Centre for Studies in Economics and Finance

WORKING PAPER NO. 205

An Egg Today and a Chicken Tomorrow: A Model of Social Security with Quasi-Hyperbolic Discounting

Matteo Bassi

September 2008

University of Naples Federico II   University of Salerno   Bocconi University, Milan

CSEF - Centre for Studies in Economics and Finance – UNIVERSITY OF SALERNO
84084 FISCIANO (SA) - ITALY
Tel. +39 089 96 3167/3168 - Fax +39 089 96 3167 – e-mail: csef@unisa.it

WORKING PAPER NO. 205

An Egg Today and a Chicken Tomorrow: A Model of Social Security with Quasi-Hyperbolic Discounting

Matteo Bassi[♥]

Abstract
Strotz (1956) first suggested that individuals are more impatient when making short-run tradeoffs than long-run ones. Many experimental studies supports his conjecture. Motivated by recent evidence from the British Department of Work and Pension (2006), this paper applies this behavioral framework to retirement decisions. We propose a three-periods OLG model with quasi-hyperbolic consumers whosave for post retirement consumption in the first period and choose their retirement age in the second. We show that this behavioral assumption explains the observed drop in post retirement consumptiondue to lack of saving and the high level of voluntary (i.e. not due to disability or dismission from the firm) early exit from the labor force. When deciding about their retirement age, workers weight too much the costs of remaining at work (i.e. disutility of working, implicit tax on continued activity) and too little the benefits of postponed retirement (i.e. increase of the Bismarckian component of the pension formula), perceived as too far in the future. We investigate the implications of time inconsistent preferences for a political economy model in which voters determine simultaneously thesize and the degree of redistribution of the pension system. We show that, when voting over thepayroll tax, time inconsistent young workers, who look for a commitment device that increases boththeir saving and retirement age, form a coalition with rich in order to decrease the size of the system. When voting over the degree of redistribution, they form a coalition with poor individuals as to in-crease the at part of the pension formula. Our political model provides a political justification for the negative relationship between size and redistribution observed in most OECD countries (Disney 2004).

JEL classification: A12, D91, E21, H55, J64

Keywords: Hyperbolic Discounting, Majority Voting, Redistribution, Retirement Age, Saving Behaviour.

Acknowledgements: I wish to thank Helmuth Cremer, Georges Casamatta, Marco Pagano, Paola Profeta, Catarina Goulão and Marie-Louise Leroux for very helpful suggestions. I benefit from the comments of conference and seminar participants in Salerno (CSEF), Marseille, Budapest and Uppsala. The usual disclaimers apply.

[♥] Università di Salerno, CSEF and Toulouse School of Economics (GREMAQ).
Address: CSEF, Dipartimento di Economia, Università di Salerno, Via Ponte don Melillo, 84084 Fisciano (SA), Italy. E-mail: matteo.bassi@gmail.com

Table of contents

1 Introduction

The future financial viability of social security systems is a central issue in political debates across the world. In the last 20 years, most OECD countries have experienced a sharp increase in the population's proportion of elderly and retired individuals, and a more drastic shift in the same direction will take place in the near future, if the demographic trends continue. In particular, the increase in the number of retirees is due to the fact that increases in life expectancy has not been always followed by a corresponding increase of the minimal retirement age.

Governments have not yet understood the effects of early retirement on social security deficits, as most pension systems continue to have implicit or explicit features that allow workers to quit their workplace before the minimal age at almost no costs.

Not surprisingly, the combination of insufficient reforms and generous early retirement provisions[1] has generated severe pension crisis, especially in European countries. It is not surprising, therefore, that governments are trying to deeply reform their pension systems, either by increasing the minimal retirement age or by tightening the link between pension benefits and past earnings. Although necessary, reforming the pension system is a complicated task. For instance, the two reforms proposed above have a very different political appeal: while the former is heavily opposed by workers and unions and thus appear to be politically unfeasible, the latter receives more support, since it introduces "incentives" and "disincentives" that are supposed to induce workers to postpone efficiently and autonomously their retirement age, without any imposition by the government.

Before implementing any reform, we believe that the policymaker should be provided with a complete theoretical framework that, from one side, illustrates the determinants of early retirement and that, from the other side, explains how workers would respond to the incentives introduced by an eventual reform. Concerning the first aspect, the economic literature (Mulligan and Sala-i-Martin, 1996, and Conde Ruiz and Galasso, 2004) emphasizes the role of firms (the *push* argument): early retirement concerns primary low skilled workers, who have been dismissed by their employer either because they are more likely to receive a negative shock on their productivity level, or because they create a negative externality on young workers. Moreover, as stressed by Gruber an Wise (2000), the pension system itself induces early retirement, through an implicit tax on continued activity[2]. Explanations of early retirement based solely on exogenous factors are not entirely justified from an empirical point of view (DWP, 2006): according to recent evidence, retiring earlier than planned appears to be an voluntary choice, and depends on individuals' preferences for

[1] We define early retirement as the discrepancy between planned and effective age of retirement (Blondal and Scarpetta, 1998).

[2] To better understand the concept of implicit tax, consider, for example, a 59 years-old worker: what is the change in his pension benefits if he retires at age 60 instead of age 59? We call *accrual rate* the difference between the two pensions. If this rate is positive, working an additional year increases the total compensation; if the accrual is negative, working more reduces the pension. Thus a negative accrual rate discourages continuation in the labor force and a positive one encourages continued labor force participation. We have an *implicit tax* on continued activity whenever the ratio of the accrual rate to net wage earnings is negative, otherwise we have an *implicit subsidy*. Gruber and Wise (2000) show that the accrual rate and the associated tax are negative at older ages: continuation in the labor force reduces pension benefits, providing an incentive to leave the labor force earlier.

leisure (the *pull* argument): ex-ante, individuals overestimate their retirement age but, once they realize that accumulated savings and pension benefits are below their estimated values, regret ex-post about their lack of commitment.

Based on these observations, this paper presents a different view of early retirement, based on the *pull* argument, that complements both the push and the implicit tax explanations. More precisely, following the recent developments of the Economics and Psychology literature, and in order to fill the gap between the perfect rationality assumed in Public Economics and Political Economy and the bounded rationality observed in experiments (Kahneman and Tversky, 1979), we introduce hyperbolic discounting in individuals' preferences[3].

Hyperbolic individuals, when facing intertemporal trade-offs, change their preferences over time, such that what is preferred at one point in time far in the future is inconsistent with what is preferred today. An example by Thaler (1981) illustrates this point: if a person has to choose between an apple in 100 days or two apples in 101 days, he will probably prefer the second option. However, proposing the same trade-off between today and tomorrow, if the individual has a high preferences for today's utility, his choice may change and the first alternative becomes the preferred one. This example shows how certain agents are more impatient in the short-term than in the long-run. Laboratory and field studies confirm this intuition and find that discount rates are much greater in the short run than in the long run. It follows that present-biased individuals increase their utility level if a commitment device that force them to stick with the long run plans would have been made available to them (Laibson, 1997).

Notice that the inability of individuals to respond or to understand correctly intertemporal incentives introduced by a social security reform can seriously undermine the success of the reform itself, as stressed by Chan and Stevens (2004). In the paper, making use of repeated observation of individuals' retirement expectations, they investigate the effects of retirement incentives on those expectations. They show that, contrary to most of the literature on the topic, that, even if workers consider forward looking incentives when making their retirement plans, the magnitude of this responsiveness to pension related incentives was overestimated. Tightening the link between pension and retirement age, therefore, may be an ineffective way to overcome the financial crisis of the system, if workers are not able to respond effectively to the incentives introduced by such reform

This paper considers three sources of heterogeneity among agents: the first two, productivity level and age, are common features of most political economy model of social security (see Galasso and Profeta, 2002, for a review), whereas the third one, the difference in the degree of time inconsistency, is the main peculiarity of this work. In particular, to better match the experimental evidence, we assume that bounded rationality affects only a fraction of the population, with different degrees of awareness.

By introducing quasi-hyperbolic discounting we accomplish three main objectives: first, we complement the economic literature on early retirement by showing that the observed early exit from the workforce can

[3] In the paper, we will refer alternatively to present-biased preferences, time inconsistency and preference reversal. Although in the literature the meanings of these expressions are slightly different, for us all these terms denote a situation where people are not able to commit to future actions, and have a strict preference for the present.

be also explained through a model that focus mainly on workers' (bounded) rationality, and not only by exogenous factors. Because of the lack of commitment, hyperbolic individuals are not able to stick with their optimal plans, and when deciding whether to retire or not before the mandatory retirement age (benefiting from the early retirement provision), they weigh too much the costs associated to remain at work (foregone leisure) and less to the benefit implied by a longer working career (increase in pension benefit, if the the formula has a component related to past earnings).

Second, the model provides a behavioral justification for the drop in post retirement consumption due to inadequate saving observed among early retirees (Loewenstein, 1991 and Bernheim, 1995 and Laibson, 1997), within a model that considers explicitly social security and endogenous retirement age.

Third, we analyze the implications of hyperbolic discounting for a voting model in which individuals vote over the main characteristics of the pension system (its size and its degree of redistribution). To our knowledge, this is the first paper that studies the implication of this assumption for a political economy model with endogenous retirement age. By introducing time inconsistency as an additional source of heterogeneity, our political model provides a more complet explanation for the paradoxical observation that countries with low redistribution in social security systems have also larger public pension expenditures compared to countries with more redistributive pension systems (Disney 2004, Conde Ruiz and Profeta 2005). We show that the relevant winning coalition determining the size of the pension system is not composed simply by all poor and all retirees, as shown by traditional political models (see Galasso and Profeta, 2002, for a review). More precisely, young hyperbolic, also with low productivities, prefer to decrease the payroll tax. This counterintuitive result comes from the fact that sophisticated hyperbolic workers (who are aware of their inconsistency issue), looking for a commitment device that would increase both their retirement age and saving, form a coalition with the rich, who do not support the pension system as they prefer to finance their post retirement consumption through private savings. When voting over the Bismarckian factor, a proxy we use for the degree of redistribution of the PAYG system, time inconsistent individuals form a coaltion with the poor and the old, who are all in favor of a more redistributive system. The intuition goes as follows: besides the traditional intergenerational and intragenerational forms of redistribution, our model introduces a third one, namely from time consistent to hyperbolic individuals.

Our paper provide a behavioral justification for the negative relationship between size and redistribution observed in most OECD pension systems: the more time inconsistency represents an issue[4], the smaller will be the pension system, and the higher will be its degree of redistribution. Finally, we use the results of both the theoretical and the political model to evaluate the effectiveness of perspective reforms of the pension system; due to their bounded rationality, individuals are not able to fully understand intertemporal trade-offs, and are not able to correctly respond to incentives. Our model suggests that, in spite of the popular view, increasing the minimal retirement age would be a more effective way for the government to overcome the financial crisis of the pension system.

[4] We are not saying that certain countries are more time inconsistent than others, but only that the existence of privately-provided commitment devices differs across countries: if such instruments are not available, hyperbolic individuals look for other forms of commitment, in the form of a lower payroll tax.

The paper is organized as follows: in section 2 we shortly review of the behavioral literature and its main results; in section 3, we present the behavioral anomalies in retirement decisions (retirement age and saving for retirement) that show the empirical relevance of our hyperbolic model and motivate our research. In section 4, we discuss the relationship between social security and redistribution. In section 5, the basic features of the model are introduced. In section 6 we analyze the optimal choices of a young worker (saving) and of an old one (retirement age). In section 7 we derive the solution of an utilitarian social planner. In section 8, the majority voting equilibrium is determined. Section 9 concludes.

2 Background on Time Inconsistency

The economics and psychology literature (see Laibson, 1997, for a review) has challenged the view that individuals are perfect far-sighted optimizers. Large body of experimental and field evidence has demonstrated that the way individuals take a decision departs substantially form the perfect rationality paradigm. Examples of such departures are given by anxiety, misperceptions, dynamic inconsistency, mistakes, anticipatory feelings, reference points and loss aversion. Starting with the seminal works by Laibson (1997), most of the behavioral literature has focused on a particular form of bounded rationality: time inconsistency (or quasi-hyperbolic discounting) in intertemporal decisions.

In the hyperbolic model, each individual is represented as a collection of selves: time inconsistency arises because present selves overweight current payoffs compared to future ones, giving rise to a conflict among different intertemporal selves. Formally, an hyperbolic individual has the following intergenerational utility function (Strotz 1956, Phelps and Pollacks 1968, Laibson 1997):

$$u_0(.) + \beta \sum_{t=1}^{T} \delta^t u_t(.) \tag{1}$$

where β is the short-term psychological discount factor and δ is the time-consistent, long term one. This formulation implies that the discount factor between now and the next period is $\beta\delta$, whereas the discount factor between any two periods in the future is δ. Equivalently, the discount factor between today and the next period is declining, and constant thereafter.

Following Gruber and Köszegi (2004), we assume that individuals are heterogenous with respect their degree of time inconsistency. We model this heterogeneity as follows: each agent is characterized by a **perceived** psychological discount factor, which may or may not differ from the **true** one. The following definition clarifies this distinction.

Definition 1 *The perceived discount factor is the one individuals have in mind when making decisions. The real psychological discount factor represents their true degree of time inconsistency. Overconfidence is defined as the difference between the perceived and the true psychological discount factor. More precisely, we have:*

Individual Type	Perceived DF $\hat{\beta}$	True DF	Overconfidence
Exponential	1	1	0
Sophisticated	β (< 1)	β (< 1)	0
Naive	1	β (< 1)	$1 - \beta$

Table II: Perceived and real discount factors.

Exponential, or time consistent, individuals are rational utility maximizers. *Sophisticated* are aware of their time inconsistency, as their perceived discount factor coincides with the true one: self t knows that self $t+1$ will want to do something other than what self t would have him to do. Therefore, the best thing self t can do is to find a commitment device that force him to actually follow his (optimal) plan. Finally, *naive* agents are not able to make consistent plans through time (future selves will change systematically those plans) and are not able to actualize predicted or desired future levels of consumption, since their perceive themselves as time consistent, but they are actually hyperbolic.

This paper contributes to the recent literature on applied behavioral economics (Della Vigna and Malmendier, 2004, Della Vigna and Paserman 2005). So far, however, only few papers have considered behavioral and psychological issues into more traditional model of public and political economy.

Feldstein (1985), for instance, in a framework with completely myopic individuals who do not save at all, shows that it may be optimal to have either no social security or one with a very low replacement ratio. Imrohoroglu, Imrohoroglu and Joines (2004) extend Feldstein model, allowing for hyperbolic discounting, and show that social security is a poor substitute for private saving, even for naive consumers with high short term discount factor. Diamond and Köszegi (2003), introducing endogenous retirement age in Laibson (1997) quasi-hyperbolic consumption model, investigate the effects over saving of a social security system, seen as a commitment device for impatient individuals. Gruber and Köszegi (2004) use the quasi-hyperbolic setting to derive the optimal tax for addictive goods (cigarettes, fatty foods etc.), and find that it includes a "self-control adjustment" component, that provides to sophisticated a commitment device that reduces consumption of such goods. Cremer et al. (2005) study the optimal design of a linear pension scheme in a world of wage heterogeneity and complete naiveté for a group of workers who do not save for retirement. In a companion paper (Cremer et al., 2007), they studies the determination through majority voting of a pension scheme when there are far-sighted and myopic individuals, who tend to look for instant gratification when deciding over saving and labor supply. Their model, however, considers only two groups of consumers, fully myopic who do not save at all and far-sighted, and they do not consider endogenous retirement age, as we do.

Except for our behavioral assumption, our paper contributes also to the literature on the political economy of early retirement (Conde-Ruiz and Galasso 2004, Casamatta et al. 1999); our work is also closely related to Conde-Ruiz and Profeta (2005) and Casamatta et al. (2005)[5].

[5] In this paper, individuals, after having chosen retirement age and post-retirement saving, vote over the payroll tax and the level of the implicit tax on continued activity. They show that a biased pension system may emerge as an equilibrium of the voting game.

3 Time Inconsistency in Retirement Decisions

In the two following subsections, we present two stylized facts providing evidence on our behavioral assumption. We focus on the discrepancy between planned and real retirement age (subsection 3.1) and the reduction of post-retirement consumption due to a lack of saving (subsection 3.2).

3.1 The Retirement Age

The objective of this subsection is to illustrate the determinants of the observed high levels of unanticipated retirement.

The literature generally explains the early exit from the workforce with a redundancy/dismission argument (Sala-i-Martin 1996, Ahituv and Zeira 2000, Conde-Ruiz and Galasso 2004): older workers are excluded from the labor force by the firm, either because they have received a negative shock on their productivity or because they create a negative externality on young's productivity that decreases aggregate output. We refer to this explanation as the *push* argument.

We depart from this view by proposing the *pull* argument: the observed unexpected early exit from labor force is mainly due to voluntary quits and is a direct consequence of individual, hyperbolic, preferences: when deciding whether retiring or not, present-biased workers do not evaluate correctly the costs and the benefits associated to their choice and, when young, overestimate their future retirement age.

To give preliminary evidence of our claims, we present a report by the British Department of Work and Pension[6] (2005), in which old workers (starting from the age of 50) are asked to self-report their planned retirement age. These reports can be seen as the utility-maximizing choices of rational agents who take into account the costs and the benefits associated with the retirement age. Afterwards, expectations are compared to the true retirement age. Results are reported in Table III.

Age	**% Planned to retire**	**% Effectively retired**	**% Δ**
<65	25	59	+34
65	45	26	-19
>65	20	9	-11
N/A	10	6	-4

Table III: Planned and effective retirement age[7]. *Source*: DWP (2005)

It is immediate to see that the proportion of workers expecting to retire earlier than the mandatory age is much lower than the proportion of those who actually do. One may argue that the push argument can explain this discrepancy: workers are forced unexpectedly into early retirement, either because their low productivity (this could happen in a more or less explicit way, *i.e.* through bribes, psychological pressure etc.) or illness or disability. The report contradicts this conclusion: only 15% of early retirees reported to be made redundant by the firm (explicitly or implicitly); 35% of the sample reported health as the main

[6]Hereinafter DWP.

[7]The mandatory retirement age is 65 but people can retire (at reduced benefits) starting from 55.

reason for early retirement, and more than 50% has retired earlier simply because wanted to do, but they were not able to predict it[8]. Furthermore, the report shows that neither there is a clear pattern in the educational or income level of early retirees, nor financial considerations seem to influence the retirement choice (DWP, page 59), in contrast with the idea that early retirees have low level of human capital and income. According to the DWP, then, early retirement is mainly a voluntary choice, and do not necessarily concern only poor workers[9].

Blondal and Scarpetta (1999) shows that these conclusions can be extended to the majority of European countries (Figure 1).

	Austria	Belgium	Denmark	Finland	France	Germany	Greece	Ireland	Italy	Luxembourg	Netherlands	Portugal	Spain	Sweden	Switzerland	UK
Dismissed or make reduntant	5,1	3,7	23,4	24,1	10,7	9,5	2,5	8,8	2,0	0,0	7,9	1,0	10,2	30,2	4,9	22,0
A job of limited duration has ended	0,2	0,7	7,0	3,8	1,5	0,7	1,7	3,1	1,7	0,5	0,1	0,4	11,1	8,2	N/A	3,6
Personal of family responsibilities	0,2	1,3	0,2	0,0	0,4	0,6	0,5	1,4	1,2	0,2	1,9	0,0	0,2	2,0	2,2	1,6
Own illness or disability	2,6	7,7	9,5	25,0	7,3	22,9	4,1	15,1	5,2	16,6	15,6	2,1	18,3	7,0	28,9	22,8
Early retirement	49,0	30,6	37,2	0,0	16,9	33,1	5,1	15,9	9,2	29,1	42,9	2,3	13,0	25,9	35,1	14,7
Normal retirement	30,2	19,6	2,3	11,7	38,6	10,9	52,2	12,0	53,4	31,7	0,0	1,2	17,8	12,5	22,8	4,8
Other reasons	12,8	36,4	20,3	35,5	24,5	22,3	33,9	43,7	27,5	22,0	31,6	93,1	29,5	14,2	6,1	30,6

Figure 1: Retired males aged 55-64: Main reasons for leaving last job or business. *Source*: Blondal and Scarpetta, (1998)

From the table we see that, except for Finland and Sweden, that experienced big labor market shocks in the 90s, all countries exhibit large share of voluntary early retirees (around 30%).

These stylized facts confirm that individual are not always farsighted optimizers: even if the have strong preferences for leisure, they should be able to anticipate their retirement age. Therefore, the high level of early retirees can not be explained only by the usual models of disability and/or low human capital. Our framework justifies the discrepancy between planned and effective retirement age not by a redundancy argument but by assuming that individuals, although with different degrees, display of time inconsistency *à la* Laibson: some workers (exponential) plan to retire at a certain age and effectively do it, some others know that their retirement age will be lower than the optimal one (sophisticated) and others simply fool themselves and retire earlier than planned (naive).

3.2 Savings for Retirement

A well known result in intertemporal decision theory is that consumers prefer a smooth or increasing consumption path and save accordingly, even in presence of stochastic income changes (Loewenstein 1991).

[8] In particular, the main reason for early retirement was the "opportunity of spending more time with a partner or family" (DWP, page 55), thus confirming that leisure in old age receives great weight in retirees' preferences. Another weakness of the push argument is that most countries have implemented legislation aimed to reduce age discrimination on the workplace, making dismissal of old workforce more difficult for the firm. For instance, in 2005, the British Government approved "The Green Paper" whose objective was to limit age discrimination by employers.

[9] "*The Report shows that, among those who were working, the pull factors were clearly more important than the push factors in determining the decision to retire early*" (DWP, page 56). Another possible explanation is that a change of legislation, occurred during the period of the survey, has induced workers to unexpected early retirement. Indeed, the "Green Paper", implemented during the Blair administration, has raised the minimum pension for low income pensioners. However, in spite the fact that this reform has make early exit more attractive for this group, the data do not show a clear correlation between income level and voluntary early retirement: the reform had not modify incentives.

In reality, however, individuals, when young, consume too much and save too little for their post retirement consumption[10]. Loewenstein (1991) shows that individuals' consumption paths are declining, even in absence of liquidity constraints, and retirees experience a substantial reduction of their pre-retirement consumption levels. The sharp drop in the average consumption around the typical retirement age is due to negative innovations to the income process because workers have overestimated their retirement income and saving (Bernheim et al., 1998).

Not surprisingly, the drop in post retirement consumption is more severe for early retirees (Haussman and Paquette, 1987). According to Bernheim (1995) these behaviors can not be explained through the standard life cycle framework but are more likely justified by psychological issues, and in particular by psychological impediments to adequate planning for post retirement consumption. This intuition is confirmed by Laibson and al. (1998): not only individuals save too little, but also regret for this choice: 76% of U.S. workers near the retirement age believe that they should have saved more. The quasi-hyperbolic model (Laibson et al. 1998) provides a theoretical justification for these behaviors: simulations in Laibson and Harris (2001) show that hyperbolic calibrated simulations are able to reproduce the observed high comovement between consumption and income and the drop in post retirement consumption better than exponential calibrated simulations.

Individuals' inadequate ability to plan future consumption levels has recently induced financial intermediaries to promote new instruments aimed to provide time inconsistent individuals with a commitment device that help them to boost saving (Benjamin, 2003). In the U.S., for instance, these instruments took the form of tax-deferred savings accounts, such as IRA, 401*(k)*,403*(b)*, 457*(b)* and 457*(f)* etc.

Our model assumes that markets are incomplete, in the sense that instruments helping hyperbolic individuals to save more and retire later are not available. This assumption could appear unrealistic but can be easily justified: first, there is no general agreement on the effectiveness of such plans (Bernheim, 1999). Second, assuming the completeness of financial markets would require to take into account "counter-commitment devices" that exploit consumers' bias towards the present by reducing the commitment power of deferred saving accounts. For instance, in the U.S., the introduction of tax deferred saving account was followed by the boom of revolving credit cards.

