Life expectancy heterogeneity and the political support for collective annuities

Helmuth Cremer and Philippe De Donder

INSTITUT D'ECONOMIE INDUSTRIELLE

Toulouse
School
of Economics

Life expectancy heterogeneity and the political support for collective annuities[1]

Helmuth Cremer[2]
Philippe De Donder[3]

This version: May 2014

[2]Toulouse School of Economics (GREMAQ, IDEI and Institut universitaire de France).
[3]Toulouse School of Economics (GREMAQ-CNRS and IDEI).

Abstract

Individuals, differing in productivity and life expectancy, vote over the size and type of a collective annuity. Its type is represented by the fraction of the contributive (Bismarckian) component (based on the worker's past earnings) as opposed to the noncontributive (Beveridgean) part (based on average contribution). The equilibrium collective annuity is either a large mostly Bismarckian program, a smaller pure Beveridgean one (in accordance with empirical evidence), or nil. A larger correlation between longevity and productivity, or a larger average life expectancy, both make the equilibrium collective annuity program more Beveridgean, although at the expense of its size.

Keywords: generosity, redistributiveness, pay-as-you-go pensions, collective annuity, longevity, Kramer-Shepsle structure-induced equilibrium

JEL Codes: D78, H55

1 Introduction

A sizeable body of literature deals with the political determination of the characteristics of a public pay-as-you-go pension system. The seminal paper by Browning (1975) assumed that the only heterogeneity between voters is their age. Subsequent papers (such as Casamatta *et al.* (2000a)) have enriched this approach by assuming that agents also differ in income or in productivity. This richer set of individual traits has allowed these papers to study the determination of both the size of the pension system and of its redistributiveness across income levels. As for the latter, the literature (surveyed by Galasso and Profeta (2002)) has contrasted so-called Bismarckian systems, where the pension benefit is proportional to the individual contribution, with Beveridgean systems, where the benefit is based on the average contribution.

The main stylized fact in this domain is that Bismarckian systems tend to be larger (as measured by either the contribution rate or the share of public pensions in GDP) than Beveridgean ones: see Table 1 in Casamatta *et al.* (2000b), Disney (2004) or Conde-Ruiz and Profeta (2007). Very few papers have tried to explain this correlation. Both Casamatta *et al.* (2000b) and Koethenbuerger *et al.* (2008) study the determination by the median voter of the size of the pension system as a function of the system's type, measured by the (exogenous) relative importance of the contributive component, dubbed the Bismarckian parameter. Casamatta *et al.* (2000b) assume exogenous wage income and obtain an ambiguous impact of the Bismarckian parameter on the equilibrium size of the pension system. Koethenbuerger *et al.* (2008) introduce endogenous labor supply, and obtain that larger labor supply distortions generated by the flat-rate Beveridgean pension system versus the earnings-related Bismarckian one, may explain why more redistributive systems are smaller. Conde-Ruiz and Profeta (2007) study the simultaneous determination of the size and type of pensions system. In their model, agents differ in age, income and in ability to invest in the capital market. With only three income groups, a small Beveridgean system is supported by low-income agents, who gain from its redistributive feature, and high-income individuals, who seek to minimize their tax contribution and to invest their resources in a private scheme. Middle

income individuals, instead, favor a large Bismarckian system. The degree of inequality in earnings and the level of capital market returns determine which type of equilibrium emerges.

An important dimension of heterogeneity among voters, which may play a critical role in the determination of the pension system, is absent from these studies: longevity. It is well known empirically how people of the same age differ in life expectancy. Moreover, life expectancy is positively correlated with income or wealth, as shown by Deaton and Paxon (1999) for the US, Attanasio and Emerson (2001) for the UK, and Reil-Held (2000) for Germany. Average life expectancy has been increasing in most countries for at least half a century. But these increases have not been shared equally everywhere. For instance, in the US, the average male life expectancy at 65 has increased from 15 to 16.1 years between 1986 and 2006 for individuals in the bottom half of the earnings distribution, but from 16.5 to 21.5 years in the top half of the earnings distribution (Waldron, 2007). It is thus important to assess the impact of such variations on pension programs.

This paper studies how and to what extent differences in longevity can contribute to explaining two main questions. First, why some countries (such as France or Germany) have a mostly contributive (Bismarckian) public pension system, while others (such as the UK or Canada) develop a mainly non-contributive (Beveridgean) system. Second, why Beveridgean systems tend to be smaller than Bismarckian ones. Tackling these questions requires building a political economy model with a bidimensional type space (income or productivity and life expectancy) and a bidimensional policy space (size and degree of redistributiveness of the public pension program), where both policy dimensions are chosen at the majority. Consequently, we need to go further than providing comparative static results showing how the size depends on the exogenous type of the system.

To the best of our knowledge, no paper has yet attempted to build such a model. Few papers endogenize the public pension program when life expectancy is heterogeneous. Cremer *et al.* (2010) study the design of pension systems and the role played by "collective annuities" as currently provided by public pension systems when individuals

differ in longevity (as well as in productivity).[1] Their approach is normative and based on a utilitarian social welfare function. They contrast two schemes: a pure contributory (Bismarckian) pension and a flat rate (Beveridgean) pension. They show that the case for collective annuities is stronger when they are associated with a flat pension system.[2]

Other papers take a positive perspective. Leroux (2010) studies the case where individuals have the same income but differ in their life span. She obtains that a majority of voters are in favor of a pension system awarding the same annuity to everyone if the distribution of longevity is negatively-skewed. Borck (2007) assumes from the outset that richer individuals always live longer lives (so that heterogeneity between agents is truly one dimensional) and shows how individual preferences and equilibrium pension policies are affected by the slope of the relationship between income and life expectancy. Finally, De Donder and Hindriks (2002) assume that individuals differ both in their productivity and survival probabilities. They focus on the majority chosen size of the pension system as a function of its (exogenous) redistributiveness. They show that the equilibrium amount of distortions associated to the pension system need not decrease when the system is exogenously made less redistributive, because voters react by increasing the pension system's size.[3]

Finally, some papers take a purely empirical path and study the empirical consequences of life expectancy differences for actual pension systems: Coronado *et al.* (2000) for the US, Gil and Lopez-Casasnovas (1997) for Spain, Bommier *et al.* (2005) for France and Reil-Held (2000) for Germany. These papers take the existing characteristics of the pension system as given and assess how the joint distribution of income and life expectancy affects its redistributiveness across income levels. Not surprisingly, they find that, with public pensions not related to individual longevity, the positive correl-

---

[1] The defining characteristic of collective annuities is that they do not depend on an individual's survival probability.

