THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l' Université Toulouse 1 Capitole
Discipline : Sciences Economiques

Présentée et soutenue par

Loïc BATTÉ
Le 3 avril 2015

Titre :

Essays on the Political Economy of Intergenerational Public Transfers

---

Ecole doctorale : Toulouse School of Economics

Unité de recherche : GREMAQ – TSE

Directeur de Thèse : Gilles SAINT-PAUL

JURY

Jean-Marie LOZACHMEUR, chargé de recherches, CNRS - GREMAQ
Pierre PESTIEAU, professeur, Université de Liège
Franck PORTIER, professeur, Université Toulouse 1 Capitole
Thierry VERDIER, directeur d'études, EHESS - PSE

L'université Toulouse I Capitole n'entend donner aucune approbation, ni improbation aux opinions émises dans cette thèse; elles doivent être considérées comme propres à leur auteur.

A Lauriane, Gabrielle et Antonin

Acknowledgments Remerciements

I am grateful to Gilles Saint-Paul for his valuable advice and support throughout this thesis. Gilles did not hesitate to become my supervisor when I had to find a new one early on in my thesis, and kept me entirely free to pursue the research projects I found most interesting from this point on. With his constructive criticism, light-handed approach to supervision and overall intellectual guidance, he has been very instrumental in helping me acquire my intellectual autonomy as a researcher. I also want to thank him a lot for pointing out how little actually remained to be done, at a time when I was struggling to make progress: it greatly helped me get back to work.

I would like to express as well my gratitude to Patrick Fève, for his help and support during the last stages of this thesis. He not only contributed substantially to the completion of this manuscript through his scientific input, but he also helped me organize and get the job done by setting regular, achievable goals for me. This made the process of finishing the thesis a much less daunting task than I intially thought it would be.

I would also like to express my thanks to Georges Casamatta, for trusting me enough to ask me to join him in writing a chapter of the upcoming Handbook of the Economics of Population Ageing, to be published at Elsevier. This project led to our writing together the first chapter of this thesis. I also seize this opportunity to thank the editors of the Handbook, John Piggott and Alan Woodland, who allowed me to include this chapter in my thesis before the actual publication of the Handbook.

My work benefited also greatly from discussions with other researchers and professors from the Toulouse School of Economics, among which Philippe de Donder, Christian Hellwig, Yann Kervinio, Jean-Marie Lozachmeur, Eric Mengus, Marti Mestieri, Franck Portier and Emmanuel Thibault. I especially want to thank Jean-Marie and Franck for accepting to be members of the jury of this thesis. Finally, I want to thank Profs.

Thierry Verdier and Pierre Pestieau for doing me the honor of being referees of this thesis dissertation.

I would also like to thank Prof. Omar Licandro for his hospitality as director of the Institut d'Anàlisi Econòmica (IAE-CSIC) in Barcelona. Other researchers of the IAE and the economics department of the Universitat Autònoma de Barcelona contributed to improving the quality of this work, such as Juan Carlos Conesa, Caterina Calsamiglia or Hannes Mueller.

I would like to thank as well the administrative staff of TSE for their work, and for simplifying my professional life here through their handling of all the logistical and organizational details, a thing I am usually not good at. My special thanks in this respect go to Aude Schloesing, the pillar of the TSE doctoral school, whose retirement will make her sorely missed among PhD students.

I owe a great debt to some PhD students in Toulouse, be it for their help on academic matters or their ability to keep me in high spirits thanks to their sheer company. I have been lucky to work in a stimulating environment with nice people such as Julia Bird, Fanny Camara, Maddalena Ferranna, Mattia Girotti, Sinem Hidir, Dong-Hee Joe, Yann Kervinio, Marie Lalanne, Laura Lasio, Margaret Leighton, Yves Le Yaouanq, Eric Mengus, Guillem Roig, Ananya Sen, Anastasia Shchepetova and Bruno Ziliotto. Those whom I may have forgotten to include in this list will hopefully forgive me.

Je ne voudrais pas laisser passer cette occasion de remercier les personnes extérieures au monde académique qui m'ont accompagné au quotidien durant ces années de thèse. Merci donc aux membres des équipes de handball de l'UT1 Capitole dont j'ai fait partie dans les premières années, pour m'avoir permis de complètement déconnecter de la recherche en économie. Merci également à Matthieu et Dorothée, Florence et Romain pour les nombreux dîners, balades, sorties à Toulouse et dans les environs...

Je souhaitais également dire un immense merci à ma famille de m'avoir soutenu pendant les moments difficiles, et partagé mes joies. Merci à mes grands-parents pour leur regard bienveillant, à mes folles de sœurs pour leur bonne humeur et leur capacité à nous faire retourner en enfance aussi rapidement que nécessaire. Merci à mes parents, mes premiers enseignants, sans le soutien indéfectible desquels je n'aurais jamais pu arriver jusque-là. Enfin, merci à toi Lauriane, pour ta patience, ton énergie et ton

amour, et pour m'avoir fait les deux plus beaux cadeaux dont on puisse rêver. Gabrielle, Antonin, vos naissances ont été les deux événements les plus importants de toute la durée de cette thèse : merci à vous d'être là, tout simplement.

Contents

Acknowledgments/Remerciements **vii**

Contents **xi**

List of figures **xv**

List of tables **xvii**

Introduction **1**
0.1 Elements of context 2
0.2 Methodological aspects 5
0.3 Outline of the thesis 8

1 The Political Economy of Population Ageing: a Critical Literature Review **11**
1.1 Introduction 13
1.1.1 The political challenges of population ageing 13
1.1.2 Features of social security systems and recent reforms 16
1.1.3 Outline 20
1.2 Why individuals support PAYG social security 22
1.2.1 Dynamic inefficiency 22
1.2.2 Reduced time horizon 23
1.2.3 Effect of social security on prices 26
1.2.4 Within cohort redistribution 27
1.2.5 Risk-sharing 31
1.3 The social contract 32
1.3.1 Sustainability of the PAYG system 32
1.3.2 Dynamics of the political equilibrium 33
1.4 Endogenous retirement 34

1.4.1 Implicit taxation and early retirement 34
1.4.2 Voting on the retirement age 37
1.5 Markov perfect equilibria: Dynamics of the political equilibrium 44
1.5.1 A median voter model with capital as the state variable 44
1.5.2 Probabilistic voting 46
1.5.3 The evolution of retirement 47
1.6 The transition to a Fully Funded system 49
1.7 Quantitative analysis 52
1.7.1 Empirical estimations 52
1.7.2 Simulating the future political evolution of social security 54
1.8 The political impact of ageing on other public programs 57
1.8.1 Ageing and political support for education 58
1.8.2 Ageing and support for health care 69
1.8.3 Ageing and the environment 70
1.8.4 Capital vs. labor taxes 72
1.9 Conclusion 73

2 Demography and intergenerational public transfers 77
2.1 Introduction 79
2.1.1 Motivation and aim of this paper 79
2.1.2 Contribution to the existing literature 81
2.2 Model setup 83
2.2.1 OLG structure, and private decisions 83
2.2.2 Production 86
2.2.3 The public sector: education subsidies and pensions 87
2.2.4 Voting over education and pensions 88
2.2.5 State variables, timing of decisions and dynamics 90
2.3 The politico-economic dynamics 90
2.3.1 Dynamics of the economy under given tax policies 91
2.3.2 Additional assumptions 94
2.3.3 Taxation choices 97
2.3.4 The politico-economic dynamics, and steady state 98
2.4 Demographic change and its impact on public spending and growth . . . 99
2.4.1 Comparative statics on policy choices 100
2.4.2 Numerical example and simulation 103
2.4.3 The impact of demography on public spending in a historical perspective 107

2.5 Exploring the implications of dynamic policy expectations 109
2.5.1 Strategic behavior as a sensitivity analysis 110
2.5.2 Modelling the equilibrium using a Parameterized Expectations Algorithm with linear specification 111
2.5.3 Simulation results and interpretation 112
2.6 Conclusion and discussion 115

Appendices **119**
2.A Appendix: derivative of welfare with respect to policy instruments 119
2.B Appendix: Comparative statics results on total taxation 122
2.B.1 Comparative statics on the fertility rate 122
2.B.2 Comparative statics on the survival rate 122

3 Commitment in a political economy model of education and redistribution **125**
3.1 Introduction 127
3.1.1 Motivation 127
3.1.2 Relationship with the previous literature 128
3.1.3 Outline of this chapter 131
3.2 Basic setup of the model with self-financing of education 132
3.2.1 Overview of the model 132
3.2.2 Time structure and agents 133
3.2.3 Private behavior 134
3.3 The government 136
3.3.1 The government budget constraint 137
3.3.2 Constrained planner problem using linear instruments 137
3.4 The politico-economic equilibrium 139
3.4.1 Second-period vote on redistribution 140
3.4.2 First-period vote on subsidy s 142
3.4.3 How relevant is the timing of decisions? 144
3.5 A quantitative study of education subsidy and income tax rates 146
3.6 Conclusion and discussion 148
3.6.1 Summary of the findings 148
3.6.2 Discussion and further research 148

Appendices **151**
3.A Appendix: Simultaneous vote on s and τ 151

3.B Appendix: an OLG model with altruism and transfers to finance education 153
3.B.1 Time structure and agents 154
3.B.2 Altruism between generations 154
3.B.3 The public sector 155
3.B.4 The voting game 156

Bibliography **163**

List of figures

2.1 Dynamics of the main state and policy variables of the economy for different values of the fertility parameter n, and $\delta = 0.80$ 106
2.2 Dynamics of the main state and policy variables of the economy for different values of the longevity parameter δ, and $n = 0.10$ 108
2.3 True (in red) and expected future (in blue) policy rules as a function of output 113
2.4 Policy rules as a function of output in the strategic (solid line) and non-strategic (dashed line) cases, for different values of n, and $\delta = 0.80$ 114
2.5 Policy rules as a function of output in the strategic (solid line) and non-strategic (dashed line) cases, for different values of δ, and $n = 0.10$ 116

List of Tables

2.1 Baseline parameters for the simulation of the politico-economic dynamics 104

3.1 Labor income tax rates and higher education subsidies (in %) for several OECD countries, in 2011 147

Introduction

In his classic analysis of democracy "De la démocratie en Amérique", French political philosopher Alexis de Tocqueville expressed his concern that democracy, by shaking down the previous rigid, hierarchical structure of aristocratic societies, would bring about a new social order in which individuals would increasingly isolate themselves from the rest of society:[1]

> Chez les peuples démocratiques [...], on oublie aisément ceux qui ont précédé et ceux qui vous suivront. Les proches seuls intéressent [...]. Ainsi, non seulement la démocratie fait oublier à chaque homme ses aïeux, mais elle lui cache ses descendants et le sépare de ses contemporains : elle le ramène sans cesse vers lui seul et menace de le refermer enfin tout entier dans la solitude de son propre coeur. (de Tocqueville (1840))

Despite these concerns, the expansion of liberal democracy in the nineteenth and twentieth centuries, though far from a smooth process and ripe with disruptions in the old social order, allowed nonetheless for the inventions of new ways for individuals to organize solidarity between generations. As I shall recall in the first section of this introduction, the new democracies proceeded at the turn of the last century to a massive expansion of public education, together with the enactment of the first social security laws. These intergenerational public transfers came as a response to the disappearance of the local, class-based solidarities of the old political regimes, that had been instituted in exchange for the almost complete lack of social mobility.

[1] "Among democratic nations (...), those who went before are easily forgotten, as are those who will come after. Only the next of kin is of concern. (...) Thus, not only does democracy cause every man to forget his forefathers, it hides his progeny and isolates him from his contemporaries: it ceaselessly throws him back upon himself alone, and threatens in the end to confine him entirely in the solitude of his own heart." (Author's own translation).

Despite numerous academic treatments of the topic, from the points of view of economics, history, political science or sociology, many questions remain about the emergence of these new types of intergenerational transfers. In a new political environment, in which policies are the outcomes of popular choice, it is necessary to understand how the political preferences and economic behaviors combine to produce what has sometimes been characterized as a new social contract between successive generations. It is all the more essential if we want to predict the future evolution of these transfers at a time when the age structure of population is set to change dramatically, bringing about both political and economic change in its wake. This thesis aims at contributing to this understanding.

In this general introduction, I will first narrow the scope of my study, by defining the very term of intergenerational public transfers. I will also briefly recall the historical timeline of the development of these transfers, as well as the current demographic context in which they are determined. I will then comment on various methodological points that are relevant to this thesis, before presenting its outline.

0.1 Elements of context

Defining what intergenerational transfers are Throughout this thesis, I wish to study the political and economic determinants of the size and features of intergenerational public transfers. By intergenerational transfers, I mean all public policy programs that entail a transfer of public resources between different age cohorts. An obvious example of intergenerational transfer is the pay-as-you-go pension system, by which currently working individuals make contributions that are used to pay out benefits to retired people. This type of transfer going from relatively younger to older cohorts is usually called a backward intergenerational transfer (Rangel, 2003), as the flow of funds goes back in time if one considers the birth order of cohorts. Conversely, a public program such as child allowances is the prime example of a forward intergenerational transfer, in which public funds flow towards the most recently born generations.[2] It should be pointed out straight away that what I define as an intergenerational public transfer is not limited to those governmental programs that aim at redistributing cash between generations:

[2]Even though the actual recipients of the funds are typically the parents, the funds are awarded on the condition of the existence of children. Throughout the thesis, I will always take into consideration who the intended beneficiary of the program is, rather than the actual recipient of the cash or in-kind transfer, when defining intergenerational transfers.

provision of private goods by the government also falls, at least partially, under this category, as long as the consumers of this private good are to be found disproportionately in a specific demographic segment of the general population. This is the case for instance for public education, which overwhelmingly concerns children (for primary and secondary education) and young adults (for higher education). It is also potentially partially the case for public health care, of which the elderly are heavy consumers compared to the rest of the population.

Historical trends in intergenerational public transfers The first compulsory schooling laws can be traced back to the seventeenth century, in several protestant states in Germany and the American colonies. But real enforcement and involvement by the state to fund schooling of the masses started in the late eighteenth and early nineteenth century, in the most advanced economies of the day (Prussia, Austria-Hungary, France, England). In the United States, public funding of education, initially dependent on local jurisdictions, expanded very quickly in the second half of the nineteenth century through state-level compulsory schooling laws.[3] The twentieth century then saw the steady expansion of secondary schooling in all developed economies, in a move to prepare the workforce for the new white-collar jobs that were created. Today, public education represents a massive share of public expenditures in most advanced economies, and a substantial share of national income is devoted to education funding (6.1% of the GDP of OECD countries was spent in 2011 on educational institutions, according to OECD (2014a)). Additionally, public funding seems to be the norm rather than the exception.[4]

Pensions and family allowances, two substantial public programs by both their size and the amount of redistribution between generations they produce, have been gradually introduced in advanced capitalist economies between the late nineteenth and the early twentieth century. Germany introduced under Bismarck the first old age insurance program in 1889, to be imitated by most Western industrialised economies by the mid-1930s. As for family allowances, they were mostly implemented between the years 1920 and 1950. After an initial period of consolidation (extension of coverage to include workers in more and more sectors of the economy, maturing of the accumulated rights to pensions by the first beneficiaries) up to the end of the second World War, advanced capitalist economies have witnessed until 1975 what has sometimes been termed the "Golden Age of the Welfare State" (Pierson, 2007). This golden age has been charac-

[3] Source: Gradstein et al. (2004).

[4] See chapter 4 of this thesis for some recent figures on the share of education expenditures funded by the state on OECD countries.

terized by a massive increase of the benefits awarded to all social program recipients, and a sustained growth in the amount of public resources devoted to these types of transfers: social expenditures increased from 9.9% of the U.S. GDP in 1960, to 18.7% in 1975. Finally, the period opened by the oil shocks of the 1970s was marked by continual talks of a crisis of the Welfare State, amidst growing public indebtedness and reduced economic growth. As a consequence, the rate of growth of expenditures in most public programs pertaining to the Welfare State, including pensions and family allowances, has been radically reduced in this period compared to previous experience: indeed, the share of GDP devoted to social expenditures across OECD countries rose by only 2% during the 1980s (Pierson, 2007). This containment in the expansion of these transfers was made possible by benefit retrenchment measures that were implemented in most countries for the first time in the 1980s decade, to be continued later on. Concerning pensions for instance, the 1990s saw the implementation of raises in retirement age (for instance in the UK , Japan or Italy); increases in the qualifying contributory period for a full pension (France, Ireland, Finland); a lower adjustment of benefits relatively to inflation (UK, France, Spain); or even the income testing of public pension (in Austria, Denmark or Australia). This decade also witnessed the reduction of the real values of child allowances in the UK, Spain or the Netherlands, or the implementation of tighter eligibility criteria.

The demographic context today If one is to analyse the way in which public transfers between generations are determined politically, it is unavoidable to take into account the demographic structure of the population, and its forecastable evolution for the future. In particular, ageing of population, which is the overwhelming demographic trend of at least the last half century, and likely to accelerate in the coming decades (Lutz et al., 2008), will undeniably have a profound impact on the direction and size of flows of public funds between generations. Recent projections such as United Nations (2013) indicate that the share of the elderly (defined as individuals aged 65 or more) in the world population will rise from 7.7% in 2010 to 15.6% in 2050, a fact that will naturally increase the relative power of the older generations in every policy-making process. This ageing of the world population is due to three main factors: a decline in fertility rates (the total fertility rate having declined from about 5 children per woman in 1950 to just over 2.5 in 2005 for the whole world, according to United Nations (2013)); an increase in life expectancy (from 47 to 69 years at birth, for the same period and sample); and the coming of age of the large generation of the baby boomers in the developed countries.

One of the aims of the chapters 2 and 3 of this thesis is to untangle the contributions of both economic and political mechanisms to the relationship between the changing age structure of population and the flows of public funds between its various age segments. Concerning the purely economic impact of ageing, it is useful to stress first the accounting effects of a changing age structure, which capture what would happen to these various transfers if each generation kept a fixed behaviour with respect to savings, labor supply, investment into education, and so on...[5] As the relative size of age cohort changes, and absent some changes in either contributions or benefits, some programs may run a surplus while others will have to face looming deficits, triggering the need for a painful adjustment. Let us consider next the behavioral effects of demographic change, which explicitly take into account the changes in individual behaviour brought about by a lengthened time horizon, the need to provide for fewer dependent children, etc. These changes in behavior will induce changed individual preferences for the public transfers I wish to study: for example, individuals who anticipate they will live longer will most likely start saving more to provide for their old days, thus becoming more wary of capital taxation. In turn, a changeing age structure will have a purely political, rather than economic, impact, in the sense that larger age segments of the population will be able to weigh more heavily on the direction of public policy. In a context of ageing, this increase in political sway held by the elderly may come about through their growing importance in terms of the number of voters they represent. It is also the product of a process by which old people start recognizing their power to influence policymakers as a group, and start organizing as lobbies with that intent in mind, just as class-based organizations such as trade unions do (Pampel and Williamson, 1989).

0.2 Methodological aspects

Normative foundations of the existence of intergenerational transfers It is out of the scope of this thesis to study normative justifications of state intervention in making cash or in-kind transfers between generations. However, I shall briefly mention the landmark work of Becker and Murphy (1988), who remark that the welfare state replicates some functions that could be performed by intra-family arrangements, such as educating the children and providing for the parents during old age. The superiority of the state, the argument goes, lies in the fact that it does so in a way that is not dependent

[5]This distinction between the accounting and behavioral effects of ageing is borrowed from Bloom et al. (2010).

anymore on the level of resources available in each family, nor on the degree of altruism existing between the family members. According to the authors, state involvement is also deemed necessary due to the impossibility for some family members, such as very young children, to enter enforceable contracts that would be beneficial to them. It should be noted that this reasoning alleviates the concerns voiced by de Tocqueville (1840) at the onset of modern democracy, by showing that the state is perfectly capable of (and justified in) providing support to those in need, should a more individualistic society dissolve the traditional links of solidarity between individuals. Furthering the argument in a more formal and specific setting, to which I will come back at several points in this thesis, Boldrin and Montes (2005) show that public financing of education and pensions can replicate an efficient complete market allocation even in the absence of a market for human capital acquisition.

Data and measurement issues Having recognized that the state is heavily involved in intergenerational transfers, and the reasons for this involvement, there remains to know which age segments of the population are contributing, and which are benefitting, to and from these transfers on average. Determining the direction and size of these flows of funds in cross-section is relatively easy, as data and publications on the absolute and relative expenditures on these large governmental programs are readily available. Knowing which generations benefit most from these transfers across their lifetime is a more difficult exercise. Luckily, Auerbach et al. (1994) have designed a method to evaluate the net contributions of each cohort to the burden of fiscal policy.[6] They construct the net present value of all contributions made by each cohort of the population to all levels of government, minus the transfers received. Projections of future taxes and transfers are obviously needed to obtain these generational accounts. Such methods are useful both to the economist interested in normative and political analysis, as it allows to assess the fairness of these state-sposored transfers of income, and to determine who are the winners of the political game to direct these flows to their benefit. Even though I shall not make a direct use of their methodology, as I am not performing any empirical analysis myself in this thesis, it is worth mentioning that these tools exist, should the models developed henceforth be put to the data during further research.

Tools used from the political economy literature As the title of this thesis makes it obvious, I will use a political economy perspective when studying how some of the

[6] An important difference with my definition of the intergenerational transfers lies in their non-imputation of government spending on goods and services that could be assigned to specific generations.

most important intergenerational public transfers are determined.[7] This approach considers that the heterogeneous agents populating the economy have potentially divergent interests on matters of public policy, based on their preferences or endowments, and posits the existence of political institutions to resolve this conflict of interest and reach a collective decision. The first task of the political economist is to find how to model the way in which individual preferences aggregate into a collective decision on a given policy. Out of the many tools available to represent the political institutions (direct and representative democracy, interest group and lobbying models, models of bargaining inside a legislature), the best solution has to balance plausibility and analytical tractability, as well as being well-suited to the policy and situation considered. In the original research contributions of this thesis, I systematically use models of electoral competition between purely office-motivated parties to represent the way in which transfer policies are selected. Yet, depending on the timing and nature of policy decisions, the way into which policy proposals will translate into votes will be modelled differently, which translates into the use of different equilibrium concepts in chapters 3 (probabilistic voting) and 4 (Condorcet winner).

Another element worth mentioning rightaway is that I adopt throughout this thesis an applied theory approach, in the sense that my intent is to model the different policy programs, individual behaviors and characteristics, and political institutions in an abstract enough fashion. Doing so ensures that the models that are developed in this thesis lend themselves to analytical or quasi-analytical treatment, and allow to formulate clear qualitative predictions whenever possible. By deliberately ignoring some salient features, I have traded away the possibility to make quantitative assessments, for instance of the impact of population ageing on various aspects of pension programs (eligibility conditions, degree of redistributiveness...) that are not represented in the model used in chapter 3. This would naturally constitute an interesting way forward for research on the matters covered here.

Cross-sectional and intergenerational redistribution At this stage of analysis, it should be noted that all public programs transferring resources between generations inevitably entail some degree of redistribution within cohorts of the population as well. Indeed, these programs are never financed in practice through lump-sum contributions, so that individuals belonging to a given contributory generation do not share equally the

[7] I refer the reader to the first chapter of Drazen (2000) for an excellent overview of the specific methodological approach of political economy.

burden of financing the program. In a similar fashion, no program consists in making the same transfer to all members of a given cohort, whatever their individual circumstances. Recipients of transfers in a given program are usually different than non-recipients, especially when the program consists in the delivery of a private good and not income. For instance, it can be argued that individuals enrolled in higher education come from a better-off background than the average of the population: therefore, public financing of higher education entails some redistribution from the well-endowed to the less-endowed individuals, a fact pointed out already in Arrow (1971). Income transfer programs are also affected, even when the benefit formula aims at replicating an insurance scheme: one may for instance mention the existence of a cap on pay-as-you-go public pensions in France, which breaks the actuarial fairness of the system.[8] Finally, the means-testing of certain pure transfers such as child allowances necessarily entails some redistribution of income within cohorts, even though they also fall under the definition of a downward intergenerational transfer.

All these examples demonstrate the need to investigate the within-cohort aspect of redistribution that these programs operate, on top of the between-cohort one. Given the questions I consider and the type of tools I am using in this thesis, this constitutes an undeniable methodological challenge, as I need to take into account the fact that economic agents are heterogeneous in several dimensions (age and a measure of innate skill in chapters 3 and 4). Finding the equilibrium of the political game then automatically becomes harder, as there is no simple median voter to be found, as would be the case if heterogeneity were unidimensional. What is more, the usual median voter result will also fail to apply as soon as the set of policies considered is multidimensional: this will be the case in both chapters 3 and 4, but also in section 8 of chapter 2 where I review the impact of population ageing on programs other than social security.

0.3 Outline of the thesis

The remainder of this thesis is divided in three chapters, of which two are original research contributions and one consists in a critical review of previous work.

First chapter The first chapter of this thesis, jointly written with Georges Casamatta, reviews the latest developments in the political economy literature dedicated to the study

[8] Source: U.S. Social Security Administration (2014).

of the consequences of population ageing, with a primary focus on the threat posed by ageing to the continued existence of public pension programs in developed countries. After briefly recalling why pay-as-you-go public pensions are supported by a political majority in the first place, we turn to reviewing works that analyse how a drop in fertility or mortality rates will change the contribution rates and pension sizes at the political equilibrium, by first assuming a constant retirement age. Other theoretical works are discussed, that are mainly concerned with endogenizing the retirement age choices, and exploring the opportunity to transition to a fully-funded system. We also present empirical assessments of the relationship between a population age structure and the size of its pension programs, as tests of the various theories outlined above.

Additionally, we explore the impact of population ageing on the political support for other public programs, such as education, heath care, capital taxation, or environment protection. We also give an account of some empirical analyses of the joint determination of education and pension programs.

Second chapter The second chapter of this thesis 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 a three-overlapping-generation model with production and intragenerational heterogeneity, in which the two intergenerational transfers are funded through taxes on the working generation, and benefit respectively the young (for education) and the elderly (for pensions). 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 to sustain the transfers between generations in equilibrium, I show that the emergence and continuation of intergenerational transfers 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; it is also forecasted to lead to higher pension and education spending per recipient, under plausible parameter values. I put these predictions into perspective with historical data on the secular evolution of public spending. From a methodological point of view, I also study how important the assumptions regarding expectations are, by considering expectations of future contribution rates that are linear in future levels of production per capita, as opposed to non-strategic voting where individuals don't take into account the impact of their current choices on future policy choices.

Third chapter In the third chapter of this thesis, I study under which conditions political processes might lead to an efficient level of public subsidies to education acquisition, when the space of available policies also includes labor income redistribution, which distorts said education acquisition. To this end, I set up a two-period political economy model where individuals differing in their intrinsic abilities privately choose their levels of education expenditure and labor supply, and decide collectively on a linear education subsidy and a linear labor income tax.

The flows of public funds for both programs are shown to be larger when the distribution of pre-education intrinsic abilities is more unequal and the elasticity of labor supply to taxation is lower. They also depend on the wage returns to education. When the government is unable to commit to a labor income tax rate before individuals select their level of education, I show that the electorate collectively selects too high a level of education subsidies in the first stage of the voting game, even when compared to the high level of labor taxation selected in the second stage. In that respect, the existence of high degrees of public subsidies to education does not necessarily signal the existence of virtuous institutions promoting long-term growth, but may rather be understood as proof of the existence of time inconsistency at the level of the electorate, when setting taxes and subsidies.

I also extend the analysis by setting up an OLG model with dynastic altruism where agents publicly invest in their children's education. I show that the politico-economic equilibrium can be characterized in the same way as the sequential voting game described above, in the limit case of perfect altruism. Even though the two policies are decided simultaneously in this case, the relevant policies for a given generation are voted upon in successive periods, thus leading to the same equilibrium characterization.

1

The Political Economy of Population Ageing: a Critical Review of the Recent Literature (joint with Georges Casamatta)

1.1 Introduction

1.1.1 The political challenges of population ageing

1.1.1.1 Facts on ageing, and political consequences

Population is aging rapidly in all regions of the world, and even faster if one considers the most developed countries. Individuals aged 65 years or more represented 5.1% of the world population in 1950 and 7.7% in 2010. According to recent projections (United Nations, 2013), this ratio should rise to 15.6% in 2050.

What are the reasons for this evolution? Population ageing finds its roots in two contemporaneous phenomena: a drop in mortality and fertility rates. In developed countries, the drop in fertility has mainly occurred before the 1980s and 1990s, and the fertility rate seems to have stabilized since then. In the least developed countries, however, there remains considerable room for further fertility reductions. The global total fertility rate fell from approximately 5 children per woman in 1950 to just over 2.5 in 2010, and the UN projects that it will fall to 2.2 children per woman by 2050 (United Nations, 2013). Most of the yet-to-come decline will occur in the developing world; this will contribute to a near halving of the share of children in the population of developing countries between 1965 and 2050. The drop in mortality, or increase in longevity, is however a long-lasting trend. Average world life expectancy at birth has risen from 46.5 years in 2050 to 65.4 percent in 2000. It is forecasted to continue to increase up to 74.3 years by 2050.

The aging of the world population introduces several major policy challenges (Bloom et al., 2010). People aged 60 or above usually have different needs and behaviors than younger individuals. Older individuals tend to work and save less, meaning that they offer less labor and capital to economies. They also require more health care and, in many countries, rely on social pensions for a large part of their income. As older populations become larger and politically stronger, adopting certain policies (such as cutting health and pension benefits) will prove difficult, as the older generations will be more able to impose their views on political agendas, be it through their sheer weight in the electorate, or by mobilizing themselves through lobbys or interest groups (Hanley, 2012; Mulligan and Sala-i Martin , 1999). Sinn and Uebelmesser (2003) even estimate that in one of the countries most concerned by ageing, Germany, the age of the median voter will increase so quickly that it will soon become extremely hard to secure a political majority to pass pension reforms: according to them, "gerontocracy" will be a fact in Germany as soon as 2016. Individuals aged 80 or over also have different needs. With declining health the need for full-time, long-term care increases. In many cases, this also increases the need for financial support, as private savings tend to vanish rapidly for individuals with particularly long lifespans. As their numbers increase, they place further demands on government resources, familial resources, and personal savings.

