Life expectancy heterogeneity and the political support for collective annuities 1

Helmuth Cremer² Philippe De Donder³

8th April 2013

¹A previous version of this paper has circulated under the title of "Longevity, annuities and the political support for public pensions". We thank P. Belan, B. Heijdra and L. Meijdam for their comments. Financial support from Netspar through a research grant is gratefully acknowledged.

²Toulouse School of Economics (GREMAQ, IDEI and Institut universitaire de France).

³Toulouse 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, 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.

A third dimension of heterogeneity among voters may play a critical role in the determination of the pension system, namely 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 distribution (Waldron, 2007). It is thus important to assess the impact of such variations on pension programs.

The empirical consequences of life expectancy differences for actual pension systems have been extensively studied by e.g. 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 correlation between income and longevity reduces significantly the amount of redistribution across income levels.

However, little is known on the impact of the joint distribution of life expectancy and income on the majority-chosen characteristics of the public pension system. Tackling this question 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). As we now show, and 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 when individuals differ in longevity (as well as in productivity). Their approach is normative and based on a utilitarian social welfare function. 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. Their focus is on the majority chosen size of the pension system as a function of its (exogenous) redistributiveness. They show that the amount of distortions associated to the pension system need not decrease when the system is made less redistributive, because voters favor a larger pension size.¹

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*, which does not depend

¹Galasso 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.

on an individual's own survival probability.² This corresponds to the practice of public pensions, which redistribute ex ante from agents with short life expectancies to those with long life expectancies. This type of redistribution comes on top of the income redistribution that the pension system may achieve through the benefit formula. More precisely, the collective annuity received by any individual has both a contributive, or Bismarckian component (based on the individual's own tax payments when young) and a non-contributive, or Beveridgean component (based on the average contribution in the economy). 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, dubbed the Bismarckian parameter).

One can argue that the size of the pension program is more easily and more often modified than its type.³ We then first take a short run, or partial political equilibrium approach, where individuals vote over the contribution rate while keeping the Bismarckian parameter unchanged. We then study the long run, or general political equilibrium characterized by the joint determination of the size and type of the collective annuity program.

The partial equilibrium results constitute a necessary step to construct the general equilibrium allocation, but are also interesting by themselves. Both an increase in average life expectancy (assuming no correlation with income) and in the correlation between income and life expectancy weakly decrease the majority-chosen size of the pension system, whatever its type. This reinforces the observation made by the empirical literature that a positive correlation between income and life expectancy decreases the redistributiveness (across income levels) of collective annuity schemes of given size

²We 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.

³Casamatta et al (2000b, p.505) state that "Its redistributive character is, to a large extent, an integral part of the very definition of the [retirement] system itself. Regardless of whether systems are Bismarckian or Beveridgean, they imply specific institutional and administrative arrangements which cannot be overturned in the short run."

and type. Interestingly, pure pension programs react very differently to these variations. While the equilibrium size of a pure Bismarckian annuity either is not affected or collapses to zero once a threshold is reached, the equilibrium size of the pure Beveridgean annuity decreases smoothly with the average life expectancy and is not affected by the covariance between income and longevity. The intuition for this result is that the income redistribution enacted by the Beveridgean pension generates some political support for this program even when its average return (determined by the average life expectancy) falls below the saving rate of return, which is not true for the pure Bismarckian program.

Moving to 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 show the existence of a unique Kramer-Shepsle equilibrium, whose characteristics vary as follows. 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 smaller Beveridgean ones (see Conde-Ruiz and Profeta (2007)).⁴

We finally study how this long run/general 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 gen-

⁴In Conde-Ruiz and Profeta (2007), 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 arises.

erate a more redistributive equilibrium pension program, although sometimes at the expense of its size. They thus do not always decrease the equilibrium amount of income redistribution in collective annuities. This is in stark contrast with the results obtained either by the empirical literature (which takes the size and type of the pension program as exogenous) and by the partial/short run equilibrium approach and illustrates the importance of adopting such a long run/general equilibrium political approach.

The paper is structured as follows. Section 2 presents the model. Section 3 analyzes the private (labor supply and saving) decisions taken by individuals. Section 4 analyzes the short run/partial political equilibrium, while section 5 studies the long run/general political equilibrium. Section 6 concludes.

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]. We make no assumption at the outset on the correlation between the two characteristics w and p. The average productivity is denoted by \bar{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

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

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).