4 Social Security and Redistribution

It is well known that a pension system contributes to redistribute income both within and between generations[11]. Depending on the link between benefits and workers' contributions, PAYG systems are classified either has *Bismarckian*, if benefits are proportional to contributions paid or *Beveridgean*, if benefits are flat.

[10]The problem is particularly severe in U.K., as stressed by British newspapers: "Not enough of British are saving for retirement and many of those woh are investing for the future are underestimating the cost of retirement ". (*The Guardian,* November 17, 2005). "The gap between what the nation should be saving each year for our retirement and what we are actually putting away is £27bn" (*Association of British Insurers*, 2005). "44% of the workforce (12 million workers) did not have pensions provision beyond those on offer form the state" (*Adair Turner, president of Britsh Pensions Commission,* 2005).

[11]A pensions system serves two other fundamental functions: it forces individuals to save for post retirement consumption and insures workers against disability risks.

Since contributions paid are proportional to earnings, the former implies, *ex-ante*, less intergenerational redistribution than the latter.

Quite counterintuitively, as shown by Conde-Ruiz and Profeta (2005), more redistribution does not necessarily imply that the share of the GDP devoted to the pension system must be higher; in reality, it exists a negative correlation between the degree of intragenerational redistribution and the size of the system: Bismarckian systems tend to be bigger than Beveridgean pension schemes, in spite of the fact the latter is explicitly designed to redistribute income and thus it should receive more support from low income individuals.

A natural question arises: are Beveridgean systems effectively able to redistribute income from who has earned more in his working years to those who earned less? If not, in which direction Some papers have recently tried to answer this question, focusing principally on the (Beveridgean) U.S. system.

Liberman (2001) shows that redistribution seems not to be related to lifetime income but to other factors: the system redistributes from people with high life expectancies to people with low ones, from single workers and from couples with high earnings by the secondary earner to one-earner couples and from workers with an earning history longer than 35 years to workers who works less than 35 years. Moreover, not only redistribution is not related to income but sometimes the system appears to be regressive: for instance, 19% of workers in the top lifetime *income* quantile receive net transfers that are greater than the average transfer for people in the lowest lifetime *income* quantile.

Coronado, Fullerton and Glass (2000) show that the claimed progressiveness of the pension system depends on a series of simplifying assumptions: looking only at redistribution on an annual basis (from workers to retirees), the system appears to be highly progressive. But, when more realistic assumptions are considered (differences in mortality probabilities etc.), the progressiveness of the system decreases and it may even become negative.

These conclusion seems to be valid not only for the U.S. but also for other social security systems across the world: some Beveridgean systems have features that reduce the explicit level of redistribution. The reverse appears also to be true: Disney (2001) show that also some Bismarckian systems redistribute income from rich to poor, in spite of the fact that they are modeled to guarantee high replacement rates.

Starting from these observations. we can classify OECD social security systems into four groups, on the basis of type (Bismarckian and Beveridgean), the level of redistribution (high or low), and the generosity, expressed as a fraction of GDP devoted to pensions. The first classification (Disney 2001) is made on the basis of the replacement rates guaranteed by the pension system (higher replacement rates make the system more Bismarckian), whereas the second is calculated taking into account the progressivity index[12] calculated by OECD (2005). What are the political determinants of cross-country differences in terms of generosity, early retirement provision, degree of redistribution, type of the system? Political economy models of social security tries to answer by referring either to economic or to behavioral explanation. The first group of

[12] The index is based on the Gini coefficient and it considers only mandatory parts of public pension programs. We consider Beveridgean and Bismarkian pension system with progressivity index above, respectively, 60% and 20%, as "highly redistributive". See Table V in the Appendix for details.

paper appeals to the political power of the lower class, that is decisive in the political process and is able to determine the main characteristics of the system (generosity, early retirement provision etc.). Meltzer and Richards (1982) and Tabellini (1992) show that the equilibrium size of the pension system is determined by a coalition of old people and young poor workers.

I - Bismarckian High Redistribution High Expenditure	**II - Bismarckian Low Redistribution High Expenditure**
Belgium, Austria Germany	Greece Italy
III - Beveridgean High Redistribution Low Expenditure	**IV - Beveridgean Low Redistribution Low Expenditure**
Canada, UK Denmark, New Zealand	US, The Netherlands Japan, Switzerland

Table IV: Classification of pension systems of OECD countries.

Conde-Ruiz and Galasso (2003, 2004) and Sala-i-Martin (1996) show that an early retirement provision is introduced in the system by a coalition of old workers with incomplete working history and low-income young. The first group, dismissed by their firm because of a negative productivity shock has came, favors early retirement. The second group agrees to introduce an early retirement clause since they are likely to become, in the future, early retirees. Conde-Ruiz and Profeta (2005) propose a political economy model in which individuals vote over the size and degree of redistribution of the pension formula. It is shown that a Beveridgean system is supported by low-income agents, who gain from its redistributive feature, and rich individuals, who seek to minimize their tax contribution and to invest their resources in a private pension scheme.

A second group of papers considers behavioral and psychological factors to explains institutional differences among pension systems. Benabou and Ok (2001) show that poor may not support high levels of redistribution today because they hope to be rich in the future. Alesina and Angeletos (2005) and Benabou and Tirole (2005) develop models that consider explicitly fairness and equality of opportunities in individuals' preferences: redistribution is low when economic success is perceived as driven by effort and not by luck. Finally, other papers (see Alesina and La Ferrara 2005 for a review), consider identity: individuals care about who benefit from redistribution (for instance, they do not want to subsidize people from different ethnic groups). Thus, very heterogeneous societies may have lower preferences for redistribution.

This paper merges behavioral and economic explanations by assuming, at the same time, heterogeneity in productivity levels, age and degree of time inconsistency. Anticipating the results, our model shows that some agents, independently of their income, prefer to anticipate their retirement and prefer to consume more when young instead of saving for post-retirement consumption. Furthermore, our political model, in which the size of the social security system and the degree of redistribution are chose by direct majority voting, shows that a winning coalition of hyperbolic individuals is able to determine both the generosity and the degree of redistribution of the PAYG system. In particular, our model explains why low level of redistribution are often associated with big pension programs: hyperbolic young workers, looking for privately-provided commitment devices that increase both saving and retirement age, form a coalition with the rich in order to *decrease* the payroll tax. On the other hand, when voting over the degree of the redistribution of the PAYG system, the winning coalition includes poor and hyperbolic individuals both in favor of a more redistributive system.

5 Social Security, Retirement and Time Inconsistency: A Model

In this section we present a model of intertemporal consumption with endogenous retirement age (as in Casamatta et al., 2005 and Conde-Ruiz and Galasso, 2004), with the assumption of quasi-hyperbolic preferences.

5.1 Basic Setup

Timing The model is set in discrete time. There are three periods and two generations, *young* and *old*. At period 1, an individual is young: he supplies inelastically labor and saves for post-retirement consumption. At period 2 (*pre-retirement*), an individual is old: he supplies labor at an effort cost; we interpret labor supply as the choice of retirement age, as in Casamatta et al. (2005). In period 3 (*post-retirement*), the individual is retired and consumes accumulated savings and a pension transfer P[13]. Moreover, he enjoys some additional leisure, which is inversely related to the retirement age. The length of period 1 is normalized to 1, whereas the length of periods 2 and 3 are endogenous, since they both depend on the retirement choice of the agent. The total length of period 2 and 3 is normalized to 2.

Heterogeneities Besides age, individuals differ for their degree of time inconsistency and their productivity level. For the former, we refer to Table III and we consider three type of agents: exponential, sophisticated and naive, who differ on the basis of their perceived short-term discount factor, $\hat{\beta}_j$. The long run discount factor, δ, is assumed to be the same for all individuals.

Concerning the second source of heterogeneity, we assume that:

- **A1**: Each worker is assigned with a random productivity level ω, distributed over the support $[\omega_-, \omega_+]$ according to a density function $f(.)$ and a c.d.f. $F(.)$. The median productivity, ω^M, is lower than the mean, $\bar{\omega}$.

To simplify, we assume that both the productivity level and the behavioral type remain unchanged across periods.

[13]Savings can be interpreted as a voluntary payment to a integrative pension plan, whose benefits are paid only when the beneficiary is retired. The fact that $P(\hat{z}_j)$ is paid only after retirement is indeed a realistic assumption: Sala-i-Martin (1996), shows that, for 70 out of 108 countries where this information is available, the elderly must show that they do not get labor income from any other source in order to collect old age pensions.

We denote with n the exogenous rate of population growth; N^r is the number of old retirees in the economy, N^o the number of old workers and $N^y = (1+n)(N^o + N^r)$ the number of young. The total population is therefore $N = N^y + N^o + N^r$ or, equivalently, $(N^o + N^r)(2+n)$. Each behavioral type represent a fraction of the whole population, such that $N^i = N^{i,e} + N^{i,n} + N^{i,s}$, with $i = y, o$, where the subscripts e, n and s stand for exponential, naive and sophisticated individuals[14]. None of them represent 1/2 of the population.

Utility The multi-selves formulation of (1) implies that each individual of type $j = e, n, s$ maximizes the following intertemporal utility function:

$$U\left(c_y^j, c_o^j\right) = u(c_y^j) + \hat{\beta}_j \delta u(c_o^j) \tag{2}$$

where $u(.)$ is the instantaneous utility function, assumed to be the same for all individuals, increasing and concave, and satisfying Inada conditions.

We denote with, respectively c_y and c_o consumption levels when young and old (with include both period two and three).

As in Casamatta et al. (2005), the variable $z \in [0,1]$ denotes the fraction of the second period an old agent spent working. It follows that the lengths of periods two and three are, respectively, z and $(2-z)$.

Consumption when old is net of the disutility of working, $m(z)$ and the utility from leisure when retired, $l(z)$, both expressed in unit of consumption.

We assume specific form for functions $m(z)$ and $l(z)$:

- **A2**: $m(z) = \frac{\gamma z^2}{2}$, where γ measures the intensity of the disutility of effort.
- **A3**: $l(z) = \frac{\psi(1-z)^2}{2}$, where ψ represents the intensity of the utility from leisure. We assume that $\gamma > \omega_+ > \omega_- > \delta\psi$.

Each worker (young and old) pays a proportional payroll tax $\tau \in [0,1]$ on his wage that finances the PAYG system. Therefore, consumption levels for young and old of type $(\omega,\, j)$ are given by:

$$\begin{aligned} c_y^j &= (\omega(1-\tau) - \widehat{s_j}) + \hat{\beta}_j \delta \left(c_o^j\right) \\ c_o^j &= \underbrace{\omega \widehat{z_j}(1-\tau\theta) - \frac{\gamma\left(\widehat{z_j}\right)^2}{2}}_{\text{Pre-Retirement Consumption}} + \hat{\beta}_j \delta \underbrace{\left(P(\widehat{z_j}) + \frac{\psi(1-\widehat{z_j})^2}{2} + (1+r)\widehat{s_j}\right)}_{\text{Post Retirement Consumption}} \end{aligned} \tag{3}$$

where $\widehat{s_j} \geq 0$ and $P(\widehat{z_j})$ represent saving and pension benefits, both depending on perceived retirement age, $\widehat{z_j}$. The "hat" stresses the fact that, due to our behavioral assumption, consumption levels, retirement age and saving depend all on the perceive discount factor $\hat{\beta}_j$.

Finally, r is the exogenous interest rate paid on accumulated saving; to simplify notation, we impose $1/(1+r) = \delta$.

[14] Notice that we the distribution of behavioral types among old retirees does not matter for our purposes, since all economic decisions are already taken.

From the intertemporal utility function (2) and the old budget constraint (3), it is easy to see that individuals are not only time inconsistent *between* generations (equation 2) but also *within* generations (equation 3). It follows that, from the point of view of an hyperbolic young who is evaluating the trade-off between working and retiring when old, the discount factor is δ, whereas for an old is $\hat{\beta}_j\delta$. This discrepancy captures the peculiarity of time inconsistency: when deciding over the retirement age, hyperbolic agents change their preferences, giving too much weight to the present costs of staying at work (the disutility of effort), and less to the future benefits (increase of pension benefits). Moreover, this discrepancy affects also savings: hyperbolic young, when choosing whether to consume or not, discount the post-retirement period only by $\hat{\beta}_j\delta^2$.

The Pension System The PAYG system is assumed to be balanced every period, so the sum of all awarded pensions is equal to the sum of all contributions paid. The pension received by an individual of type (ω, j) is given by:

$$P(z) = \underbrace{\alpha\tau\omega(1+n+\theta z)}_{I} + \underbrace{(1-\alpha)\tau(1+n)\bar{\omega}}_{II} \tag{4}$$

To introduce distributional concerns in our model, we assume that (4) includes two components[15]: a *Bismarckian* (I, related to contributions paid by the worker and its retirement age) and a *Beveridgean* one (II, flat, related to the mean wage of the economy, $\bar{\omega}$); both components have a rate of return equal to rate of population growth, n. The parameter $\theta \in [0,1]$ represents the implicit tax on continued activity (Gruber an Wise, 2000), and takes into account that the pension system itself induces workers to leave earlier their job: for $\theta = 0$, the pension system is neutral, *i.e.* the marginal benefit of working one more year is ω. For $\theta > 0$, individuals have incentive to retire earlier[16], since the marginal benefit of working more is reduced to $\omega(1-\theta\tau)$.

The weight $\alpha \in [0,1]$ in (4) is the Bismarckian factor (Conde Ruiz and Profeta, 2005). If $\alpha = 1$, the system is purely Bismarckian, and benefits are proportional to the contributions paid. For $\alpha = 0$, the system is purely Beveridgean: benefits are flat and not related to the worker's wage history. If $\alpha < 1$, the pension has two tiers: a flat one, which provides a minimum amount of income, and a second one that relates pension benefits to the history of previous wage earnings.

For the rest of the paper, we make two technical assumptions about α; these assumptions will guarantee the existence of our political equilibrium.

- **A4**:

$$\alpha \leq \frac{1}{(1+\theta)\delta} \quad \text{and} \quad \alpha \leq \frac{1}{1+n}$$

[15] The dependence of $P(z)$ on ω is justified by Sala-i-Martin (1999): in most countries, benefits are typically an increasing function of the workers' previous wage history.

[16] Cremer, Lozachmeur and Pestieau (2004) show that such distortions are optimal in a second-best setting with a government concerned by redistributive issues and asymmetric information on individuals' productivity and health status.

The first inequality implies that θ is not too high[17]. The second inequality, *i.e.* $\alpha \leq \frac{1}{1+n}$ is reasonable, given the low rate of population growth ($n \approx 0$) observed in most industrialized countries.

The Political Process The voting process works as follows: elections take place every period. Young and old (at the beginning of period 2) vote simultaneously over the payroll tax τ and the Bismarckian factor, α[18]. The impact of time inconsistency on voting behavior is twofold: from one hand, hyperbolic young vote having in mind their perceived discount factor, $\hat{\beta}_j$, since do not fully internalize how their present vote affects future utility. On the other hand, old (naive) realize *ex-post* their true β and their oveconfidence, which has lead to save and retire suboptimally and vote according to their true β_j, after having chosen s_j and z_j on the basis of the perceived discount factor. Formally, we have:

- **A5:** Old's indirect utility function $V^o(\tau_j^o, \alpha; \beta_j, \omega)$ and voting behavior depend on β_j. Young's indirect utility function $V^y(\tau_j^y, \alpha; \hat{\beta}_j, \omega)$ and voting behavior depend on $\hat{\beta}_j$.

This assumption implies also that, when young, naive voters like exponential, whereas, when old, like sophisticated. It follows that the total number of time consistent young voters is $N^{y,TC} = N^{y,e} + N^{y,n}$, whereas for old is: $N^{o,TC} = N^{o,e}$.

Figure 2 summarizes the basic features of our model.

6 The Optimization Problems

In this section we characterize optimal choices for young (saving) and old (retirement age).

6.1 Retirement Age

A type (ω, j) old solves the following program:

$$\max_{\widehat{z}_j} \left[\omega \widehat{z}_j (1 - \tau\theta) - \frac{\gamma(\widehat{z}_j)^2}{2}\right] + \hat{\beta}_j \delta \left[P(\widehat{z}_j) + \frac{\psi(1-\widehat{z}_j)^2}{2} + (1+r)s^j\right] \tag{5}$$

subject to:

$$0 \leq \widehat{z}_j \leq 1$$

Notice that an old worker discounts post retirement consumption by $\hat{\beta}_j \delta$. The first order condition associated with problem (5) is:

$$\omega(1 - \tau\theta) - \gamma\widehat{z}_j + \hat{\beta}_j \delta \left[\alpha\tau\omega\theta - (1 - \widehat{z}_j)\psi\right] = 0$$

[17] According to Gruber and Wise (1997), in a sample of 11 developed countries, the value of θ is approximately around 50%, making $\theta \in [1/2, 1]$.

[18] We assume, without loss of generality, that old retirees do not vote. The reason for our assumption is the following: retirees' objective is to increase their consumption as much as possible: thus, *all* retirees will vote for the maximum payroll tax, $\tau = 1$. Moreover, when voting over α, all agents with $\omega < \bar{\omega}$ prefer a completely flat pension, $\alpha = 0$, whereas all individuals with $\omega > \bar{\omega}$ vote for $\alpha = 1$. Since we assume that the income distribution is skewed to the left, the second group will always represent a minority. Therefore, considering that also retirees vote would change only the level of equilibrium functions τ^{mv} and α^{mv}, but not their shape.

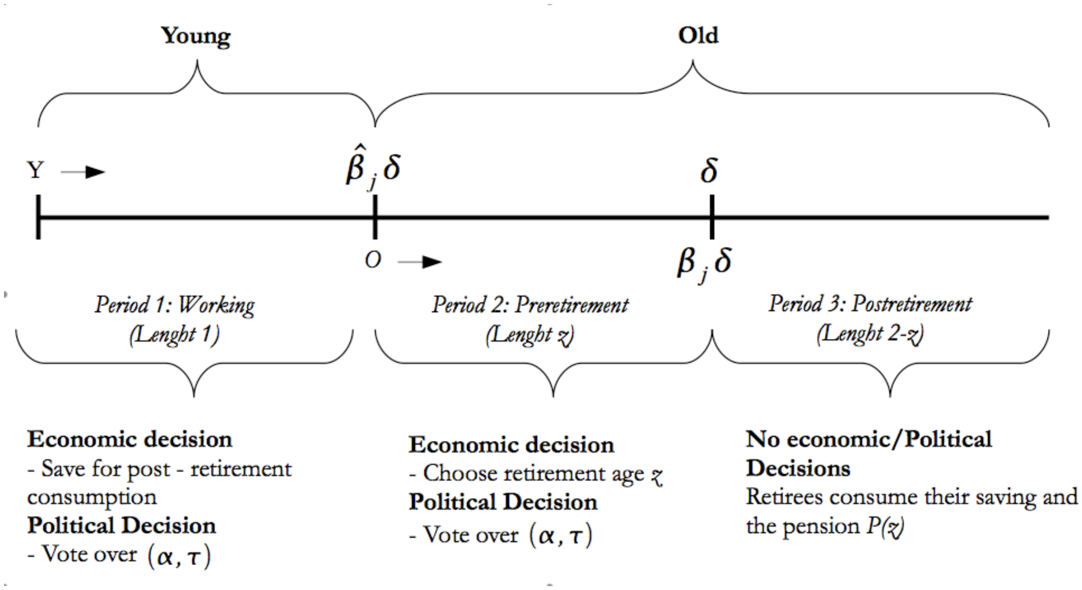

Figure 2: Timing of the Model

After some rearrangements, optimal retirement age is given by:

$$\widehat{z}_j = \frac{\omega(1-\tau\theta) + \hat{\beta}_j\delta(\alpha\tau\omega\theta - \psi)}{\gamma - \hat{\beta}_j\delta\psi} \quad \forall j = e, s, n \tag{6}$$

The solution is interior $\Leftrightarrow \omega^+ \left[1 - \tau\theta(1 - \hat{\beta}_j\delta\alpha)\right] \leq \gamma$, as implied by A3. All individuals choose to work in the second period, except if $\tau\theta = 1$ and $\alpha\omega = \psi$. The intuition is the following: nobody work when the system is completely non neutral and when the increase in the Bismarckian part of the pension due to an additional year spent working $(\alpha\omega)$ is exactly equal to the marginal benefit of quitting the job at the beginning of the old age (ψ).

Comparative statics on (6) leads to the following lemma.

Lemma 1 *Optimal perceived retirement age, $\widehat{z}_j$ is:*
(i) increasing in ω, θ and ψ;
(ii) decreasing in payroll tax, τ, and increasing in the Bismarckian factor, α;
(iii) increasing in $\hat{\beta}_j$.

Result *(i)* is in line with the literature (Casamatta et al. 2005 and Conde-Ruiz and Galasso 2004): rich workers, having a lower price of consumption/leisure in period 2 than poor, retire later. In part *(ii)* we show, first, higher τ decreases perceived retirement age:

$$\frac{\partial \widehat{z}_j}{\partial \tau} = -\frac{\omega\theta(1 - \hat{\beta}_j\delta\alpha)}{\gamma - \hat{\beta}_j\delta\psi} < 0 \tag{7}$$

and this effect is more marked for high productivity workers, since $\frac{\partial_2 \widehat{z}_j}{\partial \tau\omega} < 0$.

Intuitively, a higher Bismarckian factor increases the retirement age:

$$\frac{\partial \widehat{z_j}}{\partial \alpha} = \frac{\omega\theta\hat{\beta}_j\delta\tau}{\gamma - \hat{\beta}_j\delta\psi} > 0$$

Notice that this effect is more marked for individuals with high $\hat{\beta}_j$: intuitively, individuals who perceive themselves as time consistent think they can work more if the pension is more earning-related.

More interesting is result *(iii)*: $\partial\widehat{z_j}/\partial\hat{\beta}_j > 0$ implies that the higher is the worker's perceived degree of time inconsistency, the later he *believes* he will retire. This conclusion is in line with the stylized facts presented before: some workers anticipates correctly their retirement age, whereas others overestimate it. Depending on the behavioral type, three cases are possible.

Exponential These individuals have $\hat{\beta}_e = \beta = 1$. Replacing the true discount factor in (6), an exponential worker ($\widehat{z_e}$) will retire at:

$$\widehat{z_e} = \frac{\omega(1-\tau\theta) + \delta(\alpha\tau\omega\theta - \psi)}{\gamma - \delta\psi} \tag{8}$$

which also coincides with the actual one, $\widehat{z_e} = z_e$. In the following, we refer to z_e as the *normative retirement age*, defined as the retirement age to which a young individual would commit himself.[19].

Sophisticated They are aware of their time inconsistency: their perceived discount factor is equal to the real one: $\hat{\beta}_s = \beta < 1$; planned $\widehat{z_s}$ and actual retirement age, z_s, coincide:

$$z_s \equiv \widehat{z_s} = \frac{\omega(1-\tau\theta) + \beta\delta(\alpha\tau\omega\theta - \psi)}{\gamma - \beta\delta\psi} \tag{9}$$

For these agents, however, the lack of commitment leads to a welfare loss; they would be better, in a Pareto sense, if their retirement age would be the one preferred by exponential, as shown in the following proposition.

Proposition 1 *For sophisticated workers, the normative retirement age, z_e, is greater than the equilibrium retirement age z_s. If β is not too small, switching from z_s to z_e will create a Pareto-improvement.*

Proof. All Proofs are in Appendix B. ■

Naive Their true discount factor, $\beta(< 1)$, is lower than the perceived one, $\hat{\beta}_n(= 1)$. Their planned retirement age $\widehat{z_n}$, is:

$$\widehat{z_n} = \frac{\omega(1-\tau\theta) + \delta(\alpha\tau\omega\theta - \psi)}{\gamma - \delta\psi} \tag{10}$$

but the true one is:

$$z_n = \frac{\omega(1-\tau\theta) + \beta\delta(\alpha\tau\omega\theta - \psi)}{\gamma - \beta\delta\psi} \tag{11}$$

[19]Rigorously, we should define normative retirement age as the z_j to which self 0 *and* all future selves would commit. However, there is no loss in generality to use this definition of normative retirement age (see Laibson 1998).

