Demography and intergenerational public transfers: a political economy approach

Loïc Batté*

Work in progress, please do not cite.
This version: March 25, 2013

Abstract

This paper aims at studying how demographic changes impact the public provision of social security and publicly-funded education, when these policies are determined as the outcome of a vote that involves both contributory and beneficiary generations. To this end, I set up an OLG model with production and intragenerational heterogeneity, in which these two intergenerational transfers are funded through taxes on the working generation. Individual preferences for taxation are aggregated through probabilistic voting. Contrary to previous studies in the field, which need to posit the existence of successive contracts between generations, the emergence and continuation of intergenerational tranfers only results here from the ability of each age group to tip the scales of redistribution to its side at each given period.

Under the assumption of non-strategic voting, and picking specific functional forms, I derive predictions on the impact of fertility and mortality rate changes on the level and composition of public spending, as well as on the potential of the economy to accumulate physical and human capital. In particular, population ageing points out to a rising tax burden in the future. I put these predictions into perspective with historical data on the secular evolution of public spending.

JEL Classification: E24, E62, H42, H52, H55, J11.

Keywords: Demographic change, education, pensions, political economy of taxation.

---

*PhD candidate, Toulouse School of Economics & GREMAQ. Contact: loic.batte@m4x.org. I wish to thank warmly Gilles Saint-Paul for his very valuable advice and support. This work also benefited a lot from comments by Georges Casamatta, Philippe de Donder, Eric Mengus, Christian Hellwig, Marti Mestieri, Juan Carlos Conesa, as well as seminar participants at Toulouse School of Economics, Universitat Autonoma de Barcelona and the ENTER Jamboree 2013 in Université Libre de Bruxelles. All remaining errors are of course mine.

1 Introduction

1.1 Motivation and aim of this paper

The recent debates on social security reforms in many advanced economies have shed new light on the relevance of population ageing on long-term growth or public finance sustainability. More generally, they compel us to think about the impact of public policies on intra- and intergenerational redistribution, human and physical capital accumulation, and productivity growth. Crucial to these issues is the way in which the size of government and the composition of public spending are collectively decided upon. If, as is the case in many OECD countries, public policies are determined through a voting process, we need to know both how individual preferences about the policies aggregate into collective choice, and how this collective choice affects private decisions. The relative political power of each generation then becomes determinant when policies that involve intergenerational tranfers are considered.

In this paper, I focus on two particular public policies with strong generational contents: education and retirement.[1] I seek to address specific questions that are relevant to practitioners of political economy, public economics or even growth theory. First, I tackle the problem of explaining the emergence and continuation of PAYGO pension systems, viewing it here primarily as an inter- and intra-generational redistribution tool. Similarly, I investigate the reasons that underlie the public funding of education, in a positive rather than a normative way. Second, I look at how demography (summed up in the model by two proxies of the fertility and mortality rates) can impact total public spending, as well as its allocation between pensions and education subsidies. Finally, I try to shed light on how demographic change may alter the processes of physical and human capital accumulation. In particular, I wish to separate the direct effects of demographic change on economic behaviour (e.g. the fact that increased longevity makes saving more desirable, or the capital dilution effect of increased fertility) from the effects that stem from modified political decisions on public spending. All these issues are adressed with a political economy lens, as public spending decisions are assumed to be the outcome of popular vote.

To these ends, I set up a model in which I divide the general population in three generations (young, middle-aged and old individuals) that coexist in a closed-economy, overlapping-generations setting with production, physical and human capital accumulation. The middle-aged (i.e. working-age or active) generation finances education subsidies[2] to the youngs, as well as pensions for the elderly, through two earmarked flat

[1] The focus on these two policies stems from the large share of GDP they represent in most advanced economies, and the fact that part of the literature considers them to exhibit complementarity features. I do not study health care in this paper: although the healthcare system operates some degree of redistribution towards the old, it also exerts a function of insurance within each generation.

[2] These education subsidies are also referred to as tuition help.

taxes levied on labor income. The political process, modelled as probabilistic voting on the two tax rates, determines the size of the public sector and the allocation of public spending. Because the education policy may have a redistributive impact within each generation, I assume some degree of intragenerational heterogeneity, alongside the intergenerational one. This takes the form of an initial private ability owned by young individuals, who can use it (and the tuition help) to acquire education.

I find that pensions and education subsidies come into existence only as the result of a direct conflict between beneficiaries of the two policies, and that there is no complementarity between the two types of intergenerational transfers. The relative importance of the two transfers comes from the political influence each generation exerts on policymakers, be it through the higher responsiveness of each member of the age groups to policy or through the sheer numbers of voters in each age group. Additionally, the model delivers the prediction that the ageing of population should translate into an increasing taxation burden: a decrease in fertility and an increase in longevity both force the total level of taxes on the working segment of the population upwards.

1.2 Literature review

A part of the literature on the political economy of pensions considers pensions as one of the two pillars of an intergenerational pact (the second pillar being education transfers), which has the ability to restore efficiency even when some of the generations lack commitment or cannot access capital markets.[3] For instance in Boldrin & Montes (2005), authors consider the system of transfers to be the outcome of majority voting: since recipients of the transfers fail to constitute a majority, these tranfers are sustained in equilibrium only through reputational considerations and specific out-of-equilibrium strategies that can be interpreted as a contract between all successive generations.[4]

This paper departs from this idea by considering a different mechanism for preference aggregation, namely probabilistic voting (as first introduced by Lindbeck & Weibull (1987)): the outcome of the voting process then maximises a weighted sum of the utility of voters, and ultimately depends on current state variables only, provided voters behave in a non-strategic way (i.e. if they fail to internalise the effect of current policy choices on future ones, even if future policy rates are correctly anticipated). In this setting, no reputation mechanism, commitment technology nor altruism is needed to explain the outcome of positive education and pension transfers. In this respect, my

[3]See for instance Boldrin & Montes (2005), or Rangel (2003) for a more general discussion of what is meant by intergenerational contracts.

[4]The set of strategies and beliefs sustaining the equilibrium amounts indeed to a contract that stipulates that each generation accepts to provide pensions and education to the dependent generations, expecting repayment when old, and where breaching the pact once results in the termination of all transfers for the whole history of the game. See also Belletini & Berti Ceroni (1999) for an example of this in a slightly different setting.

model is closest to Gonzalez-Eiras & Niepelt (2007) and Lancia (2010).

Another advantage of this formulation is that taxation levels (and hence the share of output devoted to public spending) depend on the parameters of the model in a richer way than if majority voting was used: indeed, parametrical changes fail to change the identity of the median voter (which is usually one individual inside a large age and/or income group) in the setups presented above, so that they influence policy only insofar as they have an impact on the median voter's preferred policy. On the opposite, parametrical changes in my model impact both the policy preferences of each agent and their weight in the decision process. This allows to analyse how the age structure of the population (summed up by two parameters that proxy the fertility and mortality rates) influences the size and composition of public spending. These results can then be interpreted in a historical perspective: we are able to confront the results of our comparative statics exercises to empirical estimates in Lindert (94), which show that the shift in the age structure explains to a large extent the development of social spending (encompassing social security and educational spending) between the end of the 19th century and 1930. Lindert (96) confirms this role for subsequent years.

Additionally, this model incorporates general equilibrium effects of taxation (as in Boldrin & Rustichini (2000)), since the factor prices (wages and interest rates here) are endogenously determined and depend on the levels of taxation which are voted for in the political process: in turn, they influence the private behaviour of agents (on capital accumulation and education decisions), and the tax rate preferences of each category of individuals. A last contribution of this model is to introduce intragenerational heterogeneity, which allows to investigate distributional issues as well as efficiency. Although the education subsidy is not a purely redistributive policy tool, it allows for a potential reduction of initial inequality within a given generation. To the best of our knowledge, introducing within-cohort heterogeneity has not been attempted before, in a context where education and pension policies are jointly determined by popular vote. One thing we do not consider in this version of the paper[5] is a (within-a-generation) redistributive pension system: it has been shown elsewhere (see Conde Ruiz & Galasso (2005)) that this brings additional support to pensions, from several age segments of the general population.

Outline The rest of the paper is organised as follows. Section 2 presents the structure of the model and its main assumptions. In section 3, I solve for the general dynamics of the model in two steps: first by considering private decisions under given tax rates, then by analysing political decision-making on the level of the two taxes, under additional assumptions. Section 4 derives predictions about the impact of demographic change on the levels of both tax rates, and the overall trajectory of the economy: these predictions

[5] A previous version of this paper allowed for redistributive pensions; further research is needed to solve the model with this specific amendment.

are confronted with available empirical evidence. Finally, section 5 discusses some of the assumptions and results, and concludes.