[2] With positive correlation between income and life expectancy, collective annuitization implies an undesirable redistribution from low incomes to high incomes and a desirable redistribution from short to long lived individuals. This is true both in Beveridgean and in Bismarckian systems. However, the undesirable redistribution is mitigated under a Beveridgean system, which in turn redistributes from high to low income individuals.

[3] Galasso and Profeta (2007) study the impact of ageing on the political determination of the size and income redistributiveness of pension systems. In their model, agents live two periods and differ in income, ageing being modeled as a decrease in the ratio of young (workers) over old (retirees) individuals.

ation between income and longevity reduces significantly the amount of redistribution across income levels.

We assume that individuals live at most two periods, and differ in productivity and in probability to be alive in the second period. In the first period, they choose how much to work and to save. In the second period, they retire, consume their saving and the pension benefit (if any), which is financed by a linear payroll tax on labor income. Pension benefits are paid out as a *collective annuity*[4] with both a (Bismarckian) contributive and a (Beveridgean) non-contributive component. Voters choose both the generosity (or size) of the pension system (the value of the proportional income tax rate) and its degree of redistributiveness (or type, measured by the relative importance of the contributive component, *i.e.,* the Bismarckian parameter).

We study the joint political determination of the size and type of the collective annuity program. It is well know that simultaneous voting over a bidimensional policy space has generically no equilibrium (see De Donder *et al.* (2012) for instance). We adopt the voting procedure first proposed by Kramer (1972) and Shepsle (1979), where each policy dimension is a majority voting equilibrium given the other dimension. We obtain that the median productivity voter is decisive in the choice of both size and type of pension system. We show the existence of a unique Kramer-Shepsle equilibrium, with the following properties. If the Bismarckian return is larger than the interest rate, the unique equilibrium is a large, mostly but not always exclusively Bismarckian program. If the Bismarckian return is smaller than the interest rate (because of a large correlation between income and life expectancy, for instance), the unique equilibrium depends on the median productivity level in the economy. If this productivity level is small, the unique equilibrium is a smaller and purely Beveridgean pension, while there is no collective annuity program at equilibrium with a large median productivity. These Shepsle equilibria correspond to the empirically observed large Bismarckian systems and

---

[4]We concentrate on the case where there is no private (individual) annuity market. This is in line with the empirical evidence, since most retirees are reluctant to buy an annuity, so much so that this behavioral pattern is often referred to as the "annuity puzzle"; see Brown *et al.* (2005). Finkelstein and Poterba (2002, 2004) and Mitchell (1996) show that where they exist, rates of return of individual annuities are much below actuarially fair levels and often significantly less attractive than the implicit return of collective annuities.

smaller Beveridgean ones.

Next we study how this political equilibrium is affected by variations in the longevity distribution. Both a larger positive correlation between income and longevity, and a larger average life expectancy when uncorrelated with income generate a more redistributive equilibrium pension program, although sometimes at the expense of its size. Consequently, taking theses induced changes in the system into account, their net impact on the amount of redistribution operated by collective annuities is not necessarily negative.

Our analytical results show that the political equilibrium crucially depends on key parameters such as the interest rate, the productivity and longevity distributions and their correlation. In Section 7 we illustrate these results with simulations calibrated to reflect the stylized facts in two countries, namely France and the UK. The results show that our model is consistent with the relevant empirical findings, namely a Bismarckian system for France and a Beveridgean system for the UK. Though admittedly highly stylized, our simple model thus provides a possible explanation for the main features of the retirement systems in these two countries. We also use the calibrated example to study some comparative statics properties of the political equilibrium. Part of the results provide mainly an illustration of our analytical findings. In addition, the simulations also allow us to study scenarios for which no analytical results could be obtained. For instance, we examine the impact of an increase in average life expectancy when wage and longevity are positively correlated.

2   The model

Consider an economy consisting of a continuum of agents who live (at most) two periods, working in the first period and retiring in the second one. Individuals differ in productivity w and in life expectancy, measured as the probability p to be alive in period 2.[5] The joint distribution of these two characteristics is denoted by $H(w,p)$, with marginal distributions $F(w)$ over $[0,\infty[$ and $G(p)$ over $[0,1]$. At this point, we do

---

[5] We will use without distinction the terms life expectancy, longevity and survival probability when referring to p.

not assume a specific pattern of correlation between the two characteristics w and p. However, to interpret our results we will concentrate on the empirically most relevant case where productivity and life expectancy are positively correlated. The average productivity is denoted by $\overline{w}$ while the average survival probability is $\bar{p}$. We assume as usual that the productivity distribution is positively skewed, so that the median productivity, w^{med} is lower than the average, $\bar{w}$.

Individual preferences are given by

$$u(c - h(z)) + \beta p u(d),$$

where c is first-period consumption, d is second-period consumption, $h(z)$ measures the disutility of supplying the labor quantity z and β is the discount rate. The function u is increasing and concave while h is increasing and convex with $h(0) = 0$. First-period consumption net of labor supply disutility is denoted by $x = c - h(z)$.

Individuals take two private decisions, both in their first period of life: labor supply z and saving, s. The labor supply choice can be interpreted as either at the intensive (number of hours worked) or extensive margin (such as the retirement age).[6] We assume away any borrowing constraint, so that saving can be negative.[7] Saving has a gross return of $1+r$, where $r > 0$ denotes the exogenous world interest rate.[8] The first-period individual budget constraint is given by

$$c = (1 - \tau)zw - s, \tag{1}$$

where $\tau \in [0, 1]$ is the payroll tax rate.

In their second period, individuals retire and consume their private saving and a

---

[6] If agents retire before the end of the first period, they do not collect any benefit before the beginning of the second period.

[7] The assumption that agents can borrow against future retirement income is of course a strong assumption, made to simplify the algebra. As we explain in the concluding section, imposing borrowing constraints would decrease the most-preferred tax rate of some voters, but would not affect the qualitative results we obtain.