1.1.1.2 Policies at stake with ageing

In this chapter, we will focus on the impact of population ageing on public programs such as social security, health care and education, that entail substantial transfers of resources between generations. It should be noted from the outset that our survey will not explore the diversity of institutional frameworks under which these programs are shaped, nor the interplay between political parties, interest groups, social movements or expert advice in defining and reforming them. We feel that the study of these important subjects is best left to political scientists (see Pierson (2007), Goldstone et al. (2011) and Vanhuysse and Goerres (2012)), and that the need to make quantitative predictions on the future evolution of these key programs requires that the political economy literature should stick to models where individual preferences aggregate into a collective vote on a limited number of features of the programs, and abstract away from the particularities of each country's and each policy's institutional features.

The main part of the chapter will be devoted to social security. Ageing has a direct and dramatic impact on the functioning of unfunded, or Pay-As-You-Go (PAYG), pension systems. In these systems, pension benefits are financed with the contributions of the workers. Population ageing implies that the proportion of recipients increases while the proportion of contributors decrease, and thus threatens the financial viability of the system. The magnitude of this phenomenon is captured by the change in the old-age dependency ratio, which is defined as the ratio of elderly (aged 65 years or more) to adult individuals (aged between 18 and 64 years): according to the United Nations (2013), it will jump from 11.7% in 2010 to 24.7% in 2050 for the world taken as a whole, and from 23.8% to 44.4% for developed economies[1] over the same time frame. Meier and Werding (2010) give an order of magnitude of the increased burden that would represent

[1] As defined in the report cited above: Europe, Northern America, Australia, New-Zealand and Japan.

if the system (and notably the replacement rate)[2] was left unchanged: for a subset of OECD economies, the increase in pension spending to GDP would spread from 3% (for the UK) to 21% for Poland, over the 2000-2050 period. In many cases, this pension-spending-to-GDP ratio would be more than doubled over the period, should no reform be enacted.

Together with these changes in the age composition of the population, most OECD countries have also experienced a large drop in the labor force participation of middle-aged and elderly workers, which contributes to aggravating the financial difficulties of pension systems. As we will see later in this chapter, this drop is mainly due to the design of social security systems, which induces people to retire early, as well as other programs of the welfare state (amongst which unemployment benefits and disability insurance).

Confronted to this demographic evolution, pension systems thus need to be reformed. A substantial part of this chapter will be devoted to the political sustainability of such reforms. In a democratic society, this amounts to identifying the reforms which are likely to receive the support of a majority of the voters. Two kind of reforms of PAYG systems can be envisioned. The first one consists in keeping the system unchanged, but simply adjust its parameters, that is the contribution rate, the pension benefit level or the retirement age. This is called a *parametric* reform. The other possibility consists in changing the system and move to a fully funded (FF) system.

An analysis of the political sustainability of social security under ageing needs to examine the individuals' position on social security. Preferences over social security typically depend on an individual's age - since different cohorts of people have different remaining periods of contributions and benefits, but also on the redistributive design of the system. To assess the political sustainability of social security, one has to aggregate these preferences into a collective choice procedure. In this chapter, we focus almost exclusively on majority voting.

Once we have determined how these individual preferences shape the collective choice that arises out of majority voting, we can study how population ageing affects this democratic choice. Demographic dynamics impact the majority voting equilibrium over social security in essentially two ways. It modifies individuals' preferences - through for example the change in the rate of return of the pension system or equilibrium prices.

[2] The replacement rate of a pension system is defined as the average ratio of individual pensions to wages before retirement.

But it also changes the *identity* of the decisive voter: when population ages, the median voter becomes older and therefore individuals at or close to retirement get more political power.

1.1.2 Features of social security systems and recent reforms

Most industrialized countries feature an unfunded social security system that collects contributions from the workers' labor income and uses the revenue raised to provide pension benefits to current retirees. The first unfunded public program of retirement income was introduced in Germany in the late nineteenth century; several other countries followed at the beginning of the twentieth century - often with the creation of small funded systems targeted to workers in specific sectors. By the end of World War II, most systems had become unfunded. Until the beginning of the nineties, these systems have constantly grown, either because of the extension of the coverage among workers or of the increase in the generosity of pension benefits. Since then, governments, recognizing the long-term financial effect of the ageing process, have started to adopt retrenching measures.

1.1.2.1 Features of social security systems

Social security systems can be classified into four broad categories, as a combination of two key features: (a) pension systems are either fully funded (FF) or unfunded (henceforth referred to as pay-as-you-go or PAYG), (b) a system can provide payments based on either defined benefits or defined contributions. A system is deemed "defined benefits" (henceforth DB) if the benefits accrued at retirement are predetermined based on a formula that takes into account mainly life earnings, years of contribution and age, while a "defined contribution" system does not guarantee future pension levels, which are calculated by applying a (market-based or fictitious) rate of return on contributions.

Funded defined-contribution system Chile's social security reform in 1981 remains the best known international example of a FF defined contribution system. Under the Chilean system, all workers are required to contribute 10 percent of their salary into a savings plan of their choice, which is administered and regulated by the Administradora de Fondos de Pensiones. Eligibility for retirement is based on age and early retirement is available to those with sufficient accumulated savings. At retirement, workers can choose monthly withdrawals or purchase an annuity. Furthermore, workers are guaranteed a minimum pension paid from the general revenue fund. The benefits of such a system include reduced exposure to political and demographic risks.

Several other countries, including most of Latin America, have a funded defined-contribution pillar that follows Chile's example. Valdes-Prieto (1998) presents a summary of the reforms in Peru, Colombia, Argentina, Bolivia, Mexico, El Salvador, and Uruguay. He offers five reasons why Chile's model is so successful, including low levels of private-sector corruption, little political pressure on investment options, and successful implementation of a redistributive means-tested benefit to workers not covered by the Administradora de Fondos de Pensiones.

In Australia also the system is funded with defined contributions. The peculiarity of this system is that is that it offers the choice of either a lump-sum payment or an annuity at retirement. The U.K. system also offers a privatized, funded defined-contribution system but a unique one, in that it allows workers to opt out of their public, unfunded, defined-benefit system.

Funded defined-benefit system More traditional pensions, similar to those awarded to older U.S. workers during previous decades, are good examples of funded defined-benefit systems. Workers pay into the pension system, and the corporation manages how these contributions are invested. Workers then receive a defined benefit at retirement, which is usually based on years of service or some other related measure. Switzerland currently offers a hybrid system: a funded defined-contribution system with a guaranteed return.

Unfunded defined-benefit system A publicly operated, unfunded defined-benefit plan constitutes the first pillar of social security among most countries. Pension benefits may be granted to every individual that complies with the age and contribution range requirements, or may be means-tested (such that only workers below an income threshold are eligible).

Unfunded defined-contribution system Sweden and Italy are concrete examples of countries with an unfunded defined-contribution social security system. In recent years, both countries have switched to a so-called NDC ("notional defined contribution") plan. The government credits each worker for the taxes he or she and the employer contribute, and then pays upon retirement a benefit equal to the worker's contributions plus a notional (i.e. not market-based) interest rate.

Pension benefits and eligibility Benefits are computed based on the number of years of contributions and on a reference wage, which typically depends on the worker's past wages. However, even countries with a defined-benefit system differ in how pensions relate to the reference wage and in how this reference wage is obtained. France, Germany and Spain feature a tight link between wages and benefits. In these so-called Bismarckian systems, the benefit formula is constructed so as to entitle the retirees to a pension income that replaces a certain share (called the replacement rate) of their previous labor income. On the contrary, the United Kingdom exhibits an essentially redistributive system (often referred to as Beveridgian). The basic state pension is not related to any reference wage, and depends mostly on the number of years of contributions only.

Benefits are most of the time indexed on inflation or to the net wage growth, to reflect the increase in the cost of living. Eligibility to these benefits may depend on the years of contributions and/or on a minimum retirement age.

Retirement age All countries feature an official retirement age, when people are allowed, if not forced, to exit the labor market and receive their pension benefits. Most countries also have early retirement provisions that allow workers to retire before the official age on a reduced pension benefit. On top of that, some countries allow workers to enjoy full pensions even if they did not contribute to the system during the required amount of years, once they reach an age that is slightly higher than the official retirement age.

1.1.2.2 Recent reforms

It is outside the scope of this chapter to provide a detailed account of the state of pension program reforms in any economy taken individually, as every country exhibits specific provisions linked to the history of its public pension system development, the national public debate on the subject, and other idiosyncrasies such as how much the recent crisis hit public finances. We refer the reader interested by very up-to-date information on any particular country to periodic reports made by international or specialized institutions to cover the specifics of recent reforms on all aspects of pension programs. Notably, the "Social Security Programs Throughout the World" reports, issued by the American Social Security Administration jointly with the International Social Security Association, give detailed accounts of most features of the pension system of every country in a given area of the world, every six months on a rotating basis for the different regions (see e.g. the latest report for Europe: U.S. Social Security Administration (2014)).

Other international organizations are also publishing reports on the situation of pension programs in their member states for policy coordination purposes, see for instance OECD (2012), OECD (2013) for OECD countries, or European Commission (2010) for the EU. OECD (2013), for instance, examines pension reforms in its member countries between 2009 and 2013 against a set of six objectives: modification of coverage and eligibility criteria, adequacy of benefits for retirees to cover their basic needs, financial sustainability of the system, incentives to work longer and save more for retirement while active, administrative efficiency, and finally diversification of retirement income sources across providers, pillars and financing forms.[3] It stresses that all 34 OECD countries have undertaken reforms to further one or several of these objectives across the five-year study period. Coverage is shown to be quasi-universal in the surveyed countries (except Mexico), with certain countries granting extended pension coverage to subsets of the population such as mothers on maternity leave or family-carers. Adequacy has generally been tackled by reevaluating minimum and survivor pensions, and furthering the progressive nature of pension systems, either through means-tested components or by changing the benefits formulas away from flat-rate income replacement. The trend for financial sustainability is unequivocal: pension indexation has been either freezed or drastically slowed down, minimum retirement age is now tied to life expectancy almost everywhere, and benefits levels are increasingly tied to the whole contribution history rather than the last years (or best years in terms of wage) of an individual's career. European Commission (2010) also adds that eligibility criteria now more and more include a statutory number of contribution years on top of the minimum age. Recent reforms are also shown to be aimed at lengthening working lives, either through increases in statutory retirement age, provision of financial incentives to work while in retirement, or the retrenchment of generous early retirement schemes. These reforms are sometimes coupled with labor market reforms aimed at promoting employment of older workers from the labor demand side, with incentives given to firms that hire workers close to

[3]In this document, the term "provider" designates whether the subpart of a pension system is funded by the public or private sector, while the term "pillar" refers to whether the system is universal, personal or industry-based. The "financing form" naturally refers to the funded or unfunded (i.e. PAYG) nature of the system.

retirement. Finally, furthering diversification and security of the pension plans available to workers has been brought through the establishment of voluntary pension plans and increase in competition between private providers, or by relaxing restrictive regulations over investment choices made by the individuals (when saving to a private fund) and pension funds (to increase diversification of their portfolios). In this respect, European Commission (2010) notes an "increased complexity of pension systems [meaning] a transfer of risk from pension scheme sponsors to the beneficiaries".

Additionally, European Commission (2010) notices a trend towards the pre-funding of future pension outlays, which consists in frontloading parts of the adjustments costs linked to ageing in order to distribute them over a longer period and aver several generations. This pre-funding typically is made by establishing a pension reserve fund, paying down national debt,[4] or reforming the systems from defined benefits to defined contributions.

1.1.3 Outline

In a first step, we consider that the retirement age is fixed, so that a PAYG system is characterized by its contribution rate and its level of benefits, and examine in section 1.2 the reasons why a majority of the population may sustain a PAYG pension system. These reasons are the following. First of all, the economy may be dynamically inefficient. This implies that the rate of population growth exceeds the interest rate. In other words, the rate of return of the PAYG system is larger than the rate of return of a FF system (we assume for simplicity that there is no wage growth). It follows that individuals find the PAYG system to be a better "investment" opportunity than the FF system. Browning (1975), in his seminal paper, noted that the economy need not be dynamically inefficient for a majority of the population to support the PAYG system. His argument relies on the idea that past contributions to the social security system are a sunk cost for individuals. These latter, in evaluating the relative return of a PAYG and a FF system, compare the contributions that remain to be paid to the future pension benefits they expect to receive. It is then clear that an individual close to retirement receives a very high return from the PAYG system. If the median voter (who is the median age individual) is close enough to retirement, he will therefore vote for a PAYG system.

[4] Once we take into account the fact that the state is financing the deficits of universal unfunded systems in many countries, it becomes clear why reducing public debt is as effective as building up a pension reserve fund.

Previous arguments were developed in a partial equilibrium setting. In a general equilibrium, the introduction of a PAYG system depresses savings and thus makes the interest rate increase. This provides an additional reason for the voters to support PAYG. Besides, pension systems not only redistribute wealth across generations, they also redistribute within cohorts. As a consequence, a sufficiently redistributive PAYG system may be sustained by a coalition of the retirees and the poor workers. Finally, a PAYG system may constitute a device to insure individuals against the fluctuations of the interest rate.

The previous arguments were made under the once-and-for-all voting assumption: when voting on the contribution rate, individuals anticipate that the chosen tax rate will apply to their retirement period. A strand of the literature has shown that this assumption can be rationalized as the equilibrium outcome of an infinitely repeated game. In each period, the young voters sustain the system because of the fear to be punished by the subsequent generations. We present this game in section 1.3 .

We then turn in section 1.4 to the political determination of the retirement age. Two different frameworks are considered. In the first one, individuals freely decide when to retire. They however vote on some features of the PAYG system that affect their retirement decision. In the second framework, we consider that the majority vote applies to the retirement age, which is unique and common to all individuals.

Section 1.5 develops a recent literature on social security as a Markov perfect equilibrium. The payroll tax rate is assumed to be a function of some state variables, for example the capital stock. This creates a link between generations: if some generation decides to alter the tax rate in a given period, it will impact the capital stock in the next period and, through the Markov link, the tax rate in this next period.

A widely discussed reform of the social security system consists in moving from a PAYG to a FF system. This latter having a larger rate of return, many observers propose to change the system. Some authors have however shown that this transition cannot be Pareto improving, as it only results in making the implicit debt of the PAYG system explicit. Only when the PAYG system generates some distortions could the transition to a FF system be Pareto improving. We discuss this reform in section 1.6 .

Section 1.7 presents quantitative, rather than theoretical, analysis of the consequences of population ageing, by first reporting empirical evidence and then turning to simulation work. Section 1.8 deals with other public programs that are likely to be affected by population ageing, mainly health care and education. Lastly, section 3.6

concludes.

1.2 Why individuals support PAYG social security

1.2.1 Dynamic inefficiency

We present the basic two-period overlapping generations model that is used for analyzing pension policy. We consider a small open economy, so that the wages and the interest rate are given. Individuals live two periods. In any given period, two generations thus coexist: the young y and the old o. The size of each cohort at time t is denoted N_t^y and N_t^o respectively. Let assume that population grows at a constant rate n, so that $N_t^y = (1+n)N_t^o$. Lifetime utility of the generation born in t depends on consumptions in young and old ages and is assumed to be additively separable: $U(c_t^y, c_{t+1}^o) = u(c_t^y) + \beta u(c_{t+1}^o)$, where β is the discount factor. Assuming each individual earns the same income, normalized to 1, when young,[5] he chooses savings s_t in order to solve:

$$\begin{aligned} \max_{s_t} \quad & U(c_t^y, c_{t+1}^o) \\ st \quad & \\ & c_t^y = 1 - s_t \\ & c_{t+1}^o = s_t(1 + r_{t+1}) \end{aligned}$$

where r_t is the interest rate. This leads to the first-order condition on savings:

$$u'(c_t^y) = \beta(1 + r_{t+1})u'(c_{t+1}^o). \tag{1.1}$$

Consider now the introduction of a Pay-As-You-Go social security system. Each individual contributes a fraction τ of his income when young and receives a pension benefit p when old. The budget constraint of the system in each period is:

$$\begin{aligned} \tau N_t^y &= p N_t^o \\ \Leftrightarrow \quad p &= \tau(1+n). \end{aligned}$$

With the PAYG system, consumptions in both periods are $c_t^y = 1 - \tau - s_t$ and $c_{t+1}^o =$

[5] Wage growth is thus ruled out by assumption.

$s_t(1 + r_{t+1}) + \tau(1 + n)$. The effect of introducing a PAYG system at any time t_0 on the welfare of generations born after t_0 is given by:

$$\left.\frac{\partial V}{\partial \tau}\right|_{\tau=0} = -u'(c_t^y) + \beta(1+n)u'(c_{t+1}^o), \tag{1.2}$$

where V is the indirect utility function. From (1.1), we see that this expression is positive when $n > r_t$, a condition known as *dynamic inefficiency*. Noting that the introduction of the PAYG system constitutes a windfall for the old of the initial period, it is found that the PAYG system is *Pareto improving* when the economy is dynamically inefficient (Aaron, 1966 ; Samuelson , 1958). Whether the economy is dynamically efficient or not is a debated issue (Abel et al., 1989 ; Homburg , 1991). However there exist other reasons, that we examine in the next sections, why individuals support social security even when the economy is dynamically efficient.

1.2.2 Reduced time horizon

1.2.2.1 Central argument

Browning (1975) provided the first analysis of a majority vote over pensions. He considers a small open economy (in which the interest rate is given) with individuals differentiated according to age only (in particular there is no income heterogeneity). They live three periods, meaning that three generations coexist in each period, the young y, the middle-aged m and the old o. Under this setup, preferences are single-peaked over the payroll tax rate, implying that a Condorcet winner exists (Black, 1948).[6] This policy is the majority-voting equilibrium of a standard two-party Downsian electoral competition game (Roemer, 2001).

With three generations, noting that $N_t^o = (1+n)^2 N_t^y$ and $N_t^m = (1+n)N_t^y$, the budget constraint of the PAYG system writes:

$$\begin{aligned} N_t^o p_t &= N_t^y \tau_t + N_t^m \tau_t \\ \Leftrightarrow \quad p_t &= \tau_t(1+n)(2+n). \end{aligned}$$

It is assumed that there are no future re-voting opportunities. In other words,

[6] The Condorcet winner is the tax rate that is preferred by more than one half of the population, when confronted to any other possible tax rate.

individuals vote with the belief that the contribution rate chosen today will not be modified in the future ($\tau_{t+1} = \tau_t \equiv \tau$). Under this assumption, the optimal payroll tax rate of the decisive voter (who is a median-aged individual) solves:

$$\max_{\tau} \quad V_t^m(\tau) \equiv u(c_t^m) + \beta u(c_{t+1}^o)$$
$$st$$
$$c_t^m = 1 + s_t^y(1+r) - s_t^m - \tau$$
$$c_t^o = s_t^m(1+r) + \tau(1+n)(2+n).$$

He chooses a positive payroll tax rate if $\partial V_t^m / \partial \tau > 0$. Using the optimality condition on private savings, (1.1), this will be the case when $(1+n)(2+n) > 1+r$. In words, this implies that he may sustain the PAYG system even when the economy is dynamically efficient. The reason for this is that past contributions to the system are sunk cost. At the time of the vote, he compares the benefit of the system, $\tau(1+n)(2+n)$, to its cost, τ. The rate of return of the system is thus $(1+n)(2+n)$, which is larger than the rate of return of the young, $1+n$.[7] More generally, the closer individuals are to retirement, the higher the rate of return of the system. As a consequence, *optimal tax rates are increasing with age*. Noting that the steady state socially optimal tax rate is the one maximizing life-cycle utility at birth, Browning reaches the conclusion that the voted tax rate, that corresponds to the preferences of the middle-aged, is too high. In other words, *voting leads to a pension system excessively generous*.

In this stylized economy, middle-aged individuals vote for a 100% tax rate. They prefer to "invest" all their income in the social security system and finance current consumption by borrowing. There exist however obvious limits to the size of the system. First of all, when borrowing is constrained, agents need to keep resources for consumption in the working period (Boadway and Wildasin, 1989). The presence of uncertainty about future voting outcomes may also dissuade them to adopt too high tax rates (Hu, 1982). Finally, distortions caused by taxation are absent from Browning's analysis. With distortionary taxation, the voted tax rate lies strictly between 0 and 1 (Breyer, 1994).

[7] These latter vote for a zero tax rate in a dynamically efficient economy.

1.2.2.2 Population ageing in the Browning model

We consider the impact of population ageing in the framework just described. In this purpose we formulate a simple continuous version of this model. Assuming no time-discounting and no savings for simplicity, life-cycle utility of an aged a worker writes:

$$\int_a^R u(1-\tau)dt + \int_R^T u(p)dt = (R-a)u(1-\tau) + (T-R)u(p),$$

where R is the age of retirement and T the length of life.

The budget constraint of the PAYG system is:

$$p = \frac{F(R)}{1-F(R)}\tau,$$

where $F(.)$ is the c.d.f. of the age distribution and we denote $N^y = F(R)$ the number of workers (recall that the retirement age is assumed to be fixed). The optimal contribution rate of an aged a worker solves the following first-order condition:

$$-(R-a)u'(c^y) + (T-R)\frac{N^y}{1-N^y}u'(c^o) = 0. \tag{1.3}$$

Obviously, and for the reason explained in the previous section, $\partial\tau^y/\partial a > 0$, so that the majority voting tax rate is the preferred tax rate of the individuals with median age a^m.

Now consider a change in the fertility rate that makes the ratio $\eta \equiv N^y/(1-N^y)$ decrease. As a consequence the median age increases and the total effect on the majority voting tax rate is:[8]

$$\frac{d\tau^*}{d\eta} = \frac{\partial\tau^y}{\partial\eta} + \frac{\partial\tau^y}{\partial a^m}\frac{da^m}{d\eta}.$$

As explained before, the second term is negative. The first term is obtained by differentiating (1.3):

$$\frac{\partial\tau^y}{\partial\eta} = \frac{(T-R)u'(c^o)(1-\varepsilon)}{-D_\tau},$$

where $\varepsilon = -xu''(x)/u'(x)$ is the coefficient of relative risk aversion and $D_\tau < 0$ is the derivative of the first-order condition (1.3) with respect to τ. It is equal to the inverse of the inter-temporal elasticity of substitution.

Optimal tax rates increase with η *when* $\varepsilon < 1$. When η decreases, the rate of return

[8] We are interested here in the comparison of steady states and do not address the question of the demographic transition.

of the PAYG system becomes lower. In other words, the price of second-period relatively to first-period consumption increases and individuals are induced to substitute first- for second-period consumption. This is achieved by decreasing the tax rate. But the income effect goes into the opposite direction. With a high enough inter-temporal elasticity of substitution, the substitution effect dominates so that individuals react to a drop in the fertility rate by increasing the tax rate.

In this case, we see that the total effect of an increase in the dependency ratio (equal to $1/\eta$) on the equilibrium tax rate is ambiguous. The direct, economic effect calls for a reduction in the size of the system. But at the same time, in a greying society the increased political power of the old pushes toward a higher tax rate (see in subsection 1.7.2.1 a discussion of Galasso and Profeta (2004), who assess the magnitude of these two effects).

What is the consequence of an increase in life expectancy? This change should lead to a reduction in η and therefore have similar effects as previously. There is however an additional economic effect. Because individuals live longer, they should invest more resources in PAYG to ensure a decent standard of living at retirement:

$$\frac{\partial \tau^y}{\partial T} = \frac{\eta u'(c^o)}{-D_\tau} > 0.$$

Both the political and the economic effects go into the same direction. When population ageing follows from an increase in life expectancy, the size of the PAYG system increases.

1.2.3 Effect of social security on prices

Diamond (1965) extended the analysis of Samuelson (1958) to a model with production. In this context, the introduction of PAYG social security crowds out private savings and thus lowers the capital stock. This in turn implies lower wages and a higher interest rate. It follows that social security may, through its positive effect on the interest rate, be sustained by a majority of voters in a dynamically efficient economy (Boldrin and Rustichini, 2000 ; Cooley and Soares , 1999).

Consider the extension of the simple model in subsection 1.2.2.1 to a closed economy with wage in period t denoted w_t. The program of the representative young agent in period t remains the same with consumptions in both periods equal to $c_t^y = w_t(1-\tau)-s_t$ and $c_{t+1}^o = s_t(1+r_{t+1}) + \tau w_{t+1}(1+n)$. Denoting f the production function from per capita capital k, the equilibrium conditions on the production side are

$$\begin{aligned} w_t &= w(k_t) = f(k_t) - k_t f'(k_t) \\ 1 + r_t &= r(k_t) = f'(k_t) \\ k_{t+1} &= s_t/(1+n). \end{aligned}$$

The effect of the introduction of a PAYG system is now given by:

$$\left.\frac{\partial V}{\partial \tau}\right|_{\tau=0} = -w_t u'(c_t^y) + \beta((w_{t+1} + \frac{\partial w_{t+1}}{\partial k_{t+1}}\frac{\partial k_{t+1}}{\partial \tau})(1+n) + \frac{\partial r_{t+1}}{\partial k_{t+1}}\frac{\partial k_{t+1}}{\partial \tau}s_t)u'(c_{t+1}^o).$$

There are two new terms with respect to (1.2). The introduction of the PAYG system depresses savings and therefore $\partial k_{t+1}/\partial \tau < 0$. With less capital, the wage rate in period $t+1$ decreases (first term negative), but at the same time the interest rate increases (second term positive). If this second effect dominates, introducing a PAYG system has a positive effect on prices from a welfare perspective.

1.2.4 Within cohort redistribution

Even in a dynamically efficient economy, poor young individuals may support the PAYG system because its (intra-generational) redistributive properties (Casamatta et al., 2000; Tabellini , 2000). We consider a two-period OLG model of a small open economy. There is wage heterogeneity within cohorts: $w \in [w-, w+]$, with the mean larger than the median ($\bar{w} > w_m$) and agents are assumed to be credit constrained.[9] The pension formula is

$$p(w) = (1+n)\tau(\alpha w + (1-\alpha)\bar{w}),$$

where $\alpha \in [0,1]$ is a parameter that measures the redistributiveness of the pension system. A low value of α implies a highly redistributive, or Beveridgian, pension system (for $\alpha = 0$, everyone receives the same pension, whatever the level of contributions). Large values of α correspond to a contributory, or Bismarckian, system.

The analysis is conducted at steady state under the once-and-for-all voting assumption.

[9] In order to have interior optimal tax rates. See discussion at the end of section 1.2.1 .

1.2.4.1 Individually optimal tax rates

We first determine the tax rates that maximize the life-cycle utility of the different agents, considering first the retirees and then the workers.

The retirees Private saving is the result of past decision. They choose the value of τ, τ^o, that maximizes their consumption: $c^o = (1+r)s + (1+n)\tau(\alpha w + (1-\alpha)\bar{w})$. The solution is straightforward: $\tau^o = 1$, the same tax rate for all retirees.

The workers A worker with earning w chooses $\tau^y(w)$ in order to maximize

$$V(\tau, w) \quad = \quad u(w(1-\tau) - s^y) + \beta u((1+r)s^y + (1+n)\tau(\alpha w + (1-\alpha)\bar{w}))$$

where $s^y \geq 0$ is the optimal level of private saving.

If $1 + r >$(resp.$<$) $(\alpha + (1-\alpha)\bar{w}/w)(1+n)$, the individual optimal policy involves $\tau^y = 0$ (resp. > 0) and $s^y > 0$ (resp. $= 0$). Put differently, for an individual to prefer private saving to PAYG pensions, his wage must be strictly higher than $\hat{w}$ defined as:

$$\hat{w} = \frac{1-\alpha}{\frac{1+r}{1+n} - \alpha}\bar{w} \leq \bar{w}.$$

One easily checks that $\hat{w} = \bar{w}$ if $n = r$, $\partial\hat{w}/\partial n > 0$ and $\partial\hat{w}/\partial\alpha < 0$. One can also check that the indirect utility function $V(.)$ is concave in τ, meaning that preferences are single-peaked and therefore that the median voter theorem applies.

1.2.4.2 Variation of the optimal tax rates with the wage level

Let us now take a look at how the optimal tax rate changes with individual wage $w \leq \bar{w}$. Differentiating the first-order condition on τ^y,

$$-wu'(c^y) + \beta(1+n)(\alpha w + (1-\alpha)\bar{w})u'(c^o) = 0, \tag{1.4}$$

we obtain

$$\frac{\partial \tau^y}{\partial w} = \frac{\beta(1+n)(1-\alpha)\frac{\bar{w}}{w}u'(c^o)(1-\varepsilon)}{V''(\tau)}$$

where ε is the coefficient of relative risk aversion. Optimal tax rates increase with wages if and only if $\varepsilon < 1$. A variation of the wage level generates an income and a substitution effect. When the wage increases:

1) The individual is richer. Because consumption in the second period is a normal good, he wants to increase it. This is achieved by increasing the tax rate (income effect).

2) The price of first-period consumption with respect to second-period consumption, $(1+n)(\alpha w+(1-\alpha)\bar{w})/w$, decreases. By this effect, richer individuals are induced to buy more first period consumption and this is achieved by reducing the tax rate (substitution effect).

When the inter-temporal elasticity of substitution is small ($\varepsilon > 1$), the income effect dominates: individually optimal tax rates are increasing with wages.

1.2.4.3 Majority voting tax rate

The workers are divided into two classes, those who prefer a zero tax rate and positive savings and those who prefer a positive tax rate and no saving. A fraction $1/(2+n)$ of citizens, the retirees, is in favor of $\tau^R = 1$. Furthermore, all workers with earnings above $\hat{w}$ are in favor of a zero tax and the preferred tax rate for the workers with earnings below $\hat{w}$ increases (resp. decreases) with w when $\varepsilon > 1$ (resp. < 1). For the majority voting tax rate to be positive, it must be that the number of individuals who want a positive rate is larger than half the total population:

$$\begin{array}{lcl} & N_t^y \int_{w_-}^{\hat{w}} f(w)dw + N_t^o & \geq \dfrac{N_t^y + N_t^o}{2} \\ \Leftrightarrow & \int_{w_-}^{\hat{w}} f(w)dw & \geq \dfrac{n}{2(1+n)}. \end{array}$$

Under this condition,[10] the majority voting equilibrium tax rate is the rate preferred by the workers with earning $\tilde{w}$, such that the number of people who prefer a higher tax rate is exactly half the total population. For $\varepsilon > 1$ (increasing tax rates), it solves:

$$\begin{array}{lcl} & N_t^y \int_{\tilde{w}}^{\hat{w}} f(w)dw + N_t^o & = \dfrac{N_t^y + N_t^o}{2} \\ \Leftrightarrow & \int_{\tilde{w}}^{\hat{w}} f(w)dw & = \dfrac{n}{2(1+n)}. \end{array} \tag{1.5}$$

This corresponds to an *ends-against-the-middle* equilibrium (Epple and Romano, 1996): a coalition of the middle-class and the retirees support a larger PAYG system whereas

[10]Note that this condition is always satisfied when $n \geq r$. It could however be violated in the converse case.

the poor and rich workers would like to downsize it.