Thus, naive agents exhibit *overconfidence*, which is given by:

$$\Delta z \equiv \widehat{z_n} - z_n = \frac{(1-\beta)\delta[\psi\omega(1-\tau\theta)+\gamma(\alpha\tau\omega\theta-\psi)]}{(\gamma-\delta\psi)(\gamma-\beta\delta\psi)} > 0 \tag{12}$$

Notice that the overconfidence effect justify why we some individuals systematically overestimate their retirement age: they believe will retire at $\widehat{z_n}$, but their bias will lead them to retire at z_n. Moreover, it is easy to see that $\partial\Delta z/\partial\beta < 0$: more time consistent individuals display less overconfidence; in the limiting case, $\Delta z = 0$ if $\beta = 1$.

6.2 Savings

Young of type (ω, j) decide how much to save for post retirement consumption anticipating that retirement age will be (6). The maximization program is the following:

$$\max_{\widehat{s_j}} u(\omega(1-\tau)-\widehat{s_j}) + \hat{\beta}_j\delta u\left[\omega\widehat{z_j}(1-\tau\theta) - \tfrac{\gamma(\widehat{z_j})^2}{2} + \delta\left(P(\widehat{z_j}) + \tfrac{\psi(1-\widehat{z_j})^2}{2} + (1+r)\widehat{s_j}\right)\right] \tag{13}$$

subject to:

$$0 \le \widehat{s_j} \le \omega(1-\tau)$$

The first order condition associated with an interior solution of $\widehat{s_j}$ is:

$$-u'\left(c_y^j\right) + \hat{\beta}_j\delta^2 u'\left(c_o^j\right)(1+r) = 0 \tag{14}$$

Since we assume $(1+r)\delta = 1$, saving are implicitly defined by:

$$u'\left(c_y^j\right) = \hat{\beta}_j\delta u'\left(c_o^j\right) \tag{15}$$

It follows that, $\forall j$, $c_y^j > c_o^j$, since $\hat{\beta}_j\delta < 1$ and the concavity of the utility function: hyperbolic discounting leads to overconsumption in the young age. Individuals are not able to smooth consumption over time and the lower is the true β, the more relevant is this effect[20]. In the following, we will refer to s_e as *normative saving*, *i.e.* saving chosen by an exponential individual. The following proposition shows that sophisticated would increase their welfare if a commitment device that helps them to save s_e is made available.

Proposition 2 *For sophisticated, the normative saving s_e are greater than equilibrium saving s_s. If β^s is not too small, switching from s_s to s_e will induce a Pareto-improvement.*

Proof. See Appendix B. ■

7 The Planner's Problem

Although the aim of this work is mainly positive, it is worth to establish as a benchmark the solution an utilitarian social planner would choose. We restrict policies in the same way as in the voting process

[20] This Euler equation reduces to the one in Casamatta et al. (2005) for exponential and naive ($\hat{\beta}_e = \hat{\beta}_n = 1$), and to the Hyperbolic Euler Equation of Laibson and Harris (2001) for sophisticated ($\hat{\beta}_s = \beta$).

considered in the next section: instruments are limited to α and τ. In our framework the planner's problem is not standard: in general, according to Kahneman (1994), each individual maximizes his "decision utility", *i.e.* the utility function that reflects his choices, whereas the government maximizes the agent's "experience utility", the utility function that reflects his welfare; usually, the two concepts naturally coincides, but here they differ, since the planner and the individual do not agree on the weight attributed to future utilities: we have seen that naive individuals are biased toward the present when deciding their saving and their retirement age.

7.1 First Best

In a first best setting, where every source of heterogeneity is observable by the the government, even individuals' behavioral type, time inconsistency does not represent an issue. The planner solves:

$$\max \int_{\omega_-}^{\omega_+} u(c_y(\omega)) + \delta u \left[x(\omega) - \tfrac{\gamma(z(\omega))^2}{2} + \delta \left(\tfrac{\psi(1-z(\omega))^2}{2} + (1+r)s^y\right)\right] f(\omega)d\omega$$

subject to:

$$\int_{\omega_-}^{\omega_+} \left[c_y(\omega) + \frac{x(\omega)}{1+n} - \frac{\omega}{1+n}(1+n+z(\omega))\right] f(\omega)d\omega = 0 \tag{16}$$

where $x(\omega) = \omega z(\omega)(1-\tau\theta) + \hat{\beta}_j\delta\left(P(z(\omega)) + (1+r)\widehat{s_j}\right)$ is consumption when old. Notice that we are considering the true retirement age, $z(\omega)$ since, in a first best setting, we suppose that the government is able to force individuals to retire at the optimal age. From the first order conditions, we have:

$$u'(c_y(\omega)) = \delta(1+n)u'(c_o(\omega)) = \lambda$$
$$z(w) = \frac{\omega - \delta\psi}{\gamma - \delta\psi}$$

where λ is the Lagrange multiplier associated to the budget constraint. The first FOC implies that consumption is equated across individuals and across time, if $\delta(1+n) = 1$. The second condition tells us that optimal labor supply in second period is such that the net marginal disutility of working an additional year, $(\gamma - \delta\psi)z(w)$, equates the marginal benefit of continued activity, ω, net of the forgone leisure $\delta\psi$ enjoyed if the worker would have retired. This first best can be decentralized through a system of lump sum taxes and transfers: however, given the fiscal instruments we are restricting to, this solution can not be achieved.

7.2 Second Best

In a second best setting, time inconsistency is not observable, but the planner knows the distribution of naive and sophisticated among the population[21]. The planner's objective is to set τ and α as to minimize the welfare loss resulting from overconfidence. More precisely, we follow the approach developed by Rabin

[21]Notice that we are implicitely assuming that the planner can not, in the second best, force individuals to retire at the optimum retirement age.

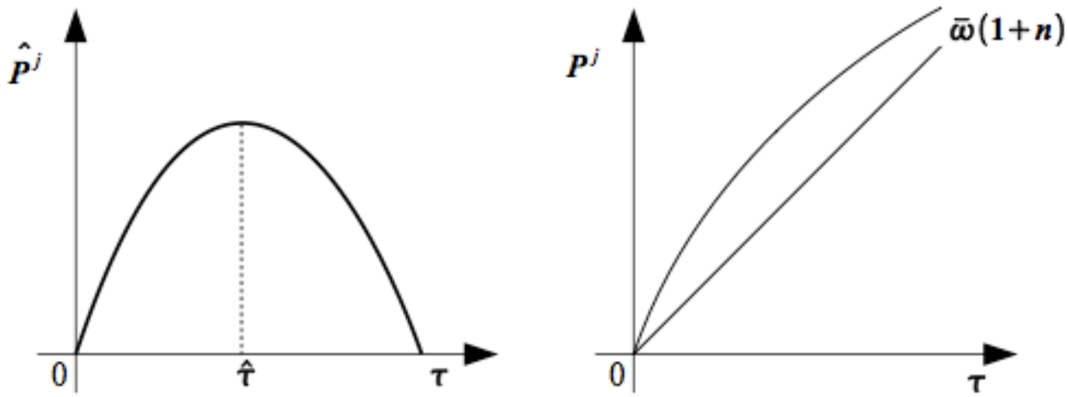

Figure 3: The perceived (left) and the true (right) pension function

and O'Donoughe (2001, 2007) and Gruber and Köszegi (2004): the social welfare function is time consistent ($\beta^j = 1$), but the budget constraint take into consideration that hyperbolic individuals retire according to their true discount factor.

7.2.1 The Government's Budget Constraint

Let us define the *perceived* pension, $P(\widehat{z_j})$, as the benefits each type think to receive after retirement, which differs from the *true* one, $P(z_j)$. The perceived function is also the pension young individuals of type $j = e, s, n$ have in mind when they cast their vote over τ and α. Replacing (6) into (4), we get:

$$P(\widehat{z_j}) = \left(1 + n + \frac{\theta\omega(1-\theta\tau) + \hat{\beta}_j\theta\delta(\alpha\tau\omega\theta - \psi)}{\gamma - \hat{\beta}_j\psi\delta}\right)\alpha\tau\omega + (1+n)\tau(1-\alpha)\bar{\omega} \tag{17}$$

Whereas for sophisticated and naive we have $P(\widehat{z_s}) = P(z_s)$ and $P(\widehat{z_e}) = P(z_e)$, for naive $P(\widehat{z_n}) > P(z_n)$. The difference between the two pensions, ΔP, reflects the overconfidence effect that naive experience:

$$\Delta P \equiv P(\widehat{z_n}) - P(z_n) = \frac{\theta\delta(1-\beta)(\omega(1-\theta\tau)\psi + \gamma(\alpha\tau\omega\theta - \psi)}{(\gamma - \beta\psi\delta)(\gamma - \psi\delta)} > 0$$

The next Lemma establishes an important property of the perceived pension function.

Lemma 2 *The pension function is concave in τ, with a maximum in $\hat{\tau} = \frac{\theta\alpha\omega^2 + (1+n)(\gamma - \hat{\beta}_j\psi\delta)(\alpha\omega + (1-\alpha)\bar{\omega})}{2\theta^2\alpha\omega^2(1-\hat{\beta}_j\delta\alpha)}$.*

Proof. See Appendix B. ■

Increasing τ has two effects on the pension function: *(i)* for given $z_j(\hat{\beta}_j)$, it increases the future pension; *(ii)* it reduces the worker's retirement age. The lemma shows that, for $\tau \leq \hat{\tau}$, the first effect prevails. Because of the overconfidence effect, $P(\widehat{z_j})$ differs from the pension scheme that satisfies the government's budget constraint: the discrepancy between the two is due to the negative externality exerted by naive individuals on exponential and sophisticated. The former do not internalize that retiring earlier than planned leads to a drop in total tax proceeds that, consequently, reduce also total benefits.

$$\sum_j N^{o,j} \int_{\omega_-}^{\omega_+} P(z_j) dF(\omega) = N^y \tau \int_{\omega_-}^{\omega_+} \omega dF(\omega) + N^{o,e} \tau\theta \int_{\omega_-}^{\omega_+} \omega z_e dF(\omega) + N^{o,s} \tau\theta \int_{\omega_-}^{\omega_+} \omega z_s dF(\omega) + \tag{18}$$
$$N^{o,n} \tau\theta \int_{\omega_-}^{\omega_+} \omega z_n dF(\omega)$$

Taking into account that naive and sophisticated retire according to (9) and (11)), the average pension for each group of retirees is:

$$\bar{P}_j(\tau,\alpha) = (1+n)\bar{\omega}\tau + \frac{\tau\theta}{\gamma - \beta_j \delta\psi}\left[(1-\tau\theta)E(\omega^2) + \beta_j \delta\alpha\tau\theta(E(\omega^2) - \psi)\right] \text{ for } j = n, s$$
$$\bar{P}_e(\tau,\alpha) = (1+n)\bar{\omega}\tau + \frac{\tau\theta}{\gamma - \delta\psi}\left[(1-\tau\theta)E(\omega^2) + \delta\alpha\tau\theta(E(\omega^2) - \psi)\right] \tag{19}$$

where $E(\omega^2) = \int_{\omega_-}^{\omega_+} \omega^2 dF(\omega)$ and $\beta_j = \{\beta, 1\}$. In the first period, the tax base is fixed, $(1+n)\bar{\omega}$, and depends on the true retirement age in the second period. Differentiating twice (19), we get:

$$\bar{P}'(\tau) = (1+n)\bar{\omega} + \frac{(\theta - 2\theta^2\tau)E(\omega^2) + 2\beta_j \delta\alpha\tau\theta^2(E(\omega^2) - \psi)}{\gamma - \beta_j \delta\psi} > 0$$
$$\bar{P}''(\tau) = -\frac{2\theta^2 E(\omega^2)(1 - \beta_j \alpha\delta) - \beta_j \delta\alpha\tau\theta^2\psi}{\gamma - \beta_j \delta\psi} < 0$$

The budget curve, represented in the Figure 3, is concave, always above the line $\tau(1+n)\bar{\omega}$ and equal to it when $\tau\theta = 0$ or $\tau\theta = \frac{E(\omega^2) - \beta_j \delta\alpha\psi}{E(\omega^2)(1-\beta_j \alpha\delta)}$.

7.2.2 Second Best Levels of α and τ

Second best value for α and τ are the solution to the following maximization:

$$\max_{\alpha,\tau} \ \Lambda \equiv \sum_j N^{y,j} \int_{\omega_-}^{\omega_+} u(\omega(1-\tau) - s_j) + \delta u\left[\omega z_j(\omega) - \tfrac{\gamma(z_j)^2}{2} + \delta\left(\bar{P}_j(z_j) + \tfrac{\psi(1-z_j)^2}{2} + (1+r)s_j\right)\right] f(\omega)d\omega$$

subject to:

$$\bar{P}_j(\tau) = (1+n)\bar{\omega}\tau + \frac{\tau\theta}{\gamma - \beta_j \delta\psi}\left[(1-\tau\theta)E(\omega^2) + \beta_j \delta\alpha\tau\theta(E(\omega^2) - \psi)\right] \quad \text{for } j = n, s$$
$$\bar{P}_e(\tau) = (1+n)\bar{\omega}\tau + \frac{\tau\theta}{\gamma - \delta\psi}\left[(1-\tau\theta)E(\omega^2) + \delta\alpha\tau\theta(E(\omega^2) - \psi)\right]$$

The first order conditions for, respectively, α and τ, are:

$$\frac{\partial\Lambda}{\partial\tau} : -\sum_j N^{y,j} \int_{\omega_-}^{\omega_+} \left[\omega u'(c_y^j) - \delta u'(c_o^j)\left(-\omega\theta z_j + \delta\frac{\partial\bar{P}_j(\tau)}{\partial\tau}\right)\right] = 0 \tag{20}$$

$$\frac{\partial\Lambda}{\partial\alpha} : -\sum_j N^{y,j} \int_{\omega_-}^{\omega_+} \left[u'(c_o^j)\left(\frac{\partial z_j}{\partial\alpha}\left[\omega(1-\tau\theta) - z_j(\gamma - \psi\delta) - \psi\delta\right] + \delta\frac{\partial\bar{P}_j}{\partial\alpha}\right)\right] = 0 \tag{21}$$

where:

$$\frac{\partial \bar{P}_j}{\partial \alpha} = \frac{\beta_j \delta \left(\tau\theta\right)^2}{\gamma - \beta_j \delta\psi} \left[(E(\omega^2) - \psi)\right] \text{ for } j = s, n$$
$$\frac{\partial \bar{P}^e}{\partial \alpha} = \frac{\delta \left(\tau\theta\right)^2}{\gamma - \delta\psi} \left[(E(\omega^2) - \psi)\right]$$

Following Casamatta et al. (2005), the first FOC has a straightforward interpretation: if there are not liquidity constraints, standard results in linear income taxation apply: it is optimal to set $\tau^{opt} = 1$ in the first period, since young's labor supply is inelastic, and to redistribute through lump-sum transfers in the second period, setting, for example, $\alpha^{opt} = 0$, *i.e.* the pension includes only a Beveridgean component. However, with liquidity constraints, it is no longer true that it is optimal to impose a 100% income tax in the first period: consumption in young age must be positive. Therefore, optimality requires $\tau^{opt} < 1$.

Replacing the expressions for $\frac{\partial \bar{P}_j}{\partial \alpha}$ into (21) and rearranging, we get the expression for the optimal Bismarckian factor α^{opt}:

$$\alpha^{opt} = M \frac{\int_{\omega_-}^{\omega_+} u'(c_o)\tau\theta(E(\omega^2) - \psi) f(\omega)\, d\omega + \frac{N^{y,TI}}{\gamma - \beta\delta\psi} \int_{\omega_-}^{\omega_+} u'(c_o)\omega^2 \left(1 - \tau\theta\right)\left(1 - \beta + \omega\right) f(\omega)\, d\omega}{\delta\tau\theta \left(\frac{N^{y,e}}{\gamma - \delta\psi} E\left(u'(c_o)\omega^2\right) - \frac{N^{y,TI}}{\gamma - \beta\delta\psi} E\left(u'(c_o)\omega\right) \right)} \tag{22}$$

where $M = \frac{\tau\theta N(\gamma - \delta\psi) + \delta\psi N^{y,e}(1-\beta)}{(\gamma - \delta\psi)(\gamma - \beta\delta\psi)}$ is a constant term and $N^{y,TI} = N^{y,n} + N^{y,s}$ is the number of time inconsistent individuals. Expression (22) implies that the planner uses that the Bismarckian parameter α as a commitment device the helps time inconsistent workers to increase their welfare. To understand why, first notice that α^{opt} is an increasing function of $N^{y,TI}$: the higher is the number of time inconsistent individuals, the higher is the Bismarckian component of the pension formula. The intuition for this result is simple: from Lemma 1, we know that naive's retirement age is increasing in α. Since in the second best setting the planner can not force individuals to retire at the optimal z_j, establishing a tighter link between working career through a higher α^{opt} is a way to induce workers not to quit earlier their job and total welfare (Proposition 1)[22]. Moreover, the term M is increasing in overconfidence ($(1-\beta)$): the more naive overestimate their retirement age, ΔZ, and pension benefits, ΔP, the tighter the link between pension and retirement age should be.

8 The Political Equilibrium

In this section we consider a voting model in which all individuals vote simultaneously over the payroll tax $\tau \in [0, 1]$, and the Bismarckian factor, $\alpha \in [0, 1]$. The political game works as follows: elections take place every period. Since each individual has zero mass, everyone vote sincerely. Because of the bidimensionality of the voting space, a Condorcet winner may not exist. We handle this problem by using the notion of structure induced equilibrium defined by Shepsle (1979), which ensures the existence of an equilibrium if

[22] Notice that, if all individuals were exponential $\left(N^{y,TI} = 0\right)$, the second term at the numerator and the second one at the denominator would disappear, and (22) would be lower.

the multidimensional voting game is transformed into an issue-by-issue voting game. We first determine the majority voting equilibrium payroll tax, for a given level of redistribution: $\tau^{mv}(\alpha)$. Then, for every value of the tax rate, we compute the majority voting Bismarckian factor, for given payroll tax, $\alpha^{mv}(\tau)$. The (structure induced) equilibrium of our game (τ^s, α^s), if any, is the point at which these two functions intersect.

8.1 Voting over τ

Because of A5, preferred tax rates for young and old are the solutions to the following maximizations:

$$\max_{\tau\in[0,1]} V^y(\tau_j^y, \alpha; \hat{\beta}_j, \omega)$$
$$\max_{\tau\in[0,1]} V^o(\tau_j^o, \alpha; \beta_j, \omega)$$

where α is considered as given. Let us define, respectively, the indirect utility functions for young of type $(\hat{\beta}_j,\omega)$ and old of type (β_j, ω), as follows:

$$\begin{aligned} V^y(\tau_j^y, \alpha; \hat{\beta}_j, \omega) = \quad & u(\omega(1-\tau_j^y) - \widehat{s}_j) + \\ & + \hat{\beta}_j \delta u \left(\omega \widehat{z}_j (1-\theta\tau_j^y) - \tfrac{\gamma(\widehat{z}_j)^2}{2} + \delta\left[P(\widehat{z}_j, \tau_j^y, \alpha) + \tfrac{\psi(1-\widehat{z}_j)^2}{2} + (1+r)\widehat{s}_j\right]\right) \end{aligned} \tag{23}$$

$$V^o(\tau_j^o, \alpha; \beta_j, \omega) = \quad u\left(\omega z_j (1-\theta\tau_j^o) - \tfrac{\gamma(z_j)^2}{2} + \beta_j\delta\left[P(z_j, \tau_j^o, \alpha) + \tfrac{\psi(1-z_j)^2}{2} + (1+r)s_j\right]\right) \tag{24}$$

8.1.1 The Young

Since the tax rate chosen by majority voting not only influences the size of the PAYG system, but also individuals' decision about retirement age: it is possible that, for some values of τ_j^y, an individual prefers the pension system to private saving and for some others not. To take into account this possibility, the maximization problem should be modified as follows:

$$\max_{\widehat{s}_j, \tau_j^y} u(\omega(1-\tau_j^y) - \widehat{s}_j) + \hat{\beta}_j \delta u \left[\omega \widehat{z}_j (1-\tau\theta) - \frac{\gamma(\widehat{z}_j)^2}{2} + \delta\left(P(\widehat{z}_j, \tau_j^y, \alpha) + \frac{\psi(1-\widehat{z}_j)^2}{2}\right)\right]$$

subject to, $\forall j = e, n, s$:

$$\tau_j^y \geq 0$$
$$\widehat{s}_j = c_y - \omega(1-\tau_j^y) \leq 0$$

The solution to the above problem give us young's most preferred tax rate, $\tau_y^+(\omega, \alpha; \hat{\beta}_j)$. In the appendix we show that results depend on the value of $\alpha\delta$. The following two propositions summarize our findings. We first consider the case $\alpha\delta < 1/2$.

Proposition 3 *Suppose $\alpha\delta > 1/2$. Then:*
(i) Preferred tax rates $\tau_y^+(\omega, \alpha; \hat{\beta}_j)$ are positive for individuals with income $\omega_- \leq \omega < \tilde{\omega}$ and decreasing with ω;

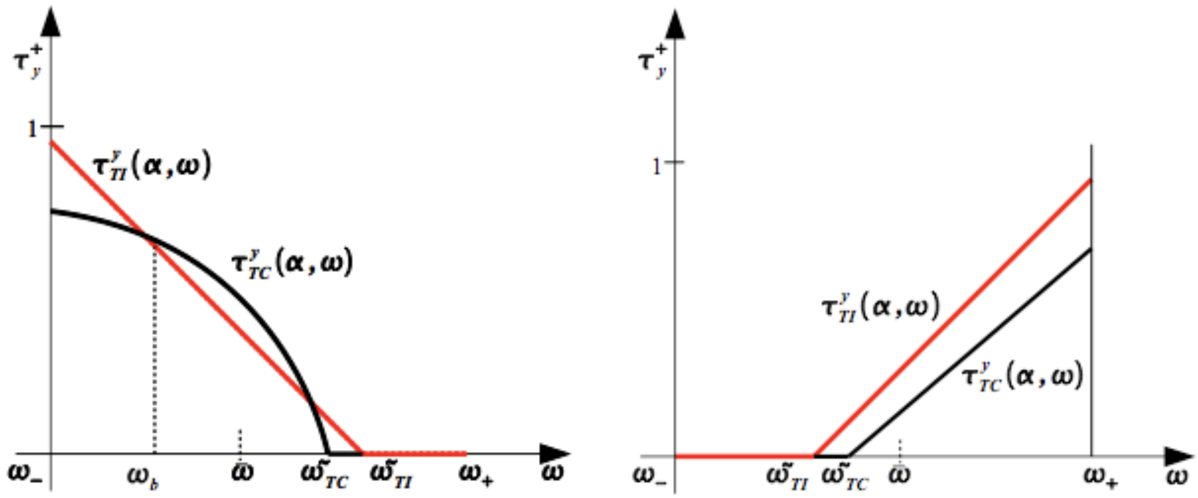

Figure 4: Preferred τ for time consistent and hyperbolic young: cases $\alpha\delta < 1/2$ (left) and $\alpha\delta > 1/2$ (right).