[8] In other words we consider a "small" open economy. Because of capital mobility this assumption appears to be reasonable for essentially all European countries, but it may be debatable for the US. An alternative interpretation of this assumption, in line with Atkinson and Sandmo (1980), is to say that the interest rate is controlled by policy instruments (like public debt) which are not accounted for in our setting.

public pension benefit b (if any), so that

$$d = (1 + r)s + b.$$

Pension benefits take the form of a collective annuity. They are financed on a pay-as-you go basis so that tax proceeds finance pensions paid to current retirees, as is most often the case in reality. The annuity consists of a contributive (or Bismarckian) part based on the individual's contributions, and of a non-contributive (or Beveridgean) part linked to the average contribution. The contributive share of the benefit is denoted by $\alpha \in [0, 1]$ and referred to as the Bismarckian parameter. Making use of the government budget constraint and assuming away demographic and economic growth for simplicity, the pension benefit b is given by

$$b = \tau \left[(1 - \alpha) \frac{Ewz}{\bar{p}} + \alpha \frac{wz}{\hat{p}} \right],$$

where Ewz is the average first-period income,[9] $1/\bar{p}$ is the internal rate of return of the non-contributory (Beveridgean) collective annuity while $1/\hat{p}$ with

$$\hat{p} = \frac{Epwz}{Ewz}$$

is the internal rate of return of the contributory (Bismarckian) collective annuity. The two components of the collective annuity differ in both their internal rate of return and the basis on which this return is applied. Both components redistribute from short-lived to long-lived agents (since both are based on some aggregate rather than individual longevity), while the non-contributory part also redistributes across income levels.

All decisions (public and private) are taken by agents in the first period of their life: they first vote over the size (τ) and type (α) of the collective annuity program, observe the result of the vote and then decide how much to work (z) and to save (s) privately.[10]

---

[9] Throughout the paper, Ef denotes $\int f(w, p) dH(w, p)$ for any function f. Similarly, $\mathrm{cov}(f, g)$ denotes $E(fg) - E(f)E(g)$ for any functions f and g.

[10] Throughout the paper, we assume that only young people vote, and that the majority-chosen policy remains in place when they retire. With a pay-as-you-go collective annuity program, the voting behavior of retirees is well known. They favor the proceeds-maximizing contribution rate since their past contributions are sunk while they enjoy the tax proceeds from the current workers. As for the system's type, it is easy to see that they have the same preferences as a younger agent of the same characteristics. Allowing older people to vote then would not bring any new insight.

As usual, we proceed by backward induction and we first solve for the individual labor supply and saving decisions, before moving to the analysis of majority voting over the characteristics of the public system.

3 Individual choices of labor supply and saving

The first-order condition (FOC) with respect to private saving s is given by

$$p\beta u'(d)(1+r) = u'(x). \tag{2}$$

The FOC with respect to labor supply z is

$$-u'(x)\left[h'(z) - (1-\tau)w\right] + \alpha\beta\frac{p}{\hat{p}}\tau w u'(d) = 0. \tag{3}$$

Using (2) and the fact that $u'(x) > 0$, equation (3) simplifies to

$$(1 + \gamma\tau)w = h'(z),$$

where

$$\gamma = \frac{\alpha}{\hat{p}(1+r)} - 1 \tag{4}$$

measures the discounted value of the extra benefit to which an individual is entitled when his tax contribution increases at the margin, net of its cost.

The sign of γ depends on the comparison between the gross individual marginal return of pension, $\alpha/\hat{p}$, and the private saving return, $1 + r$. If they are equal, γ is nil and the FOC for labor supply simplifies to $w = h'(z)$, so that the contribution rate τ does not affect the labor supply decision. If $\alpha/\hat{p} < 1 + r$, γ is negative and labor supply decreases with the tax rate, while a positive value of γ means that labor supply increases with τ.[11] In all cases, labor supply increases with both the share α and return $1/\hat{p}$ of the contributive part, since both increase the individuals' return from their own tax contributions. Labor supply is not affected by individual or average survival probability, thanks to the absence of both income effect in preferences (see (1)) and of

---

[11]When z is interpreted as the retirement age, a negative γ corresponds to the "implicit tax on continued activity" studied in the pension literature.

borrowing constraints. Labor supply increases with productivity w irrespective of the sign of γ.

The indirect utility (incorporating the optimal choices z^* and s^* of all individuals) is given by

$$V(\alpha, \tau, w, p) = u\left[(1-\tau)wz^* - s^* - h(z^*)\right] + \beta pu\left[(1+r)s^* + \tau\left((1-\alpha)\frac{Ewz^*}{\bar{p}} + \alpha\frac{wz^*}{\hat{p}}\right)\right].$$

(5)

In the remainder of the paper, we assume that the disutility of labor is given by

$$h(z) = \frac{z^2}{2},$$

(6)

so that labor supply becomes

$$z = (1 + \gamma\tau)w,$$

with[12]

$$\hat{p} = \frac{Epwz}{Ewz} = \frac{Epw^2}{Ew^2}.$$

For future reference, note that

$$
\begin{aligned}
\text{cov}(w^2, p) &= Epw^2 - \bar{p}Ew^2 \\
\Leftrightarrow \hat{p} &= \bar{p} + \frac{\text{cov}(w^2, p)}{Ew^2},
\end{aligned}
$$

(7)

so that increasing the covariance between life expectancy and productivity while keeping the marginal distributions of p and w unchanged increases $\hat{p}$ and decreases the labor supply of all agents (when $\alpha > 0$), with $\hat{p} = \bar{p}$ if $\text{cov}(w^2, p) = 0$, and $\hat{p} > \bar{p}$ in the empirically relevant case where $\text{cov}(w^2, p) > 0$. Intuitively, if more productive agents live longer, the internal rate of return of the Bismarckian public annuity decreases below the Beveridgean return, and incentives to supply labor decrease as well.

We now study the joint determination by majority voting of the generosity and of the type of the pension system. We model the joint determination procedure first suggested independently by Kramer (1972) and Shepsle (1979). A policy pair (α, τ) is a Kramer-Shepsle equilibrium, denoted by (α^{KS}, τ^{KS}), if each element in the pair corresponds to a majority voting equilibrium given the value taken by the other element –i.e., if

---

[12]The main advantage of this specification is that $\hat{p}$ does not depend on τ or α.