2 Model setup

2.1 OLG structure, and private decisions

2.1.1 Agents

Time is discrete and indexed by t. At every period, three generations (young, middle-aged workers, old pensioners) coexist. I denote N_{t2}^{t1} the number of individuals born at $t1$ alive at $t2$ [6]: N_t^t (respectively N_t^{t-1} and N_t^{t-2}) then denotes the number of young (resp. middle-aged workers and old people) alive at t. All individuals live for the first two periods. To proxy longevity, we assume that only an exogenous fraction $\delta_t \in (0;1]$ of the middle-aged workers of period $t-1$ reached old age in t. I also denote n_t the fertility rate of the middle-aged generation of t. It follows from the definitions of the fertility and death rates that in period t, the young generation is $(1+n_t)$ time as numerous as the middle-aged one, and the middle-aged workers are $(1+n_{t-1})/\delta_t$ time as numerous as the old.[7]

2.1.2 Private decisions

Utility Young agents born at t with innate ability ω_i seek to maximise their lifetime utility U with respect to consumption at all periods:

$$U_{i,t}^t = \ln(c_{i,t}^t) + \beta \ln(c_{i,t+1}^t) + \beta^2 \delta_{t+2} \ln(c_{i,t+2}^t) \tag{1}$$

Middle-aged agents also maximise a reduced version of this lifetime utility, their consumption when young being fixed once they reach working age. Here, β represents a one-period discount rate on future utility flows. The one-period utility is chosen to be logarithmic, at each period. The following paragraphs describe the actions available to agents, and the constraints they face, in each period.

Youngs Young individuals born at t (indexed by their type i) have an innate ability $\omega_i \sim F(\omega)$, which is independently and identically distributed (F denoting the cumulative distribution function of abilities). They do not consume in t. The only decision

[6]In order to be consistent, the superscript over any variable will henceforth refer to the date of birth of the individual, while the subscript will indicate the current period (possibly along another subscript designating the type of an individual, whenever relevant), unless otherwise mentioned.

[7]This last relationship stems from $N_t^{t-2} = \delta_t N_{t-1}^{t-2}$ and $N_t^{t-1} = N_{t-1}^{t-1} = (1+n_{t-1})N_{t-1}^{t-2}$.

young individuals need to take is whether to educate or not, education being a zero-one decision.

Along with help (or tuition services) provided by the state g_t, which is determined as described in subsection 2.3, the innate ability ω_i determines the actual welfare cost to educate as follows:

$$f(\omega_i; g_t) = \ln\left(1 + \frac{1}{g_t + \omega_i}\right) \tag{2}$$

It is noteworthy that the tuition help provided to each young g_t is unaffected by the individual educational decision, and does not depend on ability: I assume abilities to be unobservable, and that the education decision is taken after receiving the governmental help only, so that government has to provide universal education and to allocate educational resources equally between individuals. Educating is here a zero-one decision: either the individual pays the full cost of education f, and he will become an educated worker in the next period, or he does not pay any cost and remains unskilled. For analytical tractability reasons, the welfare cost is assumed to be logarithmic in the inverse of $(g_t + \omega_i)$, which I interpret as a (monetary-equivalent) cost to educate. I also wish to stress here the fact that ω_i, the innate ability, is infinitely substitutable with government help in education: this is naturally an extreme view, as individual ability is undeniably complementary to some degree to tuition services proposed by the state. This is nonetheless also warranted for analytical tractability.[8]

Middle-aged workers At this stage, every middle-aged worker (born at $t-1$) is either educated or non-educated, depending on his decision in the previous period. Every worker provides an inelastic labor supply and receives a wage. Let $w_{e,t}$ (resp. $w_{0,t}$) the wage received by a skilled (resp. unskilled) worker. A fraction $\tau_{e,t}$ of this wage is taxed by the government to provide a transfer to the young; similarly, a fraction $\tau_{p,t}$ is taxed to fund pension benefits for old individuals. The two tax rates are later on referred to as the education tax rate and the pension tax rate. The remaining disposable income $(1 - \tau_{e,t} - \tau_{p,t})w_{j,t}$ (where $j \in \{0, e\}$) can be spent for consumption or saved for next period, which individuals will reach with the exogenous probability $\delta_{t+1} < 1$ only. Let $c_{j,t}^{t-1}$ denote the consumption at time t of a middle-aged worker of type j and $s_{j,t}$ his savings, the private budget constraint for date t then writes:

$$(1 - \tau_{e,t} - \tau_{p,t})w_{j,t} = c_{j,t}^{t-1} + s_{j,t} \tag{3}$$

Pensioners The surviving old individuals receive pension benefits $p_{j,t}$ depending on their type $j \in \{0, e\}$, and the product of their savings from last period $s_{j,t-1}$. For tractability reasons, it is assumed that on top of their own savings, pensioners also

[8]The analytical form chosen here ensures that it is always costly to educate (namely $f(\omega, g) > 0,\ \forall g, \omega$), that education is infinitely costly when $\omega = g = 0$, and that an individual with infinite ability will incur no welfare cost to educate whatever the help provided (since $\lim_{\omega \to +\infty} f(\omega, g) = 0,\ \forall g$).

receive a share of the savings that were made by the deceased members of their own generation.[9] An interpretation is that middle-aged workers have access to perfectly competitive life insurance providers, which sell policies that collect one's savings in case of death and rebate the product of the savings of those who deceased in case of life. Between periods t and $(t+1)$, the insurance companies invest the premia they collected (i.e. the totality of savings) in the production process and get a gross rate of return R_t. Under a zero-profit condition for life insurance companies, each surviving individual of type j then receives an amount equal to $(1-\delta_t)/\delta_t \cdot R_t s_{j,t-1}$.[10]

The private budget constraint for date t of an old individual is then:

$$c_{j,t}^{t-2} = R_t(s_{j,t-1} + \frac{1-\delta_t}{\delta_t} s_{j,t-1}) + p_{j,t} = R_t \frac{s_{j,t-1}}{\delta_t} + p_{j,t} \tag{4}$$

2.2 Production

Production Y_t uses capital K_t and labor as inputs. Let $N_{e,t}$ (resp. $N_{0,t}$) denote the efficient labor supplied by educated (resp. non-educated) workers at time t. I assume that skilled labor is fully substitutable with unskilled labor, at a rate $\eta > 1$[11]. I call skill premium this exogenous η (as it will be clear that $w_{e,t} = \eta w_{0,t}$). It is assumed that capital fully depreciates after every period of production, so that the capital at each period is the sum of all savings made by workers at the previous period. The production function is Cobb-Douglas in capital and efficient labor, with $\alpha < 1$ the elasticity of production with respect to efficient labor:

$$Y_t = (N_{0,t} + \eta N_{e,t})^\alpha K_t^{1-\alpha}$$

Let $y_t = \frac{Y_t}{N_t^{t-1}}$ and $k_t = \frac{K_t}{N_t^{t-1}}$ the output and capital per worker ratios, and let μ_t the fraction of educated individuals in the workforce at time t (i.e. the fraction of middle-aged workers at time t that educated in period $t-1$). This last definition allows to reexpress each type of labor as follows: $N_{e,t} = \mu_t N_t^{t-1}$ and $N_{0,t} = (1-\mu_t) N_t^{t-1}$. Output per worker then writes:

$$y_t = (\mu_t \eta + 1 - \mu_t)^\alpha k_t^{1-\alpha} \tag{5}$$

Given this functional form for production, and assuming the labor and capital markets to operate under perfect competition, one gets:

$$w_{0,t} = \alpha(\mu_t \eta + 1 - \mu_t)^{\alpha-1} k_t^{1-\alpha} \tag{6}$$

[9] An alternative assumption would be to consider accidental bequests: this is unwarranted for our story here, and needlessly complicates the computations.

[10] This type of contract could be considered as an annuity contract, except there is only one period on which to pay the annuity in case of life.

[11] Empirical studies seem to indicate that skilled and unskilled labor are imperfect substitutes at best. Yet, the point of this assumption is to abstract from considerations of the relative supply of skilled versus unskilled labor when trying to determine the skill premium η, which is exogenous under the previous assumption.

$$w_{e,t} = \eta\alpha(\mu_t\eta + 1 - \mu_t)^{\alpha-1}k_t^{1-\alpha} = \eta w_{0,t} \tag{7}$$

$$R_t = (1-\alpha)(\mu_t\eta + 1 - \mu_t)^{\alpha}k_t^{-\alpha} \tag{8}$$

Equations (6) and (7) show why the paremeter η is called skill premium, as it represents the relative contribution to production and wage of a skilled worker as compared to an unskilled one.