 $^{^5\}mathrm{We}$ will use without distinction the terms life expectancy, longevity and survival probability when referring to p.

Individuals take two private decisions, both in their first period of life: how much labor z to supply and how much to save, denoted by 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. Saving has a gross return of 1 + r, with r > 0 the exogenous world interest rate.⁷ 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 of life, individuals retire and consume their private saving and a public pension benefit b (if any), so that

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

The pension benefit is modelled as a pay-as-you go collective annuity, where 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,⁸ $1/\bar{p}$ is the internal rate of return of the non-contributory (Beveridgean) collective annuity while $1/\hat{p}$ with

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

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

⁷The assumption that agents can borrow against future income at this exogenous interest rate 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.

⁸Throughout the paper, Ef denotes $\int f(w, p)dH(w, p)$ for any function f. Similarly, cov(f, g) denotes E(fg) - E(f)E(g) for any functions f and g.

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 shortlived 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.⁹ 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

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

The FOC with respect to labor supply z is

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

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

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

where

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

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.

⁹Throughout 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.

The sign of γ depends on the comparison between the gross individual marginal return of pension, α/\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 .¹⁰ 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 borrowing constraints.

Labor supply increases with productivity w irrespective of the sign of γ . For future reference, note 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 $\operatorname{cov}(w, p) = 0$, and $\hat{p} > \bar{p}$ in the empirically relevant case where $\operatorname{cov}(w, p) > 0$. 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.

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

$$V(\alpha, \tau, w, p) = u \left[(1 - \tau)wz^* - s^* - h(z^*) \right] + \beta p u \left[(1 + r)s^* + \tau ((1 - \alpha)\frac{Ewz^*}{\bar{p}} + \alpha\frac{wz^*}{\hat{p}}) \right]$$
(4)

From now on, we assume that the disutility of labor is given by

$$h(z) = \frac{z^2}{2},\tag{5}$$

so that labor supply becomes

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

¹⁰When z is interpreted as the retirement age, a negative γ corresponds to the "implicit tax on continued activity" studied in the pension literature.

with¹¹

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

We now move to the political determination of the collective annuity. As explained in the introduction, we first take a short run/partial equilibrium approach where agents vote over τ while considering that the type α of program is exogenous.

4 The short run/partial political equilibrium

We first compute, and comment on, the first-order condition for the individual mostpreferred value of τ , before turning to the majority chosen level. We then perform the comparative statics analysis of this level with respect to changes in the joint distribution of life expectancy and income, and in the type of collective annuity program.

4.1 Individuals' most-preferred contribution rate

Differentiating a voter's utility (4) with respect to τ while using (5) yields the following first-order condition

$$\gamma w^2 (1 + \gamma \tau) + \frac{1 - \alpha}{(1 + r)\bar{p}} (1 + 2\gamma \tau) E w^2 = 0.$$
(6)

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.¹² The first term of (6) 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.

¹¹The main advantage of this specification is that \hat{p} does not depend on τ or α . Observe that results throughout the paper hold for any increasing and concave utility function u(.).

¹²We thank Pascal Belan for pointing this to our attention.

We now introduce the following notation. Since the individual productivity plays no role in the FOC (6), we summarize the type of an agent by

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

(with $\theta^{med} < E\theta = \overline{\theta} = 1$ since $w^{med} < \overline{w}$). We denote by $\tau^*(\theta, \alpha)$ individual θ 's mostpreferred tax rate for any given value of α , and by $\tau^V(\alpha)$ the majority chosen value of τ as a function of α .

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

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 (6) 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 (6) 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 (6) 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 turn to how the majority chosen value of τ is affected by variations in the joint distribution of life expectancies and income, and in the type of pension program.

4.2 Comparative statics analysis

We have seen at the end of section 3 that increasing the covariance between income and life expectancy while keeping the marginal distributions unchanged results in an increase in \hat{p} . To assess the impact of a larger covariance between income and longevity, we thus look at how results are affected when \hat{p} is varied while maintaining \bar{p} constant. By contrast, to isolate the impact of average longevity from the impact of covariance, we assume that w and p are not correlated so that $\bar{p} = \hat{p}$ when we study the impact of a larger average life expectancy, \bar{p} . We start with the following comparative statics analysis of the majority-chosen size of the collective annuity for any $0 \le \alpha \le 1$.