$\tau^{KS} = \tau^V(\alpha^{KS})$ and $\alpha^{KS} = \alpha^V(\tau^{KS})$, where $\tau^V(\alpha)$ denotes the majority-chosen value of τ given α and where $\alpha^V(\tau)$ denotes the majority-chosen value of α given τ. We first study the choice of τ for a given value of α.

4 Voting over the size τ of the pension system

We first study the first-order condition for the individual most-preferred value of τ, before turning to the majority chosen level. Differentiating a voter's utility (5) with respect to τ while using (6) yields the following first-order condition

$$\gamma w^2(1 + \gamma\tau) + \frac{1 - \alpha}{(1 + r)\bar{p}}(1 + 2\gamma\tau)Ew^2 = 0. \tag{8}$$

This condition also corresponds to the maximization of the individual's lifetime income— i.e., $(1 - \tau)wz + b/(1 + r)$: in the absence of borrowing constraints, individuals choose τ to maximize their discounted lifetime income (with labor supply z optimally chosen) and s to reach their optimal allocation across periods.[13] The first term of (8) measures the marginal impact of increasing τ on the discounted Bismarckian (contributive) part of the pension benefit, net of the first period decrease in disposable income. It has the same sign as γ. The second term is the marginal variation in the non-contributive (Beveridgean) part of the pension benefit. Observe that $\hat{p}$ also affects the Beveridgean term (through γ) because it impacts the (dis)incentive to work of all agents, and hence the return of the non-contributive pension.

The FOC (8) shows that an individual's most preferred level of τ depends on

$$\theta = \frac{w^2}{Ew^2}.$$

Consequently, it is convenient to consider a type space over θ rather than w (with $\theta^{med} < E\theta = \bar{\theta} = 1$ since $w^{med} < \bar{w}$). We denote by $\tau^*(\theta, \alpha)$ individual θ's most-preferred tax rate for any given value of α.

We obtain the following proposition (proofs are relegated to the appendix).

---

[13]We thank Pascal Belan for pointing this to our attention. Observe that (since a similar result holds for the choice of α given τ), results throughout the paper are not affected by the precise shape of the utility function but hold for any increasing and concave utility function $u(.)$.

Proposition 1 *When agents vote over τ for a given α, we obtain that*

i) there is unanimity in favor of $\tau = 0$ if $\gamma < 0$ and $\alpha = 1$, and in favor of $\tau = 1$ if $\gamma > 0$;

ii) if $\gamma < 0$ and $\alpha < 1$, $\tau^*(\theta, \alpha)$ is positive for $\theta = 0$, decreases with θ and is zero above some productivity threshold level. Moreover, $\tau^V(\alpha) = \tau^*(\theta^{med}, \alpha)$.

If the Bismarckian internal rate of return $1/\hat{p}$ is large enough, compared with the private savings return, then the individual's discounted contributive part of the annuity increases more than his tax bill when the tax rate is increased (*i.e.*, the first term of (8) is positive for any τ when $\gamma > 0$), even though only a part α of the collective annuity is contributory. Moreover, increasing τ also affects the non-contributory part of the pension: recall that labor supply is increasing in τ when $\gamma > 0$, so that the second term in (8) is also positive for any value of τ. This is the incentive effect created by the Bismarckian part of the annuity on the return of the Beveridgean part. As the two terms of (8) are positive, all individuals favor $\tau = 1$.

If $\gamma < 0$ and $\alpha = 1$, the pension system is purely contributive with a return lower than the interest rate. All agents then prefer saving to any positive amount of Bismarckian collective annuity.

There are then two conditions to be satisfied for an individual to have an interior most-preferred size of the collective annuity program: that the system not be purely contributive ($\alpha < 1$) and that the net discounted individual marginal return of the collective annuity, γ, be negative. When both conditions are satisfied, individuals below a threshold productivity level face a trade-off between the redistribution embedded in the non-contributive component of the collective annuity and the low individual return of its contributive component. The poorest agent ($\theta = 0$) cares only about the non-contributive part and favors a positive value of τ, while individuals with larger productivities gain less from the redistributive component and most-prefer a lower size of the overall collective annuity program.

We now look at the choice of α given τ, and then move to the Kramer-Shepsle equilibrium.

5   Voting over the type α of the pension system

Assume for the time being that τ is given exogenously. We proceed as in the previous section, studying first the individually-optimal type of collective annuity, and then the majority-chosen one.

Differentiating the utility function (5) with respect to α while using (6) yields the following first-order condition

$$(1 + \gamma\tau)\left(\frac{\bar{p}}{\hat{p}}\theta - 1\right) + \frac{1 - \alpha}{\hat{p}(1 + r)}\tau = 0. \tag{9}$$

By the envelope theorem, the only first-order impact of α on the utility of voters is via the induced variation of the collective annuity. The first term in (9) measures the composition effect of α, increasing the share of the contributive component at the expense of the non-contributive one. Not surprisingly, it is positive for high productivity individuals $(\theta > \hat{p}/\bar{p})$ and negative otherwise. The second term represents the incentive impact of a higher α on the return of the Beveridgean component; it is always positive since a higher α increases labor supply.

Let $\alpha^*(\theta, \tau)$ denote individual θ's most-preferred value of α for any given τ. Combining the two effects we obtain that all individuals with $\theta > \hat{p}/\bar{p}$ most-prefer Bismarck $(\alpha = 1)$ whatever the value of $\tau > 0$. When interior, the most-preferred value of α is given by

$$\alpha^*(\theta, \tau) = \frac{\hat{p}}{2\hat{p} - \bar{p}\theta} + \frac{1 - \tau}{\tau}\frac{\hat{p}(1 + r)(\theta\bar{p} - \hat{p})}{2\hat{p} - \bar{p}\theta}, \tag{10}$$

where $\theta\bar{p} - \hat{p} < 0 < 2\hat{p} - \bar{p}\theta$ so that the first term is positive and the second term negative. The most-preferred value of α increases with θ: the composition effect of a larger value of α (the first term in (9)) increases with the individual's productivity, while its incentive effect on the Beveridgean tax base is independent of θ. Since preferences are single-peaked in α, we can apply the median voter theorem to obtain that the majority voting value of α is the one most-preferred by the median ability individual:

$$\alpha^V(\tau) = \alpha^*(\theta^{med}, \tau).$$

Equation (10) shows that the most-preferred value of α increases with τ, because a higher τ increases the labor supply distortions generated by the non-contributive part

of the annuity. We then obtain that the majority-chosen level of α increases with τ as well.