2.3 The public sector: education subsidies and pensions

The public sector is designed to implement a level of taxes $\tau_{e,t}$ and $\tau_{p,t}$, levied on middle-aged workers' payroll income, to fund education subsidies to the youngs and pensions to the old. Subsection 2.4 presents the way in which these two linear taxes are decided at each period; for now, treating these tax levels as given is sufficient.

Education subsidies g_t are provided to every young individual, irrespectively of their final decision to educate or not. As stated earlier, I make the assumption that the government cannot observe the type of each individual (represented by ω_i), and that the provision of tuition help occurs before the decision to educate or not is taken. The public budget constraint of the education system is:

$$\tau_{e,t}N_t^{t-1}(\mu_t w_{e,t} + (1-\mu_t)w_{0,t}) = g_t N_t^t$$

$$\Leftrightarrow g_t = \frac{\tau_{e,t}}{1+n}(\mu_t w_{e,t} + (1-\mu_t)w_{0,t}) = \frac{\tau_{e,t}}{1+n}\alpha y_t \tag{9}$$

Contrary to the educational transfers, pension benefits $p_{j,t}$ are awarded to retirees according to their level of education. More precisely, the relative pension of an educated retiree vs. a non-educated one is constrained to match the relative wage (which makes this pension system non redistributive within a generation, or Bismarckian). The public budget constraint of the pension system is written as follows:

$$\mu_{t-1}p_{e,t} + (1-\mu_{t-1})p_{0,t} = \frac{\tau_{p,t}(1+n)}{\delta_t}(\mu_t w_{e,t} + (1-\mu_t)w_{0,t}) = \frac{\tau_{p,t}(1+n)}{\delta_t}\alpha y_t \tag{10}$$

with $p_{e,t} = \eta p_{0,t}$ to match relative wages.[12]

[12] See equation (7) for a derivation of relative wages.

2.4 Voting over education and pensions

In order to determine the level of the education and pension taxes, a way of aggregating the preferences of heterogeneous agents must be defined.

I define first the indirect utility of an individual as the maximal utility the individual can get from his private decisions, as a function of the two taxes. For instance, the indirect utility in period t of an educated middle-aged individual born in period $(t-1)$, denoted $V_{t,e}^{t-1}$, writes:

$$V_{t,e}^{t-1}(\tau_{e,t}, \tau_{p,t}, \mu_t, k_t) = \max \quad \begin{array}{l} \ln(c_{t,e}^{t-1}) + \beta\delta \ln(c_{t+1,e}^{t-1}) \\ \text{s.t. } (3), (4), (10) \end{array} \tag{11}$$

Then this indirect utility is the criterion by which each individual evaluates any proposed level of taxes.

In what follows, we consider the levels of the two taxes to be the outcome of a political competition involving two parties, which propose a policy platform which they are committed to implement if elected. Since the policy space $(\tau_{e,t}, \tau_{p,t})$ is bidimensional, a Condorcet winner may fail to exist, so that one has to use another political equilibrium concept to describe the outcome of such a process. Among the options suggested in the literature, I choose to implement here probabilistic voting as in Lindbeck & Weibull (1987).

In a two-party electoral competition setting, it is assumed that voters have intrinsic ideological preferences for parties, on top of the preferences for their policy platform. In any constituency with shared characteristics (here, in any age-income group), some of the voters have strong ideological preference for a party, so that they are inclined to support it unless it proposes a very detrimental platform, while others have low ideological attachment and are assumed to be very responsive to a change of platform. Consequently, a policy platform shift by one party induces a gradual, smooth change in the proportion of voters supporting it in every age and income-related constituency. As constituencies may be more or less responsive to the proposed policy, a politician seeking election will be more or less willing to cater to its special interests. Because politicians are assumed to propose the platform which attracts the highest number of voters, both parties seek to balance the interests of the different groups. These considerations are translated analytically by attributing political weights to each constituency (the weight being higher for constituencies more attached to the policy platform and less concerned about ideology). Any office-motivated politician will then propose the tax policy platform that maximises the political-power weighted sum of each constituency's indirect utility under the tax policy.

In the present case, I choose to attribute age-related political weights only, avoiding to discriminate constituencies by their level of education (hence of income as well). Denoting ρ (respectively χ and ψ) the per individual political power of the young (respectively middle-aged and old) generation, the politicians propose the policy platform

(τ_e, τ_p) that maximises:

$$W_t = \rho(1+n)\left(\int_{i=0}^{\infty} V_{t,i}^t dF(i)\right) + \chi(\mu_t V_{t,e}^{t-1} + (1-\mu_t)V_{t,0}^{t-1}) \\ + \frac{\psi\delta_t}{1+n_{t-1}}(\mu_{t-1}V_{t,e}^{t-2} + (1-\mu_{t-1})V_{t,0}^{t-2}) \tag{12}$$

Probabilistic voting was chosen in this context because of its main property of taking into account the utility of each individual in the population. Indeed, under the median-voter setting, segments of the population which have preferences far away from those of the median voter fail to be represented, and may end up with a very low indirect utility under the chosen policy. While it might be argued that this is a realistic feature of general elections in a democracy, it does not fit well in a setting where large fractions of the voters share a special interest and might get their voice heard in the democratic process even if they do not form a majority. For instance, in this model, because the youngs (respectively the elderly) are direct recipients of educational subsidies (respectively pensions) and enjoy a direct, sizeable benefit from a positive τ_e (respectively τ_p), any configuration of the electorate where the median voter would choose a zero tax rate would substantively damage the utilities of the group. Hence, it seems plausible to include every individual's indirect utility into the politician's objective function at the time of policy proposal.

Moreover, as the economic conditions change, the identity of the median voter will change in continuous distributions of voters, but not necessarily in settings where the general population is divided into large homogenous categories. In our model, while the young generation is indeed distributed continuously with respect to private abilities, the middle-aged and old generations are divided into two broad categories inside which individuals are homogenous. Hence, provided that the median voter (if it exists at all in two dimensions, else in one of the two dimensions) does not belong to the young generation under a set of parameters of the model, one cannot expect a slight change in a parameter to change the identity of the median voter (while it may still change the median voter's preferred policy). This is an undesirable feature if, as in this paper one wishes to study the influence of some parameters on equilibrium (for instance, the impact of the fertility rate n in our case)[13].

2.5 State variables, timing of decisions and dynamics

Following the desciption of this economy, it is straightforward to see that the state of the economy at t is entirely described by the two state variables μ_t, the fraction of educated workers, and k_t, the capital per head. The dynamics of this economy is then

[13]For a more detailed discussion of probabilistic voting, I refer the reader to Persson & Tabellini (2000).

twofold: the level of taxes is first determined by a vote at the beginning of each period, given the state of the economy. Then given the tax rates that have been voted upon, the value of the state variables evolves according to the education decisions of young individuals and the savings decisions of middle-aged workers (who take into account their expectations of the future tax rates as well). We turn to analysing this dynamics in the next section.

3 The politico-economic dynamics

In this section, I study how the dynamics of the economy unfolds, using backward induction. I first derive the behaviour of private agents under given taxes, which yields a law of motion for the state variables μ_t and k_t when a specific tax policy τ_t is chosen. Thereafter, under several additional assumptions, I am able to characterise the policy outcome for a given state of the economy. The law of motion of the state variables, together with the implicit policy rule, yields the entire politico-economic dynamics of the economy, and allows to characterise its steady state.

3.1 Dynamics of the economy under given tax policies

In this subsection, I derive explicit solutions for the private decisions made by individuals given the state of the economy, the current value of the tax rates $\tau_{e,t}$ and $\tau_{p,t}$, and the expected values of tax rates in the next period, $\tau_{e,t+1}$ and $\tau_{p,t+1}$. It is noteworthy that the agents are assumed to take their private decisions under rational expectations: since there is no uncertainty in this model, the expected values of future tax rates are equal to the true ones.

3.1.1 Optimizing over educational choice and savings

The educational choice made by a young individual born at date t is solved backward, first obtaining the savings decision made by middle-aged workers of education $j \in \{0, e\}$ at date $t+1$.