For $\varepsilon < 1$ (decreasing tax rates), $\tilde{w}$ is implicitly defined by:

$$\int_{w_-}^{\tilde{w}} f(w)dw = \frac{n}{2(1+n)}. \tag{1.6}$$

1.2.4.4 Effect of a drop in the fertility rate

A change in the fertility rate has both a direct and an indirect effects on the majority voting tax rate. The direct effect is obtained by differentiating the first-order condition (1.4):

$$\frac{\partial \tau^y}{\partial n} = \frac{\beta(\alpha w + (1-\alpha)\bar{w})u'(c^o)(1-\varepsilon)}{-V''(\tau)}.$$

Optimal tax rates decrease with n when $\varepsilon > 1$. When n decreases, the rate of return of the PAYG system becomes lower. In other words, the price of second period relative to first period consumption being larger, individuals are induced to substitute first for second period consumption. This is achieved by decreasing the tax rate. But the income effect goes into the opposite direction. With a low inter-temporal elasticity of substitution, the income effect dominates so that individuals react to a drop in the fertility rate by increasing the tax rate.

Following a change in the fertility rate, not only optimal tax rates are modified but also the *identity* of the decisive voter. For a low inter-temporal elasticity, we differentiate (1.5):

$$\frac{d\tilde{w}}{dn} f(\tilde{w}) = \frac{1}{2(1+n)^2} - \frac{d\hat{w}}{dn} f(\hat{w}).$$

Noting that $\partial\hat{w}/\partial n > 0$, this effect has an ambiguous sign.

In the case of a high inter temporal elasticity, we differentiate (1.6) to obtain:

$$\frac{d\tilde{w}}{dn} f(\tilde{w}) = \frac{1}{2(1+n)^2} > 0.$$

The total effect on the majority voting tax rate is given by:

$$\frac{d\tau^*}{dn} = \frac{\partial \tau^y}{\partial n} + \frac{\partial \tau^y}{\partial w}\frac{d\tilde{w}}{dn}.$$

When $\varepsilon > 1$ the direct effect is negative and the indirect one ambiguous. When $\varepsilon < 1$

the direct effect is positive and the indirect negative. Therefore, one cannot conclude unambiguously about the effect of a drop in fertility. The direct and indirect effect go in opposite directions and one needs to evaluate the magnitude of these countervailing effects to be able to conclude about the direction of the change.

1.2.5 Risk-sharing

In a dynamically efficient economy, a Fully-Funded system provides a higher return than PAYG in expectation. However fluctuations in the interest rates make investments in the FF risky. In the absence of private markets for insurance, the PAYG may constitute an insurance device. This occurs when returns to capital and wages are imperfectly correlated. This argument has been formalized by Bohn (2001) and Krueger and Kubler (2006). We present here a simple version of the model by Krueger and Kubler (2006). In this model, agents live two periods but value consumption in the second period only, according to the utility function $u(.)$. The indirect utility function writes:

$$V(\tau) = E\left(u[(l-\tau)wR + \tau wG]\right),$$

where R and G are the (stochastic) returns of PAYG and savings respectively, and $E(.)$ is the expectation operator. It is welfare improving to introduce a PAYG system if $dV/d\tau > 0$. With a logarithmic utility function and a joint log-normal distribution for R and G, this condition reduces to:

$$E\left(\frac{G}{R}\right) = \frac{E(G)}{E(R)} \frac{[cv(R)^2 + 1]}{[\rho_{G,R}.cv(G).cv(R) + 1]} > 1,$$

where cv designates the coefficient of variation and ρ the correlation coefficient. This condition can be met even when the expected return from PAYG is lower than savings. This occurs if the stochastic savings returns are very volatile ($cv(R)$ is large) or the correlation between G and R is small.

1.3 The social contract

1.3.1 Sustainability of the PAYG system

In the Browning analysis and subsequent papers on the vote over pensions, it is assumed that agents vote once and for all. This assumption, while convenient, is strong. Why young individuals should support the system, if the tax rate is re-voted tomorrow and there is therefore no direct link between the contributions they make and the benefits they receive? The answer is that there exists an implicit contract between generations: the young support the system because of the threat to be punished by future generations. If they break down the system, they incur the risk of receiving no pension benefits in old age (although of course they could rely on savings). This idea has been first formalized by Hammond (1975) and then adapted to the pension game by Sjoblom (1985) and Boldrin and Rustichini (2000).

To illustrate the argument, we consider an infinitely repeated voting game with individuals living two periods. Assuming a positive rate of population growth, the young are more numerous than the old and therefore are decisive in the vote. Indirect life cycle utility is denoted $v_t^y(\tau_t, \tau_{t+1})$.

Consider the sequence of tax rates $(\tau_i^*)_{i=1}^{\infty}$ and the strategies[11]

$$\sigma_s^y = \begin{cases} \tau_s^* & if\ \tau_{s-i} = \tau_{s-i}^*\ for\ all\ i = 1, ..., s-1 \\ \\ 0 & otherwise. \end{cases}$$

The social contract prescribes the choice of τ_i^*. According to this strategy profile, young individuals comply with this implicit contract if everybody has done the same in the past. Otherwise, they break the contract by choosing a 0 tax rate.

Denote $(v_i^*)_{i=1}^{\infty}$ the resulting payoffs for the young. We have to check two conditions to obtain a subgame perfect equilibrium.

1) No player should want to deviate in any time period. The best deviation for a young in period s is to choose $\tau_s = 0$. By choosing a tax rate different from τ_s^*, he breaks the contract and receives no pension. He has therefore no interest in contributing to the

[11] Players of this game are generations of individuals. In other words, it is assumed that there exists a coordination mechanism for individuals belonging to any given generation.

system. Consequently, if $v_s^* \geq v_s^y(0,0)$, a young individual has no interest in breaking the contract.

2) Subgame perfection: if player t deviates, player $t+1$ must have an incentive to punish him. Because once a player has deviated, the system is permanently abolished, player $t+1$ will receive no pension and should thus not contribute: punishment is credible.

This leads to the conclusion that the strategies above constitute a subgame perfect equilibrium as soon as $v_s^* \geq v_s^y(0,0)$, $\forall s$. This condition simply states that the young are better off with the social security system implemented. It is satisfied in the different cases described in section 2.[12] It should however be kept in mind that there exist many other equilibria.[13] This game thus generates a high degree of indeterminacy. It should be envisaged mainly as a rationale for the use of the once-and-for-all voting assumption.

1.3.2 Dynamics of the political equilibrium

In the recent historical experience of many countries the immediate cause for the general alarm surrounding the social security system seems to be the long-run fall in the growth rates of population and labor productivity. Boldrin and Rustichini (2000) show that, even if one knows that the PAYG system will be dismantled in a finite future and therefore that some generation will contribute to it but will never receive any benefit, this system can be sustained by a majority of the voters. This follows from introducing in the previous game a stochastic process for the growth rate of population. They consider a sequence of growth rates $\{n(j)\}_{j=0}^{\infty}$ satisfying $n(j+1) < n(j)$, for all j, and $\lim_{j\to\infty} n(j) = 0$, and a transition probability

$$\begin{aligned} \Pr(n_{t+1} = n(j) | \, n_t = n(j)) &= 1-p \\ \Pr(n_{t+1} = n(j+1) | \, n_t = n(j)) &= p, \end{aligned}$$

with $0 < p < 1$.

With a linear production function, $f(k) = ak + b$, and logarithmic utility, they show that there exist an equilibrium with a sequence of tax rates $\{\tau(j)\}_{j=0}^{\infty}$, such that $\tau(j) > 0$

[12] This equilibrium could easily be extended to the case of three generations analyzed by Browning (1975).

[13] In particular, *no social security is always an equilibrium* : if a player believes that the following generation will not contribute, it is optimal not to contribute, even if it is Pareto improving that every generation contributes.

for some j and $\tau(j+1) < \tau(j)$ for all j. This means that a PAYG is sustainable even though one knows that it will be dismantled for sure in the future (but not when). The condition for having this equilibrium is $a < (1-p)(1+n_0)$. Noting that $a = 1+r$ in this model, the intuition is clear: even though there is a risk of not receiving pensions in the next period, the expected return of the PAYG system dominates private savings (as long as $a < (1-p)(1+n_t)$).

1.4 Endogenous retirement

Pensions systems are characterized by three key parameters: the payroll tax rate, the benefit level and the retirement age. So far, we have considered the retirement age to be given and common to all individuals. We address in this section endogenous retirement choice. There are two main modeling assumptions: either the retirement age is chosen collectively through a vote; or individuals decide individually when to retire.

1.4.1 Implicit taxation and early retirement

1.4.1.1 The political support for early retirement provisions

In the last thirty years, most OECD countries have experienced a dramatic drop in the labor force participation of their middle aged and elderly workers. In the OECD countries, the average labor force participation rate of male workers aged between 55 and 64 years has decreased from 84.2% in 1960 to 63.2% in 1990.

The extent to which male elderly workers have decreased their participation in the labor market may also be captured by the reduction in the average retirement age.

A comprehensive study on eleven OECD countries edited by Gruber and Wise (1999) suggests that generous early retirement provisions are largely responsible for this drop in the (male) participation rates. Gruber and Wise (1999) and a parallel study by Blondal and Scarpetta (1998) identify two features of the early retirement provisions, which display a strong correlation with the departure of the elderly workers from the labor force: the early (and normal) retirement age and the tax burden which is imposed on the labor income of the individuals who continue to work after reaching the early retirement age. Gruber and Wise (1999) and Blondal and Scarpetta (1998) argue that individuals are often induced to retire early because of the large implicit tax imposed on continuing to work after early retirement age. Agent's early retirement decision thus represents the optimal response to the economic incentives provided by the social security system.

Conde-Ruiz and Galasso (2003) develop a positive theory of why early retirement age provisions have been introduced in the first place and sustained over time. They analyze a majority voting game over two dimensions: the payroll tax rate and the decision to introduce or not an early retirement provision. Focusing on the (simultaneous) issue-by-issue voting equilibrium,[14] they show that early retirement is sustained by a coalition of the young poor and of the old with an incomplete earning history. The latter obviously sustain the system as it makes them eligible to pension benefits. The former vote for early retirement because, due to substitution effects between leisure and consumption, they tend to retire earlier than the rich. It should be noted that the equilibrium is self-sustained over time (or subgame perfect) as, by retiring earlier, the current poor young will continue to sustain the system when becoming old.

In a subsequent paper, Conde-Ruiz and Galasso (2004) note the (negative) macroeconomic consequences of early retirement provisions which, by inducing individuals to retire earlier, depress human capital accumulation and growth.

1.4.1.2 The implicit taxation on continued activity

Gruber and Wise (1999) point a second feature of social security, besides the existence of an early retirement age, that induces individuals to retire early: the implicit tax on continued activity created by the pension system. When an individual decides to work for one more year, he receives pension benefits for one year less. Furthermore he has to pay social security contributions for that year. Additional pension benefits less contributory taxes represent the *net social security accrual*. When it is negative, the pension system encourages early retirement through an implicit taxation on continued activity. In the words of Crawford and Lilien (1981), the pension system is not marginally fair.

Sheshinski (1978) and Crawford and Lilien (1981) were the first to analyze the individual retirement decision. They show that the the income and substitution effects

[14] The political process being multidimensional, there is no Condorcet winner (Plott, 1967). A possible solution is to determine the issue-by-issue voting equilibrium (Shepsle, 1979). This consists in deriving the voting equilibrium on each policy (for given the other policy). The full equilibrium is then given by the intersection point of these reaction curves.

of the implicit tax on continued activity go in opposite directions. Leisure being a normal good, the loss in income generated by the tax induces individuals to work longer. But, faced with a lower price of leisure in terms of consumption, they tend to substitute leisure for consumption, i.e. to retire earlier. Under the common assumption that the substitution dominates the income effect, the implicit taxation built in the pension system encourages individuals to retire earlier.

We here here the model by Crawford and Lilien (1981). They consider a continuous time model in which individuals work for time 0 to R and then retire from R to T. Assuming for simplicity no time discounting, life-cycle utility is:

$$\int_0^R u(c_t)dt + \int_R^T [u(c_t) + v]dt = \int_0^T u(c_t)dt + (T - R)v,$$

where v is the utility of leisure.

Normalizing wage income to 1 and assuming a zero interest rate and a perfect capital market, individuals wish to equalize consumption in all periods. Thus, the worker's problem can be written as

$$\max_{R,c} Tu(c) + (T - R)v$$

subject to

$$Tc \leq R(1 - \tau) + (T - R)p.$$

Social security benefits obey the following formula:

$$p = [(T - R(1 - B))A + RB\tau]/(T - R).$$

For $B = 1$, social security is marginally fair. For $B = 0$, benefits per period are a constant, $p = A$. In such case, the pension system implicitly taxes individuals: those who work one more year not only have to contribute for that year, but also loose one year of pension benefits. Substituting this formula into the constraint of the individual optimization problem yields:

$$c \leq (R/T)[1 - (1 - B)(A + \tau)] + A.$$

Assuming an interior solution, the conditions for a solution of the individual program are:

$$u'(c)\omega = v$$

and

$$c = (R/T)\omega + A,$$

where $\omega = 1 - (1 - B)(A + \tau)$ is the relative price of leisure in terms of consumption.

Differentiating these conditions, we obtain:

$$\frac{dR}{dB} = -\left[\frac{R}{\omega} + \frac{T}{\omega^2}\frac{u'(c)}{u''(c)}\right](A + \tau).$$

An increase of B raises the implicit price of leisure, with an ambiguous effect on R. There is a substitution effect, $-Tu'(c)(A+\tau)/\omega^2 u''(c)$, which tends to delay retirement; and an income effect, $-R(A + \tau)/\omega$, which acts to induce earlier retirement. If we make the usual labor supply assumption that the substitution effect of a shift in wages dominates the income effect, then $dR/dB > 0$.

Casamatta et al. (2005) develop a model with no income effects. The pension benefit formula is such that total benefits do not depend on the age of retirement. An implicit tax, or bias, is however present in the form of the contributory tax. They prove the existence of an issue-by-issue voting equilibrium with a positive bias. The bias is sustained by the poor (both young and old) who benefit from the (intragenerational) redistribution of the system. Numerical simulations show that the bias chosen at the political equilibrium might be excessive with respect to the second-best optimum.

Casamatta et al. (2006) distinguish between the two components of the implicit tax: the payroll tax rate τ and the reduction in pension rights, represented by a parameter α. When $\alpha > 0$ (resp. < 0), pension rights decrease (resp. increase) with the retirement age. They show that individuals prefer $\alpha < 0$, as well as a Rawlsian or Utilitarian planner. *A positive tax is however possible* when the planner maximizes a weighted sum of utilities and the weight on the rich is sufficiently high, which corresponds to a political process biased towards the rich. This is indeed the outcome of a “bargaining” between the poor and the rich. These latter ask for a low τ. In exchange they accept to set a positive α, which is a second-best instrument for the poor to redistribute income.

1.4.2 Voting on the retirement age

1.4.2.1 One dimensional voting

Lacomba and Lagos (2007) study the one dimensional vote on the retirement age, mandatory and common to all individuals, while other parameters of the pension system (payroll tax rate, benefit level) are held fixed. The main insight of their analysis is to emphasize the dependency of individually optimal retirement ages on the status quo retirement age. If the status quo retirement age is low, individuals anticipate to receive a small pension for a long period of time and therefore tend to save a lot. Therefore an individual who is close to the (current) retirement age at the time of the vote has accumulated a large wealth. Due to the associated income effect, he will tend to favor a lower retirement age than a young individual at the start of his working life. In the words of Lacomba and Lagos, the status quo retirement age “acts like a magnet”.

While Lacomba and Lagos (2007) consider individuals with different income levels, we present here a simplified version with only age heterogeneity (wage income is normalized to 1). Population is assumed to be constant and age uniformly distributed on $[0, T]$. The interest rate is assumed to be 0 and individuals do not discount the future. The vote on the legal retirement age takes place at an arbitrary moment of time t. It is unanticipated and the newly voted retirement age is believed by everybody to remain permanent. Denoting R^{sq} the retirement age before the vote, accumulated wealth of an aged a worker at the time of the vote is

$$\pi(a) = a(1 - \tau - c(R^{sq}; 0)),$$

where $c(R^{sq}; 0)$ is planned consumption at the beginning of the life-cycle when the anticipated retirement age is R^{sq}:

$$c(R^{sq}; 0) = \frac{R^{sq}(1-\tau) + (T - R^{sq})p(R^{sq})}{T}.$$

Substituting this expression in the former, we get:

$$\pi(a) = a(\frac{T - R^{sq}}{T})(1 - \tau - p(R^{sq})).$$

The level of pension benefits, for given τ and R, is determined by the budget constraint of the pension system:

$$p(R) = \frac{\tau R}{T - R}.$$

Replacing $p(R^{sq})$ in the previous expression gives

$$\pi(a) = a(1 - \tau - \frac{R^{sq}}{T}).$$

Accumulated wealth thus depends on age, life expectancy, the payroll tax rate and the status quo retirement age.

We now turn to the preferred retirement ages of the different individuals.

The retirees It is assumed that the retirees do not return to work if the newly voted retirement age is lower than their age. Therefore postponing the retirement age is always favorable to the retirees, as the number of contributors increases and there are fewer retirees who receive a pension. This implies that the *retirees always vote for the largest possible retirement age.*

The workers First of all, it is clear that the workers never prefer a retirement age lower than their own age: $R^*(a) \geq a$. Indeed choosing $R^*(a) < a$ or $R^*(a) = a$ in both cases result in these individuals being on retirement. But in the second case the dependency ratio is lower, meaning a higher level of pension benefits.

Now remind that, with a 0 interest and discount rates, individuals optimally choose a constant stream of consumption. The optimal consumption decision of an aged a individual when the retirement age is R is thus:

$$\begin{aligned} c(R;a) &= \frac{(R-a)(1-\tau)+(T-R)p(R)+\pi(a)}{T-a} \\ &= \frac{R-a(1-\tau)+\pi(a)}{T-a} \\ &= \frac{R-(a/T)R^{sq}}{T-a}, \end{aligned}$$

and his indirect utility function:

$$V(R;a) = (T-a)u(c(R;a)) + (T-R)v.$$

The first-order condition on R is

$$u'(c) - v = 0. \tag{1.7}$$

Differentiating, this implies

$$\frac{\partial R^*}{\partial a} = \frac{R^{sq}-R^*}{T-a}.$$

Then two cases are possible (see figure). Either $R^*(R^{sq}) > R^{sq}$, in which case optimal retirement ages decrease with age, or $R^*(R^{sq}) = R^{sq}$ and the converse holds.

The status quo retirement age acts as a magnet: as workers get older, their optimal legal retirement ages are closer to the status quo age. The reason is the following. Denote $\tilde{R} = R^*(0)$. In the case where $R^{sq} < \tilde{R}$, individuals have less total lifetime income than if the retirement age were , and thus less consumption per time period. It follows that an individual born with the status quo situation have *over-accumulated* with respect to a newly born and will *over-consume* for the remainder of his life. Looking at the first-order condition (1.7) on the optimal retirement age, this explains why he will prefer a lower retirement age. A symmetric reasoning can be made for the case $R^{sq} > \tilde{R}$.

Aggregation of preferences through majority voting It can be shown that preferences are single-peaked, implying the existence of a Condorcet winner. If optimal retirement ages increase with income, a coalition of the retirees and middle-age workers will join to support a higher retirement age. For decreasing optimal retirement ages, the retirees will form a coalition with the youngest workers.

Population ageing Lacomba and Lagos (2006) analyze the impact of a change in the fertility rate on the optimal retirement age of individuals at the beginning of the life-cycle. We present here a simplified two-period model, in which individuals spend a fraction $1 - z$ of the second period on retirement (z is thus interpreted as the retirement age).[15]

Life-cycle utility is

$$\begin{aligned} & u(c_t^y) + \beta \left[u(c_{t+1}^o) + v(1-z) \right] \\ = \quad & u(1-\tau-s) + \beta \left[u(s(1+r) + z(1-\tau) + p(1-z)) + v(1-z) \right], \end{aligned}$$

where $\beta = 1/(1+r)$. Assuming a perfect capital market, individuals equalize first- and second-period consumptions:

$$c = \frac{(1-\tau)(1+r+z) + p(1-z)}{2+r} \tag{1.8}$$

[15] Contrarily to the continuous time model analyzed above, we consider non-zero discount and interest rates, denoted respectively β and r.

and the optimal retirement age solves

$$\max_z \tfrac{2+r}{1+r}u(c) + \tfrac{1}{1+r}v(1-z).$$

The budget constraint of the PAYG system is

$$\begin{aligned} & N^y\tau + N^o z\tau && = N^o(1-z)p \\ \Leftrightarrow \quad & \tau(1+n+z) && = (1-z)p. \end{aligned}$$

In a defined contribution pension system, the pension benefit level depends on the value of the tax rate:

$$p(\tau) = \frac{\tau(1+n+z)}{(1-z)}$$

while in a defined benefit system, this is that the tax rate that adjusts:

$$\tau(p) = \frac{(1-z)p}{(1+n+z)}.$$

Using (1.8), we obtain

$$c^\tau = \frac{1+r+z-\tau(r-n)}{2+r}$$

and

$$c^p = \frac{1+r+z-\frac{(1-z)p}{(1+n+z)}(r-n)}{2+r}.$$

The impact of a change in the retirement age is different in DB or in a DC system, unless $r = n$. It can indeed be shown that:

$$\frac{\partial c^\tau}{\partial z} = \frac{1}{2+r}$$

and

$$\frac{\partial c^p}{\partial z} = \left[1 + p\frac{2+n}{(1+n+z)^2}(r-n)\right]/(2+r).$$

The change in consumption in a DC system is "neutral": if individuals work one more year, their lifetime consumption increases by $1/(1+r)$.[16] The change is however not neutral in a DB system. As soon as $r > n$, consumption increases by more than the change in income. The reason is that when p is fixed, total pension $p(1-z)$ is reduced when individuals postpone retirement. This means that the amount of resources devoted to the PAYG system (the "size" of the system) decreases. These resources can be invested

[16] Lifetime consumption is $c + c/(1+r) = c((2+r)/(1+r))$.

at a better return in private savings when $r > n$. Observe that the optimal retirement age differs in DC and DB. This latter is the solution of the first-order condition:

$$(2+r)\frac{\partial c}{\partial z}u'(c) - v = 0.$$

It is then clear that *the optimal retirement age is larger in a DB system.*

It should be observed that the very reason for such a difference between the two systems is that the DB system is biased in the sense that the net pension wealth $p(1-z) - \tau z$ (equal to $\tau(1+n)$ from the budget constraint) varies with z. In a DC system on the contrary it is fixed and depends on the chosen value of the contribution rate τ.

The impact of a change in the population growth rate in a DC system is straightforward:

$$\frac{\partial z^*}{\partial n} = \frac{\tau u''(c)}{-D_z} < 0.$$

The intuition is simple. Due to a drop in the fertility rate, the PAYG system has a lower rate of return. Therefore, lifetime consumption is reduced. Leisure being a normal good, individuals choose to postpone retirement to counter this (negative) income shock.

1.4.2.2 Joint determination of the retirement age and the contribution rate

Galasso (2008) considers the joint determination of the size of the PAYG system (given by the contribution rate) and the retirement age through an issue-by-issue voting procedure. He calibrates the model on economic data for the year 2000 in various OECD countries and uses forecasted values to construct the politico-economic equilibrium in year 2050.[17]

Due to population ageing and absent any political reaction (i.e. for unchanged contribution rate and retirement age), the economy has to adjust: pension benefits drop and individuals increase their savings, making the stock of capital larger. As a consequence wages increase and the interest rates fall dramatically. These aggregate effects are dampened when political reactions are taken into account.

The political equilibrium is given by the intersection of two reaction curves: the majority voting tax rate for a given retirement age and the majority voting retirement age for a given tax rate.

[17] See subsection 1.7.2.2 for a quantitative assessment of the effects of ageing when contribution rates and retirement age are voted upon at the same time, on a model very similar to the one discussed here.

How does the majority voting tax rate react to population ageing? For a given retirement age, population ageing is known to have two opposite economic and political effects on the determination of the contribution rate (see Casamatta et al. (2000), Razin et al. (2002), Galasso and Profeta (2004), Galasso and Profeta (2007)). On the one hand aging reduces the profitability of the system. Individuals then tend to ask for a lower tax rate But on the other hand, ageing makes the median voter older and thus closer to retirement. It is well-known since Browning (1975) (see section 1.2) that this latter then chooses a larger tax rate. The overall effect is hence ambiguous. Simulation results reported in Galasso suggest that *the political push dominates in all countries*: aging shifts the reaction function $\tau(R)$ upward.

How does aging modify the individual preferences over retirement — given the social security contribution rate? Again, two main effects occur. Aging reduces the average profitability from social security, which generates a reduction in the lifetime income of all generations. This negative income effect encourages individuals to postpone retirement. But for a given social security contribution rate, aging also reduces the pension benefits level, since fewer resources have to be shared among more retirees. This amounts to a negative substitution effect (retirement consumption becomes more expensive with respect to leisure), which reduces the pecuniary incentives to retire early and leads again to an increase in the retirement age. Therefore, through these negative income and substitution effects, *aging unambiguously generates a political push for postponing retirement*: the reaction functions $R(\tau)$ shifts upward.

Two main insights emerge from this simulation exercise. First, the retirement age increases dramatically in all countries. As emphasized above, the substitution and income effects associated to ageing both go in the same direction and push for delaying retirement. This is reinforced by the economic consequences of ageing: the opportunity cost of retiring, as represented by the wage rate, increases. Second, in all countries but Italy, both the contribution rate and the level of benefits increase. The rise in the contribution rate is not that surprising: in the face of ageing, individuals must invest more in social security to maintain a decent level of retirement consumption. Moreover the interest rate is lowered, reducing the gap between the profitability of savings and PAYG social security. The increase in the levels of benefits is more surprising. While it is difficult to disentangle the various effects at play in these numerical simulations, the aging of the median voter, and the associated increased demand for social security, is certainly one of the main reasons driving this result.

1.5 Markov perfect equilibria: Dynamics of the political equilibrium

The game described in section 1.3 has many equilibria. Furthermore, it is quite poor in making predictions and does not allow to do comparative statics. A way to select among these equilibria is to consider Markov strategies: the action played in a given period only depends on some fundamental state variables, such as the level of the capital stock or the ratio of the number of retirees to workers.[18] This approach has generated a considerable interest in the last years. We review in the next sections the main work in this direction.

1.5.1 A median voter model with capital as the state variable

Forni (2005) has developed a model close to Boldrin and Rustichini (2000). It is a standard two-period overlapping-generations model. Agents work, consume, save and pay taxes in the first period. They are retired, consume, and receive old age benefits in the second one.

A logarithmic utility function is assumed, as well as a Cobb-Douglas production function: $Y = K^{\alpha}L^{1-\alpha}$. In per capita terms: $y = k^{\alpha}$ with $k = K/L$. It is assumed that capital depreciates completely in one period.

The utility maximization program of a young agent in period t is:

$$\max_{c_t^y, c_{t+1}^o} u(c_t^y, c_{t+1}^o) = \ln(c_t^y) + \beta \ln(c_{t+1}^o)$$

subject to

$$\begin{aligned} c_t^y &= w_t(1-\tau_t) - s_t \\ c_{t+1}^o &= s_t(1+r_{t+1}) + p_{t+1}. \end{aligned}$$

Given the Cobb–Douglas production function, the profit maximizing conditions imply that: $w_t = (1-\alpha)k_t$ and $r_{t+1} = \alpha k_{t+1}^{\alpha-1} - 1$. The budget constraint of the PAYG system also implies:

$$p_{t+1} = \tau_{t+1} w_{t+1}(1+n).$$

[18] See Fudenberg and Tirole (1991) for a precise definition of Markov Equilibria.

With a logarithmic utility function, the utility maximization program has closed-form solutions:

$$\begin{aligned} c_t^y &= \left(\frac{1}{1+\beta}\right)\left[w_t(1-\tau_t)+\frac{p_{t+1}}{1+r_{t+1}}\right] \\ c_{t+1}^o &= \left(\frac{\beta}{1+\beta}\right)\left[w_t(1-\tau_t)(1+r_{t+1})+p_{t+1}\right] \\ s_t &= \left(\frac{1}{1+\beta}\right)\left[\beta w_t(1-\tau_t)(1+r_{t+1})-\frac{p_{t+1}}{1+r_{t+1}}\right]. \end{aligned}$$

The last equilibrium condition is the capital formation dynamic equation:

$$(1+n)k_{t+1} = s_t.$$

Collective decisions are made through majority voting. The young being more numerous than the old, they decide on the payroll tax rate in each period. Strategies are assumed to be Markov and moreover to be time-invariant: the policy function is of the type $\tau = \tau(k_t)$. In words, the payroll tax rate chosen in period t only depends on the capital stock in this period. The reaction function in period t is found by solving:

$$\tau(k_t) = \text{argmax}_{0<\tau(.)<1}\Psi(k_t, \tau(k_t), \tilde{\tau}(k_{t+1})),$$

where $k_{t+1} = \Phi(k_t, \tau(k_t))$ and $\Psi(.)$ represents the indirect utility function. An equilibrium occurs when $\tau(.) = \tilde{\tau}(.)$. Forni (2005) shows that the equilibrium policy function is:

$$\tau(k) = \left(\frac{\alpha}{1-\alpha}\right)[Ck^{-(1+\beta\alpha)/(1+\beta)} - 1],$$

where $C \geq 0$ is a constant. Therefore there is a continuum of equilibria indexed by the free parameter C. The equilibrium policy function is non-increasing in k. To see this, suppose on the contrary that the payroll tax rate increases with k. Then young individuals should strategically increase savings to attract greater old-age benefits in the next period. But this cannot be an equilibrium, as the perspective of a larger benefit when old induces the young to reduce the level of savings.