We now move to the joint determination by majority voting of size and type of the pension system.

6 The Kramer-Shepsle equilibrium

Since the median productivity individual is decisive when voting both over τ given α and over α given τ, we obtain that $\alpha^{KS} = \alpha^*(\theta^{med}, \tau^{KS})$ and $\tau^{KS} = \tau^*(\theta^{med}, \alpha^{KS})$. We then obtain the following proposition.

Proposition 2 *i) There is no Shepsle equilibrium with $0 < \alpha^{KS}, \tau^{KS} < 1$ — i.e., with interior solutions for both τ and α.*
ii) If $\hat{p}(1+r) < 1$, there is a unique Shepsle equilibrium, with $\tau^{KS} = 1$ and $\alpha^{KS} = min(\hat{p}/(2\hat{p} - \bar{p}\theta^{med}), 1) \in [1/2, 1]$.
iii) If $\hat{p}(1 + r) > 1$ and $\theta^{med} < 1/(\bar{p}(1 + r))$, there is a unique Shepsle equilibrium, with $\alpha^{KS} = 0$ and $\tau^{KS} = 1 - 1/(2 - \bar{p}(1+r)\theta^{med}) \in [0, 1/2]$.
iv) If $\hat{p}(1 + r) > 1$ and $\theta^{med} \geq 1/(\bar{p}(1 + r))$, there is a unique Shepsle equilibrium, with $\tau^{KS} = 0$.

These Kramer-Shepsle equilibria correspond to what is empirically observed: large Bismarckian systems and smaller Beveridgean ones. Moreover, Proposition 2 explains under what conditions (on the rate of return of both types of annuities and the median productivity level) each type of equilibrium arises.

Intuitively these results can be explained as follows. If $\hat{p}$ is small, for instance because richer people do not have a (much) larger life expectancy than poorer people (see (7)), the contributive annuity's intrinsic return is large and the Bismarckian system is very attractive with a large contribution rate. The reason why a purely Bismarckian system is not always chosen is that the decisive individual benefits from redistribution, if her productivity is low enough, and thus favors the introduction of some non-contributive part in the collective annuity. If $\hat{p}$ is large (with richer people living much longer

13

for instance), the contributive annuity has a low return and voters prefer a purely Beveridgean system provided that the decisive voters' productivity is not too large. The size of this Beveridgean annuity remains quite low because a purely Beveridgean system creates large distortions when its size increases. In other words, a large value of $\hat{p}$ discourages voters from introducing a contributive component into the collective annuity, which puts an upper bound on the size of the collective program because, in the absence of a Bismarckian component, the Beveridgean scheme generates large distortions (on labor supply). Finally, if the Bismarckian intrinsic return is low while the decisive voters' productivity is large, there is no political support for a collective annuity scheme, since a majority of voters would rather rely exclusively on private saving.

We now summarize how the Kramer-Shepsle equilibria are affected by variations in the joint distribution of longevity and income.

Proposition 3 *i) Increasing the covariance between income and life expectancy decreases the Bismarckian parameter of the unique Shepsle equilibrium when $\hat{p}(1+r) < 1$, and leads to a shift to a purely Beveridgean system when this threshold is crossed. Increasing the covariance has no direct impact on the equilibrium size of the (either mostly Bismarckian, or purely Beveridgean) equilibrium program.*
ii) Increasing average longevity when $cov(w^2, p) = 0$ has no impact on the size and type of a mostly Bismarckian system, but induces a shift to a purely Beveridgean system when a threshold is reached; it then decreases the equilibrium size of the pension program.

A larger $cov(w^2, p)$ decreases the return of the contributive annuity and thus moves the Kramer-Shepsle equilibrium away from Bismarck. Note that, even though the covariance does not affect the equilibrium size of the either mostly Bismarckian, or purely Beveridgean, program, a shift to pure Beveridge goes along with a discontinuous drop in the size of the program. In other words, increasing the correlation between income and longevity makes the equilibrium system more redistributive, without impacting directly its size (although a move to a purely Beveridgean system is accompanied by a lower

14

size).

As for average longevity, our model predicts that it has no impact on the (size or type) of the equilibrium program when longevity is not correlated with income, as long as average longevity is low enough for the equilibrium program to be mostly Bismarckian. With the intrinsic returns of contributive and non-contributive systems equal by assumption ($\hat{p} = \bar{p}$), the equilibrium Bismarckian parameter is only affected by the identity of the median voter, θ^{med} (with a larger equilibrium α^{KS} as θ^{med} increases and benefits less from the non-contributive annuity). As a threshold level is crossed, the equilibrium program becomes purely Beveridgean, with a size decreasing with average life expectancy (which decreases the return of the non-contributive annuity).

7 Numerical illustration

We perform two numerical calibrations, one for a Bismarckian country (France) and another for a Beveridgean country (UK) and show that our model predicts an equilibrium which is in line with the stylized facts observed in reality. That France is mostly Bismarckian while the UK is mostly Beveridgean can be seen from Conde-Ruiz and Profeta (2007, Table 2), who show that the correlation between wage and pension is 0.268 for the UK (the lowest value among the 11 European countries reported in the Table) and 0.658 for France (the third largest value after Spain and Portugal, and larger than the value for Germany).[14]

We first calibrate the distribution function of productivities. Since labor income is proportional to the square of productivity, we calibrate a lognormal distribution function of w^2. We use the two degrees of freedom available in a lognormal distribution to replicate two facts about the country income distribution function: (1) an average income (gross annual wage in 2012 USD PPPs) of 39 600 for France and 44 222 for the UK,[15] and (2) a Gini coefficient for the after-taxes and transfers distribution of 0.293 for France and 0.345 for the UK for the late 2000s.[16] We then obtain a median income

---

[14] See also Cremer *et al.* (2000b) for additional evidence.

[15] OECD data accessed at http://en.wikipedia.org/wiki/List_of_countries_by_average_wage#OECD.27s_statistics on April 30, 2014

[16] OECD data accessed at

of 34 380 for France and of 36 200 for the UK, corresponding to $\theta_{med} = 0.868$ for France and 0.819 for the UK.