Savings decision Type j worker seeks to maximise his second-period utility $V_{j,t}^{t-1} = \ln(c_{j,t}^{t-1}) + \beta\delta_{t+1}\ln(c_{t+1}^{t-1})$ subject to (3), (4). The FOC on savings $s_{j,t}$ yields:

$$s_{j,t} = \frac{\beta\delta_{t+1}}{1+\beta\delta_{t+1}}(1-\tau_{e,t}-\tau_{p,t})w_{j,t} - \frac{\delta_{t+1}p_{j,t+1}/R_{t+1}}{1+\beta\delta_{t+1}} \tag{13}$$

One then gets the level of consumptions at t and $t+1$ of the middle-aged individual of education j born at $t-1$:

$$c_{j,t}^{t} = \frac{1}{1+\beta\delta_{t+1}}(1-\tau_{e,t}-\tau_{p,t})w_{j,t} + \frac{\delta_{t+1}p_{j,t+1}/R_{t+1}}{1+\beta\delta_{t+1}} \tag{14}$$

$$c_{j,t+1}^{t} = \frac{\beta R_{t+1}}{1+\beta\delta_{t+1}}(1-\tau_{e,t}-\tau_{p,t})w_{j,t} + \frac{\beta\delta_{t+1}p_{j,t+1}}{1+\beta\delta_{t+1}} \tag{15}$$

It is noteworthy that, because instantaneous utility was set to be logarithmic, the optimal consumptions at both periods are quite simple expressions of the present value of expected net lifetime income $\gamma_{j,t} = (1-\tau_{e,t}-\tau_{p,t})w_{j,t} + \delta_{t+1}p_{j,t+1}/R_{t+1}$. In turn, $\gamma_{j,t}$ depends on the future levels of pensions, which depends itself on the expected value of $\tau_{p,t+1}$. As we will see in the next paragraph, the educational choice made by young individuals born at t is entirely determined by the comparison between the expected lifetime incomes of educated ($\gamma_{e,t+1}$) versus non-educated ($\gamma_{0,t+1}$) workers.

Education decision In order to determine which individuals will choose to educate, one needs to compare the lifetime utilities of uneducated and educated individuals. From (1), (2) and (13), one gets the lifetime utility of a young born at t with initial ability ω_i who chooses to get an education:

$$V(\omega_i, e, t) = \ln(1+\frac{1}{\omega_i+g_t}) + \beta\ln(\frac{1}{1+\beta\delta_{t+2}}\gamma_{e,t+1}) + \beta^2\delta_{t+2}\ln(\frac{\beta R_{t+2}}{1+\beta\delta_{t+2}}\gamma_{e,t+1}))$$

Similarly, if the young does not educate, his lifetime utility will be:

$$V(\omega_i, 0, t) = \beta\ln(\frac{1}{1+\beta\delta_{t+2}}\gamma_{0,t+1}) + \beta^2\delta_{t+2}\ln(\frac{\beta R_{t+2}}{1+\beta\delta_{t+2}}\gamma_{0,t+1}))$$

Let now $\hat{\omega}_t$ the level of private ability at which the young individual is indifferent between educating or not. One gets from the preceding equations and (6), (7), (9) and (10):

$$V(\hat{\omega}_t, e, t) = V(\hat{\omega}_t, 0, t)$$
$$\Leftrightarrow \ln\left(1+\frac{1}{\hat{\omega}_t+g_t}\right) = \beta(1+\beta\delta_{t+2})\ln\frac{\gamma_{e,t+1}}{\gamma_{0,t+1}}$$

where

$$\frac{\gamma_{e,t+1}}{\gamma_{0,t+1}} = \frac{w_{e,t+1}+\delta_{t+2}p_{e,t+1}/R_{t+1}}{w_{0,t+1}+\delta_{t+2}p_{0,t+1}/R_{t+1}} = \eta$$

is the ratio of the present value of the expected net lifetime income of an educated worker over that of an unskilled worker, at next period. Since the pension system is non redistributive, this ratio is simply equal to the skill premium η.

Using (9) to substitute for g_t, I then get the following expression for $\hat{\omega}_t$:

$$\hat{\omega}_t = \frac{1}{\eta^{\beta(1+\beta\delta_{t+2})} - 1} - \frac{\tau_{e,t}\alpha}{1+n}(\mu_t\eta + 1 - \mu_t)^{\alpha} k_t^{1-\alpha} \tag{16}$$

If $\omega_i < \hat{\omega}_t$, the individual's ability is too small to allow for a profitable investment in education, so the young does not educate. On the opposite, for $\omega_i > \hat{\omega}_t$, the investment in education is profitable. One can check that $\hat{\omega}_t$ decreases with η or $\tau_{e,t}$, meaning that education is easier to achieve when the skill premium or the educational tax rate is higher.

3.1.2 Dynamics of the state variables

The dynamics of the whole economy is entirely described by the dynamics of the two state variables μ_t and k_t.

Capital accumulation As stated in subsection 2.2, capital fully depreciates after each period of production[14], so that the capital in period $t+1$ is constituted by savings of period t workers:

$$k_{t+1} = \frac{1}{1+n_t}(\mu_t s_{e,t} + (1-\mu_t)s_{0,t})$$

Using (7), (8), (10), (13) and simplifying:

$$k_{t+1} = \frac{\beta\delta_{t+1}\alpha(1-\alpha)(1-\tau_{e,t}-\tau_{p,t})}{(1+n_t)[(1+\beta\delta_{t+1})(1-\alpha)+\alpha\tau_{p,t+1}]} y_t \tag{17}$$

One can observe that for given values of μ_t and k_t, any increase in τ_e or τ_p has the effect of decreasing k_{t+1}. Future expected taxation $\tau_{p,t+1}$ also decreases capital accumulation, since it increases the level of future pension benefits received, thus decreasing the workers' incentives to save. Current taxation is detrimental to capital accumulation only insofar as it reduces the disposable income of workers at t.

[14]Which seems a reasonable assumption, keeping in mind the fact that each period represents a generation, or about 25 years.

Evolution of the share of educated workers As stated in subsection 3.1.1, young individuals born at t educate if and only if their private ability ω_i is above the cutoff ability $\hat{\omega}$. Recalling that I defined F as the cumulative distribution function of private abilities, one gets:

$$\mu_{t+1} = 1 - F(\hat{\omega}_t) \tag{18}$$

3.2 Additional assumptions

In order to solve the model further, I make several assumptions on the distribution of abilities of the young, their participation in the political process, and the way in which expectations of future policy impact the current policy choice at date t.

3.2.1 Time-path of demography parameters

In order to perform simple analyses of the impact of demography on the policy choices, I will analyse only one type of path for the fertility and longevity parameters. Namely, I will consider constant fertility and longevity from the start of time ($t = 0$) up to some period t', after which one of the two parameters is allowed to change permanently to a new, constant value. It is assumed that agents do not anticipate the shock at all, but that they correctly perceive it to be permanent once it has occured. The aim of this exercise is to characterise the response to an unexpected permanent demographic shock, while keeping things reasonably simple to analyse.

Assumption 1 *The time path of demography parameters is as follows:*

$$n_t = n,\ \ \delta_t = \delta,\ \ \forall t < t'$$

$$n_t = n' \ \ or\ \delta_t = \delta',\ \ \forall t \geq t'$$

3.2.2 Distribution of abilities

For analytical tractability reasons, and additionally to the setup presented in section 2, I look at a particular case where abilities are distributed uniformly.

Assumption 2 *Abilities are distributed uniformly over the interval* $[0; \bar{\omega}]$*, where:*

$$\bar{\omega} = \frac{1}{\eta^{\beta(1+\beta\delta)} - 1}$$

Under this assumption, it obtains that:

$$\hat{\omega}_t = \bar{\omega} - \frac{\tau_{e,t}\alpha}{1+n_t}(\mu_t\eta + 1 - \mu_t)^\alpha k_t^{1-\alpha}$$

Using the formula of the pdf of a uniform distribution:

$$\mu_{t+1} = 1 - F(\hat{\omega}_t) = 1 - \frac{\hat{\omega}_t}{\bar{\omega}} = \frac{\tau_{e,t}\alpha y_t}{\bar{\omega}(1+n_t)} \tag{19}$$

Then the relationship between the policy rate $\tau_{e,t}$ and the future value of the state variable μ_{t+1} is linear, for a given level of production. Moreover, $\mu_{t+1} = 0$ when $\tau_e = 0$, and the marginal effect of τ_e on μ_{t+1} is strictly positive at $\tau_e = 0$: indeed, at $\tau_e = 0$, the young with the highest ability $\bar{\omega}$ is indifferent between educating or not, so that a marginal increase in τ_e has a positive impact on education. Such features avoid obtaining a discontinuous policy function around values of k_t and μ_t where optimal policy involves no education, which highly simplifies subsequent analysis of the impact of the education tax rate on the future state of the economy.

3.2.3 Voting rights

In the rest of this paper, young individuals are assumed to hold no political power.