(ii) There exists a threshold $\omega'(< \tilde{\omega})$ such that saving are zero for individuals with $\omega \leq \omega'$, and positive and increasing with income thereafter;
(iii) For individuals with $\tau_y^+(\omega, \alpha; \hat{\beta}_j) > 0$, there exists a threshold ω_b such that $\frac{\partial \tau_y^+(\omega,\alpha;\hat{\beta}_j)}{\partial \hat{\beta}_j} < 0$ for $\omega_- \leq \omega < \omega_b$ and $\frac{\partial \tau_y^+(\omega,\alpha;\hat{\beta}_j)}{\partial \hat{\beta}_j} \geq 0$ for $\omega_b \leq \omega < \tilde{\omega}$;
(iv) The threshold $\tilde{\omega}$ is strictly decreasing in the degree of perceived time inconsistency $\hat{\beta}_j$: $\tilde{\omega}_{TI} > \tilde{\omega}_{TC}$;
(v) No young individual chooses a corner solution at $\tau_y^+(\omega, \alpha; \hat{\beta}_j) = 1$;
(vi) It exists a threshold $\widehat{\omega}$ such that $\frac{\partial \tau_y^+(\omega,\alpha;\hat{\beta}_j)}{\partial \alpha} > 0$ for $\widehat{\omega} \leq \omega < \tilde{\omega}$, and $\frac{\partial \tau_y^+(\omega,\alpha;\hat{\beta}_j)}{\partial \alpha} < 0$ otherwise.

Proof. See Appendix B. ■

We now give a sketch of the main intuitions, with the help of Figure 4, where we have denoted with $\tau_{TC}^y \equiv \tau_y^+(\omega, \alpha; \hat{\beta}_j = 1)$ preferred tax rates for *perceived* time consistent (naive and exponential) young, and with $\tau_{TI}^y \equiv \tau_y^+(\omega, \alpha; \hat{\beta}_j = \beta)$ preferred taxes for sophisticated.

In part *(i)* we show that, for given $\hat{\beta}$, only individuals with productivity levels up to $\tilde{\omega}$ support the social security system. Rich individuals, on the other hand, finance period three consumption through private saving, and vote for $\tau = 0$. Moreover, $\tau_y^+(\omega, \alpha; \hat{\beta}_j)$ is decreasing with productivity, a result in line with the literature. Several forces contribute to this result: first, notice that current consumption, c_y^j, is increasing with ω. Also consumption when old, c_o^j, is increasing with income, provided that it is a normal good. Consequently, rich individuals would like to transfer more resource in the second period through a higher τ (income effect). Second, a substitution effect operates: if ω rises, the relative price of first and second period consumption decreases, and this effect is more marked for low income workers. For utility functions such that the coefficient of relative risk aversion[23], ϵ, is lower than 1, as we assume, the substitution effect dominates the income effect: low productivity individuals prefer larger τ than high productivity ones. Finally, assuming endogenous retirement age leads to a third effect (Casamatta et al., 2005): second period consumption increases with productivity, given that more productive individuals retire later (Lemma 1), and young rich raise their first period consumption by reducing the payroll tax.

Result *(ii)* is in line with Casamatta et al. (1999): it exists a threshold ω' such that individuals with productivity below ω' do not save and rely only on the pension system to finance third period consumption. On the other and, for productivities above ω', saving are positive and increasing with ω. Combining this result with *(i)*, we see that, for $\omega' \leq \omega \leq \tilde{\omega}$, we have interior solutions for both τ and $\widehat{s_j}$. Furthermore, since saving are increasing with income and preferred tax rate are decreasing with it, we see that workers progressively replace the pension system with private savings, up to the threshold $\tilde{\omega}$, where preferred τ is zero.

Parts *(iii)* and *(iv)* are a novelty of this work and follow from our behavioral assumption. In *(iii)* we compare, for given ω, sophisticated preferred tax rates, τ^y_{TI}, with those of exponential and naive, τ^y_{TC}. We show that it exists a non-monotonic relationship between the two tax rates: in particular, we have $\tau^y_{TC} < \tau^y_{TI}$ for $\omega_- \leq \omega \leq \omega_b$ and $\tau^y_{TC} > \tau^y_{TI}$ thereafter. The intuition for the first part of the proposition is simple: notice that, for $\omega \leq \omega_b$, saving are zero and $z_s < z_e$, for a given ω. Sophisticated find optimal to set τ higher than exponential: a higher tax entails only a second-order loss, since it further reduces retirement age, but has a first-order gain, as it increases the size of pension system. Since the former effect is less marked for sophisticated, given that z_s is already suboptimal, $\tau^y_{TC} < \tau^y_{TI}$.

The second part of *(iii)* is more counterintuitive: we show that for $\omega > \omega_b$, $\tau^y_{TC} > \tau^y_{TI}$, *i.e.* sophisticated prefer a smaller payroll tax than exponential with the same ω. The intuition for this result comes from the "commitment device" argument, typical in the Economics and Psychology literature (see Gruber and Köszegi, 2004). In absence of a publicly-provided commitment instrument that would help hyperbolic discounters to overcome their self-control issues, sophisticated look for a "personal" commitment device that increases at the same time z_s and s_s, whose values are suboptimal. In Lemma 1 and Appendix B, we have shown that both retirement age and saving are decreasing with τ. Moreover, from Propositions 1 and 2, we know that increasing z_s and s_s towards z_e and s_e (normative retirement age and saving), increases sophisticated welfare as well. Therefore, the commitment device for hyperbolic individuals takes the form of a *lower* payroll tax.

In *(iv)*, we demonstrate that $\tilde{\omega}$, the income threshold such that private savings are preferred to the pension system, is greater for sophisticated than for exponential and naive: $\widetilde{\omega}_{TI} > \widetilde{\omega}_{TC}$. This is intuitive: whereas time consistent workers with $\omega \geq \widetilde{\omega}_{TC}$ find unattractive the pension sytem, as they can transfer autonomously income in the second and third periods, sophisticated with productivity levels $\omega \in [\widetilde{\omega}_{TC}, \widetilde{\omega}_{TI}]$, still find attractive it.

[23] The coefficient of relative risk aversion, for a generic increasing and concave utility function $u(x)$, is defined as $\epsilon \equiv \frac{-xu''(x)}{u'(x)}$. Assuming $\epsilon \leq 1$ is standard in the literature (see Cremer et al. 2004).

Part *(v)* is straightforward: with $\tau_y^+(\omega, \alpha; \hat{\beta}_j) = 1$, marginal utility of consumption in the first period tends to infinite, and thus young prefers tax smaller than 1.

In *(vi)*, we show that most preferred tax rates are concave in α. The intuition for this result is the following: an increase of α has beneficial effects for rich individuals (those with with $\omega \geq \widehat{\omega}$), since it increases $\widehat{z}_j$, $P(\widehat{z}_j)$ and c_o^j and tights the link between pension and working career. Moreover, the benefits of a higher α are amplified if the size of the PAYG is further increased, given that the costs of a high τ (reduced retirement age) have only a second-order effect on rich total utility. Therefore, preferred τ must increase with α. For poor individuals, $\tau_y^+(\omega, \alpha; \hat{\beta}_j)$ is decreasing with α: when $\alpha\delta > 1/2$, and redistribution is low, a greater α further reduces it. Poor individuals have to rise their z to have more resources in period three: this can be done only if the payroll tax is set to a lower level. It follows that τ must be decreasing with α, in order to increase both labor supply and $P(\hat{z}_j)$. On top of that, by continuity of the preferred tax rates function, there exists a threshold level of α, $\hat{\alpha} \equiv \left[\frac{\left[\left(\frac{\bar{\omega}}{\omega}\right) - (1+n)\delta\right] \frac{\gamma - \hat{\beta}_j \delta\psi}{\delta\theta} - \omega(1 - \theta\tau(2+\hat{\beta}_j) - \hat{\beta}_j \delta\psi)}{4\omega\theta\tau\delta\hat{\beta}_j} \right]$ such that $\tau_y^+(\omega, \alpha; \hat{\beta}_j)$ is increasing in α for $\alpha \geq \hat{\alpha}$ and decreasing otherwise.

We move now to the second case, $\alpha\delta < 1/2$, which implies that the redistributive part in the pension formula is high.

Proposition 4 *Suppose $\alpha\delta < 1/2$. Then:*

(i) Preferred tax rates $\tau_y^+(\omega, \alpha; \hat{\beta}_j)$ are positive and increasing with productivity for individuals with $\tilde{\omega} < \omega \leq \omega^+$;

(ii) The threshold $\tilde{\omega}$ is strictly increasing in the degree of perceived time inconsistency $\hat{\beta}_j$;

(iii)For individuals with $\tau_y^+(\omega, \alpha; \hat{\beta}_j) > 0$, the preferred tax rate is decreasing in the degree of perceived time consistency $\hat{\beta}_j$, $\frac{\partial \tau_y^+(\omega, \alpha; \hat{\beta}_j)}{\partial \hat{\beta}_j} < 0$.

(iv) No young individual chooses a corner solution at $\tau_y^+(\omega, \alpha; \hat{\beta}_j) = 1$.

(v) For individuals above the average income, i.e. $(\bar{\omega} \leq) \tilde{\omega} \leq \omega < \omega^+$, we have that $\frac{\partial \tau_y^+(\omega, \alpha; \hat{\beta}_j)}{\partial \alpha} > 0$.

Proof. See Appendix B. ■

In part *(i)* we show that only rich individuals prefer a positive tax rate: by increasing τ, they can transfer more resources from first to second period, provided that consumption is a normal good. Poor workers, on the other side, will retire earlier and prefer to keep the tax rate as low as possible, in order to increase retirement age and saving.

Part *(ii)* says that the higher is the perceived time consistency of the worker, the less support the social security system receives: for sophisticated, the effect described in part *(i)* emerges for lower productivity levels, because these agents realize they are not able to respect their plans.

In part *(iii)*, we show that time-inconsistent agents prefer a higher tax rate than exponential, for a given productivity level. The intuition goes as follows: given that α is low, rich workers increase $P(\hat{z}_j)$ through a higher tax rate, and this effect is more marked for hyperbolic individuals. The commitment effect of

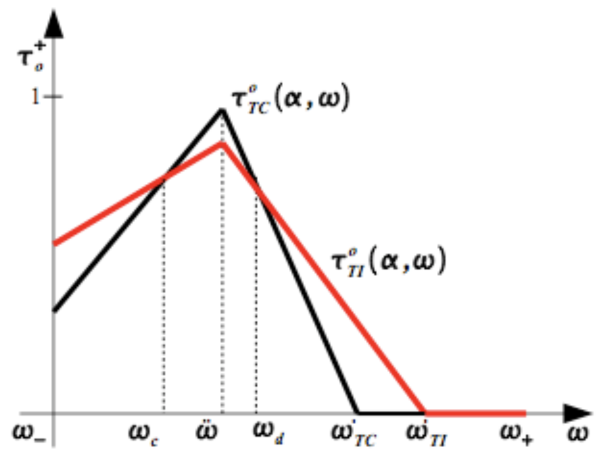

Figure 5: Preferred τ for time consistent and hyperbolic old.

proposition 3 does not operate here: those in favor of a positive τ have already positive saving and high retirement age. The increase of utility to a lower τ (which increases z and s) has a second order effect than the increase in utility due to a more generous pension, given that $\partial z/\partial\tau$ is decreasing with income. Time consistent individuals, on the other hand, are less interested in transferring resources through the PAYG system, given their optimal choices, thus preferring a lower, although positive, τ.

Part *(v)* is similar to Proposition 3: here, given that $\tilde{\omega} \geq \bar{\omega}$, only the case $\frac{\partial\tau_y^+(\omega,\alpha;\hat{\beta}_j)}{\partial\alpha} > 0$ is relevant.

8.1.2 The Old

Maximization of (24) leads to the following proposition.

Proposition 5 *Preferred tax rates for old have the following properties:*

(i) Preferred tax rates $\tau_o^+(\omega,\alpha;\beta_j)$ *are positive for individuals with productivity* $\omega_- \leq \omega \leq \dot{\omega}$ *and always lower than 1.*

(ii) Preferred tax rates $\tau_o^+(\omega,\alpha;\beta_j)$ *are increasing for productivity levels* $\omega_- \leq \omega < \ddot{\omega}$ *and decreasing in* ω *for* $\ddot{\omega} \leq \omega \leq \dot{\omega}$;

(iii) The threshold $\dot{\omega}$ *is strictly decreasing in* β_j;

(iv) Preferred tax rates $\tau_o^+(\omega,\alpha;\beta_j)$ *are increasing with* β_j *for income levels* $\omega_c < \omega < \omega_d$ *and decreasing with* β_j, $\forall\omega$ *such that* $\omega_- \leq \omega \leq \omega_c$ *and* $\omega_d \leq \omega \leq \dot{\omega}$;

(v) For individuals above the average income, i.e. $\bar{\omega} \leq \omega < \widetilde{\omega}$, *we have that* $\frac{\partial\tau_o^+(\omega,\alpha;\beta_j)}{\partial\alpha} > 0$, *whereas for very poor individuals,* $\frac{\partial\tau_o^+(\omega,\alpha;\beta_j)}{\partial\alpha} < 0$.

Proof. See Appendix B. ■

We provide now a sketch of the main intuitions, with the help of Figure 5, where $\tau^o_{TC} \equiv \tau^+_o(\omega, \alpha; \beta = 1)$ denotes preferred tax rates of time consistent old and $\tau^o_{TI} \equiv \tau^+_o(\omega, \alpha; \beta)$ those of naive and sophisticated.

In part *(i)* we demostrate that only agents with productivity levels below $\dot{\omega}$ support the social security system; this result contrasts with the literature on positive social security (Casamatta et al. 1999), in which every old votes for a positive τ and at least a fraction for $\tau = 1$. This result follows from the fact that labor supply at period two is elastic and from the concavity of the perceived pension function; to see why, notice that increasing τ has two opposite effects on individuals' utility: from one side, it reduces retirement age (Lemma 1) but, from the other, it increases pension benefits, provided that the tax rate is below the threshold $\hat{\tau}$ (Lemma 2). For individuals with productivity up to $\dot{\omega}$, the second effect overweights the first one: preferred tax rates are positive.

In *(ii)* we show that, for a given β, $\tau^+_o(\omega, \alpha; \beta_j)$ is a concave function with a maximum at $\ddot{\omega}$. This effect is due to the two effects pointed out in part *(i)* and a third effect: in Lemma 1, we show that $\frac{\partial \hat{z}_j}{\partial \tau}$ is increasing with ω, *i.e.* poor reduces more than rich z in response to an increase in the payroll tax. It follows that, for low productivity levels, the first effect prevail over the other two, and tax rates are increasing with income, up to $\ddot{\omega}$. For income levels $\omega \in [\ddot{\omega}, \dot{\omega}]$ however, the third effect prevails, and the reduction of z is less than compensated by the increase in P: τ decreases with income. Moreover, we have shown that individuals with high productivity levels start to replace the PAYG system with private saving. This continues up to $\dot{\omega}$, where an individual exclusively relies on accumulated saving and vote for $\tau^+_o = 0$.

In *(iii)* we claim that exponential old are less likely to support the pension system. The threshold such that private saving are preferred to the social security system is lower for time consistent old: $\dot{\omega}_{TI} > \dot{\omega}_{TC}$. Rich exponential have optimal private saving and a longer career, and a generous pension system is not necessary. On the other hand, naive and sophisticated, who have realized the suboptimality of their choices, support social security for higher productivity levels than exponential, as the want to raise post-retirement consumption levels.

In *(iv)* we show that, for given ω, preferred tax rates are non-monotonic in β_j; more precisely, for productivity levels below ω_c and above ω_d, time inconsistent individuals prefer higher τ than exponential. Both hyperbolic poor (those with $\omega_- \leq \omega \leq \omega_c$) and hyperbolic rich (those with $\omega_d \leq \omega \leq \dot{\omega}$) prefer a more generous system in order to increase post retirement consumption, but the reasons behind such behavior differ: since α is given, the former group would like to augment the size of Beveridgean component of $P(z_j)$, whereas the latter the size of the Bismarckian part. Hyperbolic agents, *i.e.* those with $\omega_c < \omega < \omega_d$, on the other hand, prefer a lower τ than exponential. The intuition goes as follows: for given α, these individuals are less interested in increasing the size of the Beveridgean part, since their productivity level is around the average; moreover, their lack of self control makes the Bismarckian part less attractive. It follows that a commitment device strategy operates here: lowering τ increases consumption for second and third periods, increases z, but at a cost, since it reduces the pension benefits. However, since period three utility is discounted at the hyperbolic factor $\beta\delta$, this loss has only a second order effect in total utility.

Intuition for *(v)* goes exactly as for Proposition 3, and thus is omitted.

8.2 Equilibrium Tax Rate

In the appendix we prove that preferences over τ are single-peaked, both for young and old voters, and the median voter theorem applies. In the following, to highlight our contribution with respect to traditional political economy model of social security, we compare the equilibrium tax rate of our hyperbolic voting game, $\tau_{TI}^{mv}(\alpha)$, to the tax rate that would emerge in a model with exponential discounting, $\tau_{TC}^{mv}(\alpha)$.

Since preferred tax rates depends crucially on the value of the parameter $\alpha\delta$, we will consider separately the two cases.

Case $\alpha\delta > 1/2$

Figure 6 and Proposition 6 illustrate the Condorcet winner of the voting game[24].

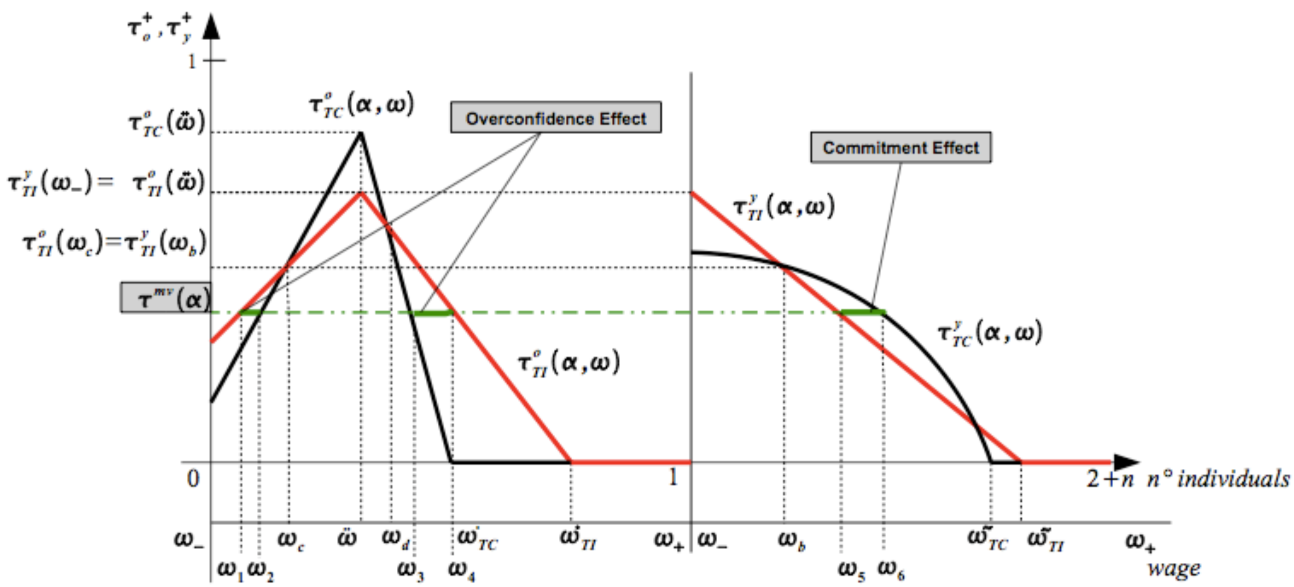

Figure 6: Equilibrium τ when $\alpha\delta > 1/2$ and $\tau_{TC}^o(\ddot{\omega}) \geq \tau_{TI}^y(\omega_-)$.

Proposition 6 *If $\alpha\delta \geq 1/2$ and $\tau_{TC}^o(\ddot{\omega}) \geq \tau_{TI}^y(\omega_-)$, the majority voting equilibrium tax rate $\tau^{mv}(\alpha)$ satisfies the following conditions:*

(i) If $(1+n)N^s \int_{\tilde{\omega}_{TI}}^{\omega_+} f(\omega)d\omega + (N^s + N^n) \int_{\dot{\omega}_{TI}}^{\omega_+} f(\omega)d\omega > N(2+n)/2$, then $\tau_{TI}^{mv}(\alpha) = 0$.

[24] In the appendix we show that $\tau_{TC}^o(\ddot{\omega})$, the maximal preferred tax rate for time consistent old, and $\tau_{TI}^y(\omega_-)$, the maximal preferred tax rate for hyperbolic young are not comparable: both cases are possible. However, the case $\tau_{TC}^o(\ddot{\omega}) < \tau_{TI}^y(\omega_-)$ gives the same results of the case $\tau_{TC}^o(\ddot{\omega}) \geq \tau_{TI}^y(\omega_-)$. Therefore, without loss of generality, we can focus only on the second one.

(ii) The majority voting equilibrium $\tau_{TI}^{mv}(\alpha)$ is positive if and only if:

$$N^o \int_{\omega_-}^{\dot{\omega}_{TC}} f(\omega)d\omega + (1+n)N^y \int_{\omega_-}^{\tilde{\omega}_{TC}} f(\omega)d\omega + \underbrace{(N^s + N^n) \int_{\dot{\omega}_{TC}}^{\dot{\omega}_{TI}} f(\omega)d\omega + (1+n)N^s \int_{\tilde{\omega}_{TC}}^{\tilde{\omega}_{TI}} f(\omega)d\omega}_{\textit{extra support for SS}} \geq \frac{N(2+n)}{2}$$

(iii) $\tau_{TI}^{mv}(\alpha)$ is the rate preferred by the workers with earnings $\omega_1, \omega_2, \omega_3, \omega_4$, ω_5 or ω_6 such that:

$$N^o \left(\int_{\omega_2}^{\omega_3} f(\omega)d\omega + (1+n) \int_{\omega_-}^{\omega_6} f(\omega)d\omega \right) + \underbrace{(N^s + N^e) \left(\int_{\omega_1}^{\omega_2} f(\omega)d\omega + \int_{\omega_3}^{\omega_4} f(\omega)d\omega \right)}_{\textit{Overconfidence Effect (+)}} +$$

$$- \underbrace{N^s(1+n) \int_{\omega_5}^{\omega_6} f(\omega)d\omega}_{\textit{Commitment Effect (-)}} = \frac{N(2+n)}{2}$$

and

$$\tau_{TI}^{mv}(\alpha) = \tau_o^+(\omega_1, \alpha; \beta_j) = \tau_o^+(\omega_2, \alpha; \beta_j) = \tau_o^+(\omega_3, \alpha; \beta_j) = \tau_o^+(\omega_4, \alpha; \beta_j) = \tau_y^+(\omega_5, \alpha; \hat{\beta}_j) = \tau_y^+(\omega_6, \alpha; \hat{\beta}_j)$$

(iv) $\tau_{TI}^{mv}(\alpha) < \tau_{TC}^{mv}(\alpha)$ if and only if:

$$(N^s + N^e) \left(\int_{\omega_1}^{\omega_2} f(\omega)d\omega + \int_{\omega_3}^{\omega_4} f(\omega)d\omega \right) < N^s(1+n) \int_{\omega_5}^{\omega_6} f(\omega)d\omega$$

We give a sketch the proof, which relies on standard arguments, through Figure 6, where we have supposed that the median income is such that the equilibrium payroll tax is $\tau_{TI}^{mv}(\alpha)$.

In *(i)*, we show that social security can not be sustained as an equilibrium, if the number of voters in favor of aτ rate does not represent at least half of the population.

Part *(ii)* follows directly from Propositions 3 *(iii)* and 5 *(iii)*. We show that, for a given ω, time inconsistent agents are more favorable to support a pension system than exponential: thresholds $\widetilde{\omega}$ and $\dot{\omega}$ (productivities level such that preferred τ are positive) are both decreasing with $\hat{\beta}_j$[25]. It follows that, if the fraction of hyperbolic in the population with productivity $\omega \in [\dot{\omega}_{TC}, \dot{\omega}_{TI}]$ and $\omega \in [\tilde{\omega}_{TC}, \tilde{\omega}_{TI}]$, the pension system is more likely to be supported as an equilibrium. The *extra-support* for social security due to time inconsistency is given by the term $(N^{o,s} + N^{o,n}) \int_{\dot{\omega}_{TC}}^{\dot{\omega}_{TI}} f(\omega)d\omega + (1+n)N^{o,s} \int_{\tilde{\omega}_{TC}}^{\tilde{\omega}_{TI}} f(\omega)d\omega$.

