Social security and retirement decision: A positive and normative approach

Helmuth Cremer[1], Jean-Marie Lozachmeur[2]
and Pierre Pestieau[3]

December 2005

[1]GREMAQ and IDEI, University of Toulouse, 21 Allée de Brienne, 31000 Toulouse, France; e-mail: helmut@cict.fr.

[2]GREMAQ, University of Toulouse.

[3]CREPP, University of Liège, CORE, PSE and CEPR.

Abstract

Social insurance for the elderly is judged responsible for the widely observed trend towards early retirement. In a world of laissez-faire or in a first-best setting, there would be no such trend. However, when first-best instruments are not available, because health and productivity are not observable, the optimal social insurance policy may imply a distortion on the retirement decision. The main point we make is that while there is no doubt that retirement systems induce an excessive bias towards early in many countries, a complete elimination of this bias (i.e., a switch to an actuarially fair system) is not the right answer. This is so and for two reasons. First, some distortions are second-best optimal. This is the normative argument. Second, and on the positive side, the elimination of the bias might be problematic from a political perspective. Depending on the political process, it may either not be feasible or alternatively it may tend to undermine the political support for the pension system itself.

Keywords: social security, early retirement, optimal income taxation, majority voting.

JEL classification: J26, H55, H21.

1 Introduction

Over the last decades one has observed a trend towards lower and lower participation rates of elderly male workers in Europe. For men aged 60–64 these rates were above 70% in the early 1960s and they have fallen to below 20% in Belgium, France, Italy and the Netherlands. It has been shown that this trend, which applies as well to women and to the younger age group 55-60, can be in great part explained by the incentive structure underlying social protection programs aimed at elderly workers: pension plans, but also unemployment and disability insurance, and early retirement schemes. As a result, these social protection programs are far from being actuarially fair as regards the retirement decision, more in countries such as Germany or France than in countries such as Sweden or Finland. In this paper we survey recent work aimed at addressing two questions. First, do such biases towards early retirement result from bad policy or can they be vindicated on equity and efficiency grounds? Second, can the political process at work in Europe explain that today in so many countries workers retire much before the "normal" age of retirement? As a corollary to the latter question, why is it so difficult to reform the system now that all governments agree that something has to be done?

This survey is organized as follows.[1] In Section 2, we present the facts and the standard explanation: even though they live longer and longer, people retire earlier and earlier mainly because of biases inherent to the social security system. Those biases are explained by the belief that retiring elderly workers fosters youth employment and by firms which want to replace overpaid elderly workers by underpaid young workers. In Section 3 we use a linear retirement scheme. We first analyze the optimal scheme from a utilitarian viewpoint. Then we turn to majority voting models. Individuals vote either for payroll taxes that induce an age of retirement or they directly vote for some uniform age of retirement. It also explains why there is so much resistance to raise the age of retirement even though such a reform is welfare improving. In Section 4 we present the optimal non-linear retirement scheme using an optimal income tax setting with two

[1] For other surveys, see Fenge and Pestieau (2005) and Cremer, Lozachmeur and Pestieau (2005a).

sources of asymmetric information, ability and health.

Even though this survey is concerned with a theoretical and conceptual question, the underlying issue has obvious policy implications. To understand these, our results have to be put in a proper perspective. The paper does show that some downward distortion in the retirement decision is unavoidable in a second best setting. However, this does *not* justify the ridiculous activity rates of elderly workers one observes in countries such as France or Germany. There is not doubt that in these countries raising the age of retirement is desirable not only because it generates additional resources to both the pension administration and the individuals but also because these additional public resources can be used for redistribution.

This point is particular relevant in times of aging population. With a PAYG system and increasing longevity, one normally expects an increase in the age of retirement and in the rate of contribution and a cut in pension benefits. If the age of retirement is kept constant at a rather low level relying on just the two other instruments is most likely inefficient. Compared to a reform wherein the adjustment to aging goes only through a decrease in benefits (the contribution rate being unchanged because of tax competition), one has shown that for a country like Belgium allowing for an increase in the age of retirement could be Pareto improving.(Cremer and Pestieau (2003)).

To sum up, the main point we make is that while there is no doubt that retirement systems induce an excessive bias towards early in many countries, a complete elimination of this bias (i.e., a switch to an actuarially fair system) is not the right answer. This is so and for two reasons. First, some distortions are second-best optimal. This is the normative argument. Second, and on the positive side, the elimination of the bias might be problematic from political perspective. Depending on the political process, it may either not be feasible or alternatively it may tend to undermine the political support for the pension system itself.

2 The evidence

Over the coming decades, the EU will face an acceleration of demographic ageing due to three main factors: the baby-boom generation reaching retirement age, continued increase in life expectancy and decreased fertility. All three factors jointly create a major financial challenge for pension systems when the number of pensioners will rapidly increase and the size of the working age population will diminish.[2]

By the year 2050, European will live at least four to five years longer than today. Such an increase will raise the cost of providing the same pension level by 25 to 30%. At the same time, large cohorts born after World War II will reach retirement age and subsequent cohorts are much smaller as a result of lower birth rates. As a consequence, one expects a quasi doubling of the old age dependency ratio, i.e., the number of people of statutory retirement age (65+) to the potentially working population (15–64). In the year 2000 those over 65's represented about one quarter of the working age population in the EU25. By 2050, this figure will be nearly 50%.

This is the demographic challenge that Europe is facing and has to prepare for. In a *laissez-faire* world or alternatively in a first-best optimal economy, a large part of this challenge could be addressed by adjusting the age of retirement to the increased longevity. Controlling for growth, it is rather intuitive that in a *laissez-faire* setting, if my longevity is increasing I will work longer in order to keep an appropriate rate of replacement (ratio of pension to earnings). And yet this is not what we observe in a number of European countries. Over the last decades, the rate of activity of elderly workers has steadily decreased and henceforth the effective age of retirement. Further, attempts to reverse this evolution is heavily resisted.