Assumption 3 *Young individuals hold no political power: $\rho = 0$.*

This is largely a simplifying assumption, yet it can be supported by the fact that political rights are tied with electoral majority in most countries, entailing that the youngest fraction of the population has no impact on electoral outcomes. Notice that assuming a strictly positive value of ρ would only tip further the scales towards education subsidies, since the young are the direct beneficiaries of education transfers.[15] Assumption 3 then merely underestimates the scope for public finance of education, without changing the essence of the results.

In this setting, ψ and χ then become redundant parameters (in the sense that collective preferences for taxation are left unchanged if both ψ and χ are multiplied by the same constant), so we assume $\chi = 1$ without loss of generality. ψ then becomes the relative political power of one pensioner compared to one worker: although solid empirical evidence on the value of this parameter is hard to obtain, higher electoral turnout of elderly people compared to the general population, as well as the relative salience of pension policy on pensioners' welfare, point out to a value of ψ higher than one.[16]

[15] Of course, the young want to increase τ_e only up to the point where taxation is not too detrimental to capital accumulation: I believe the direct effect of an increase in τ_e nonetheless dominates this latter effect, around the values of τ_e that are obtained when ρ is set to zero.

[16] Notice that ψ is independent from the relative mass of pensioners and workers in the population, even though sheer numbers in a generation also have an impact on the policy outcome.

3.2.4 Expectations of future policies

As may be seen from subsection 3.1.1, private decisions about savings and education are dependent on the expected future state of the economy (and expected future policy choices). For instance, the lifetime income (in net present value) of a middle-aged individual at t depends not only on his wage, but also on the net present value of the future pension he will receive at $(t+1)$, which in turn depends on the level of capital k_{t+1} and the share of educated workers μ_{t+1} of next period. But this future pension also depends on the future contribution rate $\tau_{p,t+1}$, which will be decided upon in next period. Given that the savings decision at t is based on the expected net present value of lifetime income, it is dependent on the anticipation the agent makes about $\tau_{p,t+1}$. More generally, all private decisions depend on the anticipated future policy choices. Therefore, the current policy choices $(\tau_{e,t}, \tau_{p,t})$ potentially depend as well on the anticipated future policy choices of next period $(\widetilde{\tau_{e,t+1}}, \widetilde{\tau_{p,t+1}})$, since these policy choices are made to maximise indirect utilities that are themselves the result of private decisions.

In this paper, I assume that agents have rational expectations of future policies, but do not vote strategically: when voting on their preferred policy for today, they take the (correctly anticipated) value of future ones as given and ignore the impact their choice will have on the future policy choice.

Assumption 4 *All agents are assumed to vote in a non-strategic way:*

$$(\widetilde{\tau_{e,t+1}}, \widetilde{\tau_{p,t+1}}) = (\tau_{e,t+1}, \tau_{p,t+1})$$

$$\frac{\partial W_t}{\partial \tau_{e,t+1}} = \frac{\partial W_t}{\partial \tau_{p,t+1}} = 0$$

Another assumption about the formation of expectations is to consider that agents vote in a sophisticated way. In this case, agents not only know the level of future tax rates but also the true policy rule of the economy: this knowledge allows to compute the impact of a change of policy now on future policies, which are not considered as given anymore at the time of the current political decision. The main motivation for considering the agents to vote non-strategically is the fact that each agent is atomistic and thus cannot change the outcome of the vote by himself; then future policy itself can be considered to be invariant to the current individual policy choice.

3.3 Taxation choices

Under assumption 3, the two policy rates now maximise the following welfare function:

$$W_t = (\mu_t V_{e,t}^{t-1} + (1-\mu_t) V_{0,t}^{t-1}) + \frac{\psi \delta_t}{1+n_{t-1}} (\mu_{t-1} V_{e,t}^{t-2} + (1-\mu_{t-1}) V_{0,t}^{t-2})$$

The choice set for policies is defined as:

$$S = \left\{ (\tau_e; \tau_p) \in [0;1]^2 \mid 0 \leq \tau_e + \tau_p \leq 1, \tau_e \leq \tau_{e,max}(y_t) = \frac{\bar{\omega}(1+n_{t-1})}{\alpha y_t} \right\}$$

where the last condition on τ_e is used to rule out from the start cases in which τ_e is so large that education subsidies divert more resources than what is needed to ensure $\mu_{t+1} = 1$ (which is a pure waste). S is a compact and convex subset of $\mathbb{R}^2$.

Under assumptions 2 and 4, appendix A shows that the partial derivatives of W_t on S are as follows:

$$\frac{\partial W_t}{\partial \tau_{e,t}} = \beta\alpha\delta_{t+1}\frac{(\eta-1)\alpha y_t}{\bar{\omega}(1+n_t)+\tau_{e,t}(\eta-1)\alpha y_t} - \frac{1+\beta\delta_{t+1}-\beta\alpha\delta_{t+1}}{1-\tau_{e,t}-\tau_{p,t}} \tag{20}$$

$$\frac{\partial W_t}{\partial \tau_{p,t}} = \frac{\psi\delta_t\alpha}{(1+n_{t-1})(1-\alpha+\alpha\tau_{p,t})} - \frac{1+\beta\delta_{t+1}-\beta\alpha\delta_{t+1}}{1-\tau_{e,t}-\tau_{p,t}} \tag{21}$$

Because future levels of taxation are taken as given in the welfare maximisation program, and since per-period utility is assumed to be logarithmic, these future expected tax rates disappear in the determination of current taxes (again, see appendix A for more details). From the point of view of tax rate determination, the political economy problem is essentially a succession of static ones, unlike cases in which agents vote in a fully strategic way. Assumption 4 plays here a fundamental role on this result: since agents take future levels of taxes as given, they neglect the impact their vote today will have on the outcome of the vote tomorrow (even if this level is correctly anticipated). The consequence of this assumption is to destroy the dynamic linkage between successive policies, which means that this model does not need to rely on reputational arguments to sustain equilibria with positive pensions and education, among other things. Then the current policy choice merely balances the interests of agents, taking into account their relative political power and numbers.

By computing the Hessian matrix of W, one can easily check that W is strictly concave in $\tau = (\tau_e; \tau_p) \in S$. Given that S is compact and convex, one and only one couple of policies τ^* maximises W in S; moreover, if τ^* lies in the interior of S, both derivatives of W with respect to the tax levels (as expressed in equations (20) and (21)) are equal to zero at $\tau*$. This defines a policy rule that maps the state variables of the economy into a vector of tax rates (τ_e, τ_p). The following proposition summarizes this insight:

Proposition 1 *The policy rule of the economy $\tau(k,\mu) = (\tau_e(k,\mu); \tau_p(k,\mu))$ is implicitly defined by equations (20) and (21). It solely depends on output y, as defined in (5): in particular, it is unaffected by the future expected values of the tax rates in next period.*

3.4 The politico-economic dynamics, and steady state

Given an initial state of the economy (k_0, μ_0), the previous sections allow us to describe the politico-economic dynamics of this economy. As stated in proposition 1, equations (20) and (21) implicitly define a policy rule $\tau(k_t, \mu_t) = (\tau_e(k_t, \mu_t), \tau_p(k_t, \mu_t))$ that maps the state variables of the economy into a choice of the education and pension contribution rates.[17] In turn, equations (17) and (19) map the current values of the state variables and policy choices into the values of the state variables in next period (k_{t+1}, μ_{t+1}).[18]

A steady state of the economy is then defined by the stationarity of k and μ with respect to the policy rule and the laws of motion of both variables. Formally, for a constant sequence of fertility and mortality rates $\{\delta_t = \delta, n_t = n\}_t$, $(\bar{k}, \bar{\mu})$ is a steady state, with associated tax rates $(\bar{\tau}_e, \bar{\tau}_p)$ and output $\bar{y}$ if:

$$\tau(\bar{k}, \bar{\mu}) = (\bar{\tau}_e, \bar{\tau}_p)$$

$$\bar{\mu} = \frac{\bar{\tau}_e \alpha \bar{y}}{\bar{\omega}(1+n)}$$

$$\bar{k} = \frac{\beta\delta\alpha(1-\alpha)(1-\bar{\tau}_e-\bar{\tau}_p)}{(1+n)[(1+\beta\delta)(1-\alpha)+\alpha\bar{\tau}_p]}\bar{y}$$

Having characterised both the dynamics of the economy and its steady state, I now turn to analysing the impact of demographic change on the level and composition of public spending, and its consequences on factor accumulation and output.

4 Demographic change and its impact on public spending and growth

In this section, I first perform comparative statics on the joint policy function $\tau(k, \mu) = (\tau_e(k, \mu); \tau_p(k, \mu))$, which allows to determine how the structural parameters of the model (and in particular the demographic variables) influence the political process in each period. I then compare the predictions of the model to historical evidence on the level and composition of public spending.