In part *(iii)* we state that the majority voting equilibrium tax rate $\tau_{TI}^{mv}(\alpha)$ is determined by two opposite forces. The first one, the **overconfidence** effect, increases $\tau_{TI}^{mv}(\alpha)$: old naive have realized the suboptimality of their choices, and favor a more generous system in order to compensate the drop in post retirement consumption. The second one, the **commitment** effect, goes in the opposite direction: for a given productivity level, a sophisticated young favor a lower τ than a young exponential. As stressed above, sophisticated young, aware of their present bias, see a lower payroll tax as a commitment device that increases both z_s and s_s and mitigate the negative effects of time inconsistency.

[25] Most preferred tax rates for exponential old with productivity $\in [\dot{\omega}_{TC}, \dot{\omega}_{TI}]$ and for exponential and naive young with $\omega \in [\tilde{\omega}_{TC}, \tilde{\omega}_{TI}$ are zero, whereas for sophisticated young and naive old, these tax rates are positive.

In part *(iv)*, we compare the hyperbolic equilibrium tax rate, $\tau_{TI}^{mv}(\alpha)$, to the one that would emerge in a standard exponential model, $\tau_{TC}^{mv}(\alpha)$. We show that, if the coalition made of hyperbolic young and rich individuals outnumbers that made of poor old and time consistent young, *i.e.* the commitment effect is higher than overconfidence, we have in the hyperbolic equilibrium a smaller pension system that of the exponential model: $\tau_{TI}^{mv}(\alpha) < \tau_{TC}^{mv}(\alpha)$.

Case $\alpha\delta < 1/2$

When the Bismarckian part of the pension system is below $1/2\delta$, the characterization of the equilibrium is slightly modified. Proposition 7 summarizes our findings.

Proposition 7 *If $\alpha\delta < 1/2$, the majority voting equilibrium tax rate satisfies the following conditions:*
(i) If $(1+n)N^s \int_{\omega_-}^{\tilde{\omega}_{TI}} f(\omega)d\omega + (N^s + N^n) \int_{\dot{\omega}_{TI}}^{\omega_+} f(\omega)d\omega > N(2+n)/2$, then the majority voting equilibrium tax rate $\tau_{TI}^{mv}(\alpha)$ is 0.

(ii) The majority voting equilibrium $\tau_{TI}^{mv}(\alpha)$ is positive if:

$$N^o \int_{\omega_-}^{\dot{\omega}_{TC}} f(\omega)d\omega + (1+n)N^y \int_{\tilde{\omega}_{TC}}^{\omega_+} f(\omega)d\omega + \underbrace{(N^s + N^n) \int_{\dot{\omega}_{TC}}^{\dot{\omega}_{TI}} f(\omega)d\omega + (1+n)N^s \int_{\tilde{\omega}_{TC}}^{\tilde{\omega}_{TI}} f(\omega)d\omega}_{\textit{extra support for SS}} \geq \frac{N(2+n)}{2}$$

(iii) The majority voting equilibrium tax rate is the rate preferred by the workers with earnings $\omega_1, \omega_2, \omega_3, \omega_4$, ω_5 or ω_6 such that:

$$N^o \left(\int_{\omega_2}^{\omega_3} f(\omega)d\omega + (1+n) \int_{\omega_6}^{\omega_+} f(\omega)d\omega \right) + \underbrace{(N^s + N^e) \left(\int_{\omega_1}^{\omega_2} f(\omega)d\omega + \int_{\omega_3}^{\omega_4} f(\omega)d\omega \right)}_{\textit{Overconfidence Effect (+)}}$$
$$+ \underbrace{N^s(1+n) \int_{\omega_5}^{\omega_6} f(\omega)d\omega}_{\textit{Commitment Effect (+)}} = \frac{N(2+n)}{2}$$

and

$$\tau_{TI}^{mv}(\alpha) = \tau_o^+(\omega_1, \alpha; \beta_j) = \tau_o^+(\omega_2, \alpha; \beta_j) = \tau_o^+(\omega_3, \alpha; \beta_j) = \tau_o^+(\omega_4, \alpha; \beta_j) = \tau_y^+(\omega_5, \alpha; \hat{\beta}_j) = \tau_y^+(\omega_6, \alpha; \hat{\beta}_j)$$

(iv) If $N^{y,TI} \neq 0$ (or $N^{o,TI} \neq 0$), $\tau_{TI}^{mv}(\alpha) > \tau_{TC}^{mv}(\alpha)$.

Parts *(i)* and *(ii)* parallel Proposition 6. The only difference is given by *(iii)*, in which we determine the equilibrium tax rate, $\tau_{TI}^{mv}(\alpha)$. Contrary to the previous case, here both the overconfidence and the commitment effect go in the same direction, and contributes to increase the equilibrium payroll tax. The intuition for this result is the following: now the commitment device for sophisticated young is a *higher* τ; since redistribution is high, and only rich favor a positive tax, there is no need to further increase z_j. The only way to transfer resources in the second period is to increase the size of the system: the loss in utility due to a reduced retirement age is more the compensated by the increase in the generosity of the system.

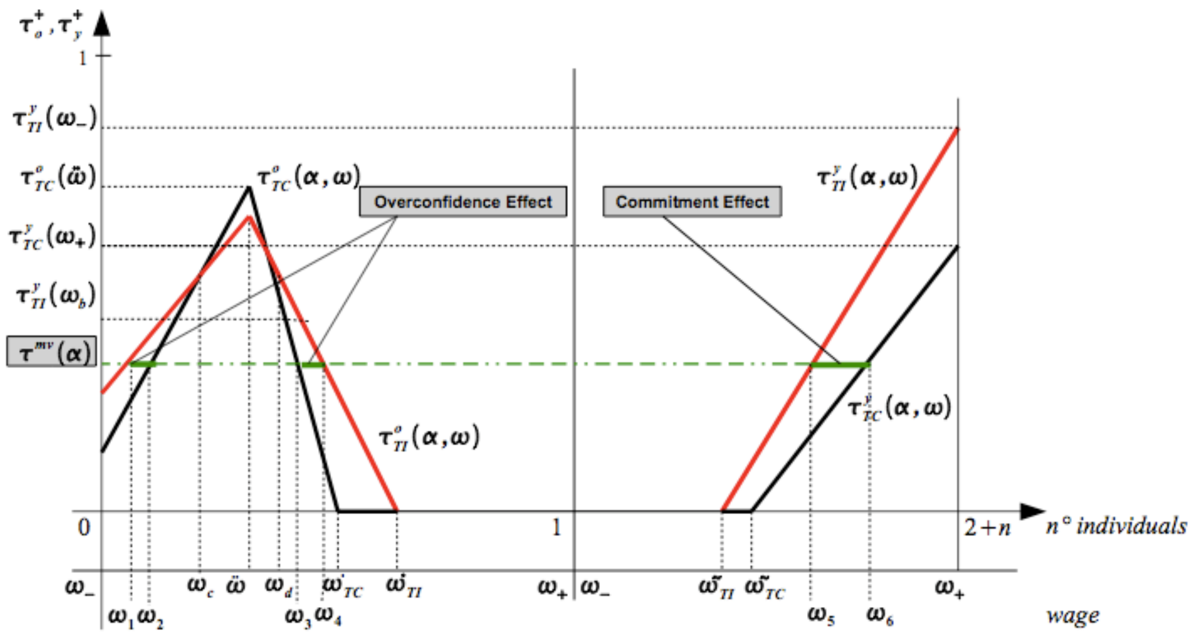

Figure 7: Equilibrium τ when $\alpha\delta < 1/2$.

It follows that, when $\alpha\delta < 1/2$, the equilibrium payroll tax $\tau_{TI}^{mv}(\alpha)$ is higher than $\tau_{TC}^{mv}(\alpha)$: the winning coalition is now composed by time inconsistent agents (old and young), who prefer a higher tax than time consistent individuals.

8.3 Voting over the Bismarckian Factor

In the appendix, we show that preferences over α are single-crossing both for young and for old: hence, a voting equilibrium on α exists. When τ is kept fixed, most preferred levels of α are the solution to the following problems:

$$\max_{\alpha \in [0,1]} V^o(\tau, \alpha; \beta_j, \omega) \tag{25}$$

$$\max_{\alpha \in [0,1]} V^y(\tau, \alpha; \hat{\beta}_j, \omega) \tag{26}$$

where the expressions for the two indirect utility functions are given, respectively, by equations (23) and (24). Because of A5, the problem is not same for the 2 generations: young evaluates future utility and consumption according to the perceived discount factor $\hat{\beta}_j\delta^2$, whereas old uses the true discount factor $\beta_j\delta$.

8.3.1 The Old

The following proposition summarizes our findings about old's preferred levels of the Bismarckian factor, $\alpha_o^+(\tau, \omega; \beta_j)$ (see Figure 8 for a graphical interpretation).

Proposition 8 *Old's preferred levels of α have the following properties:*

(i) It exists a productivity level ω_e such that $\alpha_o^+(\tau,\omega;\beta_j) > 0$ for $\omega_e \leq \omega \leq \omega_+$. Otherwise, $\alpha_o^+(\tau,\omega;\beta_j) = 0$; moreover $\frac{\partial\alpha_o^+(\tau,\omega;\hat{\beta}_j)}{\partial\omega} > 0$;

(ii) The threshold ω_e is decreasing in the degree of true time inconsistency $\beta_j : \omega_e^{TC} < \omega_e^{TI}$;

(iii) $\frac{\partial\alpha_o^+(\tau,\omega;\beta_j)}{\partial\beta_j} > 0$;

(iv) For individuals above ω_e, we have that $\frac{\partial\alpha_o^+(\tau,\omega;\beta_j)}{\partial\tau} > 0$.

Proof. See Appendix B. ■

In part *(i)* we show that, for given β, only old with $\omega > \omega_e$ favor $\alpha_o^+ > 0$, while individuals with $\omega \leq \omega_e (< \bar{\omega})$ prefer a flat pension and vote for $\alpha_o^+ = 0$. To see why poor prefer a purely Beveridgean system, we have to consider how changes in the Bismarckian factor influence individuals' utility: first, higher α increases z, as it establishes a closer link between working career and pension benefits; second, higher α reduces redistribution. For $\omega < \omega_e$, the second effect prevails, whereas for $\omega > \omega_e$ agents prefer to decrease redistribution and to make the system more Bismarckian. Moreover, as z is increasing with ω, also α is increasing with productivity.

Part *(ii)* shows that, for given ω, the income threshold such that $\alpha_o^+(\tau,\omega;\beta_j)$ is strictly positive, is lower for exponential than for hyperbolic: time consistent individuals, who have a longer career, prefer to have a more Bismarckian system than sophisticated with the same income level. This is quite intuitive, as hyperbolic, who have realized that their retirement plans are not optimal, and prefer to limit the link between pension benefit and working history. An interesting corollary is that, contrary to the literature (Casamatta et al. 2005), also individuals with productivity below the average prefer a positive α: for those with $\omega \in [\omega_e, \bar{\omega}]$, the increase in utility due to a longer career and a Bismarckian pension more than compensate the loss due to a reduction of the flat part of the transfer.

It follows that (part *iii*) hyperbolic old prefer a lower α than exponential with the same ω; intuitively, although their income level should make them in favor of a Bismarckian system, their time inconsistency, that leads them to choose lower z and c_o, makes them in favor of a smaller link between length of the working career and $P(z)$.

In *(iv)*, we show that, for productivities above ω_e, $\alpha_o^+(\tau,\omega;\beta_j)$ is increasing with τ: augmenting the payroll tax reduces z_j and, therefore, only rich individuals with $\omega \geq \omega_e$ would like to increase the weight attached to the Bismarckian part, as to increase both z and $P(z)$.

8.3.2 The Young

The following proposition summarizes our findings about $\alpha_y^+(\tau,\omega;\hat{\beta}_j)$, young's most preferred Bismarckian factor.

Proposition 9 *For young, the most preferred level of α has the following properties:*

(i) There exist a threshold income level ω_f such that $\alpha_y^+\left(\tau,\omega;\hat{\beta}_j\right) = 0$ for $\omega_- \leq \omega \leq \omega_f$ and $\alpha_y^+\left(\tau,\omega;\hat{\beta}_j\right) >$

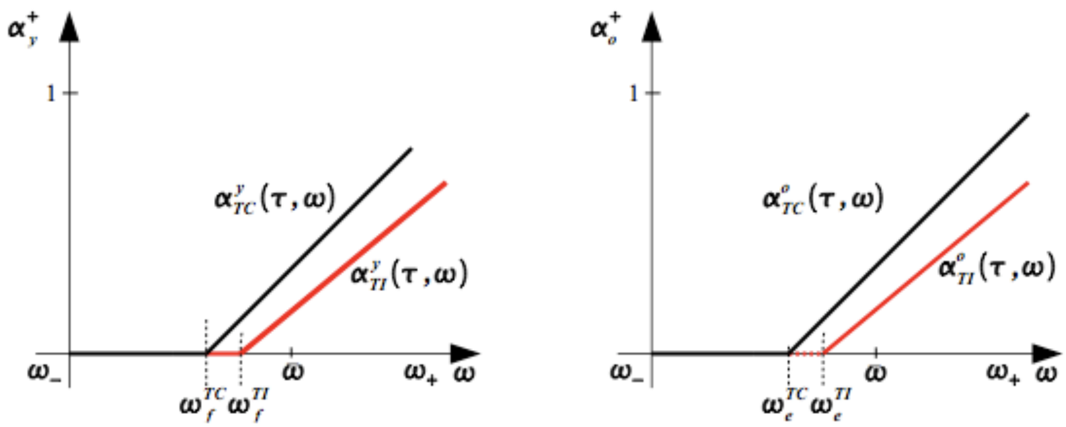

Figure 8: Preferred α for time consistent and hyperbolic old (left) and young (right).

0 otherwise; moreover, when $\alpha_y^+\left(\tau,\omega;\hat{\beta}_j\right) > 0$, *we have* $\frac{\partial \alpha_y^+(\tau,\omega;\hat{\beta}_j)}{\partial \omega} > 0$;

(ii) The threshold ω_f *is decreasing in the degree of true time inconsistency* $\beta_j : \omega_f^{TC} < \omega_f^{TI}$;

(iii) $\frac{\partial \alpha_y^+(\tau,\omega;\hat{\beta}_j)}{\partial \hat{\beta}_j} > 0$;

(iv) For $\omega \geq \omega_f$, *we have that* $\frac{\partial \alpha_y^+(\tau,\omega;\hat{\beta}_j)}{\partial \tau} > 0$

Proof. See Appendix B. ∎

Intuitions are similar to those behind Proposition 8 and are omitted.

8.4 Equilibrium Bismarckian Factor

Combining our results about $\alpha_y^+(\tau,\omega;\hat{\beta}_j)$ and $\alpha_o^+(\tau,\omega;\beta_j)$, we can determine the majority voting equilibrium value of the Bismarckian factor (see Figure 9).

Proposition 10 *The equilibrium* $\alpha_{TI}^{mv}(\tau)$ *satisfies the following conditions:*

(i)If $N^o \int_{\omega_-}^{\omega_e^{TC}} f(\omega)d\omega + (N^s+N^n)\int_{\omega_e^{TC}}^{\omega_e^{TI}} f(\omega)d\omega + (1+n)N^y \int_{\omega_-}^{\omega_f^{TC}} f(\omega)d\omega + (N^s+N^n)\int_{\omega_f^{TC}}^{\omega_f^{TI}} f(\omega)d\omega >$ $N(2+n)/2$, *then* $\alpha_{TI}^{mv}(\tau) = 0$.

(ii) The majority voting equilibrium $\alpha_{TI}^{mv}(\tau)$ *is positive if:*

$$N^o \int_{\omega_e^{TC}}^{\omega_+} f(\omega)d\omega + (1+n)N^y \int_{\omega_f^{TC}}^{\omega_+} f(\omega)d\omega - \underbrace{\left[(N^s+N^n)\int_{\omega_e^{TC}}^{\omega_e^{TI}} f(\omega)d\omega + (1+n)N^s \int_{\omega_f^{TC}}^{\omega_f^{TI}} f(\omega)d\omega\right]}_{\text{extra support for redistribution}} \geq \frac{N(2+n)}{2}$$

(iii) The majority voting equilibrium tax rate is the rate preferred by the workers with earnings $\omega_1,\omega_2,\omega_3,\omega_4$ *such that:*

$$N^o\left(\int_{\omega_1}^{\omega_+} f(\omega)d\omega + (1+n)\int_{\omega_3}^{\omega_+} f(\omega)d\omega\right) - \underbrace{(N^s+N^e)\left(\int_{\omega_1}^{\omega_2} f(\omega)d\omega\right)}_{\text{Overconfidence Effect (-)}} - \underbrace{N^s(1+n)\int_{\omega_3}^{\omega_4} f(\omega)d\omega}_{\text{Resignation Effect (-)}} = \frac{N(2+n)}{2}$$

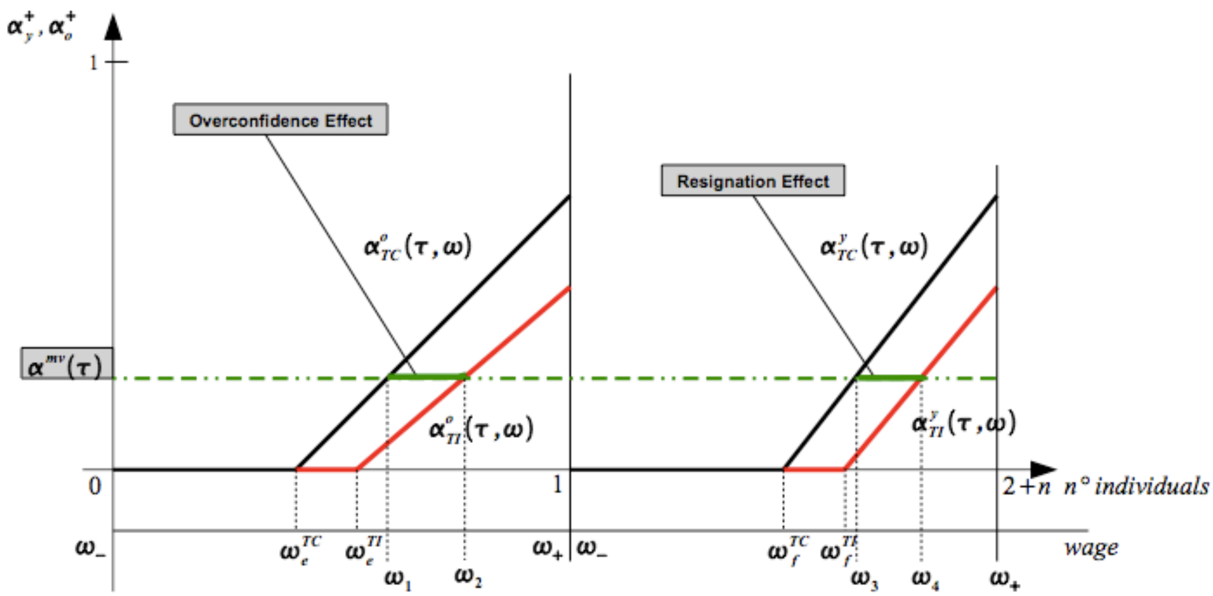

Figure 9: The majority voting equilibrium α, $\alpha_{TI}^{mv}(\alpha)$.

and

$$\alpha^{mv}(\tau) = \alpha_o^+(\omega_1, \tau; \beta_j) = \alpha_o^+(\omega_2, \tau; \beta_j) = \alpha_y^+(\omega_3, \tau; \hat{\beta}_j) = \alpha_y^+(\omega_4, \tau; \hat{\beta}_j)$$

(iv) If $N^{y,TI} \neq 0$ *(or* $N^{o,TI} \neq 0$*),* $\alpha_{TI}^{mv}(\tau) < \alpha^{opt}$.

Parts *(i)* and *(ii)* of the proposition are intuitive: if the fraction of poor and time inconsistent is high enough, the equilibrium pension system is purely Beveridgean. In particular, the extra support for $\alpha_{TI}^{mv}(\tau) = 0$ is given by middle-income hyperbolic individuals, *i.e.* those with productivity levels $\omega \in [\omega_e^{TC}, \omega_e^{TI}]$ and $\omega \in [\omega_f^{TC}, \omega_f^{TI}]$.

In *(iii)*, we compute the equilibrium level $\alpha_{TI}^{mv}(\tau)$: as for the payroll tax, time inconsistency introduces two effects in the political game: the **overconfidence** effect, displayed by old, is due to the fact that hyperbolic try to compensate the loss due to impatience by increasing the redistributive part of the pension formula, *i.e.* lowering α. Also the **resignation** effect lowers the equilibrium α: sophisticated young already know that they will not be able to stick with their optimal plans, and therefore support a more redistributive pension compared to time consistent individuals with the same ω.

In part *(iv)*, to stress the differences between our behavioral model and a standard model with exponential preferences, we distinguish between $\alpha_{TI}^{mv}(\tau)$, the Bismarckian factor emerging in the hyperbolic voting game, and $\alpha_{TC}^{mv}(\tau)$, the equilibrium when $N^{y,TI} = N^{o,TI} = 0$. If hyperbolic individuals are able to form a coalition with exponential poor, the resulting pension system is more redistributive than that an ideal, exponential, one. The reasons for this excess of redistribution are quite intuitive: hyperbolic discounting introduces a third form of redistribution, besides from young to old and from rich to poor: from time consistent to hyperbolic individuals, or from individuals with long career to early retirees, a result in line with the empirical evidence provided in Liberman (2001).

In *(v)* we compare α^{opt} with the equilibrium one; we find that, while the former is increasing in N^{TI}, the latter is actually decreasing with it: if the number of hyperbolic individuals increases, their political power increases too, thus making the pension system more Beveridgean. The intuition for this result is simple: the social planner, when setting optimal α, takes into account into the budget constraint that present-biased workers will retire earlier than planned. He therefore internalizes the negative externality between hyperbolic and exponential, and applies a corrective Pigouvian tax, in the form of a higher Bismarckian factor. By increasing α, the planner postpones workers' retirement.

8.5 Simultaneous Voting

The simultaneous equilibrium à la Shepsle is determined by aggregating the two reaction functions, $\tau_{TI}^{mv}(\alpha)$ and $\alpha_{TI}^{mv}(\alpha)$. The point(s) in which they intercepts, if any, is a candidate for being the equilibrium outcome of the simultaneous game, (α^s, τ^s). Depending on the values of the parameters and the income distribution, two equilibria are possible (see Figure 10): if the income distribution is relatively skewed to the right, the resulting $\alpha^{mv}(\tau)$ is high. We refer to this equilibrium as a Bismarckian social security system. If the income distributions skewed to the left, we have a Beveridgean pension system.

Our objective is to show how time inconsistency modifies the voting equilibrium and improves the realism of our political predictions: to do that, we compare two situations: $(\alpha_{TI}^s, \tau_{TI}^s)$, the simultaneous equilibrium that arises in our hyperbolic model and $(\alpha_{TC}^s, \tau_{TC}^s)$, the equilibrium of an ideal, time consistent, economy. Therefore, in each picture, we depict four curves: the voting functions resulting from our hyperbolic model, $\tau_{TI}^{mv}(\alpha)$ and $\alpha_{TI}^{mv}(\tau)$, and the voting functions of the ideal time-consistent economy, $\tau_{TC}^{mv}(\alpha)$ and $\alpha_{TC}^{mv}(\tau)$. The intersections of these curves will give us two equilibrium points $(\alpha_{TI}^s, \tau_{TI}^s)$ and $(\alpha_{TC}^s, \tau_{TC}^s)$. From Proposition 6 we know that, when $\alpha \leq \frac{1}{2\delta}$, the overconfidence and the commitment effects go in opposite directions. In Figure 10, we have depicted the most interesting case, $\tau_{TI}^{mv}(\alpha) < \tau_{TC}^{mv}(\alpha)$. If a Bismarckian system emerges in equilibrium, by comparing $(\alpha_{TI}^s, \tau_{TI}^s)$ and $(\alpha_{TC}^s, \tau_{TC}^s)$, we can see that time inconsistency reduces the equilibrium value of α and increases the generosity of the social security with respect to the time consistent economy: hyperbolic agents are decisive in the political process and they can, at the same time, decrease τ and soften the link between pension benefit and working history, because it represents a way to raise old-age consumption. On the other hand, whenever a Beveridgean system emerges, redistribution is higher in the hyperbolic economy than in the exponential one.