Table 1 gives the employment rate among older workers (55–64) as well as the effective retirement age in 2001. Focusing on the latter, it ranges from 57 in Belgium to 63.1 in Ireland. The evolution is even more striking. Table 2 provides the change in longevity and retirement age in 8 European countries during the period 1960–95. It shows clearly that the commonly observed increase in the length of retirement that is represented in Figure 1, is due as much to a decreasing activity rate than to an increase in life expectancy. But why have we observed such a decline in the rate of employment in old age? In the late 90's, both the NBER and the OECD documented a strong relationship across countries between social security incentives to retire and

[2] See EC (2003).

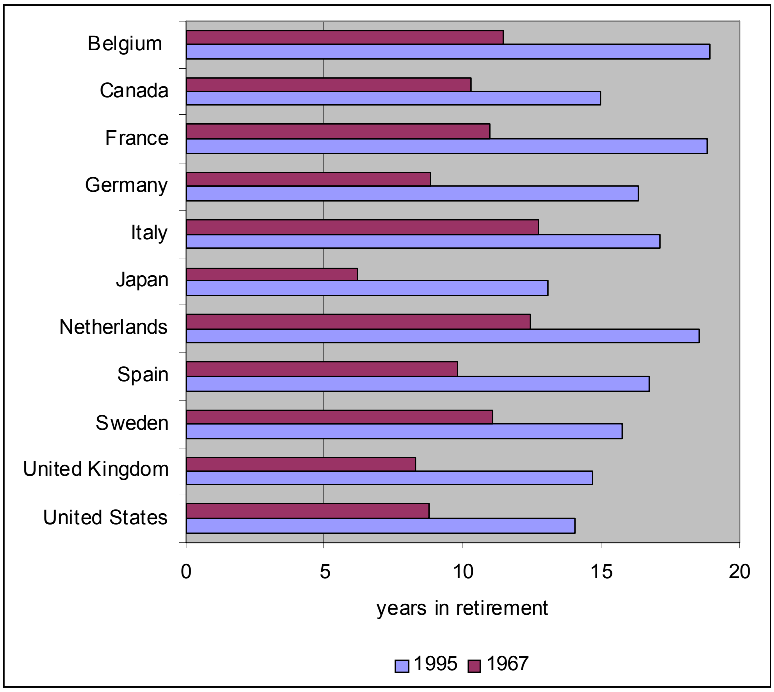

Figure 1: Years in retirement of males in 1967 and in 1995. *Source: Blöndal and Scarpetta (1998a)*

the proportion of older persons out of the labor force (Gruber and Wise (1999, 2005), Blöndal and Scarpetta (1998a,b)). In measuring incentives to retire, these authors include not only pension benefits, but also disability, unemployment and early retirement compensations as these are available to elderly workers and may induce them to retire before the first eligibility age for the standard pension system.

Suppose that though these alternative programs workers can leave the labor market as early as age 55. For each age t above 55, one computes the difference between the expected discounted value of social security benefits if retirement is age $t+1$ and the present value if retirement is at age t. This difference, called the accrual of benefit between age t and age $t+1$, is often negative and thus reduces total compensation from working an additional year.

The negative of the accrual to net wage earnings ratio is the so called implicit tax on continued work. By summing the implicit tax over age 55-65, we obtain the so called tax

Table 1: Employment rates and average labour market withdrawal age, 2001

	Average age of withdrawal from labour market 2000/01 (in brackets: standard retirement age)			Older workers employment rate (55-64)
	Women	Men	Total	2001
BE	55.9(62)	57.8(65)	57.0	25.1
DK	61.1(65)	62.2(65)	61.9	58.0
DE	60.4(65)	60.9(65)	60.7	37.7
EL	57.7(65)	61.2(65)	59.6	38.0
ES	60.2(65)	60.7(65)	60.6	39.2
FR(1)	58.0(60)	58.2(60)	58.1	31.9
IE	62.2(66)	63.2(66)	63.1	46.8
IT	59.2(65)	59.6(60)	59.4	28.1
LU	55.3(65)	57.5(65)	56.8	24.4
NL	60.3(65)	61.1(65)	60.9	39.6
AT	58.6(60)	60.0(65)	59.6	28.6
PT	61.5(65)	62.0(65)	62.0	50.1
FIN	61.4(65)	61.6(65)	61.6	45.8
SE	61.9(65)	62.1(65)	62.0	66.8
UK	61.0(65)	63.1(60)	62.1	52.3
EU15	59.1	60.5	59.9	38.8

(1) *First quarter.*
Source: Eurostat, LFS.

Table 2: Longevity and retirement age (1960-1995)

	Men				Women			
	Life expectancy		Retirement age		Life expectancy		Retirement age	
	1960-65	95-00	1960	1995	1960-65	95-00	1960	1995
Belgium	67.9	73.8	63.3	57.6	73.9	80.6	60.8	54.1
France	67.6	74.2	64.5	59.2	74.5	82.0	65.8	58.3
Germany	67.4	73.9	65.2	60.5	72.9	80.2	62.3	58.4
Ireland	68.4	73.6	68.1	63.4	72.3	79.2	70.8	60.1
Italy	67.4	75.0	64.5	60.6	72.6	81.2	62.0	57.2
Spain	67.9	74.5	67.9	61.4	72.7	81.5	68.0	58.9
Sweden	71.6	76.3	66.0	63.3	75.6	80.8	63.4	62.1
UK	67.9	74.5	66.2	62.7	73.8	79.8	62.7	59.7

Source: Cremer and Pestieau (2003)

force to retire. Figure 2 shows the relationship between the tax force and unused labor force capacity – the proportion of men between age 55 and 65 that is out of the labor force. It is clear that there is a very strong correspondence between the two. This clearly shows that the social security systems in many countries provide strong incentives to leave the labor force at older ages. As shown by Gruber and Wise (2005) such a decline in labor force participation puts enormous pressure on the financial solvency of social security system.