[17] Of course, the values of $(n_{t-1}, n_t, \delta_t, \delta_{t+1})$ are also needed to define the state of the economy in period t, a fact I overlooked in the notations above to save some space.

[18] Equation (17) shows that k_{t+1} actually depends on the expected value of the future pension contribution rate $\tau_{p,t+1}$, which itself depends on k_{t+1}: then $(k_{t+1}, \tau_{p,t+1})$ is the solution of a fixed-point equation, which happens to be unique.

4.1 Demography and the level and allocation of public spending

4.1.1 Comparative statics on policy choices

I first analyse the effect of a change in parameters on the levels of the two policy rates that are chosen in equilibrium, for a fixed value of both state variables (actually for a fixed level of y_t, which is the only relevant endogenous variable in the choice of τ_t^*). To do so, and to obtain results that are valid both when τ^* is in the interior and at the boundary of S, the monotone comparative results of Milgrom & Shannon (1995) shall be used.

Monotone comparative statics for a submodular function: general method W is strictly submodular in τ, as the cross-derivative of W is strictly negative everywhere:

$$\frac{\partial^2 W}{\partial \tau_p \partial \tau_e}_{|y=cst} = -\frac{1+\beta\delta_{t+1}-\beta\delta_{t+1}\alpha}{(1-\tau_e-\tau_p)^2} < 0$$

The method developed by Milgrom & Shannon (1995) for comparative statics is relevant in the case of supermodular functions, however. Their general method can be put into use by considering the strictly supermodular function $\hat{W}(\hat{\tau}) = W(\tau)$, where $\hat{\tau} = (\tau_e; -\tau_p)$. Now for any parameter of interest σ, if we obtain that $\frac{\partial^2 \hat{W}}{\partial \hat{\tau}_k \partial \sigma}_{|y=cst} \geq 0$ for $k = 1, 2$, then $\hat{W}$ exhibits increasing differences in $(\hat{\tau}, \sigma)$. Using Milgrom & Shannon (95), supermodularity along with increasing differences implies that the argmax of welfare $\hat{\tau}^*(\sigma)$ is a monotone nondecreasing, continuous function of σ.[19] In terms of the original choice variables, it means that $\tau_e^*(\sigma)$ would be nondecreasing while $\tau_p^*(\sigma)$ is nonincreasing in σ.

In the following paragraphs, we use this method to examine successively the effect of *ceteris paribus* increases in the two demography parameters n_t and δ_t on the two tax rates. I also consider the impact of the level of output y_t on τ.

Impact of a change in the fertility rate n_t : I first study the impact of a change in n_t on the policy decisions taken at the time of the shock t. From equations (20) and (21), we obtain that $\frac{\partial^2 W_t}{\partial n \partial \tau_p} = 0$ and $\frac{\partial^2 W_t}{\partial n \partial \tau_e} < 0$. The result on the cross-derivative with respect to τ_e and n_t translates the fact that when n_t increases, educational subsidies need to be shared between more young people, thereby diminishing the returns on the education tax rate. In what follows, this effect will be called the dilution effect of fertility on human capital, by analogy with the effect on physical capital. Besides, at t the fertility of the old generation n_{t-1} is already determined and not impacted by the

[19] The continuity property comes from the fact that τ^* maximises W on S, a convex and compact subset of $\mathbb{R}^2$, which allows us to apply the weak version of the maximum theorem.

shock, so that the fertility shock has no impact on the pension margin of decision. Then following Milgrom & Shannon (1995), right after a shock on n_t it can be predicted that $\tau_{e,t}$ will decrease and $\tau_{p,t}$ will increase. This phenomenon comes from the fact the effect of total taxation on the welfare of middle-aged voters: if education subsidies become less efficient and $\tau_{e,t}$ is reduced, then there is more fiscal space for pension financing and $\tau_{p,t}$ will go up.

Let us now consider the impact of the shock for future periods (i.e. for $t+1$ on), when the shock is permanent and perceived as such. From equations (20) and 21), we obtain that $\frac{\partial^2 W_t}{\partial n \partial \tau_p} < 0$ and $\frac{\partial^2 W_t}{\partial n \partial \tau_e} < 0$. The result on the cross-derivative with respect to τ_e and n_t is unchanged and can still be interpreted as an input dilution effect. The second inequality is due to the fact that a higher fertility rate in t (and in subsequent periods) reduces the share of pensioners in the voting population from $t+1$ on, making their interests less represented in the political process. Taken in isolation, these effects imply that an increase in the fertility parameter would lead to both lower pensions and lower education subsidies: however, and as the submodularity of W suggests, lowering one of the two tax rates has the effect of reducing the burden of total taxation, and hence leaves some space to an increase in the other tax rate. Therefore, it is so far impossible to determine whether each individual tax rate would decrease if n increased, except if one of the two tax rates lies at the boundary of S.[20] Nonetheless, both tax rates rising consecutively to an increase in n is absolutely ruled out. In fact, appendix B.1 proves that total taxation $\tau = \tau_e + \tau_p$ actually goes down when n goes up, for any given values of k_t and μ_t. The results above are summarised in the proposition that follows:

Proposition 2 *In the long run, total taxation $\tau = \tau_e + \tau_p$ is negatively impacted by increases in the fertility rate n. Additionally, if one of the two tax rates is at the boundary of the choice set S for a given state of the economy (i.e. if $\tau_e = 0$, $\tau_e = \tau_{e,max}(y_t)$ or $\tau_p = 0$), then an increase in n causes the other tax rate to decrease.*

Impact of a change in the longevity parameter δ : In order to determine how an increase in δ would modify the policy choices, the two cross-derivatives of welfare with respect to the policy rate and δ need to be computed.

$$\frac{\partial^2 W_t}{\partial \delta \partial \tau_{e,t}} = \beta\alpha \frac{(\eta-1)\alpha y_t}{\bar{\omega}(1+n) + \tau_{e,t}(\eta-1)\alpha y_t} - \frac{\beta - \beta\alpha}{1-\tau_{e,t}-\tau_{p,t}} \tag{22}$$

Concerning the determination of τ_e, equation (22) shows that two effects compete when δ increases: on the one hand, an increased probability to survive until the next period makes it more likely for current workers to reap the benefits of investing into the next generation's education (through increased pensions tomorrow). On the other hand, an increased survival rate into old age has the effect that agents will more likely experience

[20] In this case, a marginal increase in n leaves this tax rate unchanged, which means the cross-derivative effect on W can be ignored.

the adverse effect of taxation on next-period capital accumulation (which matters, as well, for tomorrow's production and ultimately pensions).

$$\frac{\partial^2 W_t}{\partial\delta\partial\tau_{p,t}} = \frac{\psi\alpha}{(1+n)(1-\alpha+\alpha\tau_{p,t})} - \frac{\beta-\beta\alpha}{1-\tau_{e,t}-\tau_{p,t}} \tag{23}$$

Concerning τ_p, equation (23) indicates that the increased likelihood of reaching old age matters for capital accumulation reasons as well, so that middle-aged voters will care more about not taxing too large a share of their income now. However, this effect is balanced by the fact that there are more old people when δ goes up, which implies a tilt towards more pensions.

To see how the two trade-offs relative to an increase in δ are usually resolved, first notice that rearranging terms in equations (22) and (23) yield, for $k \in \{e; p\}$:

$$\frac{\partial^2 W_t}{\partial\delta\partial\tau_k} = \frac{1}{\delta}\frac{\partial W_t}{\partial\tau_k} + \frac{1}{1-\tau_e-\tau_p} \tag{24}$$

Additionally, it can be easily shown that $\frac{\partial W_t}{\partial\tau_k}(\tau_e^*, \tau_p^*) \geq 0$ for $k \in \{e; p\}$ as long as $\tau_e^* > 0$ and $\tau_p^* > 0$.[21] From equation (24), we can then deduce that both cross-derivatives are strictly positive as long as the argmax of welfare is strictly positive (i.e. as long as both taxes are positive in equilibrium), meaning that the increase in the negative effects of taxations is trumped by the increase in the benefits of taxation, for both pensions and education subsidies. So both tax rates, taken in isolation, would increase following an increase in δ.