In this setting, it is shown that, for some parameters values, *there exist equilibria with social security*. The reason why the young may sustain the system is that they save less if they contribute to the system. The capital stock in the next period being lower, the payroll tax rate and thus the level of pension benefits increases.[19] This can be seen

[19] Another reason why the young may support social security is the general equilibrium effect on

as a continuous version of Boldrin and Rustichini (2000), in which the young decision consists in keeping the pension system or dismantling it. The equilibrium strategy, which implies a reduction in payroll taxes as the capital stock increases, is simply the implementation of a punishment strategy in a Markovian context.

An increase in the rate of population growth n leads to a higher rate of return of the PAYG system.[20] This in turn induces the young to reduce their savings. The steady-state capital stock is thus reduced and the size of the pension system (i.e. the contribution rate) increases.

1.5.2 Probabilistic voting

Gonzalez-Eiras and Niepelt (2008) study a model formally close to Forni (2005). The main difference lies in the modeling of the political process in each period. While Forni (2005) considers simple majority voting, Gonzalez-Eiras and Niepelt (2008) use a probabilistic voting model. This results in the government maximizing a weighted sum of the welfare of the young and the old in each period. They focus on a particular Markovian equilibrium, in which the equilibrium tax rate does not depend on the savings level. Therefore, reputational effects are absent of the model. The only reason why the young may sustain the pension system is because it allows to "manipulate" prices: a larger payroll tax rate depress savings, implying a larger interest rate.

With a logarithmic utility function and a Cobb-Douglas production function, the model predicts that, when the population growth rate decreases, the payroll tax rate and the share of pensions in GDP increase, but the social security benefit per retiree may eventually decline. The main driving force behind these results is that, with population ageing, the old gets more political power. This explains the rising tax rate. With more retirees however this may not be sufficient to give a higher pension to all of them, all the more that the tax base (the wage rate) shrinks because of a reduced capital accumulation. Calibrating the model to the US economy, it is found that payroll tax rates should increase up to 16% in 2050, accompanied by a slight reduction in the pensions level (less than 9%).

prices: a drop in the capital stock makes the interest rate increase. This effect has been pointed out in the literature by Boldrin and Rustichini (2000) or Cooley and Soares (1999).

[20] Observe that in this simple two-period model, this does not affect the "identity" of the median voter, who is always a young individual.

1.5.3 The evolution of retirement

Conde-Ruiz et al. (2005) study a three-period OLG model, with young, adult and old individuals living in each period. The young work and provide labor l on the intensive margin. The adults provide labor z on the extensive margin: this corresponds to early retirement. It is assumed for simplicity that adults do not receive pensions. The wages of the young and the adults are denoted w^y and w^a respectively. Old individuals are on retirement and receive a pension p. Population grows at a rate n.

It is assumed that individuals only consume when retired. Old age consumption of an individual born in period t is thus:

$$c_{t+2} = (1 - \tau_t) l_t w_t^y (1 + r)^2 + (1 - \tau_{t+1}) z_{t+1} w_{t+1}^a (1 + r) + p_{t+2}.$$

Individuals want to consume as much as possible. They also incur some disutility of work. Their utility function is thus:

$$U(c_{t+2}, z_{t+1}, l_t) = c_{t+2} - \frac{z_{t+1}^2}{2\gamma} w_{t+1}^a - \frac{l_t^2}{2\alpha} w_t^y.$$

Assuming $\alpha = 1/(1 + r)^2$ and $\gamma = 1/(1 + r)$, optimal labor supply functions are:

$$\hat{l}_t = 1 - \tau_t$$

and

$$\hat{z}_{t+1} = 1 - \tau_{t+1}.$$

The PAYG system is budget balanced. The pension benefit is thus:

$$p_t = \tau_t (1 - \tau_t)(1 + n_{t-1}) \bar{w}_t,$$

where $\bar{w}_t = (1 + n_t) w_t^y + w_t^a$.

Substituting the optimal labor functions and pension benefit level into the utility function, we get the indirect utility function of an adult individual at time t ,$V^a(\tau_{t-1}, \tau_t, \tau_{t+1})$.

The authors determine the Markov perfect equilibrium of the repeated voting game on payroll tax rates, in which the state variable is the share of elderly retirees in the population. In every period t, the median voter in each generation - an adult individual as long as $n_t(1 + n_{t-1}) < 1$ - decides his most favorite social security system (i.e. the tax rate τ_t). In taking his decision, he expects the future social security tax rate to depend on the current level of early retirement, according to a function $\tau_{t+1} = Q(z_t)$. Therefore the median voter's optimal decision can be obtained by maximizing his life-cycle utility with respect to τ_t and given expectations on the next period policy function $\tau_{t+1} = Q(Z(\tau_t))$:

$$\max_{\tau_t} V^a(\tau_{t-1}, \tau_t, Q(Z(\tau_t))).$$

A Markov political equilibrium is then a pair of functions (Q, Z) such that:

(i) $Q(z_{t-1}) = \arg\max_{\tau_t} V^a(\tau_{t-1}, \tau_t, Q(Z(\tau_t)))$

(ii) $Z(\tau_t) = 1 - \tau_t$

The link between generations is made though the function Q: when choosing the tax rate, the period t median voter affects the share of early retirees (through the function Z). This in turn has an influence of the period $t+1$ median voter's choice.

Conde-Ruiz et al. (2005) show that the equilibrium policy function Q satisfies:

$$\begin{aligned}
\tau_{t+1} = Q(Z(\tau_t)) &= \frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{2A - 2(1+r)w_t^a Z(\tau_t)^2}{(1+n_t)\bar{w}_{t+1}}} \\
&= \frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{2A - 2(1+r)w_t^a (1-\tau_t)^2}{(1+n_t)\bar{w}_{t+1}}},
\end{aligned}$$

where A is a free parameter.

A current increase in the social security contribution rate leads to more current early retirement - by increasing the implicit tax on continued activity - which in turn creates expectations for more social security contributions - and hence more early retirement - in the future.

The equilibrium path converges to a stable steady state with a positive social security contribution rate and early retirement.

What is the effect of aging, caused by a reduction of the fertility rate, in this model ? Suppose that the median voter at time t learns that population growth will permanently drop from the following period on: $n_{t+i} < n_t \ \forall i > 0$. Since no change occurs in the policy function at time t - $\tau_t = Q(Z(\tau_{t-1}))$ only depends on n_{t-1} and n_t - the social security contribution rate is set in accordance with the expectations of the period $t-1$ median voter. Yet the tax rate in $t+1$ depends on n_{t+1} and is thus modified, as well as the whole sequence of future tax rates. Starting from a steady state (with $\bar{\tau}$, $\bar{p}$ and $\bar{z}$), the contribution rate increases, $\tau_{t+1} > \tau_t = \bar{\tau}$, people retire earlier, $z_{t+1} < z_t = \bar{z}$, while the per-capita pension is unaffected, $p_{t+1} = p_t = \bar{p}$. At steady state, the contribution rate is larger, $\bar{\tau}' > \bar{\tau}$, implying $\bar{z}' < \bar{z}$.

This political economy analysis thus suggests that aging will lead to a larger social security system and more early retirement. This result was however obtained with a quasi-linear utility function with only substitution effects on labor supply taken into account. Conde-Ruiz et al. (2005) note that the introduction of income effects may mitigate these results: the loss of income following the drop in the PAYG rate of return induces individuals to work more and therefore to postpone retirement.

1.6 The transition to a Fully Funded system

Previous sections considered only *parametric* reforms of the PAYG system when population ages. Another type of reform, which we will call *systemic*, would consist in transitioning to a fully-funded (FF) system, where current contribution are saved to provide for benefits later in the life cycle.

Due to the drop in fertility rates, population grows at a slower rate. This implies a lower rate of return for the PAYG system. Therefore many believe that it is desirable to move to a fully funded system. The mere comparison of steady states is however misleading. It has been shown that, absent labor supply distortions, the transition to a FF system cannot be Pareto-improving (Breyer, 1989): when moving to the FF system, the implicit debt of the PAYG system is simply made explicit; the PAYG system is nothing more than an inter-generational redistribution scheme. In other words, some generations gain from the reform but others loose. Only when there are some labor market distortions can the reform yield a Pareto-improvement (Breyer and Straub, 1993 ; Homburg, 1990 ; Rangel , 1997).

We illustrate this argument with a simple two period representative agent OLG model in a small open economy. Population grows at a constant rate n and wages at a constant rate g. The budget constraint of a young individuals at time t is:

$$c_t^y + s_t \leq w_t l_t (1 - \tau_t),$$

where l_t is labor supply and w_t is the wage rate. At time $t+1$ the budget constraint of this individual is:

$$c^o_{t+1} \leq (1+r)s_t + p_{t+1},$$

where a constant interest rate r is assumed for simplicity.

Assuming a perfect capital market,[21] the intertemporal budget constraint can be written:

$$c^y_t + \frac{c^o_{t+1}}{1+r} \leq w_t l_t (1-\tau_t) + \frac{p_{t+1}}{1+r}.$$

In a PAYG system, pension benefits are

$$p_{t+1} = (1+n)\tau_{t+1} w_{t+1} l_{t+1}.$$

In a FF system, contributions are capitalized at the market rate of return:

$$p_t = (1+r)s^{ff}_{t-1}$$

With a perfect capital market, it does not matter whether contributions are compulsory or not. If they are in excess of what the individuals would freely do, these latter could dissave to achieve their optimal level of savings. Therefore the intertemporal budget constraint writes:

$$c^y_t + \frac{c^o_{t+1}}{1+r} \leq w_t l_t.$$

Considering a PAYG system with a constant contribution rate over time, $\tau_{t+1} = \tau_t = \tau$, and moreover fixed and constant labor supply, $l_{t+1} = l_t = l$, pension benefits in a PAYG system are:

$$p^{payg}_{t+1} = (1+n)(1+g)\tau w_t l$$

and therefore the rate of return of a PAYG social security system is equal to $(1+n)(1+g)$. Clearly, if $1+r > (1+n)(1+g)$, individuals will benefit more from the FF system in the steady state. Empirical data tend to confirm that, on average, the return on private investment is larger than the growth rate of a stationary economy, $(1+n)(1+g)$.

However this is not the answer to the question we are interested in. We wish to know whether, starting from a situation with a PAYG system, it is possible to operate a transition to a FF system that improves the welfare of all generations. We show now that it is possible to make the transition in a Pareto-neutral way. This suggests that no Pareto improvement is possible. The reader can refer to Breyer (1989) for a formal proof.

[21] This implies in particular that there are no borrowing constraints.

Consider a reform that occurs at time T. The mechanism to obtain a welfare-neutral transition is the following. The retirees at time T receive the promised pension p_{T+1}^{payg} so that they are indifferent between reforming the pension system or not. This pension is financed by a mix of time T workers contributions and government debt. The new contribution rate at time T is such that:

$$\tau_t^* w_t l = \tau w_t l - p_{T+1}^{payg}/(1+r)$$

and the government borrows $p_{T+1}^{payg}/(1+r)$.

Time T workers invest $p_{T+1}^{payg}/(1+r)$ in the FF system. Their pension at time $T+1$ is thus p_{T+1}^{payg} and they are also indifferent. This pension is financed by a mix of time $T+1$ workers contributions and government debt in the same way as before, so that generation $T+1$ is also indifferent. And so on until infinity. One can calculate that the present value of the government debt in period T is the same as the implicit debt of the PAYG system.

It should be noted that the tax rate in the new scheme, τ_t^*, is lower than τ. This means that, when labor supply is endogenous, the new scheme is less distortive and that it is therefore possible to use this efficiency gain to raise strictly the welfare of all generations. Homburg (1990) goes even further by noting that these efficiency gains can be used to reimburse the government debt in finite time. It is thus possible to completely abolish PAYG and to restore the economy's net wealth to the level which is associated with a pure capital reserve system.

Conesa and Garriga (2008) use these results to construct numerical simulations in which all generations benefit from the reform (transition generations receive no pension but are compensated with a debt-financed transfer).[22] They however consider a representative-agent economy. When individuals differ with respect to labor income within cohorts, efficiency gains have to be compared to the insurance benefit of the PAYG system (Conesa and Krueger, 1999 ; Nishiyama and Smetters , 2007).

[22]Before them, all quantitative analyzes have found welfare gains in the long run, but losses for the transition generations. Conesa and Garriga (2009) follow the same optimal taxation approach. They however do not address the question of the transition to a FF system, but maintain the PAYG system and look for the optimal adjustment of its parameters following a temporary demographic shock .

1.7 Quantitative analysis

1.7.1 Empirical estimations

Several empirical estimations of the relationship between the age structure of population and the share of income devoted to pension programs - and other types of social spending as well - have been conducted in recent years. We detail their most salient features hereafter.

Lindert (1994) provides a long-term, historical point of view on the link between ageing and pension transfers, by providing panel estimates on industrial countries for the period 1880-1930. The role played by demographics to explain the rise in social spending over the period is measured, including controls for income levels and growth, political determinants (indicators of democratization, women's suffrage, rate of executive turnover) and religious affiliation. He shows that 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 should induce an increase in 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 transfers in GDP. The main results are confirmed in a similar analysis of the 1960-1980 period by Lindert (1996): over this period, a higher share of older people in the 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. We also briefly mention Durevall and Henrekson (2011) as a more recent analysis of the relationship between social spending and the age structure of the population based on long-run historical data.

Breyer and Craig (1997) conduct an analysis on 20 OECD countries over the 1960-1990 period, using pension spending over GDP or average pension per capita as dependent variables. Among the explanatory variables, the age of the median voter in each country and the ratio of the population aged 40-60 to the elderly are meant to capture the effects of ageing. A strong and significantly positive effect of the age of the median voter on the size of pension programs is found in all specifications. The estimate implies that an increase in median voter age by one year would lead to a 0.5% increase in the share of social security payments to GDP, to be compared with a sample mean of 9% for this variable. The support ratio (reduced to the 40-60 age group for the numerator) has a negative effect on pension program size, when it is added into the regression. Another paper also using OECD panel data, Sanz and Velazquez (2007), indicates that public spending in several areas of government (social welfare, health, but also education and defense) increases with the share of the elderly of the population. However, benefits per capita are not reported to be affected in any significant way. Profeta (2002) reports similar findings using a large cross-country dataset which is not restricted to OECD countries: when the share of elderly in the population increases, total social security expenditure increase. The share of elderly in the overall population has more contrasted effects at the individual level, however: while the average length in retirement is reported to increase, and the labor force participation of the elderly decreases when population ages, the pension annuity is, again, not affected.

The empirical relationship between population structure and pension spending has been at the center of a controversy, started by a paper by Razin et al. (2002). They estimate a panel data of the U.S. and 12 European countries over the period 1965-1992, with the average tax rate on labor income and the per capita transfers (including pensions, unemployment and disability compensation) as dependent variables. The explanatory variables used are the dependency ratio (measured as one minus the share of labor force in the population), a measure of income skewness, real growth in GDP and other controls. Both regressions yield a negative impact of the dependency ratio on the dependent variable. Indeed, in the main specification, a 1% increase in the dependency ratio is shown to lead to a 0.4% decline in the labor tax rate: this means that the 4% decline in the dependency ratio that happened over the 1970-1991 period accounts for roughly 1.5% of the 11% increase in the labor tax rate then. Again, all specifications point to a higher dependency ratio leading to lower per capita government transfers: over the same period, the decrease in the dependency ratio accounts for 30% of the increase in transfers. Razin et al. (2002) also introduce a theoretical model to account for the non-structural relationships they find. In this two-period OLG model, a single labor tax is used to finance a lump-sum transfer to both generations, which is used by the young to acquire education and by the old as a pension. In this context, while the young want some tax to finance their education, the chosen level is limited by the fiscal "leakage" going to the old. Therefore when population growth increases (equivalent in the model to a reduction in the dependency ratio), this leakage is reduced, potentially leading to a higher chosen tax rate if the distortionary effect of taxation is not too strong on the opposite side.

The findings of Razin et al. (2002) have been questioned by Bryant (2003), Disney (2007), Simonovits (2007) and Shelton (2008), mostly for the use of the dependency ratio as a measure of population ageing. In particular, Disney (2007) proceeds to a new test of the model introduced by Razin et al. (2002) on a dataset of 21 countries for the 1970s to 1990s decades, redefining the variables used in the regression: the old-age dependency ratio (defined as the share of 65+ to the 15-64 year-old population) and the inverse of the support ratio (namely, the share of pensioners to people participating to the labor market) are used instead of the simple dependency ratio, and the share of labor taxes in GDP and the average contribution rate to social security. In all specifications, the measures of age dependency are this time positively associated with tax rates, with strong statistical significance in the case of the pension contribution rate: the typical specification indicates that the pension contribution rate paid by workers[23] would increase by around 0.6% if the old-age dependency ratio were to be increased by 1%, on average. Moreover, by introducing proxies that capture to which degree social security systems link contributions and entitlements, Disney (2007) shows that the positive relationship between ageing and tax rates becomes weaker when contributions made by individuals start to match less their future entitlements. In the limit case where current contributions bear no link to future benefits (as is the case in the model by Razin et al. (2002) where decision on tax rates is purely static), this relationship may then be reversed. In practice, though, most pension systems around the world entail some actuarial component.

1.7.2 Simulating the future political evolution of social security

1.7.2.1 Political sustainability and the evolution of the contribution rates for the current systems

Galasso (2006) is, to the best of our knowledge, the only successful attempt so far at simulating the evolution of the social security systems of several key countries[24] until 2050, by taking into account both the economic and the political effects of changes in the

[23]In this paper, the contribution rate is defined as the ratio between the social security replacement rate and the support ratio of the pension scheme, and is meant to measure the share of the average worker's labor income that needs to be paid to balance the pension scheme.

[24]The countries concerned by the simulation exercise are: the USA, the UK, Germany, France, Italy and Spain.

age structure of population.[25] To this end, the author builds a closed-economy model with several overlapping generations of agents that differ in age and education, supply work elastically, save and consume. The pension system is modelled in such a way that its budget is balanced in each period, with a fixed age of retirement. The other main institutional features of each national system considered are either held fixed across the years or not taken into account, to ensure tractability. This way, policy choices are reduced to a single dimension in each period: by voting on a pension contribution rate τ_t, all the other features of the system at date t are determined. The equilibrium of the voting game is sustained by stationary strategy profiles as in section 1.3 , with the median voter (in age but also in education) choosing the equilibrium value of τ_t in each period.

The methodology used to investigate the effects of ageing is twofold. First, the author assumes that each economy is initially at steady state at the starting point of the simulations (around the year 2000), and calibrates the model parameters to target the key variables of the economy then (such as the dependency ratio, pension contribution rate, labor force participation, capital-output ratio). In a second step, the author then feeds the model with forecasted values of the relevant economic and demographic variables (among which labor productivity growth, population growth and median voter age) for the year 2050. The output of the simulations is, for each country and for two different effective retirement ages in 2050 (either unchanged with respect to 2000 or increased to 65 years old), an equilibrium pension contribution rate (which is voted upon) and the corresponding replacement rate.

In the words of the author of the study himself, "these simulations give a gloomy picture of the future of social security: under the political constraints imposed by a graying electorate, pension expenditure is forecasted to increase in all countries" (op. cit., p. 207). In Germany, Italy and Spain, the (sometimes massive) increase in contribution rates is however not enough to yield higher replacement rates absent any increase in the effective retirement age, due to the extent of ageing happening in these countries. Higher replacement rates are achieved in all countries only when the effective retirement age is postponed to 65. Yet, even in this case, most countries see an increase of their contribution rate over the 50-year time span of the simulations, albeit less stark than if the effective retirement age is left untouched. Indeed, for instance, Galasso (2006)

[25] It should be noted that Galasso and Profeta (2004) was actually the first attempt at simulating the politico-economic dynamics of social security systems recorded in the literature. However, Galasso (2006) presents, among other things, an updated version of the simulations performed in Galasso and Profeta (2004), which is why we do not report Galasso and Profeta (2004)'s results here.

computes that postponing the retirement age in Italy or France from 58 (the effective age in both countries at the time of the study) to 65 years would limit the increase in the contribution rate by 11 and 12%, respectively. In Italy, this would even be enough to allow the contribution rate to decrease from 38% to 35,5%, instead of increasing to 46,3% if the retirement age was kept untouched.

Interestingly, as the model incorporates savings and capital, the composition of retiree income between private and public pension plans can be computed: it is forecasted that future retirees will rely more on public pension programs than before, probably given to the fact that they are able to extract more resource from the working generations under the new circumstances.

1.7.2.2 Political sustainability with endogenous retirement age

Galasso (2006) also tackles the next question that comes to mind:[26] since raising the retirement age is the only way to prevent contribution rate hikes from being massive and replacement rates from decreasing sharply, is there a political majority to raise this retirement age in the concerned countries? To answer this question, the author extends the model discussed in the previous section to allow individuals to vote both on the contribution rate and the retirement age. Given that the multi-dimensionality of the policy space may be troublesome for the existence of a Condorcet winner, the author models the voting process as an issue-by-issue simultaneous voting game, which yields as output of the model a structure-induced political equilibrium (see Shepsle (1979)).

Under this setup, it is clear that the age profile of preferences for the retirement age will be similar to the one studied in section 1.4 , holding the contribution rate fixed. Besides, if the agents are faced with a higher contribution rate, the opportunity cost of retiring will be lower, meaning older agents will vote for a lower retirement age, all else being equal. The author mentions, however, that since "social security is a dominated asset, a higher contribution rate reduces the overall income of the young, hence inducing them to postpone retirement" (op. cit., p. 215). Moreover, population ageing will affect the preferences for the retirement age in the same way as described in subsection 1.4.2.2 : the retirement age R voted in equilibrium will then inevitably rise.

Quantitatively, the simulations feature rises in the effective retirement age that match almost one-to-one the projected rises in the age of the median voter over the contribu-

[26] Galasso (2008) essentially presents the same model and results, and has already been discussed from a theoretical point of view in subsection 1.4.2.2 , to which we refer the reader.

tion rate dimension. Depending on the severity of forecasted ageing, the retirement age is projected to be increased between 5 years (for the USA) and 9 years (for France in Italy) over the study period, thus closing a big part of the initial gap between countries. Notably, the political equilibria in this case feature retirement ages that are systematically above 65 years old, the benchmark used in the previous exercise. This seems to indicate that fears over the sustainability of our pension systems and their capacity to be reformed towards a longer worklife might be overrated. A side effect of this strong increase in retirement age is to mitigate the rise of social security contribution rates (from 9.7% in 2000 to a forecasted value of 13,5% in 2050 for the USA, for instance). Concerning the replacement rate, the results are quite contrasted and depend largely on the scope of ageing projected to happen in each country:[27] while it is forecasted to decrease from 98.9% to 69.4% in Italy between 2000 and 2050, it is expected to rise by modest amounts in France (from 55% to 63.3%) and the USA (from 40.8% to 46.1%) over the same time period.

1.8 The political impact of ageing on other public programs

Of course, the most widespread concern linked to ageing is the continuous funding of the generous pension programs that have been set up in the second half of the last century, for all the reasons outlined before. Nonetheless, other public programs may also be at stake due to the way in which they are currently funded (when these programs entail transfers from the young to the old generations), or simply because these public programs are not a priority for a graying electorate. In the following section, we review recent developments in the theoretical and quantitative analysis of the political impact of ageing on these various policies, which range from education to health care, from environment conservation to the composition of taxes falling on labor and capital.

[27] Source: Galasso (2008).

1.8.1 Ageing and political support for education

1.8.1.1 Theory

With education only Kemnitz (1999) provides an early attempt at studying the impact of demographic structure on education subsidies, when the latter are politically determined. In his 2-period OLG model, the size of subsidies to privately purchased education is chosen so as to maximize a weighted sum of the utilities of the two generations (young and old) living in each period. Results are driven by a rather ad-hoc assumption made on these so-called political weights: they are assumed to be an increasing but concave function of the size of each generation: the rationale for introducing a function is that there may exist larger free-rider problems in bigger groups, when these groups try to organize themselves politically. In steady state, it is found that a decline in population growth leads to an increase in education subsidies. When the fertility rate (or the population growth rate, which is equivalent in the model) decreases, the decrease in the relative political weight of subsidy recipients is lower than the decrease in their share numbers beacause of the concavity of the political weight function. Since the relative size of the young with respect to the old is itself linked to the tax-price of subsidies, in equilibrium the political costs to the government of subsidizing education fall more than the political gains, and the government is led to increase these subsidies.

Gradstein and Kaganovich (2004) incorporate into a 2-period OLG model two opposing effects of population ageing. On the one hand, ageing increases the proportion of older voters, who cannot reap off the benefits of investing into the productivity of the future workforce. On the other hand, the currently working adults, upon which the current tax falls, can expect to benefit from their investment into future labor productivity through an increase in future PAYG pension payments, and increase interest payments on their savings. In their simple setup (logarithmic time-separable preferences for consumption when young and old, Cobb-Douglas production function, and a linear relationship between next generation's human capital and education spending), Gradstein and Kaganovich (2004) find that the second effect dominates, so that ageing should lead to increased education funding. However, by introducing a two-district version of this same economy, the authors are able to show that constituencies with a (exogenously determined) higher fraction of elderly would be less willing to finance education, at a given time period. The apparent contradiction in these two results is solved if one considers that in the two-district version of the model, education is financed locally, while productivity increases due to education are transmitted to the aggregate level (through perfect capital mobility between the districts). The discrepancy between the costs and benefits of education finance then explains why districts with more old people are more reluctant to finance public education in cross-section, as evidenced by Poterba (1997) (see subsection 1.8.1.2); however, it also shows that cross-sectional evidence has little external validity when one wishes to investigate the effect of population ageing on education spending at the aggregate level through time.

Education and general income redistribution Levy (2005) brings into the picture the interaction of education policy with general, non-age-related income redistribution. She sets up a model in which agents are not only heterogeneous in terms of age, but also in terms of income. It is shown that, all else being equal (and notably the pre-tax income distribution), a larger young group will lead to decreased education spending, which is crowded out by the untargeted income redistribution. The mechanism by which such results are attained is fairly complex, resting on endogenous party formation, and equilibrium policies are supported by coalitions of different age groups, the description of which would by far exceeds the aim of this chapter. However, this original contribution points out again, albeit in a different setting, the positive relationship between the tax-price of education spending and the share of young in the population as one of the main forces linking education spending and demography.

Joint education and social security Since Rangel (2003) and Boldrin and Montes (2005), the existence of stong linkages between publicly funded education and social security has become obvious, and analysis of these linkages has been extensive since. We believe it is important to understand precisely the nature of these linkages, if one is to be able to make predictions on the impact of ageing on both these programs that constitute a large share of public spending. Common features of the models discussed below are the necessity of publicly provided education to enhance human capital accumulation, and its beneficial impact on future PAYG pensions through the size of the wage bill, but also on fully funded ones through the increase in the return on savings.

To the best of our knowledge, Kemnitz (2000) was the first attempt at formalising the interactions between these two types of intergenerational transfers, and characterising the policy outcome as that of a game between old and active generations. The model, in turn, is used to make predictions on the impact of population ageing on the composition and size of government spending, and economic growth. Two features of the model are worth mentioning here: first, human capital accumulation (realised exclusively via public financing of education) is the sole engine of growth, and physical capital is only present in the form of life annuities workers can save into, with a rate of return that only depends on the world interest rate and probability to survive into old age. Second, policy decision is made to maximize the support politicians receive from middle-aged workers and old people, this support being a function not only of their respective numbers but also on their lobbying activities. Lobbying by individuals of a given generation is costly in terms of their current consumption. While the elderly only favor pensions, the current contributors to the system (i.e. the workers) favor education finance even though no commitment on future pension benefits is possible, since they anticipate they will in turn be in a position to extract these benefits from their children later in the game. Even though the author is unable to characterize all the forward-looking, rational-expectation equilibria of the game, he can derive the effect of ageing on the steady-state growth rate of the economy. In the steady state, population ageing leads to higher per capita income growth: the pension contribution rate increases as a consequence of a decrease in fertility or mortality (through the increased political power of pensioners), which in turn increases the incentives to fund public education to benefit from these higher future pensions. It should be noted, however, that only the amount devoted to education per child ($\tau_e/1+n$) is supposed to increase here, but not necessarily the share of national income devoted to education (τ_e): even though both are deemed to increase following an increase in life expectancy, in the case of fertility reduction education per child might only increase due to the increased effectiveness of public spending.

Poutvaara (2006) also stresses out the dynamic linkages between votes on education and PAYG pensions. In a model with heterogeneous agents, he points out how agents with higher ability would like the state to invest massively into education, while afterwards trying to renege on a contract giving sizeable pension transfers (and thus taxing their incomes heavily): this creates a time-inconsistency problem, which is evacuated by considering only subgame-perfect structure-induced equilibria that are supported with specific trigger strategies of the kind discussed in section 1.3 . Even though this paper exhibits no static analysis of the impact of demographic parameters on the size of the two policies, it is interesting in that it shows the median voter may belong to the young generation, even in the case of no population growth, because of the ability differences.

Soares (2006) also provides insights on the role played by social security in improving political support for public education. In a model in which social security funding is not subject to a vote, he shows that the level of public funding for education supported in the political equilibrium is higher for higher level of social security. He notes, however, that introducing social security is not necessarily welfare-improving, and that public education and social security do not necessarily sustain one another politically in the absence of intergenerational altruism.