The normal course of action at this stage would be to raise the age of retirement or more formally to make the overall social security system more actuarially neutral. Reforms in that direction have been conducted in some countries such as Finland and Sweden. In other countries, particularly in those which need reforms the most they have failed. In a recent EU-wide opinion survey carried out in 2001, it appears that raising the pensionable age is not a popular response to the challenge of demographic aging: fewer than a quarter of Europeans would support such a move. At the same time, the view that older workers should make room on the labor market for younger and unemployed people is still widely held in spite of evidence that there is no relation between the two. The idea that there is a given number of jobs that need to be shared, what is known as the lump-sum labor fallacy, still appears to be shared in the public opinion of many

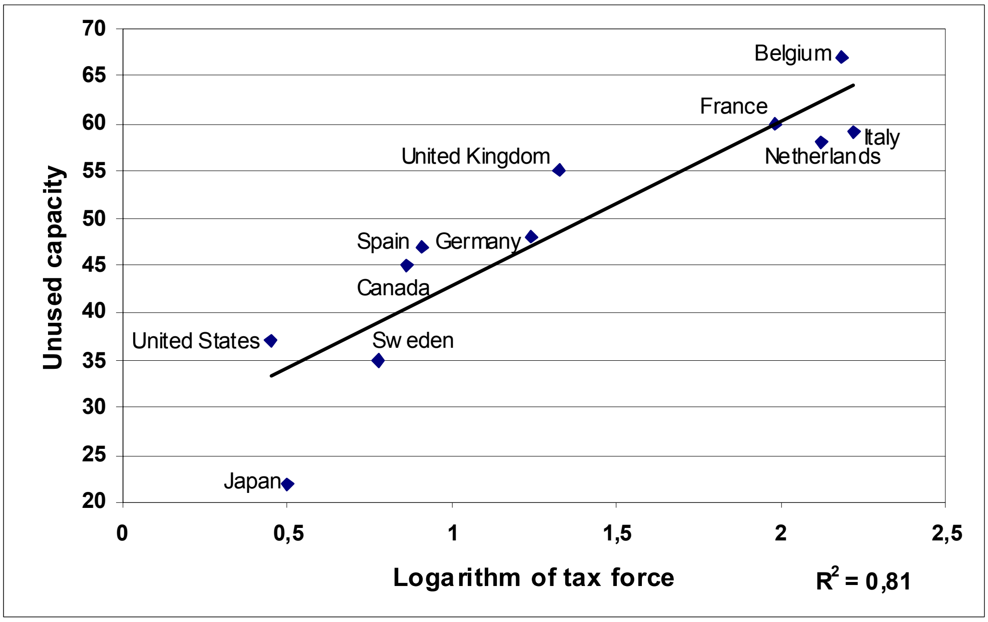

Figure 2: Unused capacity vs. tax force. *Source: Gruber and Wise (1999)*

countries, particularly those with low effective age of retirement. This appears clearly on Figure 3.

3 Linear pension scheme

3.1 Basic model

We consider a small open economy within a two overlapping generations setting. Both interest rate, r and wage $\overline{w}$ are given. At each point in time t two generations coexist: L_t young workers and L_{t-1} old workers and retirees with $L_t = L_{t-1}(1 + n_t)$. Individuals differ in two ways: the generation they belong to and their productivity, w_i with mean $\overline{w}$ and median $w_m < \overline{w}$. Individual labor supply is given and normalized to 1 in the first period; it is endogenous in the second period and equal to $z_i \leqslant 1$. This variable z_i can be viewed as the age of retirement.[3]

[3] Assume that each generation lasts 30 years. Lifetime consists of 30 years of childhood and training that are not accounted for, 30 years of full activity and a last period of 30 years the first part of which is devoted to work and lasts $30z$ years. The age of retirement is thus $60 + 30z$.

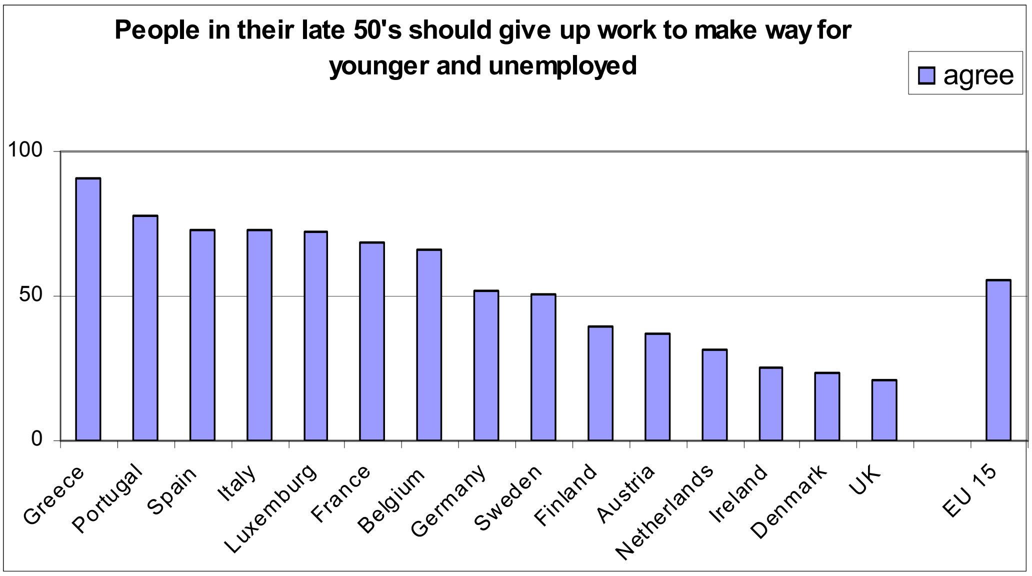

Figure 3: Perceived intergenerational redistribution. *Source: EC (2003).*

Lifetime utility is given by

$$u(c, d, z) = u(c) + \beta\ u(d - h(z)), \tag{1}$$

where c and d denote first and second period consumption, $u(\cdot)$ is strictly concave and $h(\cdot)$ strictly convex; β is a time discount factor. In the absence of any pension system, each individual solves the following problem:

$$\max_{s_i, z_i} \quad u[w_i - s_i] + \beta u[(1 + r)s_i + w_i z_i - h(z_i)], \tag{2}$$

where $s_i \geqslant 0$ denotes saving. There is a liquidity constraint so that negative saving is not possible.