As in the case of the fertility rate, however, the submodularity of W and the fact that τ_e and τ_p are jointly determined prevent us from reaching an unambiguous conclusion, except if one of the two tax rates is at the boundary of S. Both tax rates can rise, or one can rise at the expense of the other, the only case being ruled out is the one where both would decrease at the same time following an increase in δ. As before in the case of n, I prove in appendix B.2 that total taxation $\tau = \tau_e + \tau_p$ necessarily increases (weakly) when δ increases, whatever the state of the economy (k_t, μ_t). These results are summarised in the following proposition:

Proposition 3 *Total taxation $\tau = \tau_e + \tau_p$ is an increasing function of the survival rate δ. Additionally, if one of the two tax rates is at the boundary of the choice set S for a given state of the economy, then an increase in δ causes the other tax rate to increase.*

Analysing the impact of y on policy choices Although y is an endogenous variable in the model as a whole, it seems useful to analyse how the political choice made at t depends on the value of y. Since the policy choice problem is a succession of static

[21] If the argmax of welfare lies in the interior of S, both derivatives are equal to zero.

problems rather than dynamic ones, methods of comparative statics also apply here. As it has been said earlier, this political choice does not hinge on the values of k_t or μ_t taken in isolation, but rather on $y_t = (\mu_t \eta + 1 - \mu_t)^\alpha k_t^{1-\alpha}$.

Simple computations show that $\frac{\partial^2 W_t}{\partial y \partial \tau_e} > 0$ and $\frac{\partial^2 W_t}{\partial y \partial \tau_p} = 0$. As a consequence, when y is higher, the education tax rate is bigger or unchanged, and the pension contribution rate is unchanged or lower. All else being equal, a more productive economy will allow for more support for education and less support for pensions.

4.1.2 The impact of demography on public spending in a historical perspective

The late 19th and the 20th century were characterised by a secular decline in mortality and fertility rates, which can be interpreted respectively as a rise in δ and a fall in n. Notwithstanding the cross-taxation effects on welfare discussed above, it seems to be that a decrease in n or an increase in δ both lead to higher values of τ_e and τ_p: at the very least, the model predicts an increase in the level of total taxation. This seems to be in line with the historical evidence of a rising share of public spending to GDP. I now contrast this analysis to empirical studies of the links between the age structure of the population and public spending.

Cross-country estimates by Lindert (1994) suggest that for the period 1880-1930, the share of population aged 65 and more had a positive impact on the share of GDP devoted to total social spending (a notion that encompasses all social transfers such as pensions, health care, etc. and educational spending): this type of evidence concurs with the insight that longevity (proxied by δ in our model) should have a positive impact on total spending (and thus taxation). The effect on social transfers only is also positive. On the contrary, the share of population aged between 20 and 39 has a significant negative impact on the share of social tranfers in GDP. The results are globally confirmed in a similar analysis of the 1960-1980 period by Lindert (1996): over this period, a higher share of older people in total population exerts a positive impact on the share of GDP spent on pensions, while the share of school-age individuals does not affect significantly the share of educational spending on GDP.

Additionally, an analysis on U.S. states by Poterba (1997) of the demographic determinants of public education spending reveals that the share of people older than 65 in the population has a negative impact on education spending per child. Conversely, the share of public spending going to education rises with the share of population of schooling age.

There are several limits to this exercise of comparing the available evidence with the predictions of the model. First, the model so far is silent on which of the two tax rates whould increase with δ (or decrease when n increases). Second, the explanatory variables used in the aforementioned works are shares of specific age groups in the population, which do not match exactly with the broad variables of fertility and longevity that are present in the model. Indeed, in our model the share of elderly people in the population can rise either following an increase in δ, or following a decrease in n. The same reasoning also applies to the share of young pepole in the population. Fortunately, population ageing seems to have the same impact on public spending whether it comes from a drop in n or a rise in δ.

5 Conclusion and discussion

This paper takes a stance on the forces that led to the establishment of social security and publicly funded education in market economies with democratic institutions, in the period ranging from the late 19th century to the end of the 20th century. The model developed above aims at highlighting the major role played by the shift in the demographic structure in establishing and progressively growing these two types of intergenerational transfers over the period.

As in other recent developments in the literature on the political economy of pensions, I depart from the standard asssumption of majority voting as the preference aggregation mechanism, replacing it with the assumption of probabilistic voting. While the more common formulation has its advantages, it fails to take into account notions like the relative political weight of each age group, which I believe determines to some extent the tax policy pursued in each period. Since the weight of each group depends, among others, on its relative size, a change in the demographic structure of the population is bound to have some effects on the size and the composition of public spending. By virtue of this modelling of preference aggregation, there is no need anymore to interpret the pension and education systems as two pillars of a contract between generations, that is sustained by implicit punishment schemes. This model insists instead on the intra-period conflict existing between present generations, and the fact that the equilibrium policy merely balances the interest of each generation.

The model developed in this paper predicts that population ageing, whether it comes from decreased fertility or increased longevity, leads to a higher level of overall taxation and spending (including both educational and pension transfers). Evidence on the historical evolution of public spending, as well as cross-section evidence on U.S. states, seems to corroborate this analysis. Current research on this model involves finding the total effect of the demography parameters on human and physical capital accumulation, both along the transition path and at steady state.

Further research might have to include simulations of the model for different, more realistic distribution of abilities. Introducing a degree of intragenerational redistribution in the pension system is also a possibility, to take advantage of the intragenerational heterogeneity present in the model. In theses cases, the greater realism of the model comes at the expense of analytical tractability. I believe that the current formulation of the model already has the advantage of putting into evidence the main qualitative results, even if analytical computations fail to reach a conclusion on the composition of public spending between pensions and transfers to the young.

Another potential source of improvement in the model would be to endogenise the demographic parameters, especially the fertility parameter n. Indeed, several authors (see for instance Van Groezen et al. (2003), and Cremer et al. (2011)) shed light on the fact that fertility choices depend not only on the design of the educational system, but also on the pension system as well. Incorporating these insights into this model seems a promising way forwards.

References

Belletini, Giorgio; Berti Ceroni, Carlotta. “Is Social Security Really Bad for Growth ?” *Review of Economic Dynamics* 2, no.4 (1999), pp. 796-819. Elsevier for the Society for Economic Dynamics.

Boldrin, Michele; Montes, Ana. “The Intergenerational State: Education and Pensions”. *The Review of Economic Studies* 72, no.3 (jul. 2005), pp.651-664. Blackwell Publishing.

Boldrin, Michele; Rustichini, Aldo. “Political Equilibria with Social Security”. *Review of Economic Dynamics* 3, no.1 (2000), pp.41-78. Elsevier Science.

Cremer, Helmuth; Gahvari, Firouz; Pestieau, Pierre. “Fertility, human capital accumulation, and the pension system.” *Journal of Public Economics* 95 (2011), pp. 1272-1279. Elsevier Science.

Gonzalez-Eiras, Martin; Niepelt, Dirk. “Population Ageing, Government Budgets, and Productivity Growth in Politico-Economic Equilibrium”. CEPR Discussion Paper n.6581 (2007).

Lancia, Francesco. “A Dynamic Politico-Economic Model of Intergenerational Contract” *in* “Demographic Change, Intergenerational Conflict and Economic Growth” (chapter 2). Doctoral dissertation (2010). Università di Bologna.

Lindbeck, Assar; Weibull, Jörgen. “Balanced-budget redistribution as the outcome of political competition”. *Public Choice* 52 (1987), pp. 273-297. Springer Verlag.

Lindert, Peter. "The Rise of Social Spending, 1880-1930". *Explorations in Economic History* 31, no.1 (jan. 1994), pp 1-37. Elsevier Science.

Lindert, Peter. "What Limits Social Spending?" *Explorations in Economic History* 33, no.1 (jan.1996), pp 1-34. Elsevier Science.

Poterba, James. "Demographic structure and the political economy of public education", *Journal of Policy Analysis and Management* 16, no.1 (1997), pp. 48-66. John Wiley & Sons, Ltd.

Persson, Torsten; Tabellini, Guido. “Political Economy - Explaining Economic Policy.” MIT Press. Published 2000.

Rangel, Antonio. “Forward and Backward Intergenerational Goods: Why Is Social Security Good for the Environment ?” *The American Economic Review* 93, no.3 (2003), pp. 813-834. American Economic Association.

Van Groezen, Bas; Leers, Theo; Meijdam, Lex. “Social security and endogenous fertility: pensions and child allowances as Siamese twins.” *Journal of Public Economics* 87 (2003), pp. 233-251. Elsevier Science.

A Appendix: Obtaining the derivative of welfare with respect to policy instruments

In this appendix, I describe how to compute the derivative of the welfare with respect to the two policy instruments, by first using a relationship between γ_t and the state variables in $(t+1)$.

Relationship between the PV of lifetime income and state variables The present value of the net lifetime income of an educated worker is given by the following: $\gamma_{0,t} = (1-\tau_{e,t}-\tau_{p,t})w_{0,t} + \delta p_{0,t+1}/R_{t+1}$.