Gonzalez-Eiras and Niepelt (2007) study the joint determination of the size of intergenerational transfers and public education provision. They draw on a previous, normative analysis by Boldrin and Montes (2005), who specify the need for the state to set up a series of intergenerational transfers to finance education in the presence of borrowing constraints: schemes in which middle-aged workers make transfers to the youngs to finance education and are paid back in next period in the form of pension benefits are proven to restore efficiency, absent the opportunity to resort to capital markets. Gonzalez-Eiras and Niepelt (2007) complement this study by analysing to what extent a political process may be able to implement such an efficient allocation, in a dynamic growth model. They also calibrate their model with US data to study how ageing quantitatively impacts taxes, the extent and composition of government spending, and productivity, all through the political process. To this end, they set up a three-period OLG model, with individuals being successively students, workers and retirees, and homogenous in dimensions other than age. It is assumed that students only acquire human capital, while workers have to determine their labor supply and their choice of savings; a fraction δ of workers get to live to the next periods, in which they are retirees and simply consume their savings.[28] The model incorporates a production sector which uses capital (entirely depreciated at the end of each period), labor and human capital (which is accumulated in a way that allows for endogenous growth). Political competition takes place on three issues: there are two taxes on payroll labor income, which are used to finance education and pensions in a budget-balanced way, and an additional capital income tax used to finance transfers from retirees to workers. The authors assume joint probabilistic voting on the three levels of taxes, each age group being assigned a weight in the politicians' objective function which is then calibrated to match US data on the shares of education and pension expenses in GDP. Voters are assumed to form expectations upon the effect of current policy choices on future policy outcomes. Students only favor education transfers, while middle-aged workers favor education transfers (because it increases the product of the capital they will own in next period) and transfers from retirees. Retirees only favor pension benefits, and are against any taxation of their capital income. The sign of the net transfers between middle-aged workers and retirees

[28] The capital of dead individuals is shared between bequests to workers and annuitisation benefits to retirees.

depends on their relative weights in the objective function. This model is calibrated to US data to generate predictions about the effects of fertility decrease and longevity extension on labor supply, labor and capital income taxes, pensions, human and physical capital investment: the aging of population is found to induce a reallocation of public spending towards excessively large levels of pensions (with a GDP share of retirement benefits going from about 5% in 2005 to 10% in 2075), at the cost of education spending (with the education spending over GDP ratio going from 6% to 5% over the same period of time). This results in the productivity growth of the economy being severely slashed down, due to the endogenous growth component present in their model. Compared with the Ramsey allocation that would be made by a benevolent planner seeking to maximize the welfare of all alive and unborn generations, the politically implemented allocation is inefficient.

Gonzalez-Eiras and Niepelt (2012) build on a standard two-period OLG model with endogenous growth based on physical and human capital accumulation. Individuals are successively young and old adults, and homogenous in dimensions other than age. The young have to determine their labor supply and their choice of savings; a fraction δ of the young get to live to the next period, in which they consume their savings but can also work part-time[29]. The share of time spent working while old, denoted ρ (envisioned as a proxy to retirement age policy), is a policy variable as well as τ_e and τ_p, the tax rates on labor income used to finance education of the next generation young and the pensions of the current old, respectively. Preferences on those three policy choices are aggregated assuming probabilistic voting. Population ageing is shown to act on policy decisions through different channels. Firstly, population aging, by increasing the political power of the elderly, tends to raise τ_p but to decrease τ_e and ρ: elderly voters use their bigger political clout to increase the transfers they receive at the expense of productive investment, while trying to reduce their labor force participation. This effect is countered by the fact that the bigger share of elderly in the population makes it more profitable (from a government finance point of view) to raise the retirement age ρ, which allows to decrease τ_p and increase τ_e. Thirdly, increased longevity makes forward investment more profitable for young households, which in turn would tend to increase public investment in education τ_e. Finally, and since there are dynamic linkages between policies at successive periods, future anticipated demographic change can lead to deviations in current policy: this implies that permanent demographic change can lead to non-monotone dynamics in policies. Based on this analysis, the author make

[29] It is noteworthy that the effect of fertility are also taken into account here, even though the effects of fertility and longevity cannot be analytically disentangled on a two-period OLG structure.

quantitative predictions for a stylised aggregate of rich OECD economies: it turns out that projected demographic change over the 2000-2080 period would lead to a modest increase in both τ_p and τ_e (by respectively 2% and 3% of taxable labor income), with a larger increase in ρ corresponding to a 6-year increase in the retirement age (to be put into perspective with the expected 8-year increase in life expectancy over the same period). However, if one shuts down the retirement age adjustment mechanism, the contribution rate for pensions would sharply increase in equilibrium, while the tax rate to finance education would barely decrease over the 2000-2080 time range.

The second chapter of this thesis also 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 involving both contributory and beneficiary generations. To this end, an OLG model is set up with production and intragenerational heterogeneity, in which both intergenerational transfers are funded through taxes on the working generation. Contrary to other papers, the economy is assumed to be closed, and the impacts of ageing on capital accumulation and relative prices of the different inputs is taken into account explicitly. Individual preferences for taxation are aggregated through probabilistic voting. Under the assumption of non-strategic voting, and picking specific functional forms, I derive in the second chapter predictions on the impact of fertility and mortality rate changes on the level and composition of public spending: in particular, population ageing (whatever the cause) points out to a rising tax burden, and thus a bigger size of government, in the future. Analytical results fail to unambiguously state which type of policy is likely to be expanded as population ages, but simulations under plausible parameters point out to rising pension and education expenditures as a share of total production. At the same time, higher expenditures at the level of the whole economy also translate into higher pensions per capita when longevity rises.

Naito (2012) elaborates on the political linkages between public education and pensions, by setting up a tractable model in which the level of inequality in initial human capital endowments plays a role on the comparative statics with respect to demography. It considers a canonical three-overlapping-generation economy with exogenous and constant fertility rate n, but without mortality. Even though the model features physical capital as a means of production alongside aggregate human capital H, the assumptions on production technology and openness of the economy are such that human capital accumulation is the only source of growth, and provides the only dynamic linkage between policies at different periods. For each individual i, human capital is produced using parental human capital and public education: $h_{i,t+1} = h_{it}e_t^\theta$, $\theta \in (0,1)$. The only policy subject to a vote is the total labor income tax rate τ[30]: the share of tax receipts that goes to education and pensions, denoted χ, is assumed to be fixed from the outset. The conflict about policy is here both inter- and intragenerational: as individuals differ in their human capital h_{it}, and since both types of policies redistribute income intragenerationally, the policy outcome reflects the interests of a coalition of poor active workers and retirees against the higher-endowed workers. The author restricts his attention to stationary, fully forward-looking Markov equilibria; due to the multiplicity equilibria, the discussion is further restricted to equilibria where the policy function τ is a power function of the ratio of the decisive voter's human capital to mean human capital. This and other functional form assumptions allow to characterize a growth trajectory which is on a balanced growth path from period 0 onwards, with a constant tax rate over time (since the human capital distribution is only shifted, but not distorted, by the education policy from one period to the other). Interestingly, a rise in the fertility rate n has two opposite effects on the equilibrium tax rate: on the one hand, for a given decisive voter, a rise in n decreases the optimal funding rate for education (mainly because education expenditures have to be shared among more people). On the other hand, when n increases, the old generation loses weight so that the ratio between the decisive voter's human capital h_m and aggregate human capital H is higher, which in turn increases the willingness to fund education expenditures[31]. As a whole, it is shown that because of these various political effects, human capital accumulation rate increases with n for high levels of initial inequality, while it decreases when n increases for low levels of inequality compared to n.

Ono (2013) also formulates a three-period OLG model where public education funding and social security coexist as the outcome of a political process. Key features are the existence of altruism towards children, and the possibility to invest privately in education. Although the main aim of the paper is to provide an explanation for cross-country differences in social security and public education spending in the OECD, the model sheds light on the key role of longevity (proxied as usual by an exogenous survival probability into old age δ) as a determinant of these two types of spending. In this paper, homogenous individuals of the middle generation are assumed to use their disposable income, which they derive by inelastically supplying labor the efficiency of which de-

[30] Again, labor is assumed to be supplied inelastically for tractability.

[31] In the paper, it is shown that the elasticity of the marginal benefits of a tax increase with respect to h_m/H is higher than the elasticity of its marginal cost, so that the equilibrium tax rate τ increases in h_m/H.

pends on human capital h_t, to finance current consumption and private investment in education z_t. They do so to maximize expected lifetime utility:

$$\ln c_t^m + \delta\{\theta \ln h_{t+1} + \beta \ln c_t^o\}$$

where θ is the degree of altruism towards children, and c_t^m and c_t^o are middle and old consumptions, respectively. Old agents make no private decisions, and simply consume social security benefits b_t. Tax receipts $\tau_t h_t$ are used to finance public education e_t and social security benefits. Spending on these two items is decided through probabilistic voting (see Lindbeck and Weibull (1987)), assuming individuals vote on a fully forward-looking fashion (as in **?**). Finally, human capital accumulation drives the dynamics of the economy as follows:

$$h_{t+1} = h_t^\eta (z_t + (1-\gamma)e_t)^\delta$$

with γ being a parameter for the efficiency losses coming from the fact that public education is standardized, and hence less efficient than private education. For a sufficiently low value of γ, two political equilibria coexist, one featuring only private education with public spending focusing exclusively on pensions while the other one is distinguished by no private education spending and positive public spending on education (along with pensions). Population ageing (an increase in δ in this setting) induce several effects, which may or may not exist depending on the equilibrium selected:[32] firstly, greater longevity implies middle individuals attach more weight to their offsprings' acquisition of human capital, implying an increased incentive to invest in private education z. Secondly, an increase in the number of pensioners has the direct political effect of increasing their weight in the political process, hence a tendency to increase pensions at the expense of both types of human capital accumulation. Thirdly, higher survival probability into old age means middle-aged individuals have a greater incentive to enact public education funding for their children, whose human capital they have a higher chance of enjoying later in life. By comparing the two types of equilibria obtained, Ono (2013) delivers another interesting result in terms of the correlation between the two types of public spending. It states that if longevity is sufficiently high, public expenditure on education and on pensions should be negatively correlated across countries (if we accept similar countries to be instances of the realisations of the two types of equilibria mentioned above). Similarly, if longevity is low enough, the correlation between the two types of spending should be positive.

[32]Note that the author remains silent on equilibrium selection. Multiplicity of equilibria is only explored to rank policies in terms of human capital accumulation and welfare.

Kaganovich and Meier (2012) present a small, open-economy model where voting takes place only on a labor income tax to finance education, under four different pension schemes (PAYG versus fully-funded, contribution-based vs. flat). The only demography-related parameter is fertility. No specific analysis is carried out on the impact of this parameter value on the equilibrium. The main contribution consists in pointing out that PAYG system do not depress the level of physical capital present in the economy (since the economy is open to capital flows), and that redistributive schemes are better for human capital accumulation, since the median voter (who has lower-than-average ability) on the education decision can expect to reap off more future benefits when the benefit formula becomes less contribution-based. However, this needs to be balanced with the labor supply distorsions introduced by redistributive pensions, which does not seem to be done in this paper.

Additionally, recent papers (Iturbe-Ormaetxe and Valera, 2012 ; Kaganovich and Zilcha, 2012), while not focused on exploring the impact of ageing on the size of social security, deliver interesting insights into the relationship that exists between the design of social security and political support for public education. Kaganovich and Zilcha (2012) compare the two systems of fully funded and PAYG pensions, for identical levels of the pension contribution rate, and ask how these systems influence political support for public education funding. In their model, public education funding is assumed to have an effect on the current generation's rate of return on savings, and on the present value of next-period social security benefits. They show that lower-income individuals will tend to favor higher levels of funding for education under the fully funded system than under the PAYG one: in the political equilibrium, the fully funded system of pension will then tend to generate higher levels of education spending than PAYG. In a similar approach, Iturbe-Ormaetxe and Valera (2012) analyse the effect of pension reforms on publicly funded education, also taking into account the effect of these reforms on labor market incentives. They model the pension system (the level of funding, contrary to public education, is not subject to a vote) as having three components: a Beveridgian (unconditional) one, a Bismarckian (i.e. related to earnings) one, and a funded component in the form of mandatory individual retirement accounts. They consider three types of reforms that either transfer some resources from the first two components to the third one, or increase the funding of the fully funded component. Under their setup, all reforms have the effect of weakening the link between the quality of education today and pensions tomorrow, thus tipping the collective decision point away from public education.

1.8.1.2 Empirical evidence

Empirics with education only: Miller (1996) performs a panel estimation of the impact of the population structure on educational spending for U.S. states. She explicitly tests for interest-group determination of spending, by including as explanatory variables the proportion of parents of school-age children, and of retirees, in the electorate. The dependent variable is the log of real per adult expenditures on education (state and local) for the years 1960, 1970, 1980 and 1990. While most specifications show a positive effect of the proportion of parents on spending, elderly proportion seems insignificant. A similar analysis on Texas counties, conducted for the years 1970 and 1980, finds both variables to be significant with the expected sign, however.

Poterba (1997) similarly uses panel data on K-12 education spending on the U.S. states, over the 1960-1990 period, using (log of) education spending per child as the dependent variable. The proposed explanatory variables this time include the proportion of school-aged children and elderly in the state population of each state. It is estimated that per-child spending has an elasticity with respect to the over-65 population share of approximately -0.25: in other terms, "a one standard deviation change in the share of elderly in the population, a shift from 0.108 to 0.130, results in almost a 5% decline in per-pupil education spending", which is sizeable. This also needs to be put into perspective with the fact that other governmental programs are rising with the share of the elderly in the population. On the other hand, per capita education spending has a negative elasticity (varying between -0.4 and -1.0) with respect to the fraction of school-age children in the population. While this result may seem surprising, the author emphasizes that other types of public spending are more negatively affected by increases in the share of the young: thus, the share of public spending that goes to the young increases when the share of the young in the population increases. Fernandez and Rogerson (1997) perform a similar analysis on U.S. states from 1950 to 1990. In their main specification, the logarithm of schooling expenditures per student is regressed on the log of income per capita, the fraction of the population of school age and the fraction of the adult population over 65. While the coefficient on the share of students is insignificant, the one on the share of elderly is consistently negative and significant (around -0.2). On the other hand, per student education expenditures go hand in hand with income per capita.

Ladd and Murray (2001) explicitly build onto the approach in Poterba (1997), while shifting the unit of observation away from states to local counties. They start by presenting descriptive evidence showing that elderly residential choices may be related to the level of spending on education by the county. In order to try to eliminate the bias created by this Tiebout sorting of elderly with respect to local taxes, they then proceed to estimating a panel equation where the share of the elderly in a county is instrumented by the same share in previous census. By so doing, Ladd and Murray (2001) obtain a significantly smaller estimate of the elasticity of educational spending with respect to the share of pupils; more importantly, the share of the elderly now becomes statistically insignificant, suggesting that elderly actually do not exert any downward pressure on spending per child.

Harris et al. (2001) use panel data on spending and demographic structure at the school district level to reconcile the findings in Poterba (1997) and Ladd and Murray (2001). Indeed, they show that the elderly are much less likely to support state funding than funding at the local, school-district level. This, the authors remark, is consistent with the widely accepted observation that the quality of local schools is capitalised in housing values.

Grob and Wolter (2007) perform a panel analysis of spending on compulsory education for Swiss Cantons, between 1990 and 2002. It shows that under a Canton fixed-effect specification, the percentage of retired in the population has a negative, yet significant at the 10% level only, effect on educational spending per child. Using first differences, the authors find that variations in the number of pupils (resp. the number of retirees) have a positive (resp. negative) impact on total education spending.

Empirics with several public programs: Borge and Rattsø (2008) use a specification very similar to that of Poterba (1997), to evaluate the impact of population structure on three welfare programs of the 275 Danish local governments, over the 1989-1996 period. The welfare programs considered are child care, compulsory education and elderly (instituionalized) care. The demographic variables are the fraction of the potential recipients of each of the three programs (i.e. resp. the 0-6 years, 7-15 and 80+ age groups) over the total population. Using as dependent variables the spending per capita on each program, the authors first find that being part of a large age cohort is a disadvantage in terms of transfers received. They also shed into light that while the elderly exert significant crowding out on spending towards the younger generations, the opposite is not true.

Surveys of voter preferences All the aforementioned papers attempt to analyse the empirical impact of the age structure of the population on the size of these public programs, which is the outcome of the political process concerning them. Meanwhile, a complementary approach has been developed by some other papers (Berkman and Plutzer, 2004 ; Brunner and Balsdon , 2004 ; Busemeyer et al. , 2009 ; Cattaneo and Wolter , 2009; Sørensen , 2013) that have used individual surveys about the likelihood to vote for specific, age-related transfer policies, in order to analyse the role of age on the individual preferences on these policies. For instance, Cattaneo and Wolter (2009) use a representative survey of Swiss voters to show the elderly are less likely to be willing to fund education programs, and prefer spending on health and social security. They find that these result still hold after controlling for the fact that the elderly are less supportive of public spending in general. The interest of this type of analysis is to establish a possible causal link between the age structure of the population and the policy outcome that goes through individual preferences, and not only through the tax cost of spending on the different programs as the age structure varies.

1.8.2 Ageing and support for health care

Surprisingly enough, the recent political economy literature contains few articles linking the ageing of the electorate and the evolution of public health care provision, even though it may be considered as hot a topic as pensions. We will focus here on only one significant contibution, although other recent papers on the political economy of health care without direct focus on ageing may be worth mentioning here, such as Moreno-Ternero and Roemer (2007) or De Donder and Hindriks (2007): indeed, the results and insights obtained can be extended to a situation with a changing composition of the population with regards to ageing.

Bethencourt and Galasso (2008) emphasize a political complementarity between two transfer programs, namely pensions and public healthcare. It has been known for some time already (Philipson and Becker, 1998) that the existence of pensions as annuity encourages investment in healthcare. Here, expenditure in healthcare reduces the longevity differential between rich and poor individuals, thereby making pensions more attractive for the poor individuals with lower longevity. The authors support their analysis by pointing out the existence of a health gradient in income and the fact that publicly provided healthcare increases more longevity for the low-income than for the high-income individuals (in other words, the health production function is concave). Essential to the establishment of a political equilibrium supporting a large welfare system is the intragenerational redistribution element of both programs. Their model features a 2-period OLG endowment economy where young agents do not take any economic decisions (they only participate in the vote over social security and healthcare). Old agents (who are heterogeneous in the endowment they received) use their income to finance private healthcare and non-health expenditures. They are taxed linearly to finance non-contributory pensions and public healthcare. Only public healthcare raises the life expectancy, which is modeled as the fraction of the second period during which agents will enjoy their pension payments. On the political side of the model, the authors assume that a majoritarian vote takes place over the size and composition of the welfare state. They use subgame perfect structure-induced equilibrium to reduce the game to a (degenerately) dynamic, issue-by-issue voting game. Under adequate parameter restrictions, an equilibrium with positive taxation is supported by a coalition of the old and the young poor, provided the amount devoted to healthcare is sufficient to increase the young poor life expectancy up to the point where they support pensions. Concerning the determination of the composition of the system, the authors find that the preferred policy mix depends solely on income. For any transfer size, richer agents prefer a purely pension-based system, since it is the one which provides the widest longevity gap in income and thus redistributes income the least. As agents become poorer, they prefer a bigger part of the welfare system to be spent on healthcare, in order to make receiving pensions more worthwhile. The authors also find out that an increase in health care technology leads to a larger welfare state and a larger pension share. Using a calibrated version of the model for the U.S. economy in the early 80s, they find that such a technology improvement would then cause an increase in aging, and also higher per-capita welfare spending. An interesting potential development of this line of research would be to analyze the impact of a change in fertility, which enters the model here as an exogenous parameter.

1.8.3 Ageing and the environment

Although some may consider the subject of environment conservation very loosely related with our subject, we argue that its characteristics make it a subject with important intergenerational dimensions. Indeed, to use the terminology introduced by Rangel (2003), a healthy environment can be considered a “forward intergenerational good” in the sense that it is financed by a given generation to be consumed mainly by the next one (as opposed to “backward intergenerational goods” like pensions for instance), especially if one considers climate change. In this sense, the way in which a society decides on the funds to allow for environment conservation may well be impacted by the age structure of its population as well.

Ono (2005) is, to the best of our knowledge, the first significant contribution on this topic. It focuses on the impact of greater longevity and lower growth rate of population may affect the level of a politically-determined environmental tax, and through it the quality of environment over time.

This paper uses a two-period OLG model with environmental quality and uncertain lifetimes. Lifetime uncertainty is represented by an exogenous probability δ to survive into old age. It is assumed this uncertainty cannot be completely hedged against in the savings decision of young agents, as they can only save a fixed fraction γ of their savings in the form of annuities. On the production side, it is assumed that firms create environmentally harmful emissions as by-products, in the form of a fixed fraction η_y of the output of the firms. To repair the damage made by emissions, an environmental tax τ on each unit of production is levied in each period. The proceeds of the tax are used to improve environmental quality, which is modeled as a stock that depletes with emissions and is maintained by tax receipts. Finally, agents care about consumption and environmental quality when old, in a separable fashion. The political process boils down to a short-lived representative government that sets up τ_t in each period, so as to maximize a weighted sum of the utilities of currently living generations: representatives are assumed to have rational expectations about future policies, but to act in a myopic fashion with regards to the impact of current decisions on future ones. A first insight from the model is that longevity has no impact on the level of the environmental tax at steady state, while a slower level of population growth n will reduce the level of the tax. However, the final impact on environmental quality is more complex. Indeed, increased longevity leads to a lower level of environmental quality under partial annuitization of savings, while having no effect at all in the case of full annuitization. On the other hand, environmental quality increases when the rate of population growth decreases: indeed, a lower n enhances capital accumulation, in a way that increases environmental investment more than emissions for a given level of τ. This effect outweighs the negative effect of n on τ described above, hence the result. As is already the case for pensions, it is worth stressing that in this case too, one needs to take into account the effect of population change on capital accumulation decisions as well as their more politically proximate effects, if one is to get a full understanding of the consequences of political ageing.

1.8.4 Capital vs. labor taxes

Bassetto (2008) studies a setup slightly more complex than an OLG model with pensions, including capital taxes (that fall on the old) along with labor taxes (falling on the young), and tax proceeds that can be used to finance a public good and lump-sum transfers to either generation (or both). The argument for looking at this broader picture comes from generational accounting: changing demographics will probably affect what each generation pays to the system, thereby possibly offsetting the changes in what they receive as a consequence of the demographic shock. An interesting feature of the model used in this paper is the modeling of the political process as a bargaining game. In each period, citizens are drawn randomly to make proposals that have to be accepted unanimously to be implemented, with a small probability that the game terminates with a government shutdown for the period each time a proposal is rebuked. In this environment, each generation holds a veto power over the decisions, such that the outcome must somehow balance the gross contributions from and transfers to both the old and young. Moreover, concerning the public good financing, it can be seen that if a generation values less the good than the other one, this generation benefits from hold-up power (here, the threat of a shutdown), which it may use to extract larger transfers. Under a setup where the young only supply labor and the old only capital to the production process, the author points out that the young will be willing to reduce the size of both their contribution and the transfers they receive (since the labor tax base is sensitive to the tax rate), while the old will accept large contributions in exchange for big transfers (as the capital tax base is predetermined by past decision and is thus insensitive to the tax rate chosen). Let us now convey the intuition of the consequences of population ageing, under the assumption that young people value more the public good than the old[33]: in this case, as the old agents become relatively more numerous, their decision power increases and is used to decrease the overall size of the government, and in particular decrease public good financing. Moreover, and even though they manage to negotiate higher overall transfers from the young, this is not enough to counteract the fact that transfers have to be shared between more old people, so that the per capita transfer they get will decrease. This effect on the transfer side needs to be added to the fact that the capital tax rate also increases, so that the overall net transfer per capita to the old is further degraded. From the point of view of the young, ageing leads to increased labor tax rate and increased net contributions per capita, on top of the

[33] This would be the case, the author argues, if the public good is itself somehow a forward intergenerational good, like environmental quality (see subsection 1.8.3).

reduction in public good consumption, so the overall effect is unambiguously negative for them. On top of this comparative statics on the steady state of the economy, the author conducts two comparative dynamics exercises to explore what happens to these taxes and transfers when a demographic shock hits, whether anticipated or unanticipated. It is shown that the output of the model is consistent with the historical experience if one considers an anticipated ageing shock at the time of the baby boomers generation, with transfers to the old going upwards when this generation is active (i.e. young), before an eventual contraction.

Contrary to the intuitions conveyed by the preceding paper, Mateos-Planas (2010) sets up a general-equilibrium model to predict a link between ageing and increased capital taxation, in a model where both labor and capital can be taxed. For the prediction to hold, the young need to hold less capital than the old, which appears reasonable, and the ageing process needs to be large enough to shift the identity of the median voter to induce a movement of the tax rate in equilibrium.

1.9 Conclusion

Summary of previous sections In the previous sections, we have thus given an account of the recent literature on the political economy of pensions, and more precisely how population ageing is compelling electoral democracies to rethink collectively the size and design of their pension systems. A first, and essential step according to us, was to remind the reader why public PAYG pensions were supported by a majority of the electorate in the first place: arguments of reduced time horizon of the median voter, or the existence of a within-cohort redistribution component in PAYG pension schemes, are the most compelling reasons why these programs exist in the first place. Once this theoretical framework for the political sustainability of pensions has been established, we have been able to give an account on the impact of ageing on the size of the programs. The effects at play are slightly different depending on whether population ageing is caused by fertility decreases or life expectancy increases: in any case, they fall into two categories. The political effect of ageing is unambiguous: as population ages, the weight of old voters in the electorate increases, which will give incentives to politicians seeking votes to propose electoral platforms with more generous retirement packages. What we call the economic effect of ageing is less clear cut, but essentially hinges on how the internal return of the PAYG system is modified as the age structure of the economy changes. The theoretical literature essentially associates higher life expectancy to higher sizes of pension programs, while lower fertility yields more ambiguous results. Empirical evidence suggests that measures of population ageing such as the old-age dependency ratio can be estimated to cause higher pension spending. Considering other public programs alongside pensions does not seem to modify the broad sense of this result: it adds to the previous analysis the finding that education funding is generally at risk with a graying electorate.

This chapter has also studied the impact of ageing on the retirement age: to sum up, ageing acts a negative shock on the return of the PAYG pension system. Then, agents faced with ageing suffer a negative income shock, which will result in their optimal retirement age being higher than before (retirement, or leisure, being a normal good). Therefore, ageing is predicted to induce the electorate to choose a higher retirement age.

The way forward As stated in subsection 1.8.2 , further research is warranted on the political economy of health care and long-term care, especially since recent increases in life expectancy make the financing of these governmental program an ever more pressing issue.

As shown in Galasso (2006), work on the joint political determination of the size of the pension system and the retirement age seems a promising avenue of research. Galasso and Profeta (2004) also suggest that the importance of family ties (and hence altruism considerations) should be taken into consideration, if one is to make sense of the political decisions taken by individuals as regards the pension system. Indeed, they indicate that countries with relatively more frequent multi-generational living arrangements would tend to choose larger pension systems, as adult children living with their parents are more likely to be in favor of pensions than their alone-living counterparts for intra-household transfer reasons. It can be argued that this effect is sizable, especially when comparing countries: in Southern Europe or Japan, a high fraction of elderly individuals (around 40% or more) live with their adult children, compared with just less than 15% in the US.[34]

Another overlooked characteristic of pension systems is its intragenerational redistribution component. Galasso and Profeta (2004) recall the distinction between Beveridgian and Bismarckian pension systems, the former aiming at giving a base income at retirement independently from past contributions while the latter tends to provide equal replacement rates for all individuals. Indeed, recent reforms go in the direction

[34]Source:Howe and Jackson (2003), as cited in Galasso and Profeta (2004).

of extended coverage and increased minimum pensions to fight poverty among elders, while at the same time tightening the eligibility criteria and containing the increases in spending by tinkering with benefits formula for the general population. It therefore appears worthwhile to ask the question whether most pension systems are not becoming increasingly Beveridgian (or at least progressive in some dimensions), as well as investigate the political reasons behind this institutional evolution.

We would also like to add that further theoretical work should consider including the possibility for the electorate to choose to run some deficits on its pension system: indeed, most theoretical papers assume from the outset that the electorate determines the size of a system by merely setting the contribution rate on labor income, the value of pension benefits being merely set so that the pension system budget is in equilibrium at all times. Such formulations have several drawbacks, the most important one being that pension systems are never exactly in budgetary equilibrium for any fiscal year. It also significantly blurs the distinction between defined-benefit and defined-contribution systems. It would probably be a hard process, as there would be several decision dimensions (including the size of the system and its degree of financing by current contributions) to consider, and would most likely require precious insights from the political economy literature on inaction and reform delay, and debt (see Drazen (2000), chapters 3 and 10, but also Tepe and Vanhuysse (2012) for the point of view of political science). More generally, every step should be made in terms of modelling or quantitatively assessing the future evolutions of the pension systems to better explain the institutional details of these systems, and how these are likely to be impacted by the coming demographic shockwave.

2

Demography and Intergenerational Public Transfers: a Political Economy Approach

2.1 Introduction

2.1.1 Motivation and aim of this paper

The recent debates on social security reforms in many advanced economies[1] have shed new light on the relevance of population ageing on long-term growth or public finance sustainability (Bloom et al., 2010 ; Holzmann , 2012 ; Meier and Werding , 2010). 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 the OECD countries, public policies are determined as the outcome of democratic deliberation and voting on the alternatives, we need to know both how individual preferences about the policies aggregate into collective choice, and how this collective choice affects private decisions (Krusell et al., 1997). The relative political power of each generation then becomes determinant when policies that involve intergenerational tranfers are considered.

In this chapter, I focus on two particular public policies with strong generational contents: education and retirement.[2] 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 pay-as-you-go pension systems,[3] 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 behavior (e.g. the fact that increased longevity makes saving more desirable, or the capital dilution effect of increased fertility) from the ef-

[1] See for instance European Commission (2010) or OECD (2012) for overviews of the institutional debate on pension reforms.

[2] 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 chapter: although the healthcare system operates some degree of redistribution towards the old, it also exerts a function of insurance within each generation.

[3] See chapter 1 for an overview of the literature on the subject.

fects 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[4] to the youngs, as well as pensions for the elderly, through two earmarked flat taxes levied on labor income. The political process, modelled as probabilistic voting *à la* Lindbeck and Weibull (1987) 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 characterizing 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 within a given period. The only link between education transfers and pensions comes from the willingness of the current working generation to increase its children's future productivity, so that taxable labor income, and ultimately pensions, be bigger in the next period.[5] 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, I show analytically 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. Finally, using simulations for values of the main parameters of the model that are meant to capture the main characteristics of an advanced capitalist economy, I find that the process of ageing may well be conducive to more factor accumulation and growth in output per capita,[6] a fact

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

[5]Since I assume away any form of altruism between generation, the middle-aged generation does not care about increasing the young's productivity *per se*, but only for the aforementioned self-interested reasons.

[6]This result may seem similar to the standard result from neoclassical growth models that capital and output per capita increase when the rate of growth of population decreases. In this model, ageing also causes an increased taxation burden which prevents factor accumulation, so that the overall result

overlooked by proponents of the existence of a looming "grey dawn" (Peterson (1999)).