The optimality conditions are

$$w_i = h'(z_i), \tag{3}$$

and, denoting $x_i = d_i - h(z_i)$ the second period consumption net of the monetary labor disutility,

$$\begin{aligned} -u'(c_i) + \beta(1 + r)u'(x_i) &= 0 \quad (s_i > 0), \\ &< 0 \quad (s_i = 0). \end{aligned} \tag{4}$$

Turning to the utilitarian first-best, we can express the problem of the social planner as follows:

$$\max_{c_i,d_i,z_i} \quad \sum_i n_i \left[u(c_i) + \beta\, u(d_i - h(z_i)) - \mu \left(c_i + \frac{d_i}{1+r} - w_i - \frac{w_i z_i}{1+r} \right) \right], \tag{5}$$

where n_i is the fraction of workers of type i and μ the Lagrange multiplier associated with the revenue constraint. The FOCs are given by

$$u'(c_i) = u'(x_i)\,\beta\,(1+r) = \mu, \tag{6}$$

and

$$w_i = h'(z_i)\,. \tag{7}$$

In words, equation (6) shows that consumption levels are equalized across individuals. When $\beta(1+r) = 1$ they are also equalized across periods ($c_i = \alpha_i$). Equation (7) shows that retirement age is the same as in the *laissez-faire*.

We now introduce the pension scheme. It has three key features: first the payroll tax is proportional to earnings; second, it is based on the PAYGO principle; third, it has a contributory and a flat rate benefit component. Formally, the pension of a worker with wage w_i is

$$p(w_i) = (1+n)\,\tau\left[\alpha w_i + (1-\alpha)\,\overline{w}\right] + \tau\left[\alpha z_i w_i + (1-\alpha)\,\overline{wz}\right], \tag{8}$$

where $\overline{wz} = \sum_i n_i w_i z_i$ is the average second period income, while α is the contributory share. Substituting into the individual problem (2) yields the objective function

$$\begin{aligned} &u\left[w_i(1-\tau) - s_i\right] + \beta\, u\Big[w_i z_i (1 - \tau(1-\alpha) + (1+r) s_i \\ &+\tau\big((1+n)(\alpha w_i + (1-\alpha)\bar{w}) + (1-\alpha)\,\overline{wz}\big) - h(z_i)\Big], \end{aligned} \tag{9}$$

and the choice of z_i is given by

$$w_i(1 - \tau(1-\alpha)) = h'(z_i)\,. \tag{10}$$

To interpret these expressions let us compare them to their counterparts under an alternative pension system, indexed FF that is both fully funded and non redistributive. Under such a system we would have:

$$p^{FF}(w_i) = (1 + r + z_i)\,\tau w_i.$$

The individual's objective is then given by

$$u\big(w_i\,(1-\tau)-s_i\big)+\beta\ u\big(w_iz_i+(1+r)\,(\tau w_i+s_i)-h(z_i)\big),$$

and retirement age is determined by (7). With the FF and with the "contributory" system (i.e., a PAYGO scheme with $\alpha = 1$), there is no distortion and the age of retirement is at its first-best level (for every individual). With the considered PAYGO system on the other hand, it follows from (10) that as soon as we have some redistribution, which is the case when $\alpha < 1$, a distortion towards early retirement is unavoidable. This applies not only to the linear case, but also to the non linear tax as we show in Section 4. Observe that a way to avoid this distortion (even when $\alpha < 1$) would be to differentiate taxes between periods with τ_1 and τ_2 and set $\tau_2 = 0$. We shall also examine this possibility below.

3.2 Optimal linear scheme

We consider the utilitarian optimum obtained by maximizing the sum of utilities (9) subject to the revenue constraint (8). The instruments are α and τ. This problem is solved by Cremer, De Donder, Maldonado and Pestieau (2005a) who show that the level of α is crucially affected by the liquidity constraint. When this constraint is not binding (for any individual), one has $\alpha = 0$;[4] otherwise the optimal pension formula involves "targeting', i.e., one obtain $\alpha < 0$. In either case, a positive tax rate obtains under plausible assumptions and the value of τ depends on two terms: the covariance of $u'(x_i)$ and wz_i (equity term) and the elasticity of z with respect to net-of-tax wage rate $(w\,(1-\tau))$ (efficiency term). Not surprisingly, the expression for the optimal tax rate thus has the same structure as in the traditional optimal linear income tax problem à la Sheshinsky.[5]

Besides the distinction between Bismarckian and Beveridgean system, there is another question: assuming flat rate benefit, should the tax rate be the same in the two

[4] To be more precise $\alpha = 0$ is a solution, but it is not unique. What matter is the value of $\tau\,(1-\alpha)$. Any combination of τ and α that satisfy $\tau\,(1-\alpha) = \tau^*$, where $\tau^*$is the optimum tax rate with $\alpha = 1$, yields exactly the same outcome. The main point from our perspective is that this necessarily implies $\alpha < 1$.

[5] See Sheshinsky (1972).

periods. Or to put it otherwise, should there be a tax rebate in the second period in order to mitigate the tax distortion. And, as mentioned above, setting $\tau_2 = 0$ completely eliminates distortion of the retirement age, even when $\alpha = 0$. To address this issue, we derive the utilitarian optimum as above except that the instruments are now τ_1 and τ_2 (while α is set at zero). The solution is derived in Casamatta *et al.* (2005b) for the case where $h(z) = z^2/2$ and the expression for τ_2 is given by

$$\tau_2 = \frac{E\ u'(x)\left[w^2 - \overline{w^2}\right]}{E\ u'(x)\left[w^2 - 2\overline{w^2}\right]},$$

or

$$\tau_2 = \frac{\operatorname{cov}\left(u'(x), w^2\right)}{\operatorname{cov}\left(u'(x), w^2\right) - \overline{w^2}E\ u'(x)},$$

where $\overline{w^2} = \sum n_i w_i^2 = Ew^2$. The following main results emerge. First, the tax on first period earnings is less than 100%, even though first period earnings are tax-inelastic. This is because of the liquidity constraint. Second, a zero percent tax on second period earnings is not desirable because a utilitarian social planner wants to redistribute that income. Consequently, a distortion towards early retirement is called for. Whether the first period tax is higher than that of the second period will depend on the liquidity constraints, the tax elasticity of z and on the density of w.