Our theoretical results match the stylized facts presented in the introduction. First, we provide a political justification for the negative relationship between generosity of the social security system and degree of redistribution, as shown by Conde-Ruiz and Profeta (2005): in our equilibrium, Bismarckian systems are bigger than Beveridgean ones. Second, we have provided a possible explanation for the cross-country differences in the degree of redistribution observed in reality. In our model, four type of pension systems may emerge: Bismarckian with high level of redistributions (Belgium, Austria, Germany), that can

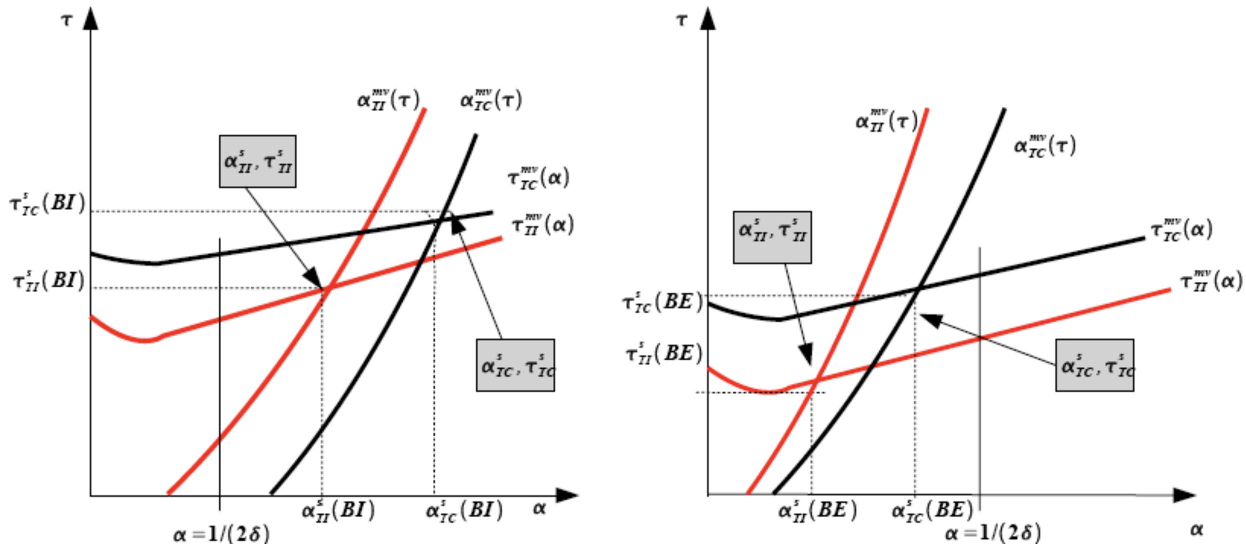

Figure 10: Simultaneous equilibrium in a Bismarckian (left) and Beveridgean system (right).

be classified as “time inconsistent”, and the winning coalition of hyperbolic agents decreases the level of α with respect to “time consistent” countries (Italy, Greece). The same happens with Beveridgean systems: from one side, we have hyperbolic countries (UK, New Zealand, Canada and Denmark) with a very high level of redistribution, and, on the other side, Beveridgean, time consistent, pension system with higher α (U.S., Japan, Switzerland).

9 Concluding Remarks

This paper has studied a model of social security with endogenous retirement age and hyperbolic discounting in individuals' preferences.

Our model provides a justification for the observed growth of voluntary early retirement and the drop in post-retirement consumption due to inadequate saving. We show that time inconsistent agents weights too much the costs associated to postponed retirement (foregone leisure) and too less the benefits (the increase in pension benefits). Early retirement and overconsumption when young are optimal from the point the view of an hyperbolic individual. A time consistent and utilitarian planner, on the other hand, to correct this externality, would like to provide a commitment device to hyperbolic workers, in order to force them to retire later. The commitment device takes the form of a high Bismarckian factor, that guarantees a tighter link between working history and benefits. This instrument plays the role of a Pigouvian tax: the planner internalize the effect of anticipated retirement on the social security budget constraint.

Our political model has shed light on three stylized facts, not yet addressed by the literature. First, as Liberman (2001) points out, redistribution in most pension system seems not to be related to lifetime income but to other factors: for instance, it goes from workers with longer careers to early retirees. Second, the classical distinction between Bismarckian and Beveridgean pension system, where the former are less redistributive than the latter, appears to be misleading, since we observe in reality very redistributive Bismarckian: other factors, besides income distribution, determine the ability of the system to redistribute income. Third, it exists a negative relationship between the size and degree of redistribution in most OECD pension systems. Bigger system (in terms of share of GDP devoted to pension transfer) are also the less redistributive.

We show that the winning coalition determining the size and the degree of redistribution of the PAYG system always include hyperbolic individuals. More precisely, time inconsistent individuals prefer to decrease the size of the system, since a lower payroll tax act as a commitment device that increases both retirement age and savings.

Morover, hyperbolic individuals prefer a more redistributive system compared to exponential. Besides the intragenerational and intergenerational redistribution, our model adds a third form of redistribution: from far-sighted to hyperbolic individuals. It follows that the equilibrium Bismarckian factor is always lower than the one chosen by an utilitarian social planner.

Therefore, whenever hyperbolic individuals have enough political power, the resulting pension system is small, *i.e.* low payroll tax but high redistribution.

The policy implications of our model are immediate: tightening the link between the length of the working career and benefits received, so that workers autonomously decide to retire later, can be an ineffective way to reduce government spending and the problem of early retirement. In this view, a "paternalistic" intervention, in the form of an increase of the minimal retirement age, appear to be more appropriate instrument to solve the pension crisis experienced by most European countries.

The next step of our research is to test empirically these predictions, and in particular to analyze whether the level of time inconsistency effectively differs among countries. If this is the case, than it would be relatively easier to compare if the presence of a high number of voluntary early retirees who regret the lack of accumulated saving leads to a particular pension system whose characteristics are in line with our predictions.

References

[1] Ahituv, A., and J. Zeira (2000), "Technical Progress and Early Retirement", *CEPR Discussion Paper* No. 2614.

[2] Benjamin, D.J. (2003), "Does 401(k) eligibility increase saving? Evidence from propensity score subclassification", *Journal of Public Economics*, Vol. 87, 1259-1290.

[3] Bernheim, B.D. (1999) "Taxation and Savings", in *Handbook of Public Economics*, Vol.3, Chapter 18, edited by A.J. Auerbach and M. Feldstein, Elsevier Science Publishers B.V. (North Holland).

[4] Bernheim, B.D. (1995) "Do Households Appreciate Their Financial Vulnerabilities? An Analysis of Actions, Perceptions, and Public Policy", in *Tax Policy and Economic Growth*, Washington, DC: American Council for Capital Formation, 1-30.

[5] Bernheim B.D., J. Skinner and S. Weinberg (1988), "What accounts for the variation in retirement wealth among U.S. households?", *NBER Working Papers* N. 6227, National Bureau of Economic Research, Inc.

[6] Blondal, S., and S. Scarpetta (1998), "The Retirement Decision in OECD Countries", *World Bank Working Paper.*

[7] Casamatta, G., H. Cremer and P. Pestieau (2005), "Voting on Pensions with Endogenous Retirement Age", *International Tax and Public Finance*, Vol. 12, 7-28.

[8] Chan and A.H Stevens (2004), "Do Changes in Pension Incentives Affect Retirement? A Longitudinal Study of Subjective Retirement Expectations ", *Journal of Public Economics*, Vol. 88, 1307-1333.

[9] Conde-Ruiz, I., and P. Profeta (2007), "The Redistributive Design of Social Security Systems", *The Economic Journal*, Vol. 117, 686-712.

[10] Conde-Ruiz, I., and V. Galasso (2004), "The Macroeconomics of Early Retirement", *Journal of Public Economics*, Vol. 88, 1849-1869.

[11] Conde-Ruiz, I., and V. Galasso (2003), "Early Retirement", *FEDEA Discussion Paper.*

[12] Cooley, T. F., and J. Soares (1999), "A Positive Theory of Social Security Based on Reputation", *Journal of Political Economy*, Vol. 107, 135-160.

[13] Coronado, J., D. Fullerton and T. Glass (2000), "The Progressivity of Social Security" *NBER Working Papers* N. 7520, National Bureau of Economic Research, Inc.

[14] Cremer, H., P. De Donder, D. Maldonado and P. Pestieu (2008), "Designing an Optimal Linear Pension Scheme with forced saving and wage heterogeneity", *forthcoming, International Tax and Public Finance.*

[15] Cremer, H., P. De Donder, D. Maldonado and P. Pestieau (2007), "Voting over type and generosity of a pension system when some individuals are myopic", *Journal of Public Economics*, Vol. 91, 2041-2061.

[16] Cremer, H., J.M. Lozachmeur and P. Pestieau (2004), "Social Security, Retirement Age and Optimal Income Taxation ", *Journal of Public Economics*, Vol. 88, 2259-2281.

[17] Della Vigna, S., and D. Paserman (2005), "Job Search and Impatience", *Journal of Labor Economics*, Vol. 23(3), 437-466.

[18] Department of Work and Pensions (2005), *DWP Departmental Report 2005,* London.

[19] Diamond, P., and B. Köszegi (2003), “Quasi-Hyperbolic Discounting and Retirement”, *Journal of Public Economics*, Vol. 87, 1839-1872.

[20] Disney, R., and P. Johnson (2001), *Pension Systems and Retirement Incomes across OECD Countries*, Edward Elgar editor.

[21] Disney, R., V. Galasso and P. Profeta (2007), “A Further Comment on Ageing and the Welfare State”, *European Journal of Political Economy*, Vol. 23, pp. 576-577.

[22] Feldstein, M. (1985), “The Optimal Level of Social Security Benefits”, *The Quarterly Journal of Economics*, Vol. C., 573-597.

[23] Frederick, S., G. Loewenstein and T. O’Donoghue (2002), “Time Discounting and Time Preferences: A Critical Review”, *mimeo.*

[24] Galasso, V. *The Political Future of Social Security in Aging Societies*, The MIT Press, Cambridge, MA.

[25] Galasso, V., and P. Profeta, (2004), “Lessons for an Aging Society: the Political Sustainability of Social Security Systems ”, Economic Policy, 63-115.

[26] Galasso, V., and P. Profeta, (2002), “The Political Economy of Social Security: A Survey”, *European Journal of Political Economy*, Vol. 18(1), 1-29.

[27] Gruber, J., and B. Köszegi (2004), “Tax Incidence When Individuals are Time Inconsistent: the Case of Cigarette Excise Taxes”, *Journal of Public Economics*, 88, pp. 1959-1987.

[28] Kahneman, D., andA. Tversky (1979), “Prospect Theory: An Analysis of Decision Under Risk”, *Econometrica*, Vol. 47, 363-391.

[29] Hammond, P. (1975), “Charity: Altruism or Cooperative Egoism”, in *Altruism, Morality and Economic Theory* (E.S. Phelps, ed.), Russel Sage Foundation, New York.

[30] Harris, C., and D. Laibson (2001), “Dynamic Choices of Hyperbolic Consumers ”, *Econometrica*, Vol. 69(4), 935-957.

[31] Hausman, J., and L. Paquette (1987), “Involuntary Early Retirement and Consumption” in Gary Burtless, editor, *Work, Health and Income among the Elderly*, 151-175. Brookings Institution, Washington, D.C.

[32] Imrohoroglu, A., S. Imrohoroglu and D.H. Joines (2004), “Time Inconsistent Preferences and Social Security”, *The Quarterly Journal of Economics*, Vol. 18(2), 745-784.

[33] Laibson, D. (1996), “Hyperbolic Discount Functions, Undersaving and the Saving Policy”, *NBER Working Papers* N. 5635, National Bureau of Economic Research, Inc.

[34] Laibson, D., A. Repetto and J. Tobacman (1998), "Self-Control and Saving for Retirement", *Brookings Paper on Economic Activity.*

[35] Liebman, J. (2001), "Redistribution in the Current U.S. Social Security System", *NBER Working Papers* N. 8625, National Bureau of Economic Research, Inc.

[36] Loewenstein, G. (1991), "Do Workers Prefer Increasing Wage Profiles?", *Journal of Labor Economics*, Vol. 9, 67-84.

[37] Loewenstein, G. (1992), "Anomalies in Intertemporal Choice: Evidence and an Interpretation", *The Quarterly Journal of Economics*, Vol. 107, 573-597.

[38] Mulligan, C., and X. Sala-i-Martin (1999), "Social Security in Theory and Practice (I): Facts and Political Theories", *NBER Working Papers* N. 7119, National Bureau of Economic Research, Inc.

[39] OECD (2005), Pension at a Glance: Public Policies across OECD countries, Paris.

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

Appendix

A Pension Programs in OECD Countries

Country	Type	Progressivity Index	Pension Expenditure (% GDP)
Australia	BE	74,8	4,7
Austria	BI	20,7	10,7
Belgium	BI	64,8	8,7
Canada	BE	86,5	4,8
Denmark	BE	91,7	8,3
France	BI	46,4	10,6
Germany	BI	22,9	11,7
Greece	BI	4,3	12,7
Ireland	BE	100	2,7
Italy	BI	4	11,3
Japan	BE	47,8	7,3
Netherlands	BE	5,7	6,4
New Zealand	BE	100	4,7
Spain	BI	13	8,3
Switzerland	BE	44,1	11,8
UK	BE	69,6	8,1
US	BE	40,6	5,3

Table V: Pension Programs in selected OECD countries (Disney 2001 and OECD 2005).

B Proofs

B.1 Proof of Proposition 1

Part *(ii)* of Lemma 1 show that $z_e > z_s$. To show the Pareto-improvement, we use the following notation. Let $U^o(z_j)$ the utility of a "pre-retirement" old, given its retirement age z_j. We want to show that:

$$f(\beta) = U(z_e) - U(z_s(\beta)) > 0$$

Note that $U(z_s(\beta))$ depends on β in two ways: β is the discount factor, so changes in β affect future consumption (when retired). Moreover, β enters into the expression that determines z_s itself. On the other hand, note that β only influences $U(z_e)$ through the first mechanism, as z_e does not depend on β. In the proof, we characterize the value of $f(.)$ in a neighborhood of $\beta = 1$. First, note that $f(1) = 0$, since $z_e = z_s$. In order to evaluate the marginal variation of $f(.)$ around $\beta = 1$, we consider $f"(1)$ and $f'(1)$; we have:

$$f'(\beta) = \frac{\partial U(z_e)}{\partial \beta} - \frac{\partial U(z_s)}{\partial \beta} - \frac{\partial U(z_s)}{\partial z_s}\frac{dz_s}{d\beta}$$

Note that $f'(1) = 0$, as $\left.\frac{\partial U(z_e)}{\partial \beta}\right|_{\beta=1} = \left.\frac{\partial U(z_s)}{\partial \beta}\right|_{\beta=1}$ and $\left.\frac{\partial U(z_s)}{\partial z_s}\right|_{\beta=1} = \left.\frac{\partial U(z_e)}{\partial z_e}\right|_{\beta=1} = 0$.
Finally, it is possible to show that $f"(1) > 0$. Given that $f(1) = 0$, $f'(1) = 0$ and $f"(1) > 0$, there exists an interval $\left(\bar{\beta}, 1\right)$ such that $f(\beta) > 0$ for $\beta \in \left(\bar{\beta}, 1\right)$. This shows that a sophisticated worker is made better off by increasing its retirement age to z_e. Pareto dominance follows from the fact that all selves (pre and post-retirement) are made better off in two ways: first, they prefer z_e that influences pre and post retirement consumption and, second, they prefer gaining more pension benefits as implied by z_e.

B.2 Proof of Proposition 2

This proof parallels the proof of Proposition 1 and is omitted.

B.3 Proof of Lemma 2

From (17), the first derivative is:

$$\frac{\partial P(\widehat{z_j})(.)}{\partial \tau} = (1+n)(\alpha\omega + (1-\alpha)\bar{\omega}] + \theta\omega\alpha\left(z + \tau\frac{\partial \widehat{z_j}}{\partial \tau}\right)$$

A necessary and sufficient condition for this expression to be positive is the following:

$$2\theta^2\alpha\omega^2\tau(1-\hat{\beta}_j\delta\alpha) \leq \theta\alpha\omega^2 + (1+n)(\gamma - \hat{\beta}_j\psi\delta)(\alpha\omega + (1-\alpha)\bar{\omega})$$

and the definition of $\hat{\tau}$ follows immediately. Differentiating again $P(\widehat{z_j})$ with respect to τ gives us:

$$\frac{\partial^2 P(\widehat{z_j})(.)}{\partial \tau^2} = -\frac{2(1-\hat{\beta}_j\alpha\delta)}{\gamma - \hat{\beta}_j\psi\delta} < 0 \qquad \text{(B-27)}$$

B.4 Proof of Proposition 3

The FOCs for the maximization problem for individuals of type j are:

$$-\omega u'\left(c_y^j\right)+\hat{\beta}_j\delta u'\left(c_o^j\right)\left[-\omega\theta\widehat{z}_j+\delta\frac{\partial P(.)}{\partial\tau}\right]+\lambda_\tau=0 \tag{B-28}$$

$$-u'\left(c_y^j\right)+\hat{\beta}_j\delta u'\left(c_o^j\right)+\lambda_s=0 \tag{B-29}$$

where λ_s and λ_τ are the Lagrange multipliers associated to the non-negativity constraints on s and τ. Depending on which constraints bind, we have different cases. Before doing that, we check whether the objective function $V^y(\tau_j^y,\alpha;\hat{\beta}_j,\omega)$ is concave in $(\tau_j^y,\widehat{s}_j)$, so that the solutions represent indeed global optima. The Hessian matrix is:

$$D^2V^y(\tau_j^y,\widehat{s}_j)=\begin{bmatrix}\frac{\partial^2V^y(\tau_j^y,\widehat{s}_j)}{\partial(\tau_j^y)^2} & \frac{\partial^2V^y(\tau_j^y,\widehat{s}_j)}{\partial\tau_j^y\partial\widehat{s}_j}\\ \frac{\partial^2V^y(\tau_j^y,\widehat{s}_j)}{\partial\tau_j^y\partial\widehat{s}_j} & \frac{\partial^2V^y(\tau_j^y,\widehat{s}_j)}{\partial(\widehat{s}_j)^2}\end{bmatrix}$$

where:

$$\begin{aligned}\frac{\partial^2V^y(\tau_j^y,\widehat{s}_j)}{\partial(\tau_j^y)^2} &= \omega^2u''\left(c_y^j\right)+\hat{\beta}_j\delta u''\left(c_o^j\right)\left[-\omega\theta\widehat{z}_j+\delta\frac{\partial P(\widehat{z}_j)}{\partial\tau_j^y}\right]^2+\\ &\quad \hat{\beta}_j\delta u'\left(c_o^j\right)\left[-\omega\theta\widehat{z}_j+\delta\frac{\partial^2P(\widehat{z}_j)}{\partial\left(\tau_j^y\right)^2}\right]<0\end{aligned} \tag{B-30}$$

$$\frac{\partial^2V^y(\tau_j^y,\widehat{s}_j)}{\partial(\widehat{s}_j)^2} = u''\left(c_y^j\right)+\hat{\beta}_j\delta u''\left(c_o^j\right)<0 \tag{B-31}$$

$$\frac{\partial^2V^y(\tau_j^y,\widehat{s}_j)}{\partial\tau_j^y\partial\widehat{s}_j} = u''\left(c_y^j\right)\omega+\hat{\beta}_j\delta u''\left(c_o^j\right)\left[-\omega\theta\widehat{z}_j+\delta\frac{\partial P(\widehat{z}_j)}{\partial\tau_j^y}\right] \tag{B-32}$$

The determinant of the Hessian is:

$$\begin{aligned}\det(D^2V^y(\tau_j^y,\widehat{s}_j)) &= \hat{\beta}_j\delta u''\left(c_y^j\right)u''\left(c_o^j\right)\left[\omega-\left(-\omega\theta\widehat{z}_j+\delta\frac{\partial P(\widehat{z}_j)}{\partial\tau_j^y}\right)\right]^2+\\ &\quad \hat{\beta}_j\delta u'\left(c_o^j\right)\left(u''\left(c_y^j\right)+\hat{\beta}_j\delta u''\left(c_o^j\right)\right)\left[-\omega\theta\widehat{z}_j+\delta\frac{\partial^2P(\widehat{z}_j)}{\partial\left(\tau_j^y\right)^2}\right]\end{aligned} \tag{B-33}$$

The sign of the determinant depends on that of the last term on the RHS of (B-33): by replacing (B-27) and (7) into (33), we find that a sufficient condition for (B-33) to be positive is $\hat{\beta}\leq\frac{1}{\alpha\delta}$, which is always satisfied. Since (B-30) and (B-31) are always negative, the Hessian matrix is negative definite and the objective function is concave in $(\tau_j^y,\widehat{s}_j)$

Case 1: $\lambda_s=\lambda_\tau=0$

None of the constraints is binding: replacing (B-29) into (B-28), we get:

$$\tau_y^+(\omega,\alpha;\hat{\beta}_j)=\frac{(\gamma-\hat{\beta}_j\delta\psi)[\omega-(1+n)(\alpha\omega+(1-\alpha)\bar{\omega})]+\omega\theta(1-\alpha)(\omega-\hat{\beta}_j\delta\psi)}{\theta^2\omega^2(1-\hat{\beta}_j\delta\alpha)(1-2\delta\alpha)} \tag{B-34}$$

The sign of this expression depends on the sign of the term $(1-2\delta\alpha)$. Since both cases are plausible, we discuss them separately. Suppose, first, that is negative ($\alpha\delta > 1/2$): a necessary and sufficient condition for this tax rate to be positive (part *(i)* of the proposition) is:

$$(\gamma - \hat{\beta}_j\delta\psi)[\omega - (1+n)(\alpha\omega + (1-\alpha)\bar{\omega}] + \omega\theta(1-\alpha)(\omega - \hat{\beta}_j\delta\psi) < 0$$

After some computations, we obtain the following second degree polynomial:

$$\omega^2 \underbrace{\left(\frac{\theta(1-\delta\alpha)}{\gamma - \hat{\beta}_j\delta\psi}\right)}_{a>0} + \ \omega \underbrace{\left(\delta(1+n)\alpha - \frac{\gamma - \hat{\beta}_j\delta\psi(1+\theta(1-\delta\alpha))}{\gamma - \hat{\beta}_j\delta\psi}\right)}_{b<0} - \underbrace{\delta(1-\alpha)(1+n)\bar{\omega}}_{c>0} < 0 \qquad \text{(B-35)}$$

The solution has the form $\omega_1 \leq \omega \leq \tilde{\omega}$, where the threshold $\tilde{\omega}$ is given by $\tilde{\omega} = \frac{-b+\sqrt{b^2-4ac}}{2a}$; given that only the second root is positive, preferred tax rates are positive for $\omega_- \leq \omega \leq \tilde{\omega}$. To prove last part of *(i)*, differentiate $\tau_y^+(\omega, \alpha; \hat{\beta}_j)$ with respect to ω :

$$\frac{\partial \tau_y^+(\omega, \alpha; \hat{\beta}_j)}{\partial \omega} = \frac{(\gamma - \hat{\beta}_j\delta\psi)(1 - \alpha\delta(1+n)) + \theta(1-\delta\alpha)(2\omega - \hat{\beta}_j\delta\psi)}{\theta^2\omega^2(1-\hat{\beta}_j\delta\alpha)(1-2\delta\alpha)} - \frac{2\,[\text{numerator of (B-34)}]}{\theta^2\omega^3(1-\hat{\beta}_j\delta\alpha)(1-2\delta\alpha)}$$

This expression is equivalent to:

$$\frac{\partial \tau_y^+(\omega, \alpha; \hat{\beta}_j)}{\partial \omega} = \frac{\delta(\gamma - \hat{\beta}_j\delta\psi)}{(1-\hat{\beta}_j\delta\alpha)(1-2\delta\alpha)}\left[\alpha(1+n) - 1 + 2(1+n)(1-\alpha)\frac{\bar{\omega}}{\omega} + \frac{\hat{\beta}_j\psi\theta(1-\delta\alpha)}{\gamma - \hat{\beta}_j\delta\psi}\right] \qquad \text{(B-36)}$$

The first term is negative, since $2\alpha\delta > 1$, and a sufficient condition for the second term to be positive is:

$$\delta\alpha(1+n) + \frac{\hat{\beta}_j\delta\psi\theta(1-\delta\alpha)}{\gamma - \hat{\beta}_j\delta\psi} > 1 \qquad \text{(B-37)}$$

which is always verified since we impose $\gamma < \alpha\psi$. Thus, $\frac{\partial \tau_y^+(\omega,\alpha;\hat{\beta}_j)}{\partial \omega} < 0$, for $\omega_- < \omega < \tilde{\omega}$.