2.1.2 Contribution to the existing literature

A part of the literature on the political economy of pensions considers pensions as one of 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 (see for instance Boldrin and Montes (2005), or Rangel (2003) for a more general discussion of intergenerational contracts). For instance in Boldrin and Montes (2005), the 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.[7]

This paper departs from this idea by considering a different mechanism for preference aggregation, namely probabilistic voting (as first introduced by Lindbeck and Weibull (1987)): the outcome of the voting process then maximises a weighted sum of the utility of voters. Assumptions about how agents form their expectations of future policies is shown to play a key role here: provided voters behave in a non-strategic way,[8], the dynamic voting game on tax rates is shown to amount to a series of repeated static choices, and the outcome of the voting game ultimately depends on current state variables only. I interpret this property of the politico-economic equilibrium as proof that 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 model is closest to Gonzalez-Eiras and Niepelt (2007) and Lancia and Russo (2013).

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, small parametrical changes in

is *a priori* ambiguous.

[7]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 section 3 of chapter 1 of this thesis for a detailed presentation of this setup, or Bellettini and Berti Ceroni (1999) for an example of this in a slightly different setting.

[8]Voters are said to behave in a non-strategic way if they fail to internalize the effect of current policy choices on future ones through changes on future state variables, even if future policy rates are correctly anticipated.

the demographic structure 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 analyze in a finer way 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, and to disentangle the various political and economic effects of ageing on factor accumulation and output.

Additionally, this model incorporates general equilibrium effects of taxation (as in Boldrin and 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 behavior 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 my 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 I do not consider in this version of the paper[9] is a (within-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.

The analytical, comparative statics results obtained can be interpreted in a historical perspective: I am able to confront the results of my comparative statics exercises to empirical estimates of the effects of the age structure on policies (Lindert (1994), Poterba (1997), Ladd and Murray (2001) or Grob and Wolter (2007)). Finally, subsection 2.4.2 presents some simulation results of the likely impact of population ageing on factor accumulation and growth, for values of the main economic and demographic parameters of the model that are in line with the structure of an advanced, OECD-like economy. Population ageing, whether it comes from a decrease in fertility or an increase in longevity, is found to bring about increased physical capital accumulation and

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

a higher share of educated workers in the labor force. These predictions come in stark contrast with those of Gonzalez-Eiras and Niepelt (2007), who reach the conclusion that the increase in pension spending caused by ageing necessarily crowds out investment in productive investment such as public education at the political equilibrium. My simulations of the model show, however, that ageing can cause an increase in both types of spending (including when considering spending per capita). I also show that the resulting increase in the taxation burden is low enough to allow for an increase in savings per capita. Indeed, the direct effect on savings per capita coming from an increase in life horizon, or the reduction in capital dilution coming from a fertility decrease, both outweigh the indirect detrimental effect on savings coming from higher taxation.

Outline The rest of the paper is organized as follows. Section 3.2 presents the structure of the model and its main assumptions. In section 2.3 , 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 2.4 derives predictions about the impact of demographic change on the levels of both tax rates, and reports the results of simulations of the overall trajectory of the economy for plausible parameters: these predictions are confronted with available empirical evidence. Section 2.5 studies how important the assumptions regarding expectations in the base model are, by considering expectations of future contribution rates that are linear in future levels of production per capita. Finally, section 2.6 discusses some of the assumptions and results, and concludes.

2.2 Model setup

2.2.1 OLG structure, and private decisions

2.2.1.1 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 $t2$ [10]: N_t^t (respectively N_t^{t-1} and N_t^{t-2}) then denotes the number of young

[10] 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.

(resp. middle-aged workers and old people) alive at t. All individuals live for the first two periods. To proxy longevity, I 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.[11]

2.2.1.2 Private decisions

Utility Young agents born at t with innate ability ω_i seek to maximize their lifetime utility U with respect to their education decision $j \in \{0, e\}$ when young, and their consumption in the two later periods of their life:

$$U^t_{i,t} = -\mathbb{1}_{\{j=e\}} f(\omega_i; g_t) + \beta \ln(c^t_{i,t+1}) + \beta^2 \delta_{t+2} \ln(c^t_{i,t+2}) \tag{2.1}$$

Middle-aged agents also maximize a reduced version of this lifetime utility, their education decision 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 to educate $\omega_i \sim F(\omega)$, which is independently and identically distributed (F denoting the cumulative distribution function of abilities). They do not consume in t. The only decision young individuals need to take is whether to educate ($j = e$) or not ($j = 0$), 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 2.2.3 , the innate ability ω_i determines the actual welfare cost to educate as follows:

$$f(\omega_i; g_t) = \ln\left(1 + \frac{1}{g_t + \omega_i}\right) \tag{2.2}$$

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 abili-

[11] This last relationship stems from $N^{t-2}_t = \delta_t N^{t-2}_{t-1}$ and $N^{t-1}_t = N^{t-1}_{t-1} = (1 + n_{t-1})N^{t-2}_{t-1}$.

ties 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.[12]

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:

$$(1 - \tau_{e,t} - \tau_{p,t})w_{j,t} = c_{j,t}^{t-1} + s_{j,t} \tag{2.3}$$

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, and following Yaari (1965) and Blanchard (1985), it is assumed that on top of their own savings, pensioners also receive a share of the savings that were made by the deceased members of their own generation.[13] The interpretation is that middle-aged workers have access to perfectly competitive life insurance providers, which

[12] 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 $\lim_{\omega \to +\infty} f(\omega, g) = 0,\ \forall g$).

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

sell annuity 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 $(1-\delta_t)/\delta_t \cdot R_t s_{j,t-1}$.[14]

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

$$c_{j,t}^{t-2} = R_t(s_{j,t-1} + \frac{1-\delta_t}{\delta_t} s_{j,t-1}) + p_{j,t} = R_t \frac{s_{j,t-1}}{\delta_t} + p_{j,t} \tag{2.4}$$

2.2.2 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 $\eta > 1$[15]. 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:

$$y_t = (\mu_t \eta + 1 - \mu_t)^{\alpha} k_t^{1-\alpha} \tag{2.5}$$

Given this functional form for production, and assuming the labor and capital mar-

[14] 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.

[15] 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.

kets to operate under perfect competition, one gets:

$$w_{0,t} = \alpha(\mu_t\eta + 1 - \mu_t)^{\alpha-1}k_t^{1-\alpha} \tag{2.6}$$

$$w_{e,t} = \eta\alpha(\mu_t\eta + 1 - \mu_t)^{\alpha-1}k_t^{1-\alpha} = \eta w_{0,t} \tag{2.7}$$

$$R_t = (1-\alpha)(\mu_t\eta + 1 - \mu_t)^{\alpha}k_t^{-\alpha} \tag{2.8}$$

Equations (2.6) and (2.7) 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.

2.2.3 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 2.2.4 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:

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

$$\Leftrightarrow g_t = \frac{\tau_{e,t}}{1+n}(\mu_t w_{e,t} + (1-\mu_t)w_{0,t}) = \frac{\tau_{e,t}}{1+n}\alpha y_t \tag{2.9}$$

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:

$$\mu_{t-1}p_{e,t} + (1-\mu_{t-1})p_{0,t} = \frac{\tau_{p,t}(1+n)}{\delta_t}(\mu_t w_{e,t} + (1-\mu_t)w_{0,t}) = \frac{\tau_{p,t}(1+n)}{\delta_t}\alpha y_t \quad (2.10)$$

with $p_{e,t} = \eta p_{0,t}$ to match relative wages.[16]

2.2.4 Voting over education and pensions

In order to determine the level of the education and pension taxes, a way of aggregating the preferences of the heterogeneous agents populating the economy must be defined. I start by defining, as is standard in the literature, 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:

$$V_{t,e}^{t-1}(\tau_{e,t}, \tau_{p,t}, \mu_t, k_t) = \max_{(c_{t,e}^{t-1}, c_{t+1,e}^{t-1})} \ln(c_{t,e}^{t-1}) + \beta\delta_{t+1}\ln(c_{t+1,e}^{t-1}) \text{ s.t. (2.3), (2.4), (2.10)}$$

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

In what follows, I 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,[17] 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 and Weibull (1987).[18]

In a two-party electoral competition setting, it is thus 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

[16]See equation (2.7) for a derivation of relative wages.
[17]See Persson and Tabellini (2000), section 2.3.
[18]Three options are generally considered. Structure-induced equilibrium (see Shepsle (1979)) has agents voting seperately on each issue, while legislative bargaining (see Baron & Ferejohn (1989)) considers legislators that represent specific constituents and may form policy-specific alliances. The third approach, probabilistic voting, is the one undertaken here.

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:

$$W_t = \rho(1+n)\left(\int_{i=0}^{\infty} V_{t,i}^{t} dF(i)\right) + \chi(\mu_t V_{t,e}^{t-1} + (1-\mu_t)V_{t,0}^{t-1}) + \frac{\psi\delta_t}{1+n_{t-1}}(\mu_{t-1}V_{t,e}^{t-2} + (1-\mu_{t-1})V_{t,0}^{t-2}) \tag{2.11}$$

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 have their preferences taken into account, 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, in Downs (1957) model of electoral competition, the identity of the median voter will change as the economic conditions are altered in settings with continuous distributions of voters, but not necessarily in settings where the general population is divided into large homogenous categories. In the present 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, should I stick to the Downsian representation of policy choice, and 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 could not 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 would be an undesirable feature if, as in this paper, one wishes to study the influence of some parameters on equilibrium.[19]

2.2.5 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 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). I turn to analysing this dynamics in the next section.

2.3 The politico-economic dynamics

In this section, I study how the dynamics of the economy unfolds, using backward induction. I first derive the behavior of private agents under given taxes, which yields

[19] For a more detailed discussion of probabilistic voting, I refer the reader to Persson and Tabellini (2000), section 3.4.

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 characterize 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 characterize its steady state.

2.3.1 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.

2.3.1.1 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 maximize 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 (2.3), (2.4). The FOC on savings $s_{j,t}$ yields:

$$s_{j,t} = \frac{\beta\delta_{t+1}}{1+\beta\delta_{t+1}}(1-\tau_{e,t}-\tau_{p,t})w_{j,t} - \frac{\delta_{t+1}p_{j,t+1}/R_{t+1}}{1+\beta\delta_{t+1}} \tag{2.12}$$

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

$$c_{j,t}^{t} = \frac{1}{1+\beta\delta_{t+1}}(1-\tau_{e,t}-\tau_{p,t})w_{j,t} + \frac{\delta_{t+1}p_{j,t+1}/R_{t+1}}{1+\beta\delta_{t+1}} \tag{2.13}$$

$$c_{j,t+1}^{t} = \frac{\beta R_{t+1}}{1+\beta\delta_{t+1}}(1-\tau_{e,t}-\tau_{p,t})w_{j,t} + \frac{\beta\delta_{t+1}p_{j,t+1}}{1+\beta\delta_{t+1}} \tag{2.14}$$

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 will become clear 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 (2.1), (2.2) and (2.12), 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 (2.6), (2.7), (2.9) and (2.10):

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

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

$$\hat{\omega}_t = \frac{1}{\eta^{\beta(1+\beta\delta_{t+2})} - 1} - \frac{\tau_{e,t}\alpha}{1+n}(\mu_t\eta + 1 - \mu_t)^{\alpha} k_t^{1-\alpha} \tag{2.15}$$

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.

2.3.1.2 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 2.2.2 , capital fully depreciates after each period of production[20], 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 (2.7), (2.8), (2.10), (2.12) and simplifying:

$$k_{t+1} = \frac{\beta\delta_{t+1}\alpha(1-\alpha)(1-\tau_{e,t}-\tau_{p,t})}{(1+n_t)[(1+\beta\delta_{t+1})(1-\alpha)+\alpha\tau_{p,t+1}]} y_t \tag{2.16}$$

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.

[20] 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 2.3.1.1 , 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:

$$\mu_{t+1} = 1 - F(\hat{\omega}_t) \tag{2.17}$$

2.3.2 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.

2.3.2.1 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 characterize 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:*

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

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

2.3.2.2 Distribution of abilities

For analytical tractability reasons, and additionally to the setup presented in section 3.2, 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:

$$\mu_{t+1} = 1 - F(\hat{\omega}_t) = 1 - \frac{\hat{\omega}_t}{\bar{\omega}} = \frac{\tau_{e,t}\alpha y_t}{\bar{\omega}(1+n_t)} \tag{2.18}$$

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.

2.3.2.3 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.[21]

[21] Of course, the young want to increase τ_e only up to the point where taxation is not too detrimental to capital accumulation: 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, since the young value more education that the middle-aged in any given period.

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 I 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 that should be slightly higher than one.[22]

2.3.2.4 Expectations of future policies

As may be seen from subsection 2.3.1.1 , 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 the 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 prospect the agent has 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 expected future policy choices of next period $(\widetilde{\tau_{e,t+1}}, \widetilde{\tau_{p,t+1}})$, since these policy choices are made to maximize indirect utilities that are themselves the result of private decisions.

With the exception of section 2.5 , it will be assumed in the rest of this chapter 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.

[22] 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.

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, or strategic way (see section 2.5). 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 them 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 that the agents 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 policy choice of one single individual.

2.3.3 Taxation choices

Under assumption 3 , the two policy rates now maximize 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:

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

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 2.A shows that the partial derivatives of W_t on S are as follows:

$$\frac{\partial W_t}{\partial \tau_{e,t}} = \beta \alpha \delta_{t+1} \frac{(\eta-1)\alpha y_t}{\bar{\omega}(1+n_t) + \tau_{e,t}(\eta-1)\alpha y_t} - \frac{1+\beta\delta_{t+1} - \beta\alpha\delta_{t+1}}{1-\tau_{e,t}-\tau_{p,t}} \tag{2.19}$$

$$\frac{\partial W_t}{\partial \tau_{p,t}} = \frac{\psi \delta_t \alpha}{(1+n_{t-1})(1-\alpha+\alpha\tau_{p,t})} - \frac{1+\beta\delta_{t+1}-\beta\alpha\delta_{t+1}}{1-\tau_{e,t}-\tau_{p,t}} \tag{2.20}$$

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 2.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 (2.19) and (2.20)) are equal to zero at $\tau*$.

Equations (2.19) and (2.20) then implicitly define a policy rule $\tau(k,\mu) = (\tau_e(k,\mu); \tau_p(k,\mu))$ that maps the state variables of the economy into a vector of tax rates (τ_e, τ_p). This policy rule solely depends on output y, as defined in (2.5): in particular, it is unaffected by the future expected values of the tax rates in next period.

2.3.4 The politico-economic dynamics, and steady state

Given an initial state of the economy (k_0, μ_0), the previous sections allow to describe the politico-economic dynamics of this economy. As stated in the previous subsection, equations (2.19) and (2.20) 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.[23] In turn, equations (2.16) and (2.18) map the current values of the

[23] 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.

state variables and policy choices into the values of the state variables in next period (k_{t+1}, μ_{t+1}).[24]

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 defined 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.

2.4 Demographic change and its impact on public spending and growth

In this section, I first perform (in subsection 2.4.1) 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. Using simulation exercises, I also report results on the impact of changes in n and δ on factor accumulation and growth.

[24] Equation (2.16) 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.

2.4.1 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 and Shannon (1994) shall be used.

2.4.1.1 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 and Shannon (1994) 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 I 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 and Shannon (1994), supermodularity along with increasing differences implies that the argmax of welfare $\hat{\tau}^*(\sigma)$ is a monotone nondecreasing, continuous function of σ.[25] 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, I 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.

2.4.1.2 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 (2.19) and (2.20), I obtain that $\frac{\partial^2 W_t}{\partial n \partial \tau_p} = 0$ and $\frac{\partial^2 W_t}{\partial n \partial \tau_e} < 0$.

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

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 shock, so that the fertility shock has no impact on the pension margin of decision. Then following Milgrom and Shannon (1994), 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 (2.19) and 2.20), I 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.[26] Nonetheless, both tax rates rising consecutively to an increase in n is absolutely ruled out. In fact, appendix 2.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 summarized in the proposition that follows:

Proposition 1 *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.*

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

2.4.1.3 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.

$$\frac{\partial^2 W_t}{\partial\delta\partial\tau_{e,t}} = \beta\alpha\frac{(\eta-1)\alpha y_t}{\bar{\omega}(1+n)+\tau_{e,t}(\eta-1)\alpha y_t} - \frac{\beta-\beta\alpha}{1-\tau_{e,t}-\tau_{p,t}} \tag{2.21}$$

Concerning the determination of τ_e, equation (2.21) 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 the adverse effect of taxation on next-period capital accumulation (which matters, as well, for tomorrow's production and ultimately pensions).

$$\frac{\partial^2 W_t}{\partial\delta\partial\tau_{p,t}} = \frac{\psi\alpha}{(1+n)(1-\alpha+\alpha\tau_{p,t})} - \frac{\beta-\beta\alpha}{1-\tau_{e,t}-\tau_{p,t}} \tag{2.22}$$

Concerning τ_p, equation (2.22) 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 (2.21) and (2.22) yield, for $k \in \{e;p\}$:

$$\frac{\partial^2 W_t}{\partial\delta\partial\tau_k} = \frac{1}{\delta}\frac{\partial W_t}{\partial\tau_k} + \frac{1}{1-\tau_e-\tau_p} \tag{2.23}$$

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 $\tau_p^* > 0$.[27] From equation (2.23), I 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

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

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 me 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 2.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 summarized in the following proposition:

Proposition 2 *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.*

2.4.1.4 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 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.

2.4.2 Numerical example and simulation

In this subsection, I report the results of simulations made to illustrate the behavior of the economy, under plausible values of the model parameters. The computing software used here is Matlab ®.

2.4.2.1 Central specification

The initial value of capital is set quite low (several orders of magnitude below the equilibrium capital level), to study the quickness of convergence of the economy from a very low level of capital. The initial share of educated individuals in the middle-aged generation is set to zero, to reflect an economy with low labor productivity. Table 2.1 shows the values that were chosen to simulate the path of the economy towards steady state. The simulations are run for $G = 40$ generations.

Parameter name	Parameter description	Value
k_0	Initial level of capital	0.001
μ_0	Initial share of educated people	0
α	Share of labor in production	2/3
β	Discount factor	0.7
η	Skill premium	3
ψ	Relative political power of the old	1
n	Fertility rate of the middle-aged	0.10
δ	Survival probability	0.80

Tableau 2.1: Baseline parameters for the simulation of the politico-economic dynamics

In order to make sense of the orders of magnitude represented here, one must bear in mind the fact that one period in the model corresponds to a generation, i.e. between 25 and 30 years. Therefore, the value of β adopted here actually corresponds to an implicit annual subjective discount rate of around $\beta^{1/25} \approx 0.986$.

The other main parameters here were taken as plausible values to represent an advanced economy in the present period. Indeed, the parameter n describes the ratio of the population of the middle-aged generation to the young population. This can naturally be interpreted in terms of TFR: indeed, the total fertility rate can be roughly interpreted as the average number of children had by a woman with a full reproductive cycle, so that an increase of 10% of the population in a generation corresponds more or less to 2.2 children per woman, a number between the TFR of the USA (2.06) and that of the entire world (2.53) for the 2005-2010 period.[28]

It is harder to find a perfect empirical counterpart for the probability to survive into old age δ as defined in the model: indeed, the model assumes that individuals either die before reaching the age when they are eligible for pensions, or live for a full period of about 25-30 years thereafter, which is a gross simplification of the real

[28] Source: United Nations (2013).

demographic process where individuals are continuously subject to mortality risk. The best estimates I can compute come from life table survivors statistics, which give the number of survivors by age of a hypothetical cohort of a given number of individuals that would be subject during their whole lives to the prevalent mortality rates of a given period. To get an order of magnitude for δ, I take from these tables the ratio of the number of survivors at age 65 (a few years into the old age in the model) to the number of survivors at age 40 (roughly in the middle of the middle-aged period in the model): I obtain 0.83 for the more developed regions of the world in the 2005-2010 period.[29]

Finally, the value for the skill premium η finds its empirical counterpart in the ratio between (some measure of) the average skilled and unskilled wage. For instance, in the U.S. in 2013, the median weekly earnings of the whole population was $827, compared to $1,714 for individuals holding a professional degree and only $472 for individuals having less than a high school diploma ($651 for those with a high school as highest diploma),[30] so that a value of $\eta = 3$ is definitely not out of range.

2.4.2.2 The impact of demography on factor accumulation and growth

Changes in the fertility rate n Figure 2.1 depicts the trajectory of the economy for three values of the fertility rate n, holding δ fixed. This comparative dynamics exercise shows unambiguously that in this range of parameters, a higher value of n is detrimental to factor accumulation and growth. Indeed, higher values of n lead to lower levels of physical and human capital per capita along the whole path towards the steady state, as evidenced by panels (e) and (f). This translates into a lower steady-state level of output per capita y (panel (d)).

As evidenced by equation (2.16), a higher level of n implies that the factor dilution effect on capital is naturally higher, which explains the lower level of capital per worker in steady state. Besides, panels (a) and (b) of the figure show that both tax rates decrease when fertility increases: this effect goes against the previous effect of fertility, by increasing the disposable after-tax income, but is too weak to completely offset it.

The effect on the share of skilled workers, μ, which is a proxy of human capital, is then unambiguous. Thanks to equation (2.18), it is easy to see that less young will educate when n increases, both because of a direct "input dilution effect" (the subsidies

[29]Source: United Nations (2013). The more developed regions are defined there as Europe, Northern America, Australia/New Zealand and Japan.

[30]Source: U.S. Bureau of Labor Statistics (2013).

need to be shared among more young people), but also through the decrease in τ_e.

Finally, the two intergenerational transfer programs are negatively impacted by an increase in n, as shown in panels (g) and (h). Education transfers g are reduced for the reasons given in the previous paragraph. Concerning pensions,[31] the lower value of τ_p is only partially compensated by the fact that a higher n means the number of contributors relative to the number of recipients is higher, so that pensions by recipient decrease.

Changes in the survival rate δ Figure 2.2 shows how the trajectory towards steady state is impacted when δ changes, for a fixed value of n. A higher value of δ is unambigu-

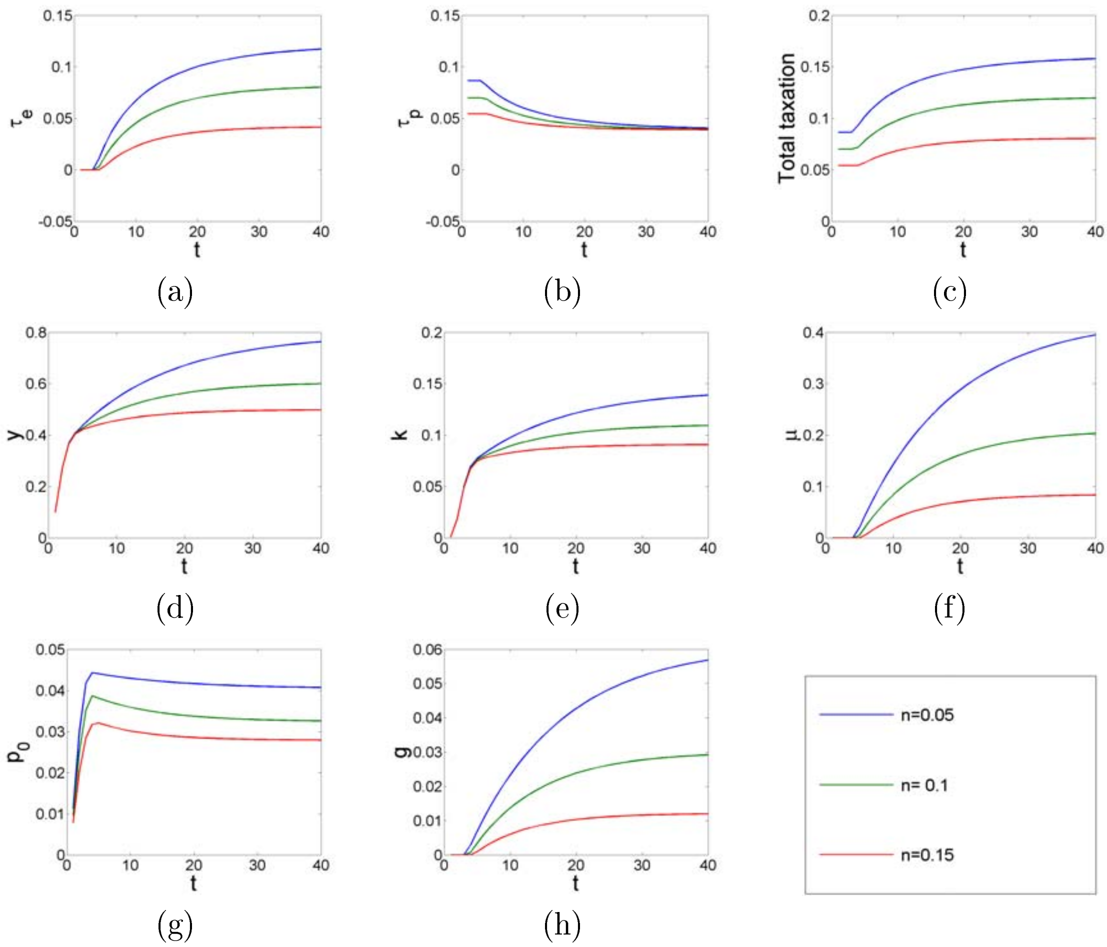


Figure 2.1: Dynamics of the main state and policy variables of the economy for different values of the fertility parameter n, and $\delta = 0.80$

[31] I represent in the figures the pensions accruing to the unskilled retirees p_0: of course, the pensions for skilled old people are equal to $p_e = \eta p_0$ since the system is assumed to be non-redistributive, so that the analysis applies to skilled pensions as well.

ously shown to be conducive to factor accumulation and growth. Indeed, higher values of δ lead to higher levels of physical and human capital along the whole path towards steady state (panels (e) and (f)). Therefore, the economy reaches a higher steady-state level of output per capita y (panel (d)) when δ is higher.

As evidenced by equation (2.16), a higher level of δ implies that individuals of working age save more, as their chances to reach old age are now higher. The fact that both tax rates increase when the survival rate increases (panels (a) and (b) of the figure) is insufficient to offset completely this first effect. The effect on the share of skilled workers, μ, which is a proxy of human capital, is also unambiguously positive, through the increase in τ_e: since current middle-aged individuals are more likely to reach old age next period, they are more enclined to investing in the human capital of their offspring.

Finally, the two intergenerational transfer programs are positively impacted by an increase in δ, as can be seen in panels (g) and (h). Education transfers g are increased for the reasons given in the previous paragraph. Concerning pensions, the higher value of τ_p is only partially compensated by the fact that a higher δ means a lower support ratio, so that pensions by recipient increase overall.

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

It is useful to compare the predictions of the model about the consequences of ageing, be it the analytical or simulation-based ones, to empirical evidence. In advanced economies, the late 19th and the 20th century were characterized by a secular decline in mortality and fertility rates, which can be interpreted respectively as a rise in δ and a fall in n. Judging from the simulations presented above, or notwithstanding the cross-taxation effects on welfare discussed in subsection 2.4.1 if I stick to the analytical comparative statics exercise, it seems to be that a decrease in n or an increase in δ should 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 when population ages. This seems to be in line with historical evidence of a rising share of spending in public transfers to GDP,[32] known as Wagner's law. 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,

[32]See Lindert (1994) and Lindert (1996).

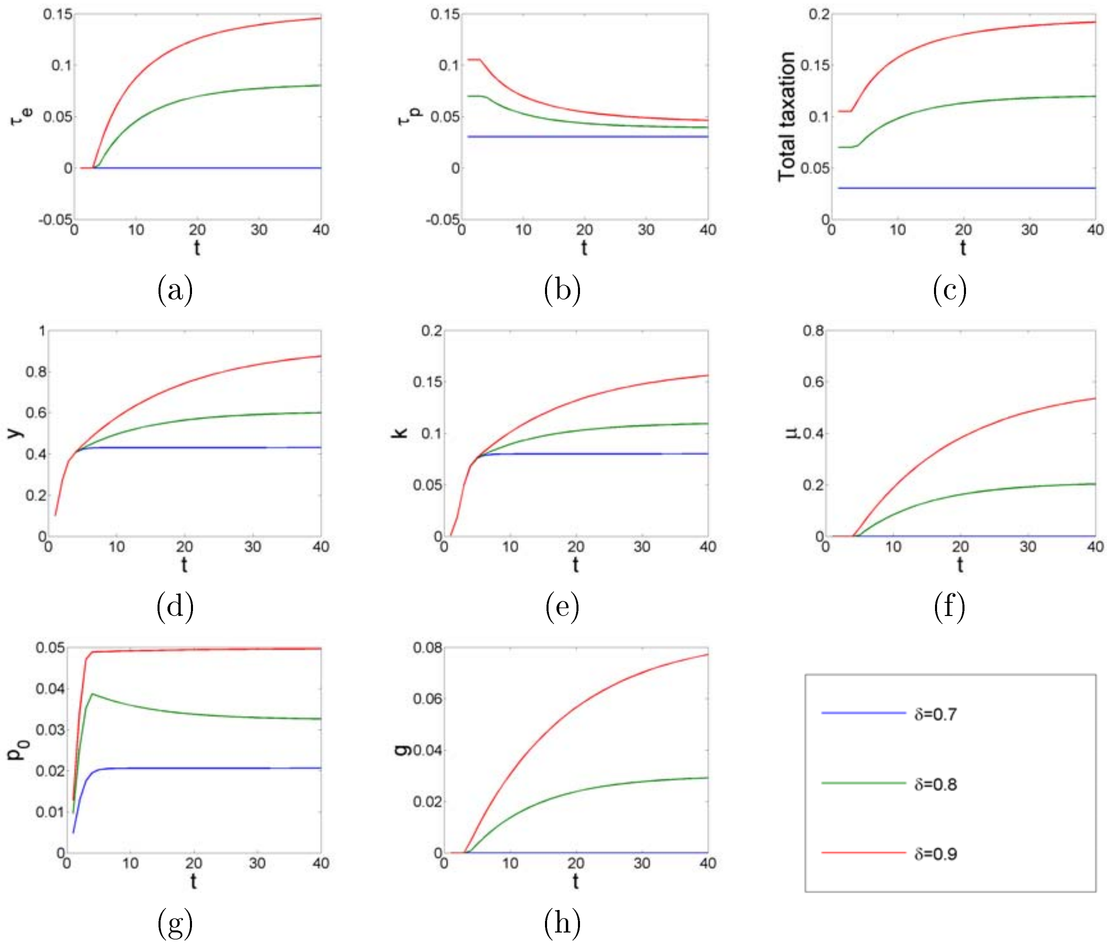


Figure 2.2: Dynamics of the main state and policy variables of the economy for different values of the longevity parameter δ, and $n = 0.10$

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 my 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. Profeta (2002), Sanz and Velazquez (2007) and other papers discussed in subsection 7.1 of the first chapter of this thesis also support in their majority the conclusion that total pension spending increases with the share of retirees, while it fails to significantly increase pensions per individual in many specifications.