To sum up, with a uniform tax we necessarily obtain $\alpha < 1$ and when the tax rate is allowed to differ, we have $\tau_2 > 0$. Consequently, whatever the specific features of the solution, it always involves a distorted retirement decision.

3.3 Voting on the age of retirement

There are two ways of looking at the positive choice of the age of retirement. First, one could think of voting on a mandatory age of retirement $\widehat{z}$ or one can vote on the tax-transfer parameters of the pension system that induce an endogenous age of retirement. For example, in the above example, the level of z chosen by an individual is a function of the net wage in the second period, $w_i(1 - \tau(1 - \alpha))$, which in turn depends on the parameters of the pension scheme.[6]

[6] This assumes away income effect, which is implied by quasi-linear utility functions.

Voting on the age of retirement implies an assumption concerning whether the tax or the benefit is defined. Using the simple flat rate benefit system (i.e., with $\alpha = 0$), we have:

$$\tau \left(\widehat{z}\overline{w} + \overline{w} \right) = \left(1 - \widehat{z} \right) p.$$

The choice of $\widehat{z}$ will not be the same if we posit τ or p as given. The traditional distinction between defined contribution and defined benefits schemes appears crucial here. For more on this, see Lacomba and Lagos (1999) and Fenge and Pestieau (2005).

The second approach, namely the vote on the parameters of a linear retirement schedule is adopted in Casamatta *et al.* (2005b), where individuals vote on the tax rate τ to be applied in the two periods and a rebate $\theta < 1$ that reduces the tax in the second period. Using the notation τ_1 and τ_2, we have $\tau_1 = \tau$ and $\tau_2 = \theta\tau$. Observe that $\theta = \tau_2 = 0$ would yield the case where the retirement decision is not distorted.

Using a sequential voting procedure (first θ, then τ) the authors show that for a given value of θ, high tax rates are supported by the old and by low productivity young individuals. From the perspective of this survey, the most important result is that the introduction of θ, namely the distortion on the retirement choice, increases the political support for the pension system. More precisely, when θ increases the number of voter who favor a positive tax (and are thus in favor of a retirement system) increases. Conversely, a decrease (or an elimination) of the bias undermines the political support of the pension system. Finally, considering the simultaneous choice of θ and τ through an issue by issue voting procedure à la Shepsle, it appears that the equilibrium, if any, always implies a downward distortion on the age of retirement.

Casamatta *et al.* (2005a) also study the determination of a (linear) pension scheme through the political process. They consider a situation in which the choice of the retirement age is subject to a double burden: the tax burden that has been studied so far and another burden that arises from the system when working one more year does not increase the social security wealth. In other words, the increase (if any) of the yearly pension is not sufficient to compensate for the foregone benefits during the extra year of activity. A simple way to introduce this double burden is to assume that total (lifetime) pension is given by $p - \delta z$ where $\delta > 0$ is the penalty from postponing retirement. With a Beveridgean system, the worker then chooses z such that it maximizes:

$$u(w(1-\tau)z + p - \delta z - h(z)).$$

The choice of z is then determined by:

$$h'(z) = w(1-\tau) - \delta$$

The authors show that such a double burden would be rejected by both a utilitarian and a Rawlsian social planner. Furthermore, each individual would reject it as a "citizen candidate".[7] Nevertheless, it can be supported by a particular structure of social (or political) weights *biased towards the more productive workers.* The idea underlying this result is that the implicit tax is the result of a compromise: whereas a low tax rate is chosen so as to favor the highly productive workers, the policy maker chooses a positive value of δ in order to compensate the low productive workers. According to this result, a double burden could emerge from the political process as long as this process implies that sufficient weight is put on the more productive workers. Similarly, a reform towards a more neutral system may not be supported by the political process.

Up to now, no mention of incomplete working history for family or unemployment reasons was made. Conde Ruiz and Galasso (2003) also try to provide a positive theory of early retirement.[8] They find that the majority which supports early retirement is composed of elderly with incomplete working history and low-ability workers.

3.4 Political resistance to reforms

In the previous subsections we have shown that both normative and positive approaches imply some distortion inducing early retirement. Another question is why reforms that appear to be desirable for a majority of people following a shock, e.g., a demographic one, cannot be implemented.

Cremer and Pestieau (2003) argue that some reforms are rejected because individuals do not understand what is the alternative non-reform scenario. Consider a society facing

[7] In other words, the double burden would not arise when the social welfare function puts all the weight on a single individual.

[8] See also Conde Ruiz and Galasso (2004).

a sudden increase in the dependency ratio. The choice offered to voters is between raising the age of retirement and keeping the same benefits and keeping the age or retirement unchanged and cutting benefits. A majority should back the increase in retirement age and yet it does not. The reason is that most people believe that the alternative to the reform is the *status quo* (same benefits and same retirement age) and not a reduction in benefits.

Another line of explanation comes from the uncertainty as to the future. Consider a society consisting of individuals distinguished according to their productivity ($w_H > w_L$) and their health ($\gamma_H > \gamma_L$). Their second period utility is

$$u\left(w_i\widehat{z}\,(1-\tau)+p-\frac{h\,(\widehat{z})}{\gamma_i}\right),$$

where $\widehat{z}$ is a mandatory age of retirement. Assume for simplicity that there are only three types of individuals: type 1 has a productivity w_L and a poor health denoted by γ_L; type 2 has the same productivity but a good health γ_H; type 3 has a higher productivity than the two other types w_H and a good health γ_H.

People with high productivity as well as people with low productivity but good health (γ_L) are in favor of a reform consisting in raising $\widehat{z}$. People with low productivity and bad health are against. Assuming that type 1 individuals represent less then half of the total population, the reform would be supported by a majority of individuals if the vote take place when voters already know their health status. However assume that the health status is only known after the reform is implemented. Thus it is not impossible that all low wage individuals would oppose the reform.[9] Consequently, if types 2 and 3 form a majority, a reform that would be favored by a majority ex post may not be adopted ex ante because voters are uncertain about the way they are individually affected by this reform.

4 Non linear schemes

Up to now we have discussed the issue of retirement and social security within a linear framework. Linearity is easier in a normative setting and yields more intuitive results.

[9] This idea comes from Fernandez and Rodrik (1991).