The impact of hyperbolic discounting on preferred tax rates (part *ii*) is given by:

$$\frac{\partial \tau_y^+(\omega, \alpha; \hat{\beta}_j)}{\partial \hat{\beta}_j} = \frac{(-\delta\psi)(\omega - \delta(1+n)[\alpha\omega + (1-\alpha)\bar{\omega}]) - \delta\psi\omega\theta(1-\alpha\delta)}{\theta^2\omega^2(1-\hat{\beta}_j\delta\alpha)(1-2\delta\alpha)} + \frac{[\text{numerator of (B-34)}]\,(1-2\alpha\delta)\omega^2\theta^2}{\left[\theta^2\omega^2(1-\hat{\beta}_j\delta\alpha)(1-2\delta\alpha)\right]^2}$$

After some rearrangements, we get:

$$\frac{\partial \tau_y^+(\omega, \alpha; \hat{\beta}_j)}{\partial \hat{\beta}_j} = \underbrace{\frac{(\gamma\alpha - \psi)}{(1-2\alpha\delta)(1-\hat{\beta}_j\delta\alpha)}}_{>0}\left[1 - \delta(1+n)\left(\alpha + \frac{\bar{\omega}}{\omega}\right) + \theta(1-\alpha\delta)(\alpha\omega - \psi)\right] \qquad \text{(B-38)}$$

The sign of (B-38) depends on the sign of the term into square brackets, that is given by the following second degree inequality:

$$\theta(1-\alpha\delta)\alpha\omega^2 - \underbrace{(1 - \delta(1+n) - \theta(1-\alpha\delta))}_{b}\omega \underbrace{-\delta(1+n)\bar{\omega}}_{c} \gtrless 0$$

This expression defines a threshold $\omega_b = \frac{-b \pm \sqrt{b^2-4ac}}{2a}$ (the second root is negative, since $c < 0$). It can be easily shown that $\omega_b < \tilde{\omega}$, and thus $\frac{\partial \tau^j}{\partial \hat{\beta}_j} \geq 0$ for $\breve{\omega} \leq \omega < \tilde{\omega}$ and $\frac{\partial \tau_y^+(\omega,\alpha;\hat{\beta}_j)}{\partial \hat{\beta}_j} < 0$ for $\omega^- \leq \omega < \breve{\omega}$.

Part *(iii)* of the proposition is obtained by differentiating $\tilde{\omega}$ with respect to $\hat{\beta}$.

$$\begin{aligned} \frac{\partial \tilde{\omega}}{\partial \hat{\beta}_j} = & -(1+n)\delta\alpha + (\theta(1-\alpha\delta)+1) + \frac{1}{k}(1+n)\alpha\left\{2\left(\theta(1-\delta\alpha)+1\right) - \gamma\left(\theta(1-\delta\alpha)+2\right)\right\} + \\ & \frac{1}{k}\left\{\left[\left(\theta(1-\delta\alpha)+1\right)\hat{\beta}_j\delta\psi - \gamma\right]\left(\theta(1-\delta\alpha)+1\right) - (\delta(1+n)\alpha)^2(\gamma - \hat{\beta}_j\delta\psi) - 2\delta(1-\alpha)\bar{\omega}(1+n)\theta(1-\delta\alpha)\right\} \end{aligned} \tag{B-39}$$

where $k = \sqrt{\Delta \ of \ expression \ (B-35)} > 0$. After some rearrangements, we obtain that the sign of $\frac{\partial \tilde{\omega}}{\partial \hat{\beta}_j}$ is equivalent to the sign of:

$$\begin{aligned} & \hat{\beta}_j\delta\psi\left[\delta(1+n)\alpha + (\theta(1-\delta\alpha)+1)\right]^2 - \gamma\left[(1+n)\delta\alpha+1\right]^2 - \\ & \gamma[\theta(1-\delta\alpha)\left[(1+n)\alpha\delta+1)\right]] - 2\delta(1-\alpha)\bar{\omega}(1+n)\theta(1-\alpha\delta) \end{aligned}$$

The last two terms of this expression are negative: therefore, a sufficient condition for the whole expression to be negative is:

$$\hat{\beta}_j\delta\psi\left[\delta(1+n)\alpha + (\theta(1-\delta\alpha)+1)\right]^2 - \gamma\left[(1+n)\delta\alpha+1\right]^2 \leq 0$$

or:

$$\begin{aligned} & \left[\sqrt{\hat{\beta}_j\delta\psi}((1+n)\delta\alpha + 1 + \theta(1-\delta\alpha)) - \sqrt{\gamma}\left[(1+n)\delta\alpha+1\right]\right] * \\ & \left[\sqrt{\hat{\beta}_j\delta\psi}((1+n)\delta\alpha + (\theta(1-\delta\alpha)+1)) + \sqrt{\gamma}\left[(1+n)\delta\alpha+1\right]\right] \leq 0 \end{aligned}$$

Notice that the first term is always negative, since $\sqrt{\hat{\beta}_j\delta\psi} < \sqrt{\gamma}$ and $\theta(1-\delta\alpha) < 1$. Therefore, $\frac{\partial \tilde{\omega}}{\partial \hat{\beta}} < 0$.

Next step is to check how saving vary with productivity (part *iv*); by (B-29), we have:

$$\begin{aligned} \frac{\partial \widehat{s}_j(\omega)}{\partial \omega} = & -\frac{\left(-(1-\tau) + \omega\frac{\partial \tau}{\partial \omega}\right)u''(c_y^j)}{u''(c_y^j) + \hat{\beta}_j\delta u''(c_o^j)} + \\ & -\frac{\hat{\beta}_j\delta u''(c_o^j)\left[\widehat{z}_j - \tau(\widehat{z}_j\theta(1-\alpha\delta) - \delta(1+n)\alpha) - \frac{\partial \tau}{\partial \omega}\left[\omega\theta z(1-\alpha\delta) - \delta(1+n)\left((1-\alpha)\bar{\omega} + \alpha\omega\right)\right]\right]}{u''(c_y^j) + \hat{\beta}_j\delta u''(c_o^j)} \end{aligned} \tag{B-40}$$

The first term is positive, since $\frac{\partial \tau_y^+(\omega,\alpha;\hat{\beta}_j)}{\partial \omega} < 0$ for $\omega\tilde{\omega}$. For the second term, we can find two sufficient conditions for this expression to be negative. We will see that this conditions are always satisfied, and then we conclude that $\frac{\partial \widehat{s}_j(\omega)}{\partial \omega} > 0$. These two conditions are:

1. $$\widehat{z}_j - \tau_y^+(\omega,\alpha;\hat{\beta}_j)(\widehat{z}_j\theta(1-\alpha\delta) - \delta(1+n)\alpha) < 0$$

2. $$\omega\theta\widehat{z}_j(1-\alpha\delta) + \delta(1+n)\left((1-\alpha)\bar{\omega} + \alpha\omega\right) < 0$$

The first condition can be rewritten as:

$$\tau > \frac{1}{\theta(1-\alpha\delta) - \frac{1}{\widehat{z_j}}\alpha\delta(1+n)} \Longleftrightarrow \alpha\delta > \frac{\theta - 1}{\frac{1}{\widehat{z_j}}(\theta+1+n)}$$

Which is always satisfied, given that $\theta \in [0,1]$.

Replacing the expression for optimal retirement age (6), condition 2 can be rewritten as:

$$\omega^2 \frac{(1-\alpha\delta)\theta(1-\tau\theta(1-\beta\delta\alpha)}{\gamma} - \delta(1+n)\alpha\omega - \delta(1+n)(1-\alpha)\bar{\omega} < 0 \Longleftrightarrow 0 \leq \omega \leq \breve{\omega}$$

It can be easily checked that this threshold $\breve{\omega}$ is always greater than $\tilde{\omega}$. Hence, $\frac{\partial \widehat{s_j}(\omega)}{\partial \omega} > 0, \forall \omega \leq \tilde{\omega}$. Moreover, from (B-29), we can see that the function $\widehat{s_j}(\omega)$ is negative when ω_- is low and positive when ω approaches to $\tilde{\omega}$. Thus, there exists a value $\omega' < \tilde{\omega}$ such that saving are zero for $\omega \leq \omega'$ and positive above. Therefore, all workers with income between ω' and $\hat{\omega}$ have an interior solution for both τ and $\widehat{s_j}$.

Part *(v)* of the proposition comes from noticing that, with $\tau_y^+(\omega, \alpha; \hat{\beta}_j) = 1, \forall j$, (B-28) becomes:

$$-\omega u'(0) + \hat{\beta}_j \delta u'(c_o^j) \left[-\omega\theta\widehat{z_j} + \delta \frac{\partial P(\widehat{z_j})}{\partial \tau} \right] \Bigg|_{\tau=1} < 0 \tag{B-41}$$

By Inada conditions, $\lim_{x \to 0} u'(x) = +\infty$.

In part *(vi)*, we show that preferred tax rates are concave in α. To see that, notice that, for $\omega_- \leq \omega \leq \tilde{\omega}$, preferred tax rates $\tau_y^+(\omega, \alpha; \hat{\beta}_j)$ are between 0 and 1. Therefore, the internal solution with respect to τ guarantees that:

$$\frac{\partial V^y \left(\tau_y^+(\omega, \alpha; \hat{\beta}_j), \alpha, \omega; \hat{\beta}_j \right)}{\partial \tau} = 0 \tag{B-42}$$

$$\frac{\partial^2 V^y \left(\tau_y^+(\omega, \alpha; \hat{\beta}_j), \alpha, \omega; \hat{\beta}_j \right)}{\partial \tau^2} < 0 \tag{B-43}$$

Differentiating (B-42) with respect to α, we obtain:

$$\frac{\partial^2 V^y \left(\tau_y^+(\omega, \alpha; \hat{\beta}_j), \alpha, \omega; \hat{\beta}_j \right)}{\partial \tau^2} \frac{\partial \tau_y^+(\omega, \alpha; \hat{\beta}_j)}{\partial \alpha} + \frac{\partial^2 V^y \left(\tau_y^+(\omega, \alpha; \hat{\beta}_j), \alpha, \omega; \hat{\beta}_j \right)}{\partial \tau \partial \alpha} = 0$$

The sign of $\frac{\partial \tau_y^+(.)}{\partial \alpha}$ is the same as $\frac{\partial V^y \left(\tau_y^+(\omega, \alpha; \hat{\beta}_j), \alpha, \omega; \hat{\beta}_j \right)}{\partial \tau \partial \alpha}$. From (23), we get:

$$\frac{\partial^2 V^y(.)}{\partial \tau \partial \alpha} = u''(c_o^j) \left(-\omega\theta\widehat{z_j} + \delta \frac{\partial P(.)}{\partial \tau} \right) \left(\frac{\partial c_o^j}{\partial \alpha} \right) + u'(c_o^j) \left(-\omega\theta \frac{\partial \widehat{z_j}}{\partial \alpha} + \delta \frac{\partial^2 P(\widehat{z_j})}{\partial \tau \partial \alpha} \right)$$

The first term is zero because of (B-28) (remember that we restrict consumption levels to be strictly positive). The second one is positive too; to see this, let us replace the expression for $\frac{\partial \widehat{z_j}}{\partial \alpha}$ and $\frac{\partial^2 P(\widehat{z_j})}{\partial \tau \partial \alpha}$. After some algebra:

$$\frac{\partial^2 V^y(.)}{\partial \tau \partial \alpha} = u'(c_o^j) \left[\delta(1+n)(\omega - \bar{\omega}) + \frac{\theta\omega\delta}{\gamma - \hat{\beta}_j \delta\psi} \left(\omega - \hat{\beta}_j \delta\psi + \omega\theta\tau B \right) \right]$$

where $B = (4\delta\hat{\beta}_j\alpha - 2 - \hat{\beta}_j)$. For the whole second term to be positive, it suffices that $|B| < 1$, which is always true. Finally, the sign of the entire expression depends crucially on the individual's income: if $\bar{\omega} \leq \omega \leq \tilde{\omega}$, $\frac{\partial \tau_y^+(\omega, \alpha; \hat{\beta}_j)}{\partial \alpha} > 0$. Otherwise, for very poor individuals, $\omega_- \leq \omega << \bar{\omega}$, we have $\frac{\partial \tau_y^+(\omega, \alpha; \hat{\beta}_j)}{\partial \alpha} < 0$.

Case 2: $\lambda_s > 0,\ \lambda_\tau = 0$

Individuals with $\omega \leq \omega'$ choose a positive tax rate $\tau_y^+(\omega, \alpha; \hat{\beta}_j)$ (given by B-28) and no saving.

Case 3: $\lambda_s = 0,\ \lambda_\tau > 0$

Individuals with productivity above $\tilde{\omega}$ rely exclusively on private saving, that are defined by (B-29).

To determine the majority voting solution, we have to know how preferred tax rates $\tau_y^+(\omega, \alpha; \hat{\beta}_j)$ vary with income for individuals below ω' (those with $\widehat{s_j} = 0$). From (B-28):

$$\frac{\partial \tau_y^+(\omega, \alpha; \hat{\beta}_j)}{\partial \omega} = -\frac{-u'\left(c_y^j\right) - \omega(1-\tau_y^+)u''\left(c_y^j\right) + \beta\delta u'\left(c_o^j\right)\left[-\theta\widehat{z_j} - \omega\theta\frac{\partial \widehat{z_j}}{\partial \omega} + \frac{\partial^2 P(\widehat{z_j})}{\partial \tau_y^+ \partial \omega}\right]}{D_{\tau_y^+}} + \tag{B-44}$$

$$+\frac{\beta\delta u''\left(c_o^j\right)\left[-\omega\theta\widehat{z_j} + \frac{\partial P(\widehat{z_j})}{\partial \tau_y^+}\right]\left[\widehat{z_j}(1-\theta\tau_y^+) + \delta\alpha\tau_y^+(1+n) + \delta\alpha\tau_y^+\widehat{z_j}\right]}{D_{\tau_y^+}} \tag{B-45}$$

where $D_{\tau_y^+}$ is the second order derivative of $V^y(\tau_y^+, \widehat{s_j})$ with respect to τ_y^+, that is negative:

$$\begin{aligned}\frac{\partial^2 V^y(\tau_y^+, \widehat{s_j})}{\partial\left(\tau_y^+\right)^2} = {} & \omega^2 u''\left(c_y^j\right) + \hat{\beta}_j\delta u''\left(c_o^j\right)\left[-\omega\theta\widehat{z_j} + \delta\frac{\partial P(\widehat{z_j})}{\partial \tau_y^+}\right]^2 + \hat{\beta}_j\delta u'\left(c_o^j\right)\left[-\omega\theta\frac{\partial \widehat{z_j}}{\partial \tau_y^+} + \delta\frac{\partial^2 P(\widehat{z_j})}{\partial\left(\tau_y^+\right)^2}\right] + \\ & + \frac{\partial \widehat{s_j}}{\partial \tau_y^+}\left[\omega u''\left(c_y^j\right) + \hat{\beta}_j\delta u''\left(c_o^j\right)\left[-\omega\theta\widehat{z_j} + \delta\frac{\partial P(\widehat{z_j})}{\partial \tau_y^+}\right]\right]\end{aligned}$$

Replacing the expression for $\frac{\partial \widehat{s_j}}{\partial \tau_y^+} = -\frac{\omega u''\left(c_y^j\right) + \hat{\beta}_j\delta u''\left(c_o^j\right)\left[-\omega\theta\widehat{z_j} + \delta\frac{\partial P(\widehat{z_j})}{\partial \tau_y^+}\right]}{u''\left(c_y^j\right) + \hat{\beta}_j\delta u''\left(c_o^j\right)}$, and after some rearrangements, we obtain:

$$\frac{\partial^2 V^j(\tau^{y,j}, \widehat{s_j})}{\partial\left(\tau_y^+\right)^2} = \frac{\hat{\beta}_j\delta u''\left(c_y^j\right)u''\left(c_o^j\right)\left[\omega - \left(-\omega\theta\widehat{z_j} + \delta\frac{\partial P(\widehat{z_j})}{\partial \tau_y^+}\right)\right]^2}{u''\left(c_y^j\right) + \hat{\beta}_j\delta u''\left(c_o^j\right)} + \hat{\beta}_j\delta u'\left(c_o^j\right)\left[-\frac{\partial \widehat{z_j}}{\partial \tau_y^+}\theta\omega + \delta\frac{\partial^2 P(\widehat{z_j})}{\partial\left(\tau_y^+\right)^2}\right] \tag{B-46}$$

where the second term on the RHS is negative, as one can check by replacing expressions (7) and (B-27).

Finally, after some rearrangements, expression (B-44) can be rewritten as:

$$\begin{aligned}\frac{\partial \tau_y^+(\omega, \alpha; \hat{\beta}_j)}{\partial \omega} = {} & -\frac{-u'\left(c_y^j\right)(1-\epsilon) + \hat{\beta}_j\delta u'\left(c_o^j\right)\left[-\theta\widehat{z_j} - \omega\theta\frac{\partial \widehat{z_j}}{\partial \omega} + \frac{\partial^2 P(\widehat{z_j})}{\partial \tau_y^+ \partial \omega}\right]}{D_{\tau_y^+}} + \\ & -\frac{\hat{\beta}_j\delta u''\left(c_o^j\right)\left[-\omega\theta\widehat{z_j} + \frac{\partial P(\widehat{z_j})}{\partial \tau_y^+}\right]\left[\widehat{z_j}(1-\theta\tau_y^+) + \delta\alpha\tau_y^+(1+n) + \delta\alpha\tau_y^+\widehat{z_j}\right]}{D_{\tau_y^+}}\end{aligned}$$

where ϵ is the coefficient of relative risk aversion, that is assumed to be lower than 1 (as in Casamatta, Cremer and Pestieau 2005). Notice that the first term in square brackets is negative because $\frac{\partial^2 P(\widehat{z_j})}{\partial \tau_y^+ \partial \omega} \leq 0$, and that the second term in square bracket is positive for $\omega \leq \omega'$. Thus, for individuals with income $\omega \leq \omega'$, who do not save privately, tax rates are decreasing with income.

B.5 Proof of Proposition 4

Case 1: $\lambda_s = \lambda_\tau = 0$

We consider now the case $\alpha\delta < 1/2$: $\tau_y^+(\omega, \alpha; \hat{\beta}_j)$ is positive only for $\tilde{\omega} \leq \omega \leq \omega_+$ (part i). For part *(ii)*, let us consider the expression (B-36). Now, the second term of the right hand side is positive and the first is positive for $\omega \geq \hat{\omega}$. Since $\hat{\omega} \geq \tilde{\omega}$, $\frac{\partial \tau_y^+(\omega,\alpha;\hat{\beta}_j)}{\partial \omega} > 0$.

For *(iii)*, we have that :

$$\frac{\partial \tau_y^+(\omega, \alpha; \hat{\beta}_j)}{\partial \hat{\beta}_j} = \underbrace{\frac{(\gamma\alpha - \psi)}{(1 - 2\alpha\delta)(1 - \hat{\beta}_j \delta\alpha)}}_{<0} \left[1 - \delta(1+n)\left(\alpha + \frac{\bar{\omega}}{\omega}\right) + \theta(1 - \alpha\delta)(\alpha\omega - \psi)\right] \tag{B-47}$$

The sign of the whole derivative depends on the sign of the term in square brackets. We have already shown (see Proposition 3) that it is positive for income levels above $\omega_b (< \tilde{\omega})$. It follows that for any ω such that $\tau_y^+(\omega, \alpha; \hat{\beta}_j) > 0$, we have $\frac{\partial \tau_y^+(\omega,\alpha;\hat{\beta}_j)}{\partial \hat{\beta}_j} < 0$.

Finally, we have to check how savings vary for productivities $\omega \geq \tilde{\omega}$:

$$\begin{aligned} \frac{\partial \widehat{s_j}(\omega)}{\partial \omega} &= -\frac{\left(-(1-\tau_y^+) + \omega \frac{\partial \widehat{z_j}}{\partial \omega}\right) u''\left(c_y^j\right)}{u''(c_y^j) + \hat{\beta}_j \delta u''(c_o^j)} \\ &-\frac{\hat{\beta}_j \delta u''(c_o^j)\left[\widehat{z_j} - \tau_y^+(\widehat{z_j}\theta(1-\alpha\delta) - \delta(1+n)\alpha) - \frac{\partial \tau_y^+(\omega,\alpha;\hat{\beta}_j)}{\partial \omega}\left[\omega\theta\widehat{z_j}(1-\alpha\delta) + \delta(1+n)\left((1-\alpha)\bar{\omega} + \alpha\omega\right)\right]\right]}{u''(c_y^j) + \hat{\beta}_j \delta u''(c_o^j)} \end{aligned} \tag{B-48}$$

The first term is negative and the second is positive: we are not able to give a clear sign to this expression: therefore, we assume that is positive. From (B-29) we can see that the function $\widehat{s_j}(\omega)$ is negative when ω_- is low and positive when ω approaches to $\tilde{\omega}$. Thus, there exists a value ω' such that saving are zero for $\omega \leq \omega'$ and positive above.

Parts *(iv)* and *(v)* follow from proposition 3. Part *(vi)* can be proved in the same way as in proposition 3.

Case 2: $\lambda_s > 0,\ \lambda_\tau > 0$

Individuals with $\omega \leq \omega'$ prefer $\tau_y^+(\omega, \alpha; \hat{\beta}_j) = 0$ and do not save.

Case 3: $\lambda_s = 0,\ \lambda_\tau > 0$

Rich individuals with income $\omega' \leq \omega \leq \tilde{\omega}$ rely exclusively on private saving, whose expression is implicitly defined by (B-29).

B.6 Proof of Proposition 5

First, we show that the object function for old is concave: differentiating (24) twice with respect to τ_o :

$$\frac{\partial^2 V^o(\tau_o, \alpha; \beta_j, \omega)}{\partial \tau_o^2} = u'\left(c_o^j\right)\left[-\omega\theta z_j + \delta \frac{\partial^2 P(z_j)}{\partial\left(\tau_o\right)^2}\right] + u''\left(c_o^j\right)\left[-\omega\theta z_j + \hat{\beta}_j \delta \frac{\partial^2 P(z_j)}{\partial\left(\tau_o\right)^2}\right]^2$$

From (B-30), the first term in square brackets at the RHS is negative and the objective function is concave.