Proposition 2 Assume that agents vote over τ for a given α , and that $\gamma < 0$ (otherwise, $\tau^{V}(\alpha) = 1$ for all α).

i) Both a larger life expectancy and a larger correlation between income and life expectancy weakly reduce the majority-chosen size of the collective annuity.

ii) $\tau^V(\alpha)$ is monotone increasing in α if $1/\hat{p} > 1 + r$, but may not be monotone in α otherwise.

An increase in \hat{p} (for \bar{p} constant) decreases the return from both the contributive part of the annuity and the non-contributive part, because it induces agents to decrease the amount of labor they supply. Both effects decrease the most-preferred contribution rate of all agents, and thus the majority-chosen level as well. When cov(w, p) = 0 so that $\bar{p} = \hat{p}$, a larger average life expectancy \bar{p} further decreases the return of the noncontributive part, reinforcing the negative impact on $\tau^*(\theta, \alpha)$ and on $\tau^V(\alpha)$. Larger average lifetime expectancy and covariance between income and longevity, by decreasing the majority-chosen size of the pension system, thus reinforce the mechanical decrease in redistribution across income levels found, for given type and size of annuity program, by the empirical literature.

As for the impact of α on τ^V , increasing α has both an incentive and a composition effect. The incentive effect induces all agents to supply more labor (since they get to keep a larger fraction of the income they produce through a higher pension benefit). This in turn makes the Beveridgean part of the pension more attractive. The composition effect consists in increasing the relative share of the non-contributory pension. When $1/\hat{p} > 1 + r$, the non-contributory benefit is a better deal than saving, and all individuals react by increasing their most-preferred value of τ , as is illustrated on Figures 1 and 2.¹³ When $1/\hat{p} < 1 + r$, the trade-off between incentive and composition effects plays in different directions for low and high productivity agents. Agents with low productivities derive most of their pension benefit from the non-contributory part and are driven by the incentive effect to favor a larger value of τ , while agents with higher productivities pay more attention to the detrimental composition effect and favor a smaller public pension, as is illustrated on Figure 3. Using the same numerical example as in Figure 3 and varying the identity of the median (decisive) individual θ^{med} , Figure 4 shows that the majority chosen level of τ may be first increasing and then decreasing in α while Figure 5 illustrates the case where $\tau^{V}(\alpha)$ is monotone decreasing in α .

Insert Figures 1 to 5 around here

Since the impact of α on $\tau^{V}(\alpha)$ is not always monotone, the comparison of the short run equilibrium size of pure Beveridgean and pure Bismarckian programs is not straight-

¹³Figures 1 to 7 assume that r = 1 while $\bar{p} = 0.8$. Observe that there is no need to specify a utility function. Figures 1 and 2 assume that $\hat{p} = 1/3$ while Figures 3 to 7 assume that $\hat{p} = 3/4$. Figures 2, 4 and 7 assume that $\theta^{med} = 3/8$ while Figure 5 assumes $\theta^{med} = 1/2$.

forward. We now look more closely at their relative sizes and at their comparative statics with respect to the joint distribution of survival probabilities and income.

Proposition 3 Assume that agents vote over τ for given $\alpha \in \{0, 1\}$.

i) While the equilibrium size of the pure Beveridgean annuity is not affected by the correlation between income and life expectancy and decreases smoothly with the average life expectancy, the equilibrium size of the pure Bismarckian annuity remains unaffected by increases in average life expectancy or in the correlation between income and life expectancy as long as $\hat{p} < 1/(1+r)$, but collapses to zero when this threshold is reached. ii) $\tau^{V}(1) > \tau^{V}(0)$ if $1 + r < 1/\hat{p}$; iii) $\tau^{V}(0) > \tau^{V}(1)$ if $1/\hat{p} < 1 + r < 1/(\bar{p}\theta^{med})$; iv) $\tau^{V}(0) = \tau^{V}(1) = 0$ if $1 + r > \max(1/\hat{p}, 1/(\bar{p}\theta^{med}))$

The intuition for part i) is that the income redistribution embedded into the Beveridgean annuity creates heterogeneity in voters preferences, so that a lower return (because of a larger average life expectancy) slowly erodes its political support. Moreover, without any contributive annuity, the correlation between income and life expectancy has no influence on the equilibrium size of the program. By contrast, because of the absence of any income redistribution element, the political support for the pure noncontributive annuity disappears when its return becomes lower than the interest saving rate, either because of a large cov(w, p), or because of a large average life expectancy when cov(w, p) = 0.