Furthermore it is necessary for a political economy treatment of the issue at hand. It is, however, limited as one knows that the government can do better using a non linear scheme. And in practice both taxes and benefits typically involve at least some nonlinearities. In this section, we study a non linear scheme for public pensions and taxes. We first present the basic model and then turn to a number of applications.

4.1 Basic model

This model comes from Cremer, Lozachmeur and Pestieau (2004a). Consider an individual with lifetime h and retirement age z. His consumption over his lifetime is equal to earnings minus taxes plus pension benefits. Formally, using a continuous time framework and zero interest rate, one writes:

$$\int_0^h c(t)\,dt = \int_0^z w(1-\tau)\,dt + \int_z^h p(z)\,dt,$$

where $w\tau$ is the amount of payroll taxes and $p(z)$ the level of benefits depending on the age of retirement. Implicit to this budget constraint, there is a flow of saving (negative or positive). Assuming a zero time discount rate, lifetime utility is given by:

$$U = \int_0^h u(c(t))\,dt - \int_0^z \rho(t)\,dt,$$

where $u(\cdot)$ is the utility for consumption and $\rho(t)$ reflects the disutility of labor, that includes both tastes and health aspects.

With these assumptions one can obtain a reduced form for both the budget constraint and the utility function:

$$hc = wz - T(z)$$

and

$$U = hu(c) - R(z)$$

where $R(z) = \int_0^z \rho(t)\,dt$ and $T(z)$ is a non linear tax-transfer scheme consisting of two terms:

$$T(z) = \int_0^z w\tau d\tau - \int_z^h p(z)\,dt = z\tau w - (h-z)\,p(z)\,. \tag{11}$$

The first term represents total (lifetime) payroll taxes and the second is overall pension. Note that differentiating $T(z)$, we obtain the implicit tax on prolonged activity consisting of three terms: payroll tax, forgone pension benefit and the change in social security wealth:

$$T'(z) = \tau w + p(z) - (h - z) p'(z).$$

We now introduce two types of heterogeneity with two values for each: w_H and w_L ($w_H > w_L$) and R_1 and R_2 ($R_1 > R_2$). We thus have four types of individuals: $H2$ (high productivity, good health or weak disutility for labor), $H1$ (high productivity and bad health), $L2$ (low productivity, good health) and $L1$ (low productivity, bad health). Let n_{ij} denote the proportion of types ij. In a first stage we assume that both w_j and R_i are common knowledge and we look at the first-best problem given by the following Lagrangian:

$$\mathcal{L} = \sum_{ij} n_{ij} \left[hu(c_{ij}) - R_i(z_{ij}) \right] - \gamma \sum_{ij} n_{ij} \left[hc_{ij} - w_j z_{ij} \right].$$

One readily obtains the following optimality conditions:

$$\begin{aligned} u'(c_{ij}) &= \gamma \\ R'_i(z_{ij}) &= \gamma w_j. \end{aligned}$$

In other words, in the utilitarian first-best all consumption levels are equal; the age of retirement is higher for the more productive and the healthier individuals.

If both health and productivity were observable such a solution could be decentralized by determining the appropriate lump-sum taxes T_{ij} that could induce everyone to consume the same amount. The optimal condition for the choice of z_{ij} is indeed the same as the one of *laissez-faire*.

We now present a number of applications of this model to the issue of retirement in case of informational asymmetry. The first one concerns the design of an optimal tax-transfer scheme when health is not observable. The second focuses also on the health variable and consider the case where a high $R(z)$ can be due to bad health or to a strong preference for leisure. Both situations are indistinguishable except through costly audit. The third application focuses also on the health issue and includes the possibility of private and public investment in medical care with the aim of lowering $R(z)$.

4.2 Health heterogeneity and implicit taxation

To simplify further the above model, we assume that only health distinguishes individuals. In other words, they all have the same productivity and only differ in health. The second-best problem can then be expressed by the following Lagrangian:

$$\begin{aligned}\mathcal{L} &= \sum_i n_i \left[hu\left(c_i\right) - R_i\left(z_i\right) - \gamma\left(hc_i - wz_i\right)\right] \\ &\quad -\lambda\left[hu\left(c_2\right) - R_2\left(z_2\right) - hu\left(c_1\right) + R_2\left(z_1\right)\right]\end{aligned}$$

where γ is the multiplier associated with the revenue constraint and λ with the self-selection constraint making sure that the individual with the better health does not mimick that with the worse health.

Cremer, Lozachmeur and Pestieau (2004a) obtain the following results:

$$c_1 < c_2 \ ;$$
$$T'\left(z_1\right) > 0 \ ;$$
$$T'\left(z_2\right) = 0 \ .$$

Compared to the first-best, type 2's individuals choice of retirement is not distorted. However, for type 1 asymmetric information results in a lower consumption and also an earlier age of retirement. Put differently, the marginal tax on the retirement age is positive. Intuitively this property can be explained by the fact that type 1 individuals have steeper indifference curves at any given point in the (z, c) space than type 2 individuals. This is because type 1 individuals must be compensated more to accept to work longer than the mimicking individual (they are less healthy and have a higher weekly labor supply). This implies that, starting from the first best tradeoff, a variation $dz_1 < 0$ along with a variation $dc_1 = (MRS^1_{cz})dz_1$ has no (first-order) effect on the utility of type 1, but it decreases the utility of type 2 mimicking type 1. Consequently, the downward distortion in z_1 is a way to relax an otherwise binding self selection constraint.

In conclusion, we have here an implicit tax on prolonged activity at least for some individuals that results from constrained social welfare maximization.

4.3 Audit on health

A natural way to resorb at least part of the lack of information is to introduce audits. If we consider bad health as a sort of disability, to avoid healthy workers to claim undeserved benefits disability tests can be conducted.

4.3.1 Perfect audit

Cremer, Lozachmeur and Pestieau (2004b) introduce a particular type of audit, namely a costly but error-proof audit. Formally any individual claiming to have type 1 is audited with probability π. This audit is perfect and has a marginal cost k such that with truthtelling the number of audits is $f_1\pi$ and total cost $f_1k\pi$. If type 2's individuals were caught lying, they would be given a minimum utility $\underline{u}$.