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. Besides, the share of public spending going to education rises with the share of population of schooling age. These findings are at odds with the predictions of the model, according to which the total expenditures of both programs, as well as spending per recipient, should increase when the population ages, whatever the reason of this ageing is. Yet, as shown in the previous chapter, the estimates of the relationship between the age structure of the population and spending on public programs are dependent on the unit of observation chosen: see the different estimates obtained by Ladd and Murray (2001) and Harris et al. (2001) on school-district and county data, respectively, as described in subsection 8.2 of the first chapter of this thesis.

There are several limits to this exercise of comparing the available evidence with the predictions of the model. First, the analytical comparative statics exercise performed in subsection 2.4.1 is silent on which of the two tax rates would increase with δ (or decrease when n increases), even if the simulations suggest both taxes should increase with ageing. 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 my 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 people 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.

2.5 An exploratory analysis of the implications of dynamic expectations of future policies

In this section, I come back to the role played by the assumption of non-strategic behavior on behalf of the voters, and take a first pass at modifying assumption 4 , which deals with the way expectations of future policies are formed. This will naturally change the policy rule obtained in equilibrium (see subsection 2.3.3 : I then examine and interpret these changes to assess to which extent the dynamics of the model are dependent on this assumption.

2.5.1 Strategic behavior as a sensitivity analysis

Equation (2.32) of appendix 2.A sheds light on the dynamic nature of the policy choice on tax rates. Indeed, the indirect utility of a middle-aged individual is shown in equation (2.32) to depend on the level of capital in the next period, which itself depends on the expected level of the future contribution rate on pensions, $\widetilde{\tau_{p,t+1}}$.[33] But it depends as well as on $\widetilde{\tau_{p,t+1}}$ directly, since a higher level of τ_p in the next period undoubtedly raises the level of pensions accruing to the individual (who will be retired then) in the next period. The dynamic nature of the problem then lies in the fact that the pension contribution rate in next period is itself determined as the outcome of next period's voting process, which itself depends on expectations of the policy two periods from now, and so on...

By making assumption 4 before, which consists in saying that individuals simply disregard the impact of their current policy choice on next period's τ_p rate, I severed the link between policies at different periods and was able to reduce this dynamic politico-economic model to a series of static choices of policy bundles $(\tau_{e,t}, \tau_{p,t})$, whose sole link between one another is the evolution of the state variables k_t and μ_t. While this assumption seems perfectly reasonable if agents of a given cohort feel that their vote changes little to the outcome (i.e. if they are atomistic in the voting game), and hence to the policy in the next period, it might not hold if agents in a well-identified generation are organized enough that they start being able to sway the policy choice collectively with their vote.[34] In that case, agents will naturally take into account the indirect effect of their current choice on future policies, which will have the effect of changing the policy outcome in the current period.

Therefore, by observing by how much the policy outcomes change when I relax assumption 4 , I will be able to determine how sensitive my initial results were to this assumption. In the next subsection, I then set out to replace assumption 4 with an assumption that agents do take into account changes in future policies when making their current choice, which I will henceforth refer to as a "strategic" behavior.

[33] This is due to the fact that savings by either type of middle-aged workers are negatively impacted by expectations of higher pensions in the future, as shown in equation (2.12).

[34] See for instance Mulligan and Sala-i Martin (1999) for evidence on the ability of voters to organize themselves to make their voice heard, especially among the elderly.

2.5.2 Modelling the equilibrium using a Parameterized Expectations Algorithm with linear specification

The previous subsection has underlined the inherent dynamic, forward-looking nature of the policy decision problem for the agents, as the choice of $(\tau_{e,t}, \tau_{p,t})$ depend on future expected values of the pension contribution rate $\widetilde{\tau_{p,t+1}}$, which is itself a function of the state variables k_{t+1} and μ_{t+1} and $\widetilde{\tau_{p,t+2}}$, and so on...

Dropping assumption 4 means that the policy function $\tau(y) = (\tau_e(y), \tau_p(y))$ now needs to solve a fixed-point equation, where expectations of future policies are correct and happen to be the policy rule itself (albeit of a different level of the state variables):

$$\begin{aligned} \tau(y_t) &= \underset{(\tau_{e,t},\tau_{p,t})}{\operatorname{argmax}} W_t(y_t, \tau_{e,t}, \tau_{p,t}, \widetilde{\tau_{p,t+1}}) \\ \widetilde{\tau_{p,t+1}} &= \tau(y_{t+1}(\tau(y_t))) \end{aligned} \tag{2.24}$$

Finding an exact, analytical solution to this functional fixed-point equation is impossible in this case. Several methods have been proposed to yield an approximation of the solution (Judd, 1998 ; Stokey et al. , 1989). I will use here the method called Parameterized Expectations Algorithm (introduced by den Haan and Marcet (1990)), henceforth PEA, which consists in positing a functional form $\mathcal{T}$ for the expectation of future policy (in the present case $\widetilde{\tau_{p,t+1}}$) as a function of the value of the state variable in the next period, with some *a priori* unknown parameters Θ :

$$\widetilde{\tau_{p,t+1}} = \mathcal{T}(y_{t+1}; \Theta) \tag{2.25}$$

The PEA proceeds as follows. Starting from an initial guess of the value of the parameters $\Theta^{(0)}$, a current policy function $\tau^{(0)}(y)$ can be computed (up to a numerical approximation) under the assumption that $\widetilde{\tau_{p,t+1}} = \mathcal{T}(y_{t+1}; \Theta^{(0)})$. The current policy function $\tau^{(0)}(y)$ is then used to estimate a new value of the parameters of the expected future policy function $\Theta^{(1)}$, using econometric methods. The method is then further iterated until two consecutive inferred values of the parameters $\Theta^{(i)}$ and $\Theta^{(i+1)}$ are deemed close enough for the algorithm to have converged.

The most common family of functional forms used to implement the PEA in practice are polynomial functions, since polynomials are dense (in L^∞ norm) in the space of all functions.[35] However, the expected future policy function can only be correctly approximated provided the order of the polynomial is high enough: depending on the case, the PEA may then become too computationally intensive (due to the high number of parameters to find and update at each iteration of the algorithm) to be a satisfactorily implementable method in practice.

In the following subsection, I present the results of simulations of the politico-economic dynamics once expectations of future policies are modelled as a linear function of the future level of production. This (admittedly rather crude) approximation of the future policy function is the only one that I managed to obtain when implementing the PEA, as higher-order polynomial approximations of the future policy rule did not allow convergence of the algorithm. The following exercise must then be understood as a first pass in relaxing assumption 4 , and replacing it with an assumption of a more elaborate behavior on behalf of the agents.

2.5.3 Simulation results and interpretation

The results reported here were obtained after running a linear specification of the Parameterized Expectations Algorithm with unknown parameter $\Theta = (\theta_1, \theta_2)$, as described in equation (2.26):

$$\widetilde{\tau_{p,t+1}} = \theta_1 + \theta_2 y_{t+1} \tag{2.26}$$

I used the baseline values of the model parameters given in table 2.1 , and updated parameters by running an OLS regression on the current policy function, for each loop of the algoritm. Whatever the intial $\Theta^{(0)}$, the PEA converges to the final value $\Theta_f = (0.0826; -0.0344)$. The fit of the policy rule on τ_p with respect to this linear approximation is admittedly not very good, with an adjusted $R^2 = 0.493$: figure 2.3 represents the policy rule for τ_p as a function of the current level of output per capita y, with the expected future policy as a function of the same variable. Looking at figure 2.3 , it seems clear that a better approximation of the true policy variable would be obtained by running at least a quadratic approximation, which unfortunately is not obtainable with a PEA in this case. Indeed, the attempts made have revealed that the algorithm keeps oscillating between two values of Θ, which I take as a sign that the model might feature multiple equilibria. Another attempted modeling of the expectation function $\mathcal{T}$ as a piecewise linear function of y has also proved unsuccessful, as the algorithm failed to deliver a fixed point in Θ.

[35] This result is known in mathematics as the Stone-Weierstrass theorem.

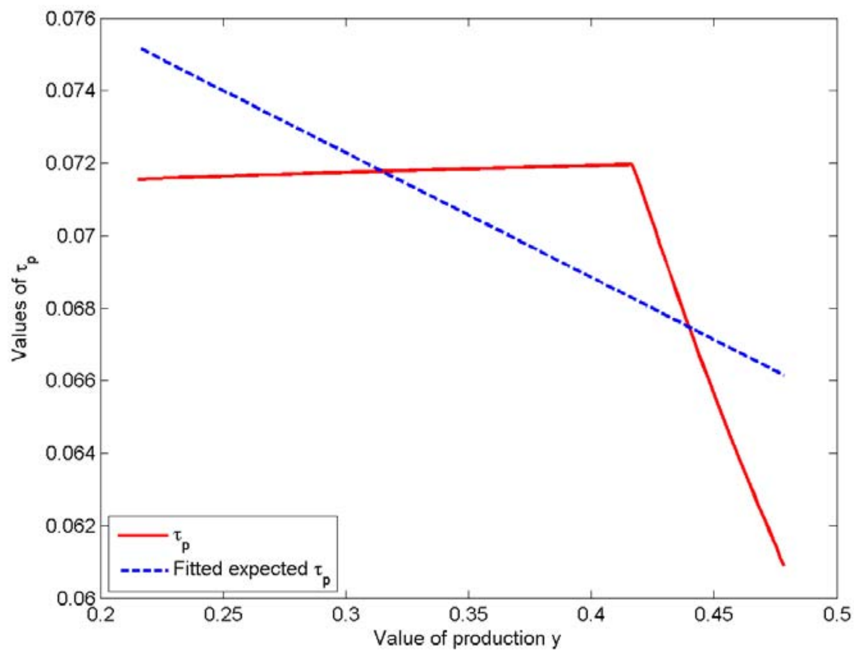


Figure 2.3: True (in red) and expected future (in blue) policy rules as a function of output

Although this linear approximation for expectations is admittedly far from perfect, it provides a benchmark with which to compare the model with static (i.e. "non strategic") expectations. To that effect, figures 2.4 and 2.5 show how the different policy rules are modified by shocks on n and δ, respectively, both in the strategic and non-strategic cases.

From a qualitative point of view, it should be noted that the shapes of each policy rule for a given value of the demographic parameters are preserved, when one switches from the strategic case to the non-strategic one. Additionally, these figures allow to compare the policy rules in the strategic and the non-strategic cases, for given values of the n and δ. It turns out that whatever the value of output per capita y, education receives more financing in the non-strategic than in the strategic cases: both the dedicated tax rate τ_e and the resulting transfer g are higher in the non-strategic case, as shown in panels (a) and (c) of figures 2.4 and 2.5 . Conversely, pensions are better financed (higher τ_p and p_0) under the strategic behavior hypothesis, as evidenced by panels (b) and (d) of the same figures.

Therefore, the political decisions are found to be more conducive to long-run growth if I assume agents to be non strategic, since the human capital-building transfer is favored more, while the purely redistributive one (which diminishes disposable income and thus

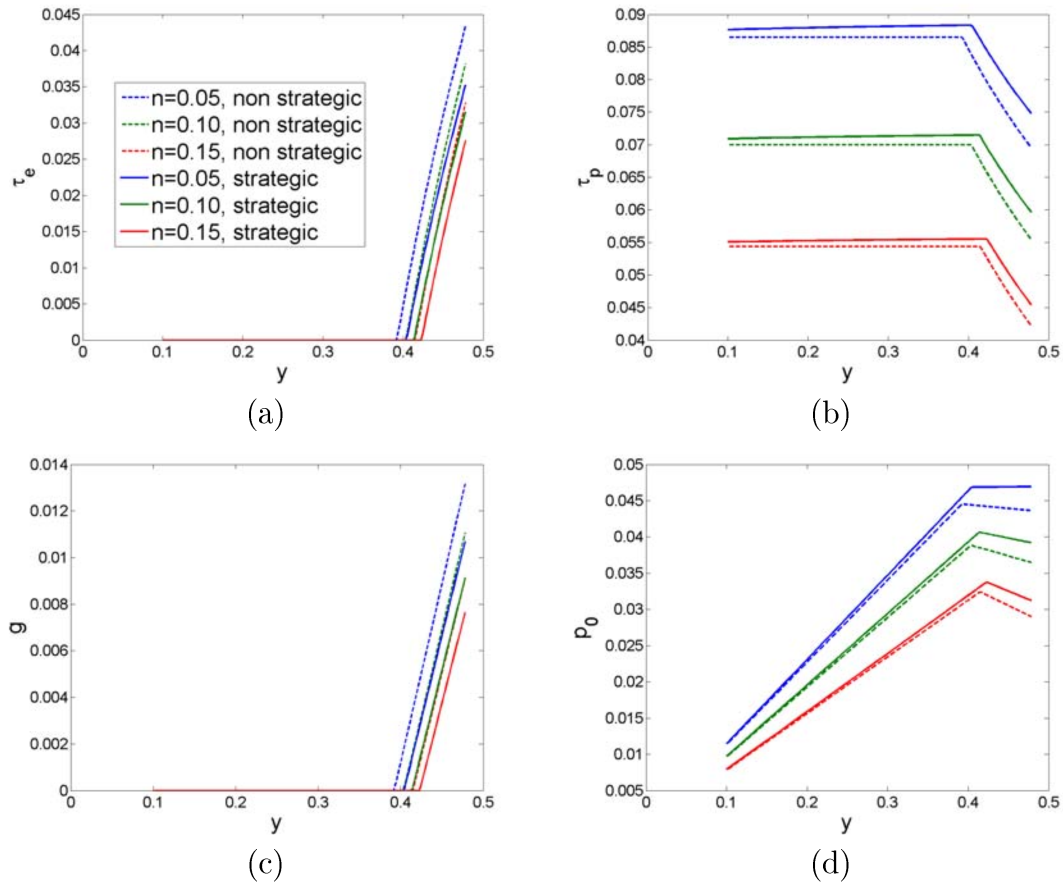


Figure 2.4: Policy rules as a function of output in the strategic (solid line) and non-strategic (dashed line) cases, for different values of n, and $\delta = 0.80$

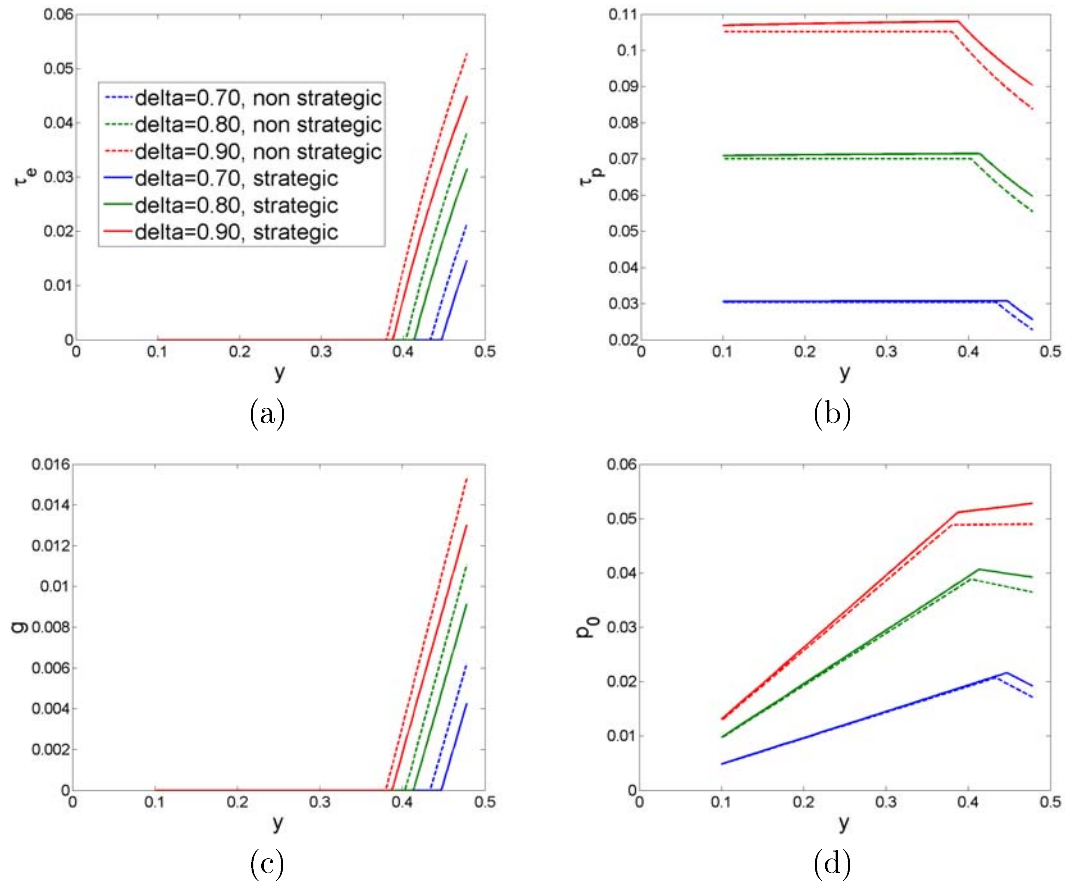
reduces physical capital accumulation) is less favored. An explanation is most likely to be found by examining figure 2.3 : under the hypothesis of strategic behavior, middle-aged agents in period t expect that an increase in future output per worker y_{t+1} will result in a lower value of $\tau_{p,t+1}$, and thus slightly less funding for their future pensions. This adds another element in the trade-off between $\tau_{e,t}$ and $\tau_{p,t}$, when compared to the case of non-strategic behavior, which explains why middle-aged agents are *ceteris paribus* less willing to undertake investment in future productivity under the assumption of strategic behavior.

Besides, the ordering of policy rules when either demographic parameter varies is left unchanged by adopting the assumption of strategic behavior instead of the non-strategic one. Indeed, figure 2.4 shows that for all values of output per capita y, an increase in n means a decrease in both tax rates τ_e and τ_p, but also in the transfers received per individual g and p_0. This statement is true both in the strategic and non-strategic case. Similarly, for all values of output per capita y, an increase in δ means an increase in both tax rates τ_e and τ_p, but also in both transfers g and p_0, as shown in figure 2.5 .

To summarize, this section has shown that it is quite difficult to implement successfully the PEA algorithm to solve the model under the assumption of strategic behavior. Despite its shortcomings, the linear specification I adopted for the expectations of the future pension contribution rates delivers a useful sensitivity analysis on the restrictiveness of the assumption of naive (or non-strategic) expectations. This analysis shows that the behavior of the policy rates, and then ultimately of the whole model, is preserved from a qualitative point of view when dropping the assumption of naive expectations in favor of the linear approximation of strategic expectations.

2.6 Conclusion and discussion

This paper takes a stance on the politico-economic links that exist between two of the most important components of state spending, social security and publicly funded education, and the age structure of population. 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.



(a) (b) (c) (d)

Figure 2.5: Policy rules as a function of output in the strategic (solid line) and non-strategic (dashed line) cases, for different values of δ, and $n = 0.10$

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. Using simulation for plausible values of the main parametters of the model, I also find that both transfers (measured both in the aggregate and on a per recipient basis) are prone to rise when population ages, and that ageing is also conducive to an increase in accumulation of physical and human capital, and ultimately to growth in output per capita.

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 endogenize 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.

Appendices

2.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:

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

$$\Rightarrow \frac{\beta\delta}{1+\beta\delta}\gamma_{0,t} = s_{0,t} + \delta\frac{p_{0,t+1}}{R_{t+1}} \tag{2.27}$$

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,

$$\frac{p_{0,t+1}}{R_{t+1}} = \frac{\alpha}{\delta(1-\alpha)} \cdot \frac{\tau_{p,t+1}(1+n)k_{t+1}}{\mu_t \eta + 1 - \mu_t} \tag{2.28}$$

Using (2.27) and (2.28), one now gets to the following equation linking next-period capital and the NPV of expected lifetime income :

$$\frac{\beta\delta}{1+\beta\delta}\gamma_{0,t} = \frac{k_{t+1}(1+n)}{(\mu_t\eta + 1 - \mu_t)} \cdot \left(1 + \tau_{p,t+1}\frac{\alpha}{1-\alpha}\right) \tag{2.29}$$

Computing the derivatives The welfare function that is maximized 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 maximize 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 (2.29):

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

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

$$\frac{\partial V_{0,t}^{t-2}}{\partial \tau_{p,t}} = \frac{\alpha}{1-\alpha+\alpha\tau_{p,t}} \tag{2.30}$$

$$\frac{\partial V_{0,t}^{t-2}}{\partial \tau_{e,t}} = 0 \tag{2.31}$$

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 (2.8) and (2.29), it obtains that:

$$
\begin{aligned}
V_{0,t}^{t-1} = \; & (1+\beta\delta-\beta\alpha\delta)\ln k_{t+1} + \beta\alpha\delta \ln((\eta-1)\mu_{t+1}+1) - (1+\beta\delta)\ln(\mu_t(\eta-1)+1) \\
& +(1+\beta\delta)\ln(1+\tau_{p,t+1}\alpha/(1-\alpha)) + c
\end{aligned}
\tag{2.32}
$$

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:

$$
\frac{\partial V_{0,t}^{t-1}}{\partial \tau_{p,t}} = (1+\beta\delta-\beta\alpha\delta)\frac{\partial \ln k_{t+1}}{\partial \tau_{p,t}} = -\frac{1+\beta\delta-\beta\alpha\delta}{1-\tau_{e,t}-\tau_{p,t}} \tag{2.33}
$$

$$
\begin{aligned}
\frac{\partial V_{0,t}^{t-1}}{\partial \tau_{e,t}} &= (1+\beta\delta-\beta\alpha\delta)\frac{\partial \ln k_{t+1}}{\partial \tau_{e,t}} + \beta\alpha\delta\frac{\partial \ln((\eta-1)\mu_{t+1}+1)}{\partial \tau_{e,t}} \\
&= -\frac{1+\beta\delta-\beta\alpha\delta}{1-\tau_{e,t}-\tau_{p,t}} + \beta\alpha\delta\frac{(\eta-1)\alpha y_t}{\bar{\omega}(1+n)+\tau_{e,t}(\eta-1)\alpha y_t}
\end{aligned}
\tag{2.34}
$$

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 (2.20) and (2.19).

2.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.

2.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 (2.19) and (2.20), are equal to zero at (τ_e, τ_p). Then total taxation cannot go up following an increase in n: indeed, equation (2.19) 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 (2.20), 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 (2.19) and (2.20) (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.[36] As a result, total taxation effectively goes down in this case following an increase in n.

2.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 (2.19) and (2.20) 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.

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

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.

3

Commitment in a Political Economy Model of Education and Redistribution

3.1 Introduction

3.1.1 Motivation

Human capital accumulation, which is a central requirement to achieving sustained growth in advanced and developing economies alike, is usually funded by the government in large amounts. For instance, in OECD countries, public funding accounted in 2011 for 84% of all funds for educational institutions on average, a share that rises as high as 92% if one considers primary and secondary institutions alone (OECD, 2014a).

Reasons for state intervention in either providing education directly or subsidizing its private acquisition are manifold, from the existence of spillovers and positive externalities (Acemoglu and Angrist, 2001 ; Heckman and Klenow , 1998 ; Lucas , 1988) to the possible existence of credit constraints bearing on education acquisition (Barham et al., 1995), even though the prevalence of these constraints is debated (Carneiro and Heckman, 2004). Other justifications for state intervention may be found in the role played by education in shaping income inequality at the very same time it spurs growth (Benabou, 2002; Galor and Zeira , 1993 ; Glomm and Ravikumar , 1992), as well as considerations of intergenerational mobility (Arawatari and Ono, 2013 ; Ichino et al. , 2011). It therefore appears particularly important, both from a theoretical and policy point of view, to study the circumstances that allow the population to support adequate funding of the educational system.[1]

Several questions arise concerning this political consent to fund education, that have already been dealt with in the literature. One of the main concerns is that public financing of education (and especially tertiary education) may be regressive, in the sense that it benefits more individuals who have already large endowments, be it in terms of social capital, or parental human capital and wealth (Arrow, 1971 ; Hare and Ulph , 1979): this may undermine the support the general population has for these policies.

Other concerns point to the detrimental effect of taxation (especially of labor income) on human capital accumulation (Milesi-Ferretti and Roubini, 1998 ; Nielsen and Sorensen, 1997). These concerns may become particularly acute once taking into account the fact that governments might face commitment problems not to expropriate parts of the human capital investment made by individuals, a problem which is likely

[1] An excellent introduction to the topic of the political support for public education systems can be found in Gradstein et al. (2004).

to come up since individuals typically need to invest in their human capital at an early stage of their life, long before entering the labor market where they can reap the benefits of their education (Andersson and Konrad, 2003 ; Boadway et al. , 1996).

In this chapter, I examine the conditions under which political processes might lead to an efficient level of public subsidies to education acquisition, when the space of available policies also includes labor income redistribution that distorts said education acquisition. To this end, I set up a two-period political economy model that builds on a previous, normative model by Bovenberg and Jacobs (2005), where individuals privately choose their level of education expenditure and their labor supply, and decide collectively on a linear education subsidy and a linear labor income tax.

When the government is unable to commit to a labor income tax rate before individuals select their level of education, I show that the electorate collectively selects too high a level of education subsidies in the first stage of the voting game, even when compared to the high level of labor taxation selected in the second stage. In that respect, the existence of high degrees of public subsidies to education does not necessarily signal the existence of virtuous institutions promoting long-term growth, but may rather be understood as proof of the existence of time inconsistency at the level of the electorate, when setting taxes and subsidies.

3.1.2 Relationship with the previous literature

This chapter is obviously closest to the first sections of Bovenberg and Jacobs (2005), in which the authors explicitly tackle the problem of selecting the optimal levels of linear education subsidies and labor income tax rates for a planner with redistributive concerns. They show that whatever the level of redistribution, the planner should aim at a zero net tax on human capital acquisition, which is obtained for a rate of education subsidy equal to the labor income tax rate. Jacobs and Bovenberg (2011) qualify this previous efficiency result by showing that, in the linear case, the optimal policy requires a zero net tax on human capital only in the case where the earnings function is weakly separable between education on the one hand, and labor supply and ability on the other hand, and when it also features a constant elasticity in education.[2] All these conditions are satisfied in the specification I adopt in this chapter. My contribution relative to these works is to show that these normative requirements cannot be decentralized through a political

[2]Maldonado (2008) also stresses the importance of the education elasticity of the wage function in determining the optimal (nonlinear) income tax and education subsidy.

process, in the case where the tax policy cannot be credibly set before individuals make their education choices.

In another work, de Fraja (2002) studies optimal education policy alongside income taxation, in a model with heterogeneity in both ability to pay and ability to benefit from education, assuming positive externalities of education. Under the assumption that the government can directly provide education in a non-uniform way to individuals, and that it has access to nonlinear income taxation, the optimal policy is found to be both input and output regressive, meaning that higher-ability individuals are subsidized by less able households, and that wealthier households contribute less to the overall education budget than poorer ones. This efficiency result rests heavily on the existence of capital market imperfections, a dual (i.e. public and private) education system and externalities in education acquisition. The intuition behind it seems quite at odds with the predictions of the model developed hereafter of a redistribution from the most able towards the least able individuals.

Other recent works have studied from a normative point of view the joint determination of income taxation and education taxation (or subsidy). Cremer and Pestieau (2006) study the opportunity of subsidizing private education with respect to directly providing public education, in a setting where tax policy is available as a redistributive tool. Individuals are assumed to differ in their productivity at work, and display (warm-glow) altruism toward their children. They develop a setting in which the Atkinson-Stiglitz proposition holds,[3] so that private education spending is subsidized or taxed not for redistributive reasons, but with the aim of correcting externalities. They also show that public provision of education may be implemented even though it is a perfect substitute to private education, in the case where the government deliberately omits the joy-of-giving component of utility that parents associate with private investments in their children's human capital.

Besides, Blumkin and Sadka (2008) seek to determine the conditions under which education acquisition might be directly taxed for redistributive purposes. They consider an economy in which agents differ in both their innate productivity at work and their cost to educate, education being a zero-one decision that multiplicatively increases the innate productivity. It is assumed that both characteristics are unobserved by the government, while labor supply and educations are observable. Therefore, the authors

[3] Atkinson and Stiglitz (1976) have shown that if the government has access to nonlinear income taxation and labor supply is separable from other commodities entering the utility function, it is unnecessary to use indirect taxes.

show that it can be optimal to directly tax education acquisition on top of labor income at the second best, so as to redistribute from the individuals with a low cost to educate toward those with a high cost. This is made possible by the fact that the education decision brings additional information on the type of individuals, on top of the labor supply decision.

The results contained in this chapter rely heavily on the notion that the government cannot commit on future taxation of earnings at the time when agents pick their level of human capital investment. To that extent, it relates to works that take into consideration issues of commitment in the taxation of human capital. For instance, Findeisen and Sachs (2014) find that if the government has limited or no commitment on the future level of income taxation, then it should act in order to compress the distribution of wages in the first stage when setting its education policy. The same line of argument is developed by Gradstein (2000) to justify the existence of public education in the first place, which is uniformly provided to agents and thus reduces pre-tax income inequality. Such a first-period compression of inequality is impossible in my model, however, given the functional form of the relationship between education and earnings adopted in my model.[4]

As stated above, I focus in this chapter on the joint political determination of income redistribution and education subsidies, which I derive from the heterogeneity in intrinsic ability of the population and the political institutions available.[5] Depending on the nature of the heterogeneity assumed and the set of available policies, various outcomes may arise at the political equilibrium.[6] I briefly describe some examples from the recent literature on the subject.

Fernandez and Rogerson (1995) set up a political economy model in which agents are *ex ante* identical except for their wealth, and credit constraints affect education decisions. They show that public subsidies to education redistribute income toward higher-income individuals. Indeed, the poor are barred from acquiring education in equilibrium, which only features partial subsidization of education expenditures, while they nonetheless contribute to financing the subsidies through an income tax falling on

[4]It actually turns out this feature is central in breaking the link between the education subsidy rate in first period and the income tax rate decided in the second one.