We then calibrate $\bar{p}$, following the same procedure (and the same length periods) as in Leroux *et al.* (2011). We assume that the first period of life goes from 25 year old to 65, and the second from 65 to 105 year old. The value of $\bar{p}$ then corresponds to average life expectancy at age 65, expressed as a proportion of the length of the second period. The average life expectancy for men at age 65 in 2007 is 17.1 years in France (Cambois *et al.* (2008)) and 18 years in the UK.[17] Given that the second period lasts 40 years, this gives

$$\bar{p} = \frac{17.1}{40} = 0.4275$$

for France, and

$$\bar{p} = \frac{18}{40} = 0.45$$

for the UK.

We then calibrate the relationship between p and income w^2. Bommier et al. (2005) estimates the elasticity of life expectancy to income for men at age 65 in France at 0.18. We then use the CES functional form for this relationship,

$$p(w^2) = p_0(w^2)^{\eta}, \tag{11}$$

with $\eta = 0.18$. We then use the definition of average p,

$$\bar{p} = \int_0^\infty p_0(w^2)^{\eta} dF(w^2),$$

to obtain that $p_0 = 0.21$.

We proceed slightly differently for the UK since we have not found estimates of the elasticity of life expectancy to income. Rather, Attanasio and Emmerson (2001) report from their computations that "for men aged 65 moving from the 40th percentile to the 60th percentile in the wealth distribution (...) increases the probability of survival by

3.4 percentage points".[18] We use this information, with the same CES functional form (11), and the definition of $\bar{p}$ to calibrate equation (11) and we obtain that $\eta = 0.10$ and that $p_0 = 0.31$.

We then obtain that

$$\hat{p} = \frac{Epw^2}{Ew^2} = 0.45 > \bar{p} = 0.427$$

and $\text{corr}(p, w^2) = 0.952$ for France, with

$$\hat{p} = \frac{Epw^2}{Ew^2} = 0.469 > \bar{p} = 0.45$$

and $\text{corr}(p, w^2) = 0.92$ for the UK.

We assume a real interest rate of 3.6% per year, with the amount of saving divided into 40 equal deposits made at the end of each working year. If i is the (constant) interest rate (in real terms) paid each year, then we obtain the following formula (see Bronshtein et al., 2004):

$$1 + r = \frac{1}{40} \frac{(1+i)^{40} - 1}{i},$$

with $i = 3.6\%$, which yields $r = 1.16$.[19]

With these calibrations, we obtain that $\hat{p}(1 + r) = 0.97 < 1$ for France, consistent with the country being Bismarckian. More precisely, we obtain an equilibrium Bismarckian value of $\alpha = 0.85$.[20] As for the UK, we obtain that $\hat{p}(1+r) = 1.01 > 1$, consistent with the country being Beveridgean. More precisely, we obtain an equilibrium Beveridgean value of $\tau = 16.9\%$.[21] We can use formula (7) to better understand why the return of a contributive system, $\hat{p}$, is lower in the UK than in France: it is not due to the fact that the covariance between income and longevity is larger (on the contrary, it is smaller in the UK than in France, and moreover divided by a larger average wage in equation (7)) but rather because average life expectancy at 65 is larger in the UK.

---

[18] We model income rather than wealth in this paper, but if we assume that wealth is increasing in income, the position of an individual in both distributions is the same.

[19] As an alternative, we have assumed an average real interest rate of 2% per year, which we compound over 40 years to compute the value of $r = 1.21$. We obtain results (available upon request) that are very similar to those reported in the text.

[20] Koethenbuerger et al (2008) estimate that the Bismarckian factor, measured as 1-progressivity index, was 0.536 in France in 2005.

[21] Koethenbuerger et al (2008) estimate that the effective pension contribution rate was 23.7% in the UK in 2004.

We perform two comparative static exercises on this data. First, we examine how the political equilibrium is affected when we increase the correlation between income and longevity without changing the average life expectancy $\bar{p}$. This corresponds to the first part of Proposition 3. Since we know that Beveridgean countries are not affected by such a move, we restrict our attention to France. More precisely, we increase the value of the elasticity of life expectancy to income in France, η, from the calibrated value of 0.18 and we report the corresponding political equilibrium in Table 1.

η	α^{KS}	τ^{KS}
0.18	0.851	1
0.2	0.849	1
0.25	0.840	1
0.27	0.836	1
0.28	0	0.165

Table 1: Political equilibrium in France as a function of η

The results obtained are in line with Proposition 3 *(i)*: for small increases in η, the country remains Bismarckian, but the larger correlation between income and life expectancy reduces the equilibrium value of the Bismarckian parameter. When a threshold value is reached for η (0.28 in our calibration), the country becomes Beveridgean, with a much lower contribution rate than in the Bismarckian setting.

The second comparative static exercise consists in increasing the average life expectancy, while keeping all other parameters unchanged. Observe that this exercise goes beyond Proposition 3 *(ii)*. The proposition assumes away any correlation between income and lifetime expectancy, while the correlation is positive in both France and the UK in our calibrated example. Table 2 reports the political equilibrium in France as a function of life expectancy at age 65.

Life expect. at 65	$\bar{p}$	$\hat{p}$	α^{KS}	τ^{KS}
17.1	0.4275	0.450	0.851	1
17.5	0.4375	0.460	0.851	1
17.6	0.44	0.463	0	0.148
17.7	0.4425	0.466	0	0.144

Table 2: Political equilibrium in France as a function of average life expectancy at 65

Increasing life expectancy while maintaining all other assumptions (including $\eta = 0.18$) constant results in an increase in both $\bar{p}$ and $\hat{p}$. As long as the increase in life expectancy is small, the political equilibrium remains Bismarckian, and the Bismarckian factor remains unaffected (because the ratio of $\bar{p}$ and $\hat{p}$ remains constant with our calibration procedure when life expectancy at 65 is changed). Once a threshold life expectancy is reached (17.6 years with our calibrations), the equilibrium system becomes Beveridgean, with a much lower contribution rate. The Beveridgean equilibrium contribution rate decreases with life expectancy, as can be seen from Proposition 2 *(iii)*.

As for the UK, the equilibrium pension system of course remains Beveridgean as life expectancy at 65 increases. The only impact of this variation is to decrease the majority chosen contribution rate (from 16.9% in the benchmark case to 16.2% when life expectancy at 65 increases from 18 to 18.2 years, for instance).