Note that without audit the optimal policy of the previous subsection implies truthtelling. Audits affect both the resource constraint and the self-selection constraint as follows:

$$\sum_i n_i\left(wz_i - hc_i - f_1k\pi\right) = 0\ ;$$
$$u\left(c_2\right) - R_2\left(z_2\right) - \left(1-\pi\right)\left(u\left(c_1\right) - R_2\left(z_1\right)\right) - \pi\underline{u} = 0.$$

Increasing the audit probability is costly but it allows for relaxing the incentive constraint by lowering the level of utility achievable by the mimicker. In the optimal solution, a compromise is found between these two considerations.

If $k = 0$, the central planner uses an audit probability that makes the self-selection constraint non binding. The first-best solution is achieved without any distortion. As k increases, audit has to be restricted to a decreasing number of type 1's individuals. The distortion on the choice of the retirement age is increased. Naturally, when the audit cost is too high we go back to the problem of the previous subsection.

One could give a premium to all those who are audited and naturally found truthtelling. Such a premium would give a plus to people who need it.

4.3.2 Indistinguishable individual

Assume now that there are three types of individuals indexed D, L and H. The first two types of individuals both have a high labor disutility $R_1\left(> R_2\right)$. Individuals of type D

are truly disabled while those of type L, are just leisure prone. Individuals of type H have a low labor disutility R_2. In a *laissez-faire* setting types D and L would retire at the same age ($z_D = z_L < z_H$) and consume a rather low level ($c_D = c_L < c_H$). Suppose that the social planner wants to favor the truly disabled at the expense of the leisure prone workers. Given that these two types are indistinguishable, the self-selection constraint implies that they must have the same utility level. The only way to differentiate them is through audits.

Cremer, Lozachmeur and Pestieau (2005c) show that when the social planner treats the disutility for work of the leisure prone workers as equal to that of the healthy workers, the solution with audits discriminates between the disabled and the leisure prone allowing the former to retire earlier but also to consume less than the latter. From a practical perspective, we can think about the payment made to individuals who are subject to audits as disability benefits.

Formally the social objective function is

$$\sum_{i=H,L,D} n_i u(c_i) - n_H R_2(z_H) - n_L R_2(z_L) - n_D R_1(z_D).$$

We present a numerical example illustrating the problem at hand. We consider two audits: audit 1 is costlier than audit 2. Table 2 gives the results for 4 settings: first-best, second-best without audit, second-best with cheap and expensive audit. In the first-best, the three types receive the same consumption and both individuals H and L work more than individual D. There is a subsidy to induce the leisure-prone individuals to retire at the same age as the healthy ones. In the second-best without audit D and L cannot be distinguished; they both face a tax on postponed retirement. Audit improves the welfare of the disabled at the expense of the two other types. With audit 1 D's utility increases and that of L decreases. The retirement age hardly changes, but the consumption levels are very different. With audit 2, the difference between L and D is sharper: lower retirement age, higher consumption level and higher utility. With audit 1 both L and D face a tax on postponed activity; individuals H must be prevented from mimicking either L or D. When audit gets less costly, H is not willing to mimick D. A tax on z_D is thus superfluous.

	First-best			Second-best without audit		
Types	D	L	H	D	L	H
c	86.74	86.74	86.74	57.72	57.72	94.04
z	0.28	1.15	1.15	0.39	0.39	1.06
U	4.29	1.80	3.79	3.74	3.74	3.97
$T'(z)$	0	-3	0	0.09	0.09	0

	Audit 1 $\pi = 0.04$			Audit 2 $\pi = 0.08$		
Types	D	L	H	D	L	H
c	65.35	54.05	91.55	68.06	53.70	90.59
z	0.38	0.37	1.09	0.36	0.41	1.10
U	3.88	3.71	3.92	3.95	3.63	3.89
$T'(z)$	0.003	0.19	0	0	0.11	0

Table 2: Numerical example with and without audit

4.3.3 Imperfect audit

In most cases audits are far from being perfect; they do not always reveal who is truly disabled and who is not. Diamond and Sheshinsky (1995) have allowed for errors in auditing.[10] Their paper analyzes optimal disability and retirement (or welfare) benefits with imperfect (but free) disability evaluation. Errors can be of both types with some able workers judged disabled and some disabled workers judged able with a certain probability. The retirement decision is binary: individuals either work a given amount of time or do not work at all. Individuals only differ with respect to their disutility of work. When this parameter is above a certain threshold value, they do not work. The optimal policy implies that society is divided into three groups: (i) healthy workers; (ii) healthy and disabled individuals who stop working and are tagged as disabled; (iii) disabled who were tagged as healthy individuals and who receive minimal benefits (welfare payment) that are lower than disability benefits. The individuals tagged as disabled can be paid more generously than the others because the benefits they get do not affect the work incentives of the healthy workers.

4.4 Health care and retirement

We have just seen that the distortion that arises from asymmetric information on health can be eliminated or at least reduced thanks to audits. We now turn to another device for reducing distortions, namely health expenditures that would mitigate health differentials.

Cremer, Lozachmeur and Pestieau (2004c) study the design of a social security scheme which can be supplemented by health care. They consider a setting where agents differ both in their productivity and in their health status. The health status determines the disutility of intensive (work week) and extensive (career length) work. Additionally, it is also affected by health expenditures. Thus, an increase in health expenditures can lower the disutility of work. Social preferences are assumed to be utilitarian and neither the health status nor the productivity of individuals are publicly observable. They consider the case for public provision of health services *as additional*

[10] See also Parssons (1996).

instrument which is used along with an optimal tax policy. They show that under plausible conditions, government intervention in the health care sector should combine subsidization of (privately provided) health care with public provision of health services.

4.5 Disability insurance in a dynamic framework

The models reviewed so far assume that consumers allocate consumption and labor supply over their life cycle. Nevertheless the models are essentially of static nature in the sense that no new information is revealed over the life cycle. In reality, however, people learn about their health status over time and this feature can be expected to affect the design of retirement and disability insurance.