The majority-chosen Bismarckian contribution rate, $\tau^V(1)$, is always larger than the Beveridgean one, $\tau^V(0)$, when the interest rate 1 + r is smaller than $1/\hat{p}$, since voters unanimously support a Bismarckian contribution rate of one, while the Beveridgean tax rate is always lower than one half for incentive reasons. A necessary condition for the majority-chosen Beveridgean size to be larger than the Bismarckian one is then that $1 + r > 1/\hat{p}$, in which case there is no support for a purely Bismarckian collective annuity. If the return from the contributive annuity is large enough compared to the interest rate, while the median productivity is not too large (i.e., if $1 + r < 1/\bar{p}\theta^{med}$), then the majority-chosen Beveridgean contribution rate is positive and thus larger than under Bismarck. If the returns of both systems are low enough compared to the interest rate, the majority chosen tax rate is zero for both schemes.

Finally, observe that a larger correlation between income and life expectancy enlarges the set of values of the interest rate r that are consistent with the majority-chosen Beveridgean contribution rate being larger than the Bismarckian one, since by increasing \hat{p} it enlarges the set of values of r for which $1/\hat{p} < 1 + r$, while the condition that $\bar{p}\theta^{med} < \hat{p}$ is satisfied whenever $cov(w, p) \ge 0$, since $\theta^{med} < 1$.

We now turn to the joint determination by majority voting of the generosity and of the type of the pension system.

5 The long run/general political equilibrium

We model the joint determination procedure first suggested independently by Kramer (1972) and Shesple (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 $\tau^{KS} = \tau^V(\alpha^{KS})$ and $\alpha^{KS} = \alpha^V(\tau^{KS})$, where $\alpha^V(\tau)$ denotes the majority-chosen value of α given τ .

We have seen that the individual with the median productivity is decisive in the choice of τ given α (Proposition 1). We now look at the choice of α given τ , and then move to the Kramer-Shepsle equilibrium.

5.1 Voting over the type of the pension system

We assume here that τ is not affected by the choice of α but is given exogenously. We proceed as in section 4, studying first the individually optimal type of collective annuity, and then the majority chosen one.

Differentiating the utility function (4) with respect to α while using (5) yields the following first-order condition

$$(1+\gamma\tau)\left(\frac{\bar{p}}{\hat{p}}\theta-1\right) + \frac{1-\alpha}{\hat{p}(1+r)}\tau = 0.$$
(7)

By the envelope theorem, the only impact of α on the utility of voters is through variations in the value of the collective annuity served in the second period. The first term in (7) measures the composition effect of α , increasing the share of the contributive component to the detriment of the non-contributive one. Intuitively, 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 and 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 . Intuitively, 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

$$\alpha^*(\theta,\tau) = \frac{\hat{p}}{2\hat{p} - \bar{p}\theta} + \frac{1 - \tau}{\tau} \frac{\hat{p}(1+r)(\theta\bar{p} - \hat{p})}{2\hat{p} - \bar{p}\theta},\tag{8}$$

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 (7)) increases with the individual's productivity, while its incentive effect on the Beveridgean tax base is independent of θ . Since preferences are concave 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 (8) 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. This means not only that the contributive part looks comparatively better, but also that a higher value of α is called for to decrease labor supply distortions. This second, incentive, impact of α explains why even low productivity individuals, who care mostly for the non-contributive part of the annuity, most-prefer a positive value of α when τ is large enough, as is illustrated on Figure 6.

Insert Figure 6

We then obtain that the majority-chosen level of α increases with τ as well, as is illustrated in Figure 7.

Insert Figure 7

5.2 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 4 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 $\theta^{med} \ge 1/(\bar{p}(1+r))$, 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 (see Conde-Ruiz and Profeta (2007) for a presentation of the evidence). Contrary to the short run/partial equilibrium analysis of section 4), which left open the possibility that the Beveridgean system could be larger than the Bismarckian system, we obtain with the long run/general equilibrium approach that Bismarckian programs are always larger than Beveridgean ones. Moreover, Proposition 4 explains under what conditions (on the rate of return of both types of annuities and the median productivity level) each type of equilibrium arises.

The intuition for these results runs as follows. If \hat{p} is small, meaning that the contributive annuity's intrinsic return is large, the Bismarckian system is very attractive and results in 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, 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 upperbound 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 5 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, 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, and then decreases the equilibrium size of the pension program.