8 Conclusion

This paper has developed a model where individuals differ in productivity and in longevity (modeled as the probability to be alive in the second period of their life). Individuals decide how much to work and to save when young. They retire and consume their saving plus any pension benefit when old. The public pension system takes the form of a collective annuity, with both a contributive (with the benefit based on the worker's own contribution) and a non-contributive (based on the average contribution in the economy) component. Voters choose both the size or generosity of the system (measured by the payroll tax rate) and its type or degree of income redistribution (measured by the relative size of the non-contributory component).

We study the joint determination of both the size and type of pension system and show that the unique (Kramer-Shesple) equilibrium is either a large (mainly) Bismarckian system, a smaller (purely) Beveridgean pension, or no public pension at all. This equilibrium pattern corresponds to what is observed in reality, with larger Bismarckian than Beveridgean systems. Moreover, our model predicts which type of equilibrium should emerge given the characteristics of the income and longevity distributions and of their correlation. We calibrate the model to the situation of France and the UK and

obtain a "large" Bismarckian system for France and a "small" Beveridgean system in the UK. Consequently, our model can contribute to explaining the stylized pattern of pension systems that is observed in reality.

Our analysis makes uses of two simplifying assumptions: we assume away borrowing constraints, so that saving can be negative, and we assume that the disutility from working can be expressed in consumption terms, independently of income. These two assumptions taken together simplify a lot the solving of the model, since preferences for collective annuities are made independent of individual longevity. The first of the two assumptions may strike the reader as especially strong, since it often (but not always) results in some voters favoring a confiscatory payroll tax (even in the presence of labor supply distortions from income taxation). The introduction of explicit borrowing constraints would complicate the model a lot without bringing much new insight. Specifically, rather than favoring confiscatory tax rates, individuals would favor the largest value of the payroll tax consistent with non-negative saving. This would prevent people from favoring extremely large values of the payroll tax, but it would not affect the qualitative results we have obtained in our simpler framework.

Finally, our political economy (median voter) model makes several assumptions that are common in that literature. These include for instance the hypotheses that voters are farsighted, selfish and vote directly on the characteristics of the pension systems, rather than for parties that propose pension policies bundled with many other policy dimensions. Additional insights would certainly be gained if these assumptions were lifted.

References

[1] Atkinson, A. and A. Sandmo, 1980, Welfare implications of the taxation of savings, *The Economic Journal 90*: 529–549.

[2] Attanasio O. and C. Emerson, 2001, Differential mortality in the UK, IFS working paper 01/16.

[3] Bommier A., T. Magnac, B. Rapoport and M. Roger, 2005, Droit à la retraite et mortalité différentielle, *Economie et Prévision* 168: 1–16.

[4] Borck R., 2007, On the choice of public pensions when income and life expectancy are correlated, *Journal of Public Economic Theory*, 9-4, 711-725.

[5] Bronshtein I, K Semendyayev, G Musiol and H Muehlig, 2004, *Handbook of Mathematics*, Springer.

[6] Brown J., T. Davidoff and P. Diamond, 2005, Annuities and individual welfare, *American Economic Review*, 95, 1573–90.

[7] Browning, E. K., 1975, Why the social insurance budget is too large in a democracy, *Economic Enquiry* 13: 373–388.

[8] Cambois E, Clavel A, Romieu I and JM Robine, 2008, Trends in disability free life expectancy at age 65 in France: consistent and divergent patterns according to the underlying disability measure, *European Journal of Ageing* 5: 285–298.

[9] Casamatta G., H. Cremer and P. Pestieau, 2000a, On the political sustainability of redistributive social insurance systems, *Journal of Public Economics* 75: 341-364.

[10] Casamatta G., H. Cremer and P. Pestieau, 2000b, The political economy of social security, *Scandinavian Journal of Economics* 102: 502–522.

[11] Conde-Ruiz I. and P. Profeta, 2007, The Redistributive Design of Social Security Systems, *The Economic Journal* 117 : 686–712.

[12] Coronado J.L., D. Fullerton and T. Glass, 2000, The progressivity of social security, NBER working paper 7520.

[13] Cremer H., J.-M. Lozachmeur and P. Pestieau, 2010, Collective Annuities and Redistribution, *Journal of Public Economic Theory*, 12(1), 23–41.

[14] De Donder Ph. and J. Hindriks, 2002, Voting over Social Security with Uncertain Lifetimes, in Institutional and Financial Incentives for Social Insurance, edited by

C. d'Aspremont, V. Ginsburgh, H. Sneessens and F. Spinnewyn, Kluwer Academic Publishers, Boston, 201–220.

[15] De Donder Ph., Le Breton M. and E. Peluso, 2012, On the (Sequential) Majority Choice of Public Good's Size and Location, *Social Choice and Welfare*, 39(2), 457–489.

[16] Deaton A. and C. Paxson, 1999, Mortality, education, income and income inequality among American cohorts, NBER working paper 7141.

[17] Disney R., 2004, Are Contributions to Public Pension Programmes a Tax on Employment, Economic Policy, 39, 267-311.

[18] Finkelstein, A. and J. Poterba, 2002, Selection effects in the UK individual annuities market, *Economic Journal*, 112, 28–50.

[19] Finkelstein, A. and J. Poterba, 2004, Adverse selection in insurance markets: policyholder evidence from the UK annuity market, *Journal of Political Economy*, 112, 183–208.

[20] Galasso V. and P. Profeta, 2002, The political economy of social security: a survey, *European Journal of Political Economy*, 18, 1–29.

[21] Galasso V. and P. Profeta, 2007 , How does ageing affect the welfare state?, *European Journal of Political Economy*, 23, 554–563.

[22] Gil J. and G. Lopez-Casasnovas, 1997, Life-time redistribution effects of the spaniels public pension system, University Pompeu Fabra working paper.

[23] Koethenbuerger M., Poutvaara P. and P. Profeta, 2008, Why are more redistributive social security systems smaller? A median voter approach, *Oxford Economic Papers,* 60(2), 275-292.

[24] Kramer, G.H., 1972, Sophisticated Voting over Multidimensional Choice Spaces, *Journal of Mathematical Sociology*, 2, 165–180.