Such dynamic properties of the optimal disability insurance scheme have been first analyzed by Diamond and Mirrlees (1978, 1986). They consider ex-ante identical individuals who face at each period of their life a probability to become disabled. When disabled, individuals definitely stop to work while non disabled individuals retire at a planned retirement age. The government aims at providing a disability insurance scheme without observing the health status of the agents. These authors show that because of asymmetric information, disability benefits should increase with the age of entry in the disability system. The tax sytem decentralizing this optimum has two basic properties. First, there is a positive tax on continued activity that decreases with age until it reaches zero at the planned retirement age. Second, there should be a marginal tax on savings. This arises because individuals can use savings as a reinsurance device. As typical in the insurance litterature, this reinforces the moral hazard problem. Consequently, the government taxes savings in order to relax moral hazard problems. The nature of this tax depends on what the government observes. If savings are observable as anonymous transactions, then the government can only use a positive linear tax on savings (see Diamond and Mirrlees (1995)). If savings are observable at the individual's level, the government can use a non linear tax on savings. As suggested by Golosov and Tsyvinski (2004), this non linear tax should be of the form of a wealth-tested transfer program. In other words, disabled individuals receive disability benefits only if their assets are sufficiently low.

5 Conclusion

In a number of countries the effective age of retirement is well below the statutory age and below what is the optimal age of retirement (at a utilitarian optimum) as well as the age that individuals would choose to retire in a *laissez-faire* economy. It seems that the main explanation for such a phenomenon lies in the way social security is organized. Pensions but also disability and unemployment insurance along with the tax system induce elderly workers to exit the labor force before they reach the age of 65 that is the statutory age in many countries.

Why do we observe those disincentives to prolong activity? Is is due to bad design or can it be justified on normative grounds? In this paper, we show that at least some distortions can be justified when public authorities use social insurance for redistribution when only distortionary tax tools are available. We consider a model with two sources of heterogeneity: health and productivity, and we show that distortions towards early retirement arise in a second-best setting. These distortions can be reduced when some uncertainty can be resorbed through audits or when some differences in health can be reduced through public health care.

We also deal with the issue of retirement from a political economy viewpoint. It appears that early retirement can result from majority voting when the pension system is partially redistributive. We also show why some reforms which are clearly welfare improving can be opposed by a majority of workers.

References

[1] Blöndal, S. and S. Scarpetta, (1998a), Falling participation rates among older workers in the OECD countries, OECD, Paris.

[2] Blöndal, S. and S. Scarpetta, (1998b), The retirement decision in OECD countries, OECD-Economic Department WP, n° 202.

[3] Casamatta, G., H. Cremer and P. Pestieau, (2005a), Is there a political support for the double burden on prolonged activity?, *Journal of Governance*, forthcoming.

[4] Casamatta, G., H. Cremer and P. Pestieau, (2005b), Voting on pension with endogenous retirement age, *International Tax and Public Finance*, 12, 7–28.

[5] Conde-Ruiz, J.I. and V. Galasso, (2003), Early retirement, *Review of Economic Dynamics*, 6, 12-36.

[6] Conde-Ruiz, J.I. and V. Galasso, (2004), The macroeconomics of early retirement, *Journal of Public Economics*, 88, 1849-1869.

[7] Cremer, H., P. De Donder, D. Maldonado and P. Pestieau, (2005), Redistribution and social security with rational and myopic individuals: linear pension schemes, mimeo.

[8] Cremer, H., J-M. Lozachmeur and P. Pestieau, (2005a), Départ à la retraite et taxation optimale, *Revue d'Economie Politique*, 115, 163–172.

[9] Cremer, H., J-M. Lozachmeur and P. Pestieau, (2005b), Retirement age and health expenditures, Annales d'Economie et de Statistique (forthcoming).

[10] Cremer, H., J-M. Lozachmeur and P. Pestieau, (2005c), Disability testing and retirement, mimeo.

[11] Cremer, H., J-M. Lozachmeur and P. Pestieau, (2004a), Social security, variable retirement age and optimal income taxation, *Journal of Public Economics*, 88, 2259–2281.

[12] Cremer, H., J-M. Lozachmeur and P. Pestieau, (2004b), Optimal retirement and disability benefits with audit, *FinanzArchiv*, 60, 278–295.

[13] Cremer, H. and P. Pestieau, (2003), The double dividend of postponing retirement, *International Tax and Public Finance*, 10, 419–434.

[14] Diamond, P. and E. Sheshinski, (1995), Economic aspects of optimal disability benefits, *Journal of Public Economics*, 57, 1–23.

[15] Diamond, P. and J. Mirrlees (1978), A Model of Social Insurance with Variable Retirement, *Journal of Public Economics*, 10, 295-336.

[16] Diamond, P. and J. Mirrlees (1986), Payroll Tax Financed Social Insurance with Variable Retirement, *Scandinavian Journal of Economics*, 88, 25–50.

[17] Diamond, P. and J. Mirrlees (1995), Social Insurance with Variable Retirement and Private Savings, unpublished.

[18] EC, (2003), Adequate and sustainable pension, joint report by the Commission and the Council, Employment and Social Affairs, Brussels.

[19] Fenge, R. and P. Pestieau, Social Security and Retirement, (2005), MIT Press, Cambridge.

[20] Fernandez, R. and D. Rodrik (1991). Resistance to reform: Status quo bias in the presence of individual-specific uncertainty, *American Economic Review*, 81, 1146–1155.

[21] Golosov, M. and A. Tsyvinski (2004), Designing Optimal Disability Insurance: A Case for Asset Testing, NBER Working Paper No. 10792.

[22] Gruber, J. and D. Wise, (2005), Social security programs and retirement around the world: fiscal implications. Introduction and summary, NBER Working paper 11290.

[23] Gruber, J. and D. Wise, eds, (1999), Social Security and Retirement around the World, University of Chicago Press, 1999.

[24] Parssons, D.O., (1996), Imperfect tagging in social insurance programs, *Journal of Public Economics*, 62, 183–207.

[25] Pestieau, P., (2001), Are we retiring too early?, *Ifo Studien*, 47, 263–281.

[26] Sheshinski, E., (1973), The optimal linear income tax, *Review of Economic Studies,* 39,297–302.