Savings are related to γ by the following:

$$s_{0,t} = \frac{\beta\delta}{1+\beta\delta}(1-\tau_{e,t}-\tau_{p,t})w_{0,t} - \frac{\delta}{1+\beta\delta}\frac{p_{0,t+1}}{R_{t+1}}$$

$$\Rightarrow \frac{\beta\delta}{1+\beta\delta}\gamma_{0,t} = s_{0,t} + \delta\frac{p_{0,t+1}}{R_{t+1}} \tag{25}$$

But next-period capital is given by:

$$k_{t+1} = \frac{1}{1+n}(\mu_t s_{e,t} + (1-\mu_t)s_{0,t})$$

$$\Rightarrow s_{0,t} = \frac{k_{t+1}(1+n)}{(\mu_t\eta + 1 - \mu_t)}$$

since $s_{e,t} = \eta s_{0,t}$.

Moreover,

$$\frac{p_{0,t+1}}{R_{t+1}} = \frac{\alpha}{\delta(1-\alpha)} \cdot \frac{\tau_{p,t+1}(1+n)k_{t+1}}{\mu_t\eta + 1 - \mu_t} \tag{26}$$

Using (25) and (26), one now gets to the following equation linking next-period capital and the NPV of expected lifetime income :

$$\frac{\beta\delta}{1+\beta\delta}\gamma_{0,t} = \frac{k_{t+1}(1+n)}{(\mu_t\eta + 1 - \mu_t)} \cdot \left(1 + \tau_{p,t+1}\frac{\alpha}{1-\alpha}\right) \tag{27}$$

Computing the derivatives The welfare function that is maximised as an outcome of the vote is the following:

$$W_t = (\mu_t V_{e,t}^{t-1} + (1-\mu_t)V_{0,t}^{t-1}) + \frac{\psi\delta}{1+n}(\mu_{t-1}V_{e,t}^{t-2} + (1-\mu_{t-1})V_{0,t}^{t-2})$$

However, since $\gamma_{e,t} = \eta\gamma_{0,t}$ for all t, it is easy to show that the welfare of educated and non-educated individuals of the same generation only differ by a constant. Hence, the tax rates $\tau_{e,t}$ and $\tau_{p,t}$ alternatively need to maximise the following function:

$$W_t = V_{0,t}^{t-1} + \frac{\psi\delta}{1+n} V_{0,t}^{t-2}$$

The indirect utility of a non-educated old individual can be further computed as follows, using equation (27):

$$V_{0,t}^{t-2} = \ln(c_{0,t}^{t-2}) = \ln(\frac{\beta}{1+\beta\delta}\gamma_{0,t-1}R_t) = \ln(R_t/\delta) + \ln\left(\frac{k_t(1+\tau_{p,t}\alpha/(1-\alpha))(1+n)}{\mu_{t-1}\eta + 1 - \mu_{t-1}}\right)$$

At t, all current and past state variables are predetermined so the welfare of old people is only sensitive to $\tau_{p,t}$:

$$\frac{\partial V_{0,t}^{t-2}}{\partial \tau_{p,t}} = \frac{\alpha}{1-\alpha+\alpha\tau_{p,t}} \tag{28}$$

$$\frac{\partial V_{0,t}^{t-2}}{\partial \tau_{e,t}} = 0 \tag{29}$$

Concerning the welfare of non-educated workers, the following is obtained:

$$\begin{aligned} V_{0,t}^{t-1} &= \ln(c_{0,t}^{t-1}) + \beta\delta\ln(c_{0,t+1}^{t-1}) = \ln(\frac{1}{1+\beta\delta}\gamma_{0,t}) + \beta\delta\ln(\frac{\beta}{1+\beta\delta}\gamma_{0,t}R_{t+1}) \\ &= (1+\beta\delta)\ln(\frac{\beta}{1+\beta\delta}\gamma_{0,t}) + \beta\delta\ln(R_{t+1}) - \ln\beta \end{aligned}$$

Using equation (8) and (27), it obtains that:

$$\begin{aligned} V_{0,t}^{t-1} = \;& (1+\beta\delta-\beta\alpha\delta)\ln k_{t+1} + \beta\alpha\delta\ln((\eta-1)\mu_{t+1}+1) - (1+\beta\delta)\ln(\mu_t(\eta-1)+1) \\ & + (1+\beta\delta)\ln(1+\tau_{p,t+1}\alpha/(1-\alpha)) + c \end{aligned}$$

where c is a constant.

Since voters are assumed to vote non-strategically, they take the future value of the pension contribution rate as given in the evaluation of their indirect welfare. The current education state variable μ_t is also given at the time of the vote. It then appears that the welfare of middle-aged individuals is affected by changes in current policy rates only insofar as it changes the future values of state variables one period ahead. As a consequence, the partial derivatives of the indirect utility of a middle-aged worker with respect to the two policy instruments are:

$$\frac{\partial V_{0,t}^{t-1}}{\partial \tau_{p,t}} = (1+\beta\delta-\beta\alpha\delta)\frac{\partial \ln k_{t+1}}{\partial \tau_{p,t}} = -\frac{1+\beta\delta-\beta\alpha\delta}{1-\tau_{e,t}-\tau_{p,t}} \tag{30}$$

$$
\begin{aligned}
\tfrac{\partial V_{0,t}^{t-1}}{\partial \tau_{e,t}} &= (1+\beta\delta-\beta\alpha\delta)\frac{\partial \ln k_{t+1}}{\partial \tau_{e,t}} + \beta\alpha\delta\frac{\partial \ln((\eta-1)\mu_{t+1}+1)}{\partial \tau_{e,t}} \\
&= -\frac{1+\beta\delta-\beta\alpha\delta}{1-\tau_{e,t}-\tau_{p,t}} + \beta\alpha\delta\frac{(\eta-1)\alpha y_t}{\bar{\omega}(1+n)+\tau_{e,t}(\eta-1)\alpha y_t} \qquad (31)
\end{aligned}
$$

Then the derivatives of total welfare with respect to the values of the two policy instruments are computed as follows:

$$
\frac{\partial W_t}{\partial \tau_{k,t}} = \frac{\psi\delta}{1+n}\frac{\partial V_{0,t}^{t-2}}{\partial \tau_{k,t}} + \frac{\partial V_{0,t}^{t-1}}{\partial \tau_{k,t}}
$$

for $k = e, p$. This yields equations (21) and (20).

B Appendix: Comparative statics results on total taxation

In this appendix, I show that total taxation $\tau = \tau_e + \tau_p$ decreases when n increases, and increases when δ increases.

B.1 Comparative statics on the fertility rate

First assume that under a given set of parameters, the chosen tax rates lie in the interior of the choice set S. Then partial derivatives of W with respect to both tax rates, the expressions of which are given by (20) and (21), are equal to zero at (τ_e, τ_p). Then total taxation cannot go up following an increase in n: indeed, equation (20) imposes that τ_e needs to decrease if n and $\tau_e+\tau_p$ are to go up simultaneously. Similarly, a simultaneous (weak) increase in $\tau_e+\tau_p$ and n imply, through (21), that τ_p needs to decrease as well. Then, a strict increase in n along with a weak increase in total taxation $\tau_e+\tau_p$ imply a strict increase in both τ_e and τ_p, hence in $\tau_e+\tau_p$, which is a logical contradiction. Hence, the effect of an increase in n is to reduce total taxation.

Now when $\tau = (\tau_e, \tau_p)$ belongs to the frontier of S, this means that one of the two derivatives in (20) and (21) (or possibly both) is strictly different from zero. Hence, a marginal increase in n in this case will fail to change one of the two tax rates, or possibly both. Then because $\partial^2 W_t/\partial\tau_e\partial n < 0$ and $\partial^2 W_t/\partial\tau_p\partial n < 0$, the one of the two tax rates that does not change following an increase in n necessarily goes down.[22] As a result, total taxation effectively goes down in this case following an increase in n.

[22] Since one of the two tax rates does not change, the submodularity of W does not matter in this case and we consider only the direct effect on the other tax rate.

B.2 Comparative statics on the survival rate

Again, assume that the optimal tax rate lies in the interior of S. In this case, the expressions in (20) and (21) are also equal to zero. Following the same logic as before, assuming total taxation $\tau = \tau_e + \tau_p$ goes down when δ increases leads to finding that τ_e and τ_p should increase, which is impossible. By contradiction, an increase in δ then has the effect of increasing total taxation.

Now when the optimal policy does not lie in the interior of S, one of the two tax rates will remain constant following an increase in δ. Then $\partial^2 W_t / \partial\tau_e \partial\delta > 0$ and $\partial^2 W_t / \partial\tau_p \partial\delta > 0$ imply that the other tax rate will rise when δ increases. Then in this case, an increase in the survival probability δ causes total taxation to go up.