[25] Leroux M.-L., 2010, The Political Economy of Social Security under Differential Mortality and Voluntary Retirement, *Journal of Public Economic Theory*, 12(1), 151–170.

[26] Leroux ML, P. Pestieau and G. Ponthière, Longevity, 2011, Genes and efforts: An optimal taxation approach to prevention, *Journal of Health Economics* 30: 62–76.

[27] Mitchell, O., 1996, Administrative Costs in Public and Private Retirement Systems, NBER Working Papers 5734.

[28] Reil-Held A., 2000, Einkommen und Sterblichkeit in Deutschland: Leben Reiche länger?, Discussion paper 00-14, SFB 504, Mannheim.

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

[30] Waldron, H., 2007, Trends in Mortality Differentials and Life Expectancy for Male Social Security–Covered Workers, by Average Relative Earnings, ORES Working Paper No. 108, U.S. Social Security Administration Office of Policy, Washington.

A Proof of Proposition 1

(i) The proof follows immediately from the FOC (8).

(ii) From (8), we obtain that

$$\tau^*(0,\alpha) = \frac{\hat{p}(1+r)}{2(\hat{p}(1+r)-\alpha)},$$

which is positive when $\gamma < 0$ and $\alpha < 1$.

From (8), we also observe that $\tau^*(\theta,\alpha)$ decreases with θ when γ defined by (4) is negative. Furthermore, there exists a threshold value of θ such that people above this threshold most prefer $\tau = 0$. We denote this threshold by

$$\tilde{\theta}(\alpha) = \frac{\hat{p}}{\bar{p}}\frac{1-\alpha}{\hat{p}(1+r)-\alpha}.$$

This threshold is positive when $\gamma < 0$. Note that the SOC need only be satisfied when τ^* is an interior solution and satisfies the FOC with equality. Straighforward computations show that this is the case when $\gamma < 0$, $\alpha < 0$ and $\theta < \tilde{\theta}(\alpha)$.

Finally, since the preferences are single-peaked over τ, we can apply the median voter theorem to obtain that $\tau^V(\alpha) = \tau^*(\theta^{med},\alpha)$.

B Proof of Proposition 2

i) This can be shown by solving simultaneously (9) and (8), the necessary conditions for an interior solution. This yields

$$\begin{aligned}
\alpha^{KS} &= 1,\\
\tau^{KS} &= 1 - \frac{1}{1-\hat{p}(1+r)},
\end{aligned}$$

which cannot be an equilibrium because it specifies a level $\tau^{KS} \notin [0,1]$ whatever the value of $\hat{p}(1+r)$.

ii) Assume that $\hat{p}(1+r) < 1$.

- If $\theta^{med} > \hat{p}/\bar{p}$, $\alpha^*(\theta^{med},\tau) = 1$ (all individuals with $\theta > \hat{p}/\bar{p}$ most prefer Bismarck whatever the value of τ) and $\tau^*(\theta^{med},1) = 1$, so that $\alpha^{KS} = 1$ and $\tau^{KS} = 1$.

- If $\theta^{med} < \hat{p}/\bar{p}$, there are two possible equilibria: $(\alpha = 0, \tau > 0)$ and $(\alpha > 0, \tau = 1)$.

  - First candidate for equilibrium: $(\alpha = 0, \tau > 0)$

    If $\alpha = 0$, solving (8) with $\theta^{med} = w_{med}^2/Ew^2$ gives the majority chosen interior value of τ. Observe that

    $$\theta^{med} < \frac{\hat{p}}{\bar{p}} < \frac{1}{\bar{p}(1+r)} = \tilde{\theta}(0),$$

    so that the majority chosen value of τ is positive. We then replace τ by this value in the first-order condition for α given by equation (9), and we solve it for $\alpha = 0$ to obtain

    $$\frac{\partial V(0, \tau, \theta)}{\partial \alpha} = \frac{-1 + \hat{p}(1+r)}{\hat{p}(1+r)(-2 + \bar{p}(1+r)\theta^{med})},$$

    which is positive because $\hat{p}(1+r) < 1$ and $\bar{p}(1+r)\theta^{med} < 1$, a contradiction with the assumption that $\alpha = 0$.

  - Second candidate for equilibrium: $(\alpha > 0, \tau = 1)$

    If $\tau = 1$, solving the first-order condition (10) for α gives

    $$\alpha^*(\theta^{med}, 1) = \frac{\hat{p}}{2\hat{p} - \bar{p}\theta^{med}} \in [1/2, 1]$$

    since $\theta^{med} < \hat{p}/\bar{p}$. We then replace α by this value in the first-order condition for τ, and we evaluate it at $\tau = 1$ to obtain

    $$\frac{\partial V(1, \tau, \theta)}{\partial \tau} = \frac{-1 + \hat{p}(1+r)}{\bar{p}(1+r)^2(-2\hat{p} + \bar{p}\theta_{med})}$$

    which is positive, confirming that $\tau^{KS} = 1$.

iii and iv) Assume that $\hat{p}(1 + r) > 1$ and that $\alpha^{KS} = 0$. From the FOC for τ measured at $\alpha = 0$, we infer that

$$\tau^{KS} = 1 - \frac{1}{2 - \bar{p}(1+r)\theta^{med}},$$

which decreases with $\bar{p}(1+r)\theta^{med}$ and is non-negative (and at most equal to $1/2$) provided that $\bar{p}(1+r)\theta^{med} < 1$. Evaluating the FOC with respect to α at this value of τ, we obtain

$$\frac{\partial V(\alpha,\tau,w,p)}{\partial \alpha} = \frac{1 - \hat{p}(1+t)}{\hat{p}(1+r)(2 - \bar{p}(1+r)\theta^{med})} < 0,$$

which proves iii). If $\bar{p}(1+r)\theta^{med} > 1$, we evaluate the FOC with respect to α for $\tau = 0$ to obtain

$$\begin{aligned}
\frac{\partial V(\alpha,\tau,w,p)}{\partial \alpha} &= \frac{\bar{p}}{\hat{p}}\theta^{med} - 1 \\
&< \frac{1}{\hat{p}(1+r)} - 1 \text{ since } \bar{p}(1+r)\theta^{med} > 1, \\
&< 0 \text{ since } \hat{p}(1+r) > 1,
\end{aligned}$$

which proves iv).