A larger cov(w, 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 impact the equilibrium size of the either mostly Bismarckian, or purely Beveridgean, program, a shift to pure Beveridge is accompanied by a discontinuous drop in the size of the program. As mentioned in the introduction, the literature has focused on the effect of the positive correlation between income and longevity on the income redistributiveness of existing pension systems, and has, not surprisingly, concluded to a negative impact. These results have been reinforced in section 4 for the short run/partial equilibrium program, whose size decreases with the covariance. Moving to a long run/general equilibrium analysis brings new elements to the fore, since 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 size). This result then calls for a reevaluation of the empirical literature, which should go beyond a mechanical assessment of the impact of positive correlation between income and longevity on existing pension systems.

As for average longevity, our model predicts that it has no impact on the (size or type) of the long run/general equilibrium program when longevity is not correlated with income, as long as average longevity is low enough that the equilibrium program is 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 (since the latter decreases the return of the non-contributive annuity).

6 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 first proceed to a short run/partial equilibrium analysis where voters choose the

size of the program given its exogenous type, and we obtain (i) that a Bismarckian system may be larger than a Beveridgean system; (ii) that while the equilibrium size of a Beveridgean system decreases smoothly with average longevity, the support for a Bismarckian system drops discontinuously to zero when the distribution of longevity is such that its return falls below the interest rate; and (iii) that increasing the covariance between income and longevity leads to smaller pension systems (provided that there is some contributive component).

We then adopt a long run/general equilibrium approach where we endogenize both the size and type of pension system using majority voting, and we obtain 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. Also, a larger correlation between income and longevity makes the collective annuity more redistributive at equilibrium, although sometimes at the expense of its size. This calls into question the results obtained by the empirical literature, which shows that, for given size and type of the collective annuity, a larger correlation reduces income redistribution.

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 this simpler framework.

References

- Attanasio O. and C. Emerson, 2001, Differential mortality in the UK, IFS working paper 01/16.
- [2] Bommier A., T. Magnac, B. Rapoport and M. Roger, 2005, Droit à la retraite et mortalité différentielle, *Economie et Prévision* 168: 1-16.
- [3] 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.
- [4] Brown J., T. Davidoff and P. Diamond, 2005, Annuities and individual welfare, American Economic Review, 95, 1573–90.
- Browning, E. K., 1975, Why the social insurance budget is too large in a democracy, *Economic Enquiry* 13: 373-388.
- [6] Casamatta G., H. Cremer and P. Pestieau, 2000a, On the political sustainability of redistributive social insurance systems, *Journal of Public Economics* 75: 341-364.
- [7] Casamatta G., H. Cremer and P. Pestieau, 2000b, The political economy of social security, *Scandinavian Journal of Economics* 102: 502-522.
- [8] Conde-Ruiz I. and P. Profeta, 2007, The Redistributive Design of Social Security Systems, *The Economic Journal* 117: 686-712.
- [9] Coronado J.L., D. Fullerton and T. Glass, 2000, The progressivity of social security, NBER working paper 7520.
- [10] Cremer H., J.-M. Lozachmeur and P. Pestieau, 2010, Collective Annuities and Redistribution, *Journal of Public Economic Theory*, 12(1), 23-41.
- [11] Deaton A. and C. Paxson, 1999, Mortality, education, income and income inequality among American cohorts, NBER working paper 7141.
- [12] 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.

- [13] 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.
- [14] Finkelstein, A. and J. Poterba, 2002, Selection effects in the UK individual annuities market, *Economic Journal*, 112, 28–50.
- [15] 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.
- [16] Galasso V. and P. Profeta, 2002, The political economy of social security: a survey, European Journal of Political Economy, 18, 1-29.
- [17] Galasso V. and P. Profeta, 2007, How does ageing affect the welfare state?, European Journal of Political Economy, 23, 554-563.
- [18] Gil J. and G. Lopez-Casasnovas, 1997, Life-time redistribution effects of the spaniels public pension system, University Pompeu Fabra working paper.
- [19] Kramer, G.H., 1972, Sophisticated Voting over Multidimensional Choice Spaces, Journal of Mathematical Sociology, 2, 165–180.
- [20] 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.
- [21] Mitchell, O., 1996, Administrative Costs in Public and Private Retirement Systems, NBER Working Papers 5734.
- [22] Reil-Held A., 2000, Einkommen und Sterblichkeit in Deutschland: Leben Reiche länger?, Discussion paper 00-14, SFB 504, Mannheim.