To prove *(i)*, we differentiate $V^o(\tau_o, \alpha; \beta_j, \omega)$ at $\tau = 0$:

$$\left.\frac{\partial V^o(\tau_o, \alpha; \beta_j, \omega)}{\partial \tau_o}\right|_{\tau=0} = u'\left(c_o^j\right)\left[-\omega\theta\left(\frac{\omega - \beta_j\delta\psi}{\gamma - \beta_j\delta\psi}\right)(1 - \beta_j\delta\alpha) + \beta_j\delta(1+n)[\alpha\omega + (1-\alpha)\bar{\omega}]\right] \qquad \text{(B-49)}$$

At $\tau = 0$, everyone works in the second period. This expression is greater than 0 only for $\omega_- \leq \omega \leq \dot{\omega}$, where the threshold $\dot{\omega}$ is the (positive) root of the following the second-degree polynomial:

$$\omega^2 \underbrace{\left(\frac{1}{\beta_j\delta}\right)}_{a} - \omega \underbrace{\left(\psi + \alpha(1+n)\frac{\theta\,(1-\beta_j\delta\alpha)}{\gamma - \beta_j\delta\psi}\right)}_{b} - \underbrace{\frac{(1+n)(1-\alpha)\bar{\omega}\theta\,(1-\beta_j\delta\alpha)}{\gamma - \beta_j\delta\psi}}_{c} < 0 \qquad \text{(B-50)}$$

The solution is $0 \leq \omega \leq \dot{\omega} = \frac{-b+\sqrt{b^2-4ac}}{2a}$; $\tau_o^+(\omega, \alpha; \beta_j)$ is positive only for individuals with income below the threshold.

Part *(ii)* can be obtained by totally differentiating (B-28) with respect to ω, for $\omega \leq \dot{\omega}$:

$$\frac{d\tau_o^+(\omega, \alpha; \beta_j)}{d\omega} = -\frac{\left[-\theta z_j\,(1-\beta_j\delta\alpha) - \omega\theta\frac{\partial z_j}{\partial\omega} + \beta_j\delta\alpha\left((1+n) + \theta z_j + \theta\omega\frac{\partial z_j}{\partial\omega}\right)\right] u'\left(c_o^j\right)}{D_\tau^2}$$
$$+ \frac{\left[-\omega\theta z_j\,(1-\beta_j\delta\alpha) + \beta_j\delta(1+n)\,(\alpha\omega + (1-\alpha)\bar{\omega})\right]\left[z_j(1-\tau\theta) + \beta_j\delta(1+n)\alpha + \beta_j\delta\alpha\tau\theta z_j\right] u''\left(c_o^j\right)}{D_\tau^2}$$

Where $D_\tau^2 < 0$ is the second derivative of the objective function. Given our assumptions on $u(.)$ and since we focus only on the case $\omega \leq \dot{\omega}$, the second term on the numerator is negative. For the first one, observe that:

$$-\theta z_j\,(1-\beta_j\delta\alpha) - \omega\theta\frac{\partial z_j}{\partial\omega} + \beta_j\delta\alpha\left((1+n) + \theta z_j + \theta\omega\frac{\partial z_j}{\partial\omega}\right) < 0$$

$$\iff \omega^2 \underbrace{\left(\frac{\theta\,(1-\beta_j\delta\alpha)\,(2-\tau\theta\,(1-\beta_j\delta\alpha))}{\gamma - \beta_j\delta\psi}\right)}_{a} - \omega \underbrace{\left(\beta_j\delta\frac{\theta\psi\,(1-\beta_j\delta\alpha)}{\gamma - \beta_j\delta\psi} + \beta_j\delta\alpha(1+n)\right)}_{b} - \underbrace{\beta_j\delta(1+n)(1-\alpha)\bar{\omega}}_{c} \qquad \text{(B-51)}$$

The solution takes the form $0 \leq \omega \leq \ddot{\omega} = \frac{-b+\sqrt{b^2-4ac}}{2a}$. It can be shown that $\ddot{\omega} \leq \dot{\omega}$; thus, preferred tax rates are decreasing with income for $\ddot{\omega} \leq \omega \leq \dot{\omega}$. Unfortunately, we can not give clear predictions for productivity levels $\omega_- \leq \omega \leq \ddot{\omega}$: however, since $\tau_o^+(\omega_-, \alpha; \beta_j) > 0$ (see part *(i)* of the proposition), we are able to compare preferred tax rates at ω_- and $\ddot{\omega}$. Equation (B-49) gives us the expression for old preferred tax rate:

$$\tau_o^+(\omega, \alpha; \beta_j) = \frac{\omega\theta(\omega - \hat{\beta}_j\delta\psi)(1-\hat{\beta}_j\alpha\delta) - \hat{\beta}_j\delta(\gamma - \hat{\beta}_j\delta\psi)(1+n)\,(\alpha\omega + (1-\alpha)\bar{\omega})}{\omega^2\theta^2(1-\hat{\beta}_j\alpha\delta)} \qquad \text{(B-52)}$$

It is easy to verify that $\tau_o^+(\omega_-, \alpha; \beta_j) < \tau_o^+(\ddot{\omega}, \alpha; \beta_j)$; therefore, for $\omega_- \leq \omega \leq \ddot{\omega}$, optimal tax rates for old are increasing with income.

For *(iii)*, let us take the derivative of $\dot{\omega}$ with respect to β_j:

$$\frac{\partial \dot{\omega}}{\partial \beta_j} = \frac{\partial h}{\partial \beta_j} + \frac{1}{2\sqrt{\Delta \text{ of expression } (B-50)}} \left(\frac{\partial (h^2)}{\partial \beta_j} + \frac{\partial \left(\frac{(1-\alpha)\bar{\omega}(1+n)(\gamma - \beta_j \delta \psi)}{\theta(1-\beta_j \alpha\delta)\beta_j \delta} \right)}{\partial \beta_j} \right) \tag{B-53}$$

where $h = \frac{\psi}{2} + \frac{\alpha(1+n)(\gamma-\beta_j\delta\psi)}{2\theta(1-\beta_j\delta\alpha)}$. We want to show that all terms of (B-53) are negative: for the first one, note that:

$$\frac{\partial h}{\partial \beta_j} = \frac{2\alpha(1+n)\theta\delta(\alpha\gamma - \psi)}{\left[2\theta\left(1-\beta_j\delta\alpha\right)\right]^2} < 0 \Leftrightarrow \alpha\gamma < \psi$$

as we assume. The parenthesis is also negative, since:

$$\frac{\partial (h^2)}{\partial \beta_j} = 2h\frac{\partial h}{\partial \beta_j} < 0$$

and:

$$\frac{\partial \left(\frac{(1-\alpha)\bar{\omega}(1+n)(\gamma - \beta_j \delta \psi)}{\theta(1-\beta_j \alpha\delta)\beta_j \delta} \right)}{\partial \beta_j} = \frac{-\delta\psi(1+n)\bar{\omega}(1-\alpha)\theta(1-\beta_j\alpha\delta)\beta_j\delta - \theta\delta(1-\alpha)\bar{\omega}(1+n)(\gamma-\beta_j\delta\psi)(1-2\alpha\delta\beta_j)}{\left(\theta(1-\beta_j\alpha\delta)\beta_j\delta\right)^2}$$

This sign of this derivative depends crucially on the term $1-2\alpha\delta\beta$: for $\alpha\delta \leq 1/2$, the whole term is negative, and thus $\frac{\partial\dot{\omega}}{\partial\beta_j} < 0$. For $\alpha\delta \geq 1/2$, rearranging the expression above, we get:

$$-(1+n)\bar{\omega}(1-\alpha)\theta\delta\left[-\gamma + \beta_j\psi\delta\frac{\beta_j\alpha\delta}{(2\alpha\delta\beta_j - 1)}\right] < 0$$

$$\Longleftrightarrow \gamma - \frac{\psi\beta_j^2\alpha\delta^2}{(2\alpha\delta\beta_j-1)} \leq 0$$

Since we assume γ being not too high, i.e. $\frac{\psi}{\alpha} > \gamma$, a sufficient condition for this inequality to be true is $\frac{\psi\beta_j^2\alpha\delta^2}{(2\alpha\delta\beta_j-1)} > \frac{\psi}{\alpha} \Longleftrightarrow (\beta_j\alpha\delta - 1)^2 > 0$, which is always true. Therefore, also for $\alpha\delta \geq 1/2$, $\frac{\partial\dot{\omega}}{\partial\beta_j} < 0$.

Part *(iv)* can be proved by differentiating (B-49) with respect to β_j :

$$\frac{d\tau_o^+(\omega,\alpha;\beta_j)}{d\beta_j} = \frac{u'\left(c_o^j\right)\left[-\omega\theta\frac{\partial z_j}{\partial\beta_j} + \delta\frac{\partial P(z_j)}{\partial\tau} + \beta_j\delta\frac{\partial^2 P(z_j)}{\partial\tau\partial\beta_j}\right]}{D_\tau^2}$$

The sign of the derivative depends on the term in square brackets; replacing the expressions for $\frac{\partial z_j}{\partial\beta_j}$, $\frac{\partial P(z_j)}{\partial\tau}$ and $\frac{\partial^2 P(z_j)}{\partial\tau\partial\beta_j}$, we get a second degree polynomial in ω such that $\frac{\partial\tau_o^+(\omega,\alpha;\beta_j)}{\partial\beta_j} \leq 0$ for $\omega_- \leq \omega \leq \omega_c$ and $\omega_d \leq \omega \leq \dot{\omega}$ and $\frac{\partial\tau_o^+(\omega,\alpha;\beta_j)}{\partial\beta_j} > 0$ otherwise.

Finally, part *(v)* can be shown in the same way as Propositions 3 and 4: for individuals with $\omega_- \leq \omega \leq \dot{\omega}$, $\tau_o^+(\omega,\alpha;\beta_j) > 0$. Therefore, we have:

$$\begin{aligned} \frac{\partial V^o\left(\tau_o^+(\omega,\alpha;\beta_j),\alpha,\omega;\beta_j\right)}{\partial\tau} &= 0 \\ \frac{\partial^2 V^o\left(\tau_o^+(\omega,\alpha;\beta_j),\alpha,\omega;\beta_j\right)}{\partial\tau^2} &< 0 \end{aligned}$$

Differentiating the first expression with respect to α, we obtain:

$$\frac{\partial^2 V^y\left(\tau^+_{yo}(\omega,\alpha;\beta_j),\alpha,\omega;\beta_j\right)}{\partial\tau^2}\frac{\partial\tau^+_o(\omega,\alpha;\beta_j)}{\partial\alpha}+\frac{\partial^2 V^o\left(\tau^+_o(\omega,\alpha;\beta_j),\alpha,\omega;\beta_j\right)}{\partial\tau\partial\alpha}=0$$

The sign of $\frac{\partial\tau^+_o(.)}{\partial\alpha}$ is the same as $\frac{\partial^2 V^o\left(\tau^+_o(\omega,\alpha;\beta_j),\alpha,\omega;\beta_j\right)}{\partial\tau\partial\alpha}$. From (24), we get:

$$\frac{\partial^2 V^y(.)}{\partial\tau\partial\alpha}=u''(c^j_o)\left(-\omega\theta\widehat{z_j}+\delta\frac{\partial P(.)}{\partial\tau}\right)\left(\frac{\partial c^j_o}{\partial\alpha}\right)+u'(c^j_o)\left(-\omega\theta\frac{\partial\widehat{z_j}}{\partial\alpha}+\beta_j\delta\frac{\partial^2 P(\widehat{z_j})}{\partial\tau\partial\alpha}\right)$$

The first term is zero because of the first order condition for τ. Moreover, replacing the expressions for $\frac{\partial\widehat{z_j}}{\partial\alpha}$ and $\frac{\partial^2 P(\widehat{z_j})}{\partial\tau\partial\alpha}$:

$$\frac{\partial^2 V^o(.)}{\partial\tau\partial\alpha}=u'(c^j_o)\left[\beta_j\delta(1+n)(\omega-\bar{\omega})+\frac{\theta\omega\delta}{\gamma-\hat{\beta}_j\delta\psi}\left(\omega-\hat{\beta}_j\delta\psi+\omega\theta\tau B\right)\right]$$

where $B=\delta\beta_j\alpha(1+3\beta_j)-1-\beta_j-\beta_j^2>0$. Therefore, this term (and also $\frac{\partial\tau^+_o(\omega,\alpha;\beta_j)}{\partial\alpha}$) is positive for $\bar{\omega}\leq\omega\leq\tilde{\omega}$. Otherwise, for very poor individuals, $\omega_-\leq\omega<<\bar{\omega}$, we have $\frac{\partial\tau^+_o(\omega,\alpha;\beta_j)}{\partial\alpha}<0$.

B.7 Proof that $\tau^+_o(\ddot{\omega},\alpha;\beta_j)$ and $\tau^+_y(\omega_-,\alpha;\hat{\beta}_j))$ are not comparable

When $\alpha\delta>1/2$, the maximal tax rate for old is given by (B-52), while the maximal rate for young is (B-34). To determine which tax rate prevails, we have to find the sign of the following expression:

$$sign\left\{\theta\left[\frac{\ddot{\omega}-\hat{\beta}_j\delta\psi}{\ddot{\omega}}-\frac{\omega_--\hat{\beta}_j\delta\psi}{\omega_-(1-2\alpha\delta)}-\frac{(\gamma-\hat{\beta}_j\delta\psi)}{\omega_-(1-2\alpha\delta)}\right]-(\gamma-\hat{\beta}_j\delta\psi)\left[\frac{\hat{\beta}_j\delta(1+n)(\alpha\ddot{\omega}_-+(1-\alpha)\bar{\omega})}{\ddot{\omega}^2(1-\beta\alpha\delta)}-\frac{\delta(1+n)(\alpha\omega_-+(1-\alpha)\bar{\omega})}{\omega_-^2(1-2\alpha\delta)}\right]\right\}$$

The first term is positive, but the second is negative: a deeper analysis does not provide any interesting conditions for the whole expression to have a clear sign.

B.8 Proof that Preferences over α are Single Crossing

Single crossing requires that, for τ^i_j $i=y,o$ fixed, $\omega_1<\omega_2$ and $\alpha_1<\alpha_2$:

$$V^y(\tau,\alpha_2;\hat{\beta}_j,\omega_1)>V^y(\tau,\alpha_1;\hat{\beta}_j,\omega_1)\Longrightarrow V^y(\tau,\alpha_2;\hat{\beta}_j,\omega_2)>V^y(\tau,\alpha_1;\hat{\beta}_j,\omega_2)$$

Let us assume, without loss of generality, that $V^y(\tau,\alpha_2;\hat{\beta}_j,\omega_1)>V^y(\tau,\alpha_1;\hat{\beta}_j,\omega_1)$; it follows that:

$$[\widehat{z_j}(\alpha_2,\omega_1)-\widehat{z_j}(\alpha_1,\omega_1)]\left[\omega_1(1-\tau\theta)-(\gamma+\delta\psi)\left(\widehat{z_j}(\alpha_2,\omega_1)+\widehat{z_j}(\alpha_1,\omega_1)\right)-\delta\psi\right]+$$

$$+\delta\tau\omega_1\left[(1+n)(\alpha_2-\alpha_1)\right]+\theta\left[\alpha_2\widehat{z_j}(\alpha_2,\omega_1)-\alpha_1\widehat{z_j}(\alpha_1,\omega_1)\right]>\bar{\omega}\delta\tau(1+n)(\alpha_2-\alpha_1)$$

After having replaced for the expressions for $\widehat{z_j}(\alpha_2,\omega_1)$ and $\widehat{z_j}(\alpha_1,\omega_1)$, we get:

$$\omega_1\underbrace{\left\{(1-\tau\theta)\left[\left(\gamma-\hat{\beta}_j\delta\psi\right)(1+\hat{\beta}_j)-\hat{\beta}_j\right]+\frac{\delta\theta\tau\hat{\beta}_j(\alpha_2+\alpha_1)(\gamma(2-\hat{\beta}_j)-3\psi\delta\hat{\beta}_j)}{2}\right\}}_{K}+A>\frac{\bar{\omega}(1+n)(\gamma-\hat{\beta}_j\delta\psi)^2}{\theta\omega_1}$$

where $A = (1+n)(\gamma - \hat{\beta}_j\delta\psi)^2 - \theta\hat{\beta}_j\delta\psi(\gamma - \delta\psi)(1+2\hat{\beta}_j)$ is a constant term that does not depend on ω. Given our initial assumption $\omega_1 < \omega_2$, we have:

$$\omega_2 K > \omega_1 K > \frac{\bar{\omega}(1+n)(\gamma - \hat{\beta}_j\delta\psi)^2}{\theta\omega_1} > \frac{\bar{\omega}(1+n)(\gamma - \hat{\beta}_j\delta\psi)^2}{\theta\omega_2}$$

which is what we want to show.

B.9 Proof of Proposition 8

For part *(i)*, maximization of (25) gives us the following FOC, evaluated at $\alpha = 0$:

$$\left.\frac{\partial V_j^o(.)}{\partial\alpha}\right|_{\alpha=0} : u'(c_o)\left[\omega(1-\tau\theta) - \gamma z_j\frac{\partial z_j}{\partial\alpha} + \beta_j\delta\left(\tau\omega(1+n+\theta z_j) + \alpha\tau\omega\theta\frac{\partial z_j}{\partial\alpha} - \bar{\omega}(1+n)\tau - \psi(1-z_j)\frac{\partial z_j}{\partial\alpha}\right)\right] = 0 \tag{B-54}$$

After some rearrangements:

$$\omega^2\underbrace{\theta(1-\tau\theta)}_{a} + \omega\underbrace{[(1+n)(\gamma - \beta_j\delta\psi) - \beta_j\theta\delta\psi]}_{b} - \underbrace{\bar{\omega}(1+n)(\gamma - \beta_j\delta\psi)}_{c}$$

We get a second degree polynomial, whose roots are given by $\omega_{e',e} = \frac{-b\pm\sqrt{b^2+4ac}}{2a}$. Therefore, $\alpha_o^+(\tau,\omega;\beta_j) > 0$ for income levels $\omega_e \leq \omega \leq \omega_+$, whereas, for productivity $\omega_- \leq \omega \leq \omega_f$, $\alpha_o^+(\tau,\omega;\beta_j) = 0$. In particular, $\omega_f = \frac{-((1+n)(\gamma-\beta_j\delta\psi)-\beta_j\theta\delta\psi)+\sqrt{((1+n)(\gamma-\beta_j\delta\psi)-\beta_j\theta\delta\psi)^2+4\theta(1-\tau\theta)\bar{\omega}(1+n)(\gamma-\beta_j\delta\psi)}}{2\theta(1-\tau\theta)}$.

To show part *(ii)* of the proposition, we derive the threshold ω_f with respect to β_j:

$$\frac{\partial\omega_f}{\partial\beta_j} = \frac{(1+n)\delta\psi + \theta\delta\psi}{2\theta(1-\tau\theta)}\left(1 - \frac{((1+n)(\gamma-\beta_j\delta\psi) - \beta_j\theta\delta\psi)}{\sqrt{X}}\right) - \frac{4\theta(1-\tau\theta)\bar{\omega}(1+n)\delta\psi}{\sqrt{X}}$$

where $X = ((1+n)(\gamma-\beta_j\delta\psi) - \beta_j\theta\delta\psi)^2 + 4\theta(1-\tau\theta)\bar{\omega}(1+n)(\gamma-\beta_j\delta\psi)$. Notice that:

$$\frac{((1+n)(\gamma-\beta_j\delta\psi)-\beta_j\theta\delta\psi)^2}{((1+n)(\gamma-\beta_j\delta\psi)-\beta_j\theta\delta\psi)^2+4\theta(1-\tau\theta)\bar{\omega}(1+n)(\gamma-\beta_j\delta\psi)} < 1$$

It follows that $\frac{\partial\omega_f}{\partial\beta_j} < 0$.

For part *(iii)*, we use the implicit function theorem to obtain:

$$\frac{\partial\alpha_o^+(\tau,\omega;\beta_j)}{\partial\hat{\beta}_j} = -\frac{u'(c_o)\left[\frac{\partial^2 z_j}{\partial\alpha\partial\beta_j}(\omega(1-\tau\theta) - z_j(\gamma-\beta_j\delta\psi) + \beta_j\delta\alpha\tau\omega\theta - \beta_j\delta\psi)\right]}{D_\alpha^2} + \\ + \frac{u'(c_o)\left[\frac{\partial z_j}{\partial\beta_j}\left(\frac{\partial z_j}{\partial\alpha}(\gamma-\beta_j\delta\psi) - \beta_j\delta\tau\omega\theta\right) + \frac{\delta}{\beta_j}\frac{\partial z_j}{\partial\alpha}(\gamma z_j - \omega(1-\tau\theta))\right]}{D_\alpha^2}$$

Replacing the expressions for $\frac{\partial^2 z_j}{\partial\alpha\partial\beta_j}, \frac{\partial z_j}{\partial\beta_j}, \frac{\partial z_j}{\partial\alpha}$ and z_j, we get:

$$\frac{\partial\alpha_o^+(\tau,\omega;\beta_j)}{\partial\hat{\beta}_j} = -\frac{u'(c_o)\delta^2\frac{\partial z_j}{\partial\alpha}(\gamma(\alpha\tau\omega\theta - \psi) + \omega(1-\tau\theta)\psi)}{D_\alpha^2} > 0$$

Since $D_\alpha^2 < 0$ and the numerator is always positive.

For *(iii),* we differentiate, with respect to α, $\frac{\partial V_j^o(.)}{\partial \alpha} = 0:$

$$\frac{\partial^2 V^o((\alpha_o^+(\omega,\tau;\beta_j),\tau,\omega;\beta_j))}{\partial \alpha^2}\frac{\partial \alpha_o^+(\tau,\omega;\beta_j)}{\partial \tau} + \frac{\partial^2 V^o((\alpha_o^+(\omega,\tau;\beta_j),\tau,\omega;\beta_j)}{\partial \tau \partial \alpha} = 0$$

Therefore, $\frac{\partial \alpha_o^+(\tau,\omega;\beta_j)}{\partial \tau}$ has the same sign as $\frac{\partial^2 V^o(.)}{\partial \tau \partial \alpha}$. Following our previous discussion, we have that for productivity above the average, $\frac{\partial \alpha_o^+(\tau,\omega;\beta_j)}{\partial \tau} > 0$.

B.10 Proof of Proposition 9

For part *(i)*, maximization of (25) gives us the following FOC, evaluated at $\alpha = 0$:

$$\left.\frac{\partial V_j^o(.)}{\partial \alpha}\right|_{\alpha=0} : u'(c_o)\left[\omega(1-\tau\theta) - \gamma \hat{z}_j \frac{\partial \hat{z}_j}{\partial \alpha} + \delta\left(\tau\omega(1+n+\theta\hat{z}_j) + \alpha\tau\omega\theta\frac{\partial \hat{z}_j}{\partial \alpha} - \bar{\omega}(1+n)\tau - \psi(1-\hat{z}_j)\frac{\partial \hat{z}_j}{\partial \alpha}\right)\right] = 0$$

After some rearrangements, we have that young vote have a preferred positive α if and only if:

$$\omega^2 \underbrace{\theta(1-\tau\theta)\left(\gamma - \hat{\beta}_j^2\delta\psi\right)}_{a} - \omega \underbrace{\left[(1+n)\left(\gamma - \hat{\beta}_j\delta\psi\right) + \frac{\hat{\beta}_j\theta\delta\psi}{\left(\gamma - \hat{\beta}_j\delta\psi\right)}\left(\gamma(2-\hat{\beta}_j) + \hat{\beta}_j\delta\psi\right)\right]}_{b} - \underbrace{\bar{\omega}(1+n)\left(\gamma - \hat{\beta}_j\delta\psi\right)}_{c} > 0$$

We get a second degree polynomial, whose roots are given by $\omega_{f',f} = \frac{-b\pm\sqrt{b^2+4ac}}{2a}$. Being the first root negative, $\alpha_o^+(\tau,\omega;\beta_j) > 0$ only for $\omega_f \leq \omega \leq \omega_+$. To prove parts *(ii)* and *(iii)*, we proceed as in Proposition 8.