Corollary 3.4: The old are more ideologically homogeneous (single-minded) than the young ($s^o > s^y$).

Proof: $\tau^o > \tau^y$ and $\psi^o \geq \psi^y$ imply that $l^o > l^y$. Since s is a positive function of l, $s^o = s(l^o) > s^y = s(l^y)$. Q.E.D.

5.4. Proof of Proposition 3.5

Proposition 3.5: There exist Social Security transfers from young to old: $b^o > 0$ $b^y < 0$

Proof.

From the first order conditions with respect to b^o and b^y, it is:

$\frac{s^o}{s^y} = \frac{1-n^y\alpha b^y}{1-n^o\alpha b^o}$

Since from the previous corollary it is $s^o > s^y$, it must be $(1 - n^y\alpha b^y) > (1 - n^o\alpha b^o)$, *i.e.*, $n^y\alpha b^y < n^o\alpha b^o$. Since $b^o b^y < 0$, and α, n^o, n^y are all positive, it must be $b^o > 0$ $b^y < 0$. Q.E.D.

5.5. Proof of Proposition 3.6

Proposition 3.6: The equilibrium levels of the transfer from young to old are the following:

$$b^y = \frac{1 - \sqrt{\frac{s^o}{s^y}}}{\alpha n^y}$$

$$b^o = \frac{1 - \sqrt{\frac{s^o}{s^y}}}{\alpha n^o}$$

Proof.
Given the budget constraint: $n^o b^o = \frac{-n^y b^y}{1-\alpha n^y b^y}$ and the equilibrium condition $\frac{s^o}{s^y} = \frac{1-n^y \alpha b^y}{1-n^o \alpha b^o}$, it is:
$\frac{s^o}{s^y} = \frac{1-n^y \alpha b^y}{1+\frac{\alpha n^y b^y}{1-\alpha n^y b^y}} = (1 - \alpha n^y b^y)^2$
Solving the second order equation, the solution is:
$b^y = \frac{1 \pm \sqrt{\frac{s^o}{s^y}}}{\alpha n^y}$
Since $b^y < 0$ it must be:

$$b^y = \frac{1 - \sqrt{\frac{s^o}{s^y}}}{\alpha n^y}$$

Substituting into the budget constraint, the equilibrium level of b^o is the following:

$$b^o = \frac{1 - \sqrt{\frac{s^o}{s^y}}}{\alpha n^o}$$

Q.E.D.

References

[1] Brennan, G. and Lomansky, L. (1983), "Institutional aspects of merit goods analysis", *Finanzarchiv.* 41(2): 183-206

[2] Brennan, G. and Walsh, C. (1977), "Pareto-Desirable Redistribution in Kind: An Impossibility Theorem, *American Economic Review* 67(5): 987-90

[3] Coughlin, P. (1992), "Probabilistic Voting Theory", Cambridge. *Cambridge University Press.*

[4] Coughlin-Nitzan (1981a), "Directional and Local Electoral Equilibria with Probabilistic Voting", *Journal of Economic Theory* 34, 1-12.

[5] Coughlin-Nitzan (1981b), "Electoral Outcomes with Probabilistic Voting and Nash Social Welfare Maxima", *Journal of Public Economics* 15, 113-121.

[6] Delli Caprini, M.X. and Scott, K. (1996), "What Americans Know about Politics and why it Matters", Yale University Press, 1996.

[7] Galasso, V. and Profeta, P. (2000), "The Political Economy of Social Security: A Survey", forthcoming in *European Journal of Political Economy*

[8] Gruber, J. amd Wise, D. (1999), "Social Security and retirement around the World", NBER

[9] Harberger, A.C. (1984), "Basic Needs versus Distributional Weights in Social Cost-Benefit Analysis", *Economic Development and Cultural Change,* 32(3): 455-74.

[10] Head, J. (1966), "On Merit Goods", *Finanzarchiv.* 25(1): 1-29

[11] Hinich, M.J., Ledyard, J.O., Ordershook, P.C., (1972), "Nonvoting and the Existence of the Equilibrium under Majority Vote", *Journal of Economic Theory* 44, 144-153.

[12] Kotlikoff, L.J. and Gokhale, J. "Estimating a Firm's Age-Productivity Profile Using the Present Value of Workers' Earnings", *Quarterly Journal of Economics* 107 (4), 1215-42.

[13] Lazear, E.P., "Why is There Mandatory Retirement?", *Journal of Political Economy* 87 (6), 1261-74.

[14] Lindbeck, A. Weibull, J.W. (1987), "Balanced-budget Redistribution as the Outcome of Political Competition", *Public Choice* 52, 273-297.

[15] Lindert, P.H. (1996), "What Limits Social Spending?", *Explorations in Economic History* 33, 1-34.

[16] Mulligan, C.B. (1998), "Substitution Over Time: Another Look at Life Cycle Labor Supply", *NBER Macroeconomics Annual,* Cambridge, MA: MIT Press

[17] Mulligan, C.B. and Philipson T.J. (1999), "Policy Design and Incidence in a Merit Good Economy", Mimeo, University of Chicago and NBER.

[18] Mulligan, C.B., and X. Sala-i-Martin (1999a), "Gerontocracy, Retirement and Social Security", NBER w.p. #7117.

[19] Mulligan C.B. and Sala-i-Martin, X. (1999b), "Retirement-Inducing Policy wears many Disguises", Mimeo University of Chicago, Columbia University and Universitat Pompeu Fabra and NBER.

[20] Mulligan C.B. and Sala-i-Martin, X. (1999c), "Social Security in Theory and Practice (II): Efficiency Theories, Narrative Theories, and Implications for Reform", NBER w.p. #7119

[21] Mueller, D. (1989), "Public Choice II". Cambridge. *Cambridge University Press.*

[22] Musgrave, R. (1957), "The Theory of Public Finance", New York-Mc Graw Hill.

[23] Musgrave, R. (1986), "On Merit Goods", in "Public Finance in a Democratic Society. Collected papers of Richard Musgrave". New York University Press.

[24] Musgrave, R. (1988), "Merit Goods", The New Palgrave: A Dictionary of Economics, Stocktone Press.

[25] Olson, M. (1971), "The Logic of Collective Action", Cambridge, MA. Harvard University Press.

[26] Persson, T. (1998), "Economic Policy and Special Interest Politics", *The Economic Journal* 108, 310-327

[27] Persson, T. Roland, G. Tabellini, G. (1998a), "Towards Micropolitical Foundations of Public Finance", *European Economic Review* 42, 685-694.

[28] Persson, T. Roland, G. Tabellini, G. (1998b), "Comparative Politics and Public Finance". W.P. 114, IGIER.

[29] Persson, T. Tabellini, G. (1999a), "Political Economics and Public Finance". Handbook of Public Economics.

[30] Persson, T. Tabellini, G. (1999b), "The Size and Scope of Government: Comparative Politics with Rational Politicians", *European Economic Review* 43, 699-735.

[31] Persson, T. Tabellini, G. (2000), "Political Economics. Explaining Economic Policy", MIT PRESS.

[32] Shepsle, K.A. (1979), "Institutional Arrangements and Equilibrium in Multidimensional Voting Models", *American Journal of Political Science*, 23 (1).

[33] Stromberg, D., (2000) "Radio's Impact on Public Spending", mimeo, Stockolm University

Figure 1: Probabilistic Voting