- [23] Shepsle, K.A., 1979, Institutional Arrangements and Equilibrium in Multidimensional Voting Models, American Journal of Political Science, 23, 27–59.
- [24] 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 (6).

(ii) From (6), 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 (6), we also observe that $\tau^*(\theta, \alpha)$ decreases with θ when $\gamma < 0$ and that 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 $\gamma < 0$.

Finally, since the preferences are concave 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) Result obtained from the straightforward differentiation of (6) with respect to \hat{p} for given \bar{p} , and to \bar{p} when $\hat{p} = \bar{p}$, using the implicit function theorem.

ii) Tedious differentiation of (6) with respect to α , using the implicit function theorem, shows that $\partial \tau^{V}(\alpha)/\partial \alpha > 0$ when $1/\hat{p} > 1 + r$. Also, observe that both $\tau^{*}(0, \alpha)$ and $\tilde{\theta}(\alpha)$ (see proof of Proposition 1) increase with α when $1/\hat{p} > 1 + r$, and that $\tau^{*}(\theta, \alpha)$ is continuous in α and θ . When $1/\hat{p} < 1 + r$, $\tau^{*}(0, \alpha)$ increases with α , while $\tilde{\theta}(\alpha)$ decreases with α . This implies that, although the median productivity individual is always decisive when voting over τ for any given α , her most-preferred value of τ may not be monotone in α .

C Proof or Proposition 3

i) It is clear from (6) that $\tau^*(\theta, 0)$ is not affected by \hat{p} and that it decreases with \bar{p} . Also, by Proposition 1(i) we have that $\tau^V(1) = 1$ if $\gamma > 0$ and $\tau^V(1) = 0$ if $\gamma < 0$.

ii) $\tau^V(1) = 1$ if $1 + r < 1/\hat{p}$, while $\tau^V(0) < 1/2$ (since $\tau^*(0,0) = 1/2$ and $\tau^*(\theta,0)$ decreases with θ : see proof of Proposition 1)

iii) $\tau^{V}(1) = 0$ if $1 + r > 1/\hat{p}$, while $\tau^{V}(0) > 0$ if $\theta^{med} < \tilde{\theta}(0) = 1/\bar{p}(1+r)$ (see proof of Proposition 1)

iv) $\tau^{V}(1) = 0$ if $1 + r > 1/\hat{p}$, while $\tau^{V}(0) = 0$ if $\theta^{med} \ge \tilde{\theta}(0) = 1/\bar{p}(1+r)$ (see proof of Proposition 1).

D Proof of Proposition 4

i) This can be shown by solving simultaneously (7) and (6), the necessary conditions for an interior solution. This yields

$$\alpha^{KS} = 1,$$

 $\tau^{KS} = 1 - \frac{1}{1 - \hat{p}(1+r)}$

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 (6) 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 (7), 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 (8) 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

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

which proves iv).

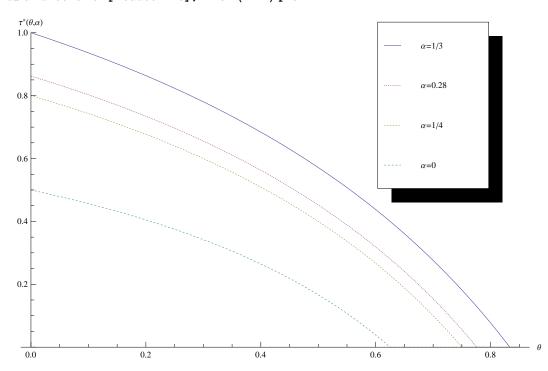


Figure 1 : Most - preferred contribution rate, as a function of productivity, when $(1+r)\ \hat{p}<1$

Figure 2 : Majority chosen value of the contribution rate, as a function of Bismarckian factor, when $(1+r)~\hat{p}<1$

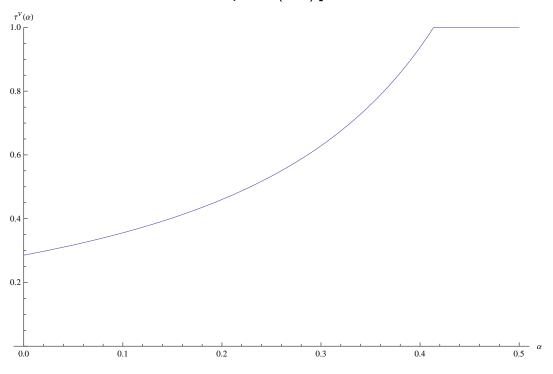


Figure 3 : Most - preferred contribution rate, as a function of productivity, when (1 + r) $\hat{p} > 1$

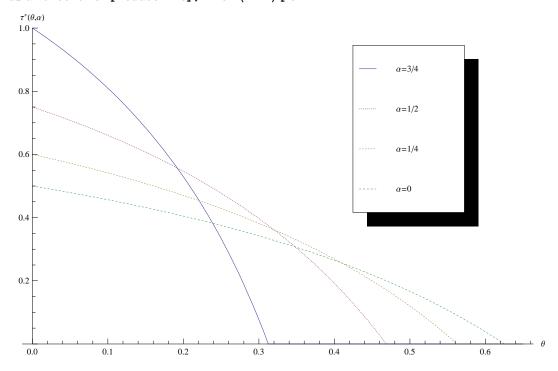
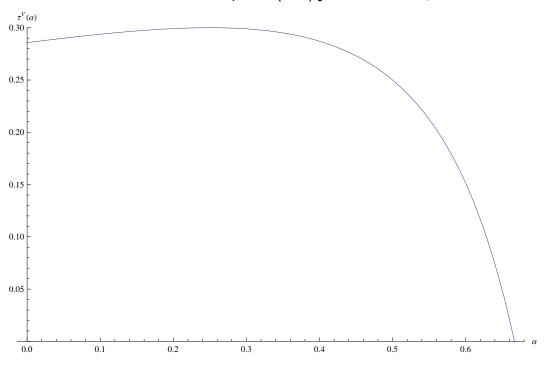


Figure 4 : Majority chosen value of the contribution rate, as a function of Bismarckian factor, when $(1 + r) \hat{p} > 1$ and $\theta med = 3 / 8$



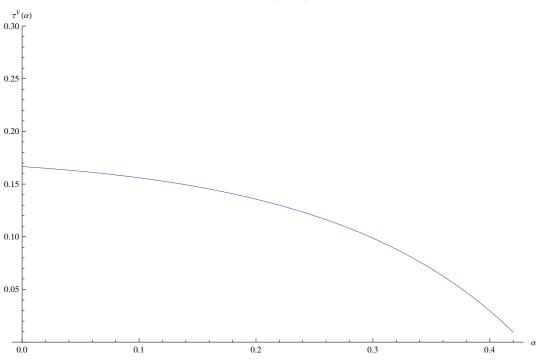
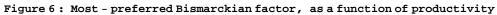
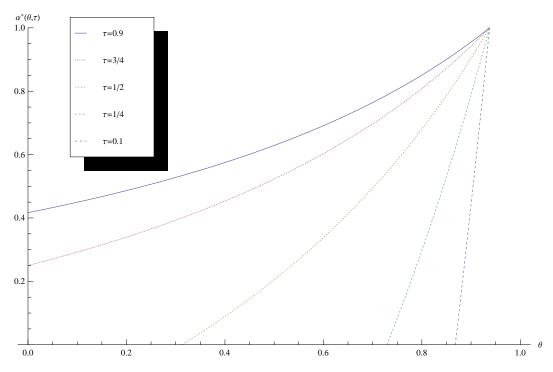


Figure 5 : Majority chosen value of the contribution rate, as a function of Bismarckian factor, when (1 + r) \hat{p} > 1 and θ med = 1 / 2





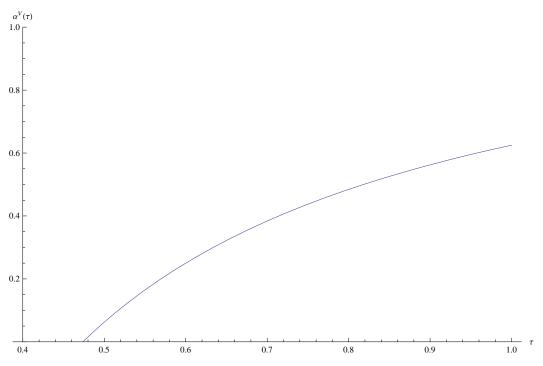


Figure 7 : Majority - chosen Bismarckian factor, as a function of contribution rate