[5]Here, by political institutions I mean the inability to commit to a tax level before education acquisition, which results in a certain timing of political and private decisions, and the availability of linear policies only.

[6]The political science literature has studied the nature of the coalitions supporting education finance policies, as well as the link between education funding and individual preferences for redistribution: see for instance Busemeyer (2013) and Busemeyer and Iversen (2014).

everyone. Their findings are at odds with other models such as Glomm and Ravikumar (1992) or Saint-Paul and Verdier (1993), in which education provision is assumed to be uniform: in these cases, the political equilibria imply a net transfer from the initially better-off individuals toward the poor, a feature shared by my model although for different reasons.

Bernasconi and Profeta (2007) feature an economy in which individuals belonging to two different social classes (rich and poor) can vote on two public transfer programs, involving respectively redistribution and education. Only education is conducive to growth, and also increases social mobility as it provides opportunities for the children of the poor to be recognized for their talent. The authors find that the poor may prefer public spending on education to direct redistribution, while the rich prefer redistribution, as education implies more competition for good jobs from the poor. Another example of a paper studying the joint determination of redistribution and education is Levy (2005). In this case, however, the political equilibrium results from a generational conflict between the young poor, who favor public education and may form a coalition with the rich to reduce redistribution of income, and the old poor who favor redistribution.

3.1.3 Outline of this chapter

The remainder of this chapter is organized as follows. In section 3.2 , I set up a two-period model of an economy which is quite close to Bovenberg and Jacobs (2005), in which agents have to choose privately both their education and labor supply, while collectively deciding on a linear education subsidy rate and linear labor income tax rate, in order to study the relationship between income redistribution and education finance in a political economy setting. A particular attention is paid to the order in which the collective and private decisions are taken. Section 3.3 describes the functioning of the government in this model, and gives an account of the optimal policy a planner with redistributive concerns would choose in this setting. This serves as a benchmark for the equilibrium of the political game, which is derived in section 3.4 . This section also compares the results obtained with a slight alteration of the timing of the collective decision on the level of the income tax rate. Section 3.5 then makes briefly the case that one of the main propositions of the chapters, the fact that the education subsidy rate is larger than the labor income tax rate at the political equilibrium, is verified in data concerning a subsample of OECD countries. Finally, section 3.6 discusses the previous results and concludes.

3.2 Basic setup of the model with self-financing of education

3.2.1 Overview of the model

In this section, I set up a two-period economy in which agents will essentially make two decisions, concerning respectively education and labor supply. Agents differ by their intrinsic productivity at work, also referred to as innate ability, which will be augmented by the education which they choose to acquire. It is assumed that the intrinsic productivity is known before the decision to educate. The monetary cost of education, which needs to be purchased by agents, and the disutility associated with labor supply are assumed to be the same across individuals. According with intuition, more able agents will choose to supply more labor, as their higher productivity translates into higher wages. Additionally, this higher labor supply will make education more profitable to acquire, since education costs do not scale up with labor supply while its benefits do. In turn, this additional education increases the marginal productivity of labor of the agents: thus is created a feedback effect between education acquisition and labor supply, which is essential to the following framework.

Education costs are linear and can be expressed in terms of the consumption good. Private expenditures on education might be subsidized by the government at a linear rate. The government funds this subsidy by taxing labor income, also at a linear rate. On top of the subsidy, the tax is used to make lump-sum transfers to agents,[7] which implies the government can have a redistributive role.

I assume the government is ruled by one of two parties that compete freely for the vote of the whole population, as in Downs (1957). Parties have no ideologies, and are only motivated by holding office. They are elected based on their policy proposals for the education subsidy and income tax rates.

[7] I do not preclude negative lump-sum transfers so far. The only requirement is that the budget of the government be in equilibrium at the end of the two periods, which fully constrains the amount of lump-sum transfers made to the agents (see subsection 3.3.1).

3.2.2 Time structure and agents

The model unfolds over two periods, labeled $t = 1, 2$. The economy is populated by heterogeneous agents who differ by their intrinsic ability type γ_i, and live both periods. Their mass is normalized to one. The government is run by one of two competing, purely office-motivated parties that we denote A and B. There is no uncertainty in the model. The distribution of ability types $\{\gamma_i\}_i$ across the population of one given generation is common knowledge. Additionally, each agent privately knows his own ability γ_i before making any decision, be it private or political.

From here on, a distinction must be made between two possible time structures, that differ by the timing at which agents vote on the two policies and the relative timing of private decisions. It can either be assumed that agents vote first on education subsidies, then educate, and lastly choose collectively the labor income tax rate, taking into account the distortions it creates on labor supply. This timing is hereafter referred to as the sequential vote case, and will be used for most of the analyses made in this chapter. On the contrary, it can be assumed that agents first vote on the two tax rates at the same time, and then perform their two private decisions to educate and supply labor: this later case will be referred to as the simultaneous vote case, and is tackled in appendix 3.A .

Timing (sequential vote case):

1. ($t = 1$) Parties A and B announce a binding policy proposal for the subsidy rate on education s. Agents elect the government for one period based on the policy proposals.

2. Private agents decide on their educational inputs purchase e_i (at an exogenous unit cost of p_e in terms of the consumption good)

3. ($t = 2$) Gross wage w_i of individual i is given by education e_i and (heterogeneous) ability γ_i. Parties A and B announce this time a binding policy proposal for the income tax rate τ. Agents elect the government for one period based on the policy proposals.

4. Agents choose their labor supply l_i. Their gross wage income $z_i = w_i l_i$ is taxed to finance redistribution, recoup subsidies and finance lump-sum transfers to all individuals. Agents pay their education costs and consume c_i.

This timing of actions can be understood as the result of the inability of the government to commit to a given level of labor income taxation at the time when agents purchase their education (see again Andersson and Konrad (2003) or Boadway et al. (1996)). Even though I assume the parties can commit to their policy proposal in each relevant period (i.e. commit on the level of subsidies in period 1, and the level of taxes in period 2), there exists no mechanism here to commit in advance to a future level of taxes when in the first period. This is a reasonable assumption if one considers that several years will elapse between the time an individual makes his education decision, and his entry in the labor market.

A fair objection to the sequentiality of the political choices I adopted here is the fact that real-world legislatures typically decide on education funding and taxation every year when voting on the budget, and that these two political choices can thus only be described as occurring simultaneously. While relevant, this objection ignores the fact that the two taxation and subsidy choices, though simultaneous, actually impact different generations, since the one being taxed is the one participating in the labor market while the one benefiting from the subsidies merely acquires skills, and does not participate in the labor market yet. As a result, from the point of view of a generation, the relevant timing of votes is sequential and not simultaneous.[8]

I thus argue that modelling the political choices as sequential rather than simulatenous is the best option so far, as I stick to analyzing the choices that are relevant for a given generation taken in isolation, and I voluntarily disregard the interaction between the choices made by different generations and issues relative to intergenerational redistribution of public funds. In this respect, appendix 3.B will extend the validity of the current analysis, by explicitly considering the interaction between successive generations in a case with altruistic preferences.

3.2.3 Private behavior

Private utility of agent i is quasilinear, and takes into account both consumption and labor supply:

$$u(c_i, l_i) = c_i - \frac{l_i^{1+1/\epsilon}}{1 + 1/\epsilon} \tag{3.1}$$

[8]Yet another interpretation consists in viewing the subsidy rate s as an entitlement to a minimal level of education which can seldom be modified (such as, in many economies, the right to free primary or secondary education), while the tax rate and lump-sum transfers can be changed every year while voting on the budget.

where ϵ is the wage elasticity of labor supply.

I assume wage to be given by the following wage formation equation:

$$w_i = \gamma_i e_i^{\beta} \tag{3.2}$$

Equation (3.2) can be understood as a simple example of a Mincerian wage equation (Mincer, 1974), withβ the elasticity of wage with respect to education, and γ_i the unobserved individual productivity of each agent, which we hereafter refer to as the intrinsic productivity or innate ability.

It is assumed agents only work and consume in the second and last period of their life. The budget constraint of agent i then writes:

$$c_i + p_e(1-s)e_i = (1-\tau)w_i l_i + g \tag{3.3}$$

where g is a lump-sum transfer received from the government.

Private decisions For given values of s and τ, utility maximisation gives the following relationships between the labor supply and education choices for individual i:

$$l_i = \left[(1-\tau)\gamma_i e_i^{\beta}\right]^{\epsilon} \tag{3.4}$$

$$(1-\tau)\beta\gamma_i e_i^{\beta-1} l_i = (1-s)p_e \tag{3.5}$$

Combining equations (3.4) and (3.5), I get the following expression for education e_i as a function of the parameters and tax rates:

$$e_i = \left[\frac{\beta}{(1-s)p_e}\right]^{1/\mu} \left[\gamma_i(1-\tau)\right]^{(1+\epsilon)/\mu} \tag{3.6}$$

where $\mu = 1 - \beta(1+\epsilon)$ is an inverse measure of the feedback effects between education and labor supply.

Similarly, labor supply is given by:

$$l_i = \left[\frac{\beta}{(1-s)p_e}\right]^{\beta\epsilon/\mu} [\gamma_i(1-\tau)]^{\epsilon/\mu} \tag{3.7}$$

It is straightforward to see that individuals with higher innate ability γ will both educate more, and supply more labor. Indeed, a higher γ raises the returns to education, while labor supply l goes up since wages become higher, due to both the higher innate ability γ and the higher education level e.

It should be noted that the model would yield inconsistent results if we did not impose the feedback effects between education and labor supply to be finite, which means there should be decreasing returns in the joint decision to educate and supply labor.

Assumption 5 *Feedback effects between the education and labor supply decision margins are finite and positive:* $\mu = 1 - \beta(1+\epsilon) > 0$

Ceteris paribus, abler agents (i.e. agents with a higher γ) are therefore more educated, supply more labor and generate more wage income $z_i = \gamma_i e_i^{\beta} l_i$.

3.3 The government

In this section, I describe the budget constraint faced by the government. I also take a step back from the actual political game I described in section 3.2 to first give an account of the optimal government policy a planner with access to the same instruments would choose, a result due to Bovenberg and Jacobs (2005).

3.3.1 The government budget constraint

The government uses the receipts from the labor tax to fund the education subsidies, along with some (possibly negative) lump-sum transfers g. Its budget constraint writes:[9]

$$\tau \int_{\underline{\gamma}}^{\overline{\gamma}} \gamma_i e_i^{\beta} l_i dF(\gamma_i) = g + s \int_{\underline{\gamma}}^{\overline{\gamma}} p_e e_i dF(\gamma_i) \tag{3.8}$$

where F is the cumulative distribution function of types γ_i over the support $(\underline{\gamma}; \overline{\gamma})$.[10]

It should be noted that equation (3.8) clearly shows that the three instruments the government can pick are tied to one another, in such a way that choosing the level of τ and s automatically results in setting a value for the lump-sum transfer g. I acknowledge this fact by not including the choice of g in the description of the political game I made in section 3.2 .

3.3.2 Constrained planner problem using linear instruments

I now proceed to describing the optimal policy that a planner with a redistribution objective and access to the two linear tax and subsidy instruments would implement.[11]

It is assumed that a benevolent government maximizes a social welfare function over agents' indirect utilities $v(i; s, \tau, g)$ subject to its budget constraint (see (3.8) above).

The Lagrangian of the problem is:

$$\begin{aligned} \mathcal{L}(s, \tau, g) &= \int_{\underline{\gamma}}^{\overline{\gamma}} \Psi(v(i; s, \tau, g)) dF(\gamma_i) \\ &+\lambda \int_{\underline{\gamma}}^{\overline{\gamma}} (\tau \gamma_i e_i^{\beta} l_i - s p_e e_i - g) dF(\gamma_i) \end{aligned}$$

with Ψ a weighting function that satisfies the following required properties for the above formula to represent a generalized utilitarian social welfare function: $\Psi'(.) > 0$, $\Psi''(.) \leq 0$.[12]

[9]It should be noted that equation (3.8) is valid only if one assumes that the government can borrow without cost the amount necessary to finance the education subsidies in first period, and then repay the corresponding amounts in the second period.
[10]I do not preclude support of F from being over the whole domain of positive reals here.
[11]All the results in this subsection are taken from Bovenberg and Jacobs (2005).
[12]In the limit case where Ψ is the identity function, then the social welfare function is purely

Let now $b_i = \Psi'(v(n; s, \tau, g))/\lambda$ be the social marginal value of consumption of individual i. Then the first-order condition of the problem with respect to g writes:

$$\bar{b} = \int_{\underline{\gamma}}^{\overline{\gamma}} b_i dF(\gamma_i) = 1$$

I now proceed to define the distributional characteristic of the government policy as the (normalized negative) covariance between the social marginal value of consumption and the pre-tax labor income, over the whole population (Atkinson and Stiglitz, 1980 ; Feldstein, 1972):

$$\begin{aligned}\xi &= -\frac{\int_{\underline{\gamma}}^{\overline{\gamma}} b_i z_i dF(\gamma_i) - \int_{\underline{\gamma}}^{\overline{\gamma}} b_i dF(\gamma_i) \int_{\underline{\gamma}}^{\overline{\gamma}} z_i dF(\gamma_i)}{\int_{\underline{\gamma}}^{\overline{\gamma}} b_i dF(\gamma_i) \int_{\underline{\gamma}}^{\overline{\gamma}} z_i dF(\gamma_i)} \\ &= \frac{\int_{\underline{\gamma}}^{\overline{\gamma}} (1 - b_i) z_i dF(\gamma_i)}{\int_{\underline{\gamma}}^{\overline{\gamma}} z_i dF(\gamma_i)}\end{aligned}$$

The optimal policy using both instruments is then as follows:

$$s^* = \tau^* \qquad (3.9)$$

$$\frac{\tau^*}{1 - \tau^*} = \frac{\xi}{\epsilon/\mu} \qquad (3.10)$$

Unsurprisingly, the higher weight the government puts on the welfare of the low-ability types (i.e. the higher ξ is), the higher the labor income tax rate will be. It is also fairly intuitive that the more elastic labor supply is, the smaller τ^* is, since a more elastic labor supply will make the tax base more sensitive to changes in τ. The dependency in μ of the optimal tax rate τ^* at the optimum can be interpreted in roughly the same way: indeed, a lower value of μ means the education decision is more sensitive to the tax rate, as evidenced by equation (3.7).

I now compute the total net tax wedge on learning, denoted Δ, which is defined as the extent to which these instruments modify the marginal net return to education

utilitarian. In that case, the planner would aim at no redistribution at all, since individuals have a quasi-linear utility and the social marginal value of consumption is the same for all individuals.

acquisition. It can be computed as follows:

$$\begin{aligned} \Delta &= \tau \frac{\partial}{\partial e_i}(\gamma_i e_i^{\beta} l_i) - s \frac{\partial}{\partial e_i}(p_e e_i) \\ &= \frac{\tau}{1-\tau}(1-\tau)\gamma_i \beta e_i^{\beta-1} l_i - \frac{s}{1-s}(1-s)p_e \\ \Leftrightarrow \Delta &= (1-s)p_e \left(\frac{\tau}{1-\tau} - \frac{s}{1-s} \right) = 0 \end{aligned} \tag{3.11}$$

where equation (3.11) was obtained using equations (3.4) and (3.9).

It follows that education subsidies restore efficiency on the education choice margin, even though income redistribution distorts the labor supply margin, by compensating individuals for taxes imposed on the returns for learning. Under this type of policy, investments in learning can be interpreted as tax-deductible. This does not mean, however, that the levels of learning chosen by the individuals at the planner's optimum are the same as under the laissez-faire equilibrium. Indeed, the labor tax means a lower labor supply for all individuals, which means the human capital accumulated by individuals will be utilized at a lower rate than under the laissez-faire situation. What this result does mean is that the level of education acquisition chosen by any individual is efficient, and undistorted, given the amount of labor they intend to supply.[13]

3.4 The politico-economic equilibrium

In this section, I solve for the politico-economic equilibrium of the model, which consists in finding the Condorcet winner in the two levels of the tax and subsidy rates. This is done by backward induction, considering first the choice of τ in second period for a given level of s, and then the level of s in first period.

[13] In other words, the existence of the education subsidies prevents the feedback effects between education and labor supply I described earlier from kicking in once the tax rate on labor decreases the labor supply of individuals.

3.4.1 Second-period vote on redistribution

At $t = 2$, the indirect utility of agent γ_i under tax policy τ is:

$$\begin{aligned} v(i; s, \tau) &= \frac{\gamma_i^{1+\epsilon}}{1+\epsilon} e_i^{\beta(1+\epsilon)} (1-\tau)^{1+\epsilon} - (1-s) p_e e_i \\ &- s \int_{\underline{\gamma}}^{\overline{\gamma}} p_e e_n dF(\gamma_n) + \tau(1-\tau)^{\epsilon} \int_{\underline{\gamma}}^{\overline{\gamma}} \gamma_n^{1+\epsilon} e_n^{\beta(1+\epsilon)} dF(\gamma_n) \end{aligned} \tag{3.12}$$

The expressions of the indirect utility consists of four terms which I explain hereafter. The first term is the aggregation of the expressions for post-tax earnings minus the disutility from work. It is, of course, a decreasing function of the labor income tax rate τ, not only because τ directly reduces disposable income, but also because it reduces the incentive to work, thereby also reducing pre-tax labor income. The second term is simply the expression for the cost of education after subsidy, so that the first two terms together reflect utility derived from consumption and leisure. The last two terms correspond to the net transfers g received by the individual: the second to last is the amount of education subsidy expenditures the government needs to finance, while the last one is the actual gross transfer received by each individual. It can easily be seen that the amount of gross transfers received as a function of the tax rate τ follows a variant of the Laffer curve.

Notice that the education level has already been decided at the time of the policy choice over τ. Taking the first-order condition for a utility-maximising policy rate τ yields the following caracterisation of the preferred tax rate of individual i, denoted $\tau^*(i)$:

$$\frac{\tau^*(i)}{1-\tau^*(i)} = \frac{1}{\epsilon} \left[1 - \frac{\gamma_i^{(1+\epsilon)} e_i^{\beta(1+\epsilon)}}{\int_{\underline{\gamma}}^{\overline{\gamma}} \gamma_n^{(1+\epsilon)} e_n^{\beta(1+\epsilon)} dF(\gamma_n)} \right] \tag{3.13}$$

The previous formula is valid as long as $\gamma_i^{(1+\epsilon)} e_i^{\beta(1+\epsilon)} < \int_{\underline{\gamma}}^{\overline{\gamma}} \gamma_n^{(1+\epsilon)} e_n^{\beta(1+\epsilon)} dF(\gamma_n)$, otherwise the preferred tax rate of individual i is simply $\tau^*(i) = 0$. Intuitively, the more the individual is educated compared to the mean, or the abler he is compared to the mean, the less the demand for redistribution.

$\tau^*(i)$ is thus decreasing in γ_i, and preferences are single-peaked. Then the chosen tax rate under majority voting is (using the median voter theorem):[14]

$$\frac{\tau}{1-\tau} = \frac{1}{\epsilon}\left[1 - \frac{\gamma_{med}^{(1+\epsilon)} e_{med}^{\beta(1+\epsilon)}}{\int_{\underline{\gamma}}^{\overline{\gamma}} \gamma_n^{(1+\epsilon)} e_n^{\beta(1+\epsilon)} dF(\gamma_n)}\right] \tag{3.14}$$

This type of result for the income redistribution tax rate is rather unsurprising given the functional forms chosen. It appears that the tax rate voted in equilibrium depends on the ratio of the median to the mean of a function of individual types (when one considers $w_i = \gamma_i e_i^{\beta}$, the hourly productivity, as the new fixed type of an individual at the beginning of the second period), as in Meltzer and Richard (1981). When the distribution of types is very unequal, this ratio will be low, leading to higher redistribution. The equilibrium tax rate τ also naturally depends on the wage elasticity of labor supply: if individuals exhibit a very high labor supply elasticity, the tax base will shrink a lot as the tax rate rises, so that it is optimal to keep the tax rate low. As a consequence, a high ϵ means that the tax rate τ will be small in equilibrium, all else being equal.

I now proceed to show that the choice of τ in the second period is independent from the choice of s in the period before. To do so, I first establish the following:

Lemma 3 *Let there be two individuals with abilities γ_i and γ_j. Then the ratio e_i/e_j is independent from both choices of policy rates.*

The proof is straightforward: using equation (3.7), one gets immediately that $(e_i/e_j)^{\mu} = (\gamma_i/\gamma_j)^{1+\epsilon}$.[15]

Then using the last equality, equation (3.14) can be rewritten in the following form:

$$\frac{\tau}{1-\tau} = \frac{1}{\epsilon}\left[1 - \frac{\gamma_{med}^{\frac{1+\epsilon}{\mu}}}{\int_{\underline{\gamma}}^{\overline{\gamma}} \gamma_n^{\frac{1+\epsilon}{\mu}} dF(\gamma_n)}\right] \tag{3.15}$$

Equation (3.15) then clearly shows that the equilibrium income tax rate is independent from the first-period education subsidy rate s. To get an intuition of this result, one needs to bear in mind that education has a multiplicative effect on an individual's hourly

[14]Notice that the monotonicity of τ^* is only required to establish the fact that the median voter is actually the individual with the median intrinsic ability γ_{med}.

[15]This result, and consequently equation (3.15) as well, obviously depend a lot on the functional forms that were chosen, more specifically on the facts that an individual's productivity is a power function of their education, and that the disutility from labor is also a power function of the labor supply.

productivity $w_i = \gamma_i e_i^{\beta}$. Additionally, equation (3.7) shows that a higher education subsidy will lead all individuals to increase their education by the same amount relative to their initial intrinsic productivity γ_i, so that a higher subsidy rate s does not change the shape of the distribution of productivities w_i.[16] Then the median-to-mean ratio present in the expression of the equilibrium income tax rate is left unaffected by changes in the first-period value of s.

Besides, it can be noted from (3.15) that $0 \leq \tau < 1$ with strict inequality as long as the distribution of $\eta_i = \gamma_i^{\frac{1+\epsilon}{\mu}}$ is skewed to the right. This condition is fulfilled, for instance, for any uniform distribution of γ_i on a finite support, or for any distribution of γ_i already skewed to the right. If this is not the case, then the median voter will prefer no redistribution at all and $\tau = 0$. In everything that follows, we restrain ousrselves to the study of distributions of γ such that $\tau > 0$.

Assumption 6 *The median of the distribution of intrinsic abilities $\{\gamma_i\}_i$, called γ_{med}, is such that $\gamma_{med}^{\frac{1+\epsilon}{\mu}} < \mathbf{E}_i(\gamma_i^{\frac{1+\epsilon}{\mu}})$.*

3.4.2 First-period vote on subsidy s

Now that the value of τ in the second period has been found, I turn to determining the equilibrium level of s in the first period. The task will be considerably easier now that we know that s has no impact on the future value of τ: indeed, the future value of τ can now be taken as given when computing the preferred level of s for any given individual. To determine this level, one needs to replace education levels e by their expressions in the expression of an individual's indirect utility function (equation (3.12)):

$$\begin{aligned} v(i;s,\tau) \;&= [\gamma_i(1-\tau)]^{\frac{1+\epsilon}{\mu}} \left[\frac{\beta}{p_e(1-s)}\right]^{\frac{\beta(1+\epsilon)}{\mu}} \frac{\mu}{\beta(1+\epsilon)} \\ &- s p_e \left[\frac{\beta}{p_e(1-s)}\right]^{1/\mu} \int_{\underline{\gamma}}^{\overline{\gamma}} [\gamma_n(1-\tau)]^{\frac{1+\epsilon}{\mu}} dF(\gamma_n) \\ &+ \tau(1-\tau)^{\frac{1+\epsilon}{\mu}-1} \left[\frac{\beta}{p_e(1-s)}\right]^{\frac{\beta(1+\epsilon)}{\mu}} \int_{\underline{\gamma}}^{\overline{\gamma}} \gamma_n^{\frac{1+\epsilon}{\mu}} dF(\gamma_n) \end{aligned} \tag{3.16}$$

Then writing the first-order condition with respect to s leads to the following rela-

[16] To be more accurate, the distributions of final productivities $\{w_i\}_i$ obtained for different values of s are homothetic to one another.

tionship between an individual's peferred subsidy rate and the anticipated future tax rate:

$$\frac{\partial v(i;s,\tau)}{\partial s}=0 \Leftrightarrow \frac{s^*(i)}{1-s^*(i)}=(1+\epsilon)\frac{\tau}{1-\tau}+\mu\left[\frac{\gamma_i^{\frac{1+\epsilon}{\mu}}}{\int_{\underline{\gamma}}^{\overline{\gamma}}\gamma_n^{\frac{1+\epsilon}{\mu}}\,dF(\gamma_n)}-1\right] \tag{3.17}$$

Then by replacing $\tau/(1-\tau)$ by its expression from (3.15) in (3.17), and taking once again as the outcome of the vote the policy favored by the individual with median ability, one gets:

$$\frac{s}{1-s}=\frac{1+\beta\epsilon(1+\epsilon)}{\epsilon}\left[1-\frac{\gamma_{med}^{\frac{1+\epsilon}{\mu}}}{\int_{\underline{\gamma}}^{\overline{\gamma}}\gamma_n^{\frac{1+\epsilon}{\mu}}\,dF(\gamma_n)}\right] \tag{3.18}$$

Under assumption 6 , it is clear that$s>0$ in equilibrium. Additionally, by comparing the expressions obtained in equations (3.15) and (3.18), it can be seen that $s>\tau$: the subsidy rate to acquire education is larger than the income redistribution tax rate. This latter result, while of limited interest in itself,[17] will allow to prove that individual decisions are distorted towards excessive human capital accumulation in this politico-economic equilibrium, as opposed to what happens under the constrained optimum.

To see this, let us compute the total net tax wedge on learning Δ created by the two policy instruments $(s;\tau)$ in this situation:

$$\begin{aligned}\Delta &= \tau\frac{\partial}{\partial e_i}(\gamma_i e_i^{\beta} l_i)-s\frac{\partial}{\partial e_i}(p_e e_i)\\ &= (1-s)p_e\left(\frac{\tau}{1-\tau}-\frac{s}{1-s}\right)\end{aligned} \tag{3.19}$$

where equation (3.19) was obtained in the same way as equation (3.11). Then $s>\tau\Rightarrow \tau/(1-\tau)-s/(1-s)<0$, which leads to the following proposition:

Proposition 4 *The net tax wedge on learning under the political equilibrium is strictly negative:* $\Delta<0$.

What this result essentially means is that the combination of both policy instruments distorts the education choices of individuals towards overaccumulation of education in

[17] I will use again this result in subsection 3.4.3 to compare education and labor supply decisions under the two time structures of simultaneous and sequential voting.

the first period, considering the high levels of the labor tax to be expected in the second period and the low levels of labor supply that follow from it. In this politico-economic equilibrium, the education subsidy more than compensates individuals for the tax imposed on the returns from learning.

Comparing with the result obtained by Bovenberg and Jacobs (2005), which I mentioned earlier in subsection 3.3.2 , it thus becomes clear that the political game cannot decentralize any optimum a central planner with redistributive concerns and access to the same linear instruments would be able to achieve, whatever the level of the distributional characteristic ξ the planner would seek to implement. In the following subsection, I proceed to show that this result is due to the difference in the timing of the decisions on the two policy rates between Bovenberg and Jacobs (2005) and the situation presented before.

3.4.3 How relevant is the timing of decisions?

This subsection is devoted to the comparison of the results obtained above with those obtained in a similar framework, that only differs from the previous one by the time at which the labor income tax rate t is voted upon. Namely, in appendix 3.A , I describe an economy in which agents get to vote simultaneously on the two policy rates s and τ before any individual decision is taken.

From (3.26), that describes the preferred policy set of an individual with a below-the-mean value of $\eta_i = \gamma_i^{\frac{1+\epsilon}{\mu}}$, it follows that under assumption 6 , the two equilibrium policy rates under simultaneous voting are:

$$\frac{s}{1-s} = \frac{\tau}{1-\tau} = \frac{\mu}{\epsilon}\left(1 - \frac{\gamma_{med}^{\frac{1+\epsilon}{\mu}}}{\int_{\underline{\gamma}}^{\overline{\gamma}} \gamma_n^{\frac{1+\epsilon}{\mu}} dF(\gamma_n)}\right) \tag{3.20}$$

Comparing the outcomes of this setup with that of Bovenberg and Jacobs (2005), and especially equations (3.9)-(3.10) with (3.20), the fundamental importance of timing becomes obvious. First of all, when the two instruments can be chosen simultaneously, the political equilibrium picks an equal value for these instruments, which is shown by Bovenberg and Jacobs (2005) to be optimal. Secondly, the actual level of the policy rates is closely related to the price-elasticity of education acquisition μ and the elasticity of labor supply ϵ in both cases. And thirdly, the political equilibrium in the simultaneous voting game (s_{sim}, τ_{sim}) replicates the optimal levels of tax rates for a carefully chosen value of ξ, related to a median-to-the mean ratio of a function of types that is quite typical of models à la Meltzer and Richard (1981):

$$\xi = 1 - \frac{\gamma_{med}^{\frac{1+\epsilon}{\mu}}}{\int_{\underline{\gamma}}^{\overline{\gamma}} \gamma_n^{\frac{1+\epsilon}{\mu}} \, dF(\gamma_n)} \Rightarrow \tau_{sim} = s_{sim} = \tau^* = s^* \tag{3.21}$$

This last result makes perfect sense, bearing in mind that ξ represents the willingness to redistribute income in the planner's problem: indeed, in the political equilibrium, the lower the median type is relative to the mean type, the higher his willigness to exert redistribution will be as well. In that sense, the political equilibrium under simultaneous voting decentralizes the planner's optimum, for a selected value of the desire to redistribute.

I now turn to the comparison between the political equilibrium under simultaneous voting to its counterpart under sequential voting (s_{seq}, τ_{seq}), as computed in equations (3.14) and (3.18). Comparing the expressions in these two equations with equation (3.20), and given that $\mu < 1$, it obtains that:

$$\tau_{sim} = s_{sim} < \tau_{seq} < s_{seq} \tag{3.22}$$

The discrepancy between the two equilibria naturally comes from the difference in timing: in the sequential game, individuals have already chosen their level of education when deciding the level of the labor income tax rate, which makes the tax base less sensitive to changes in τ than in the simultaneous game where individuals can adjust downward both their labor supply and their education following an increase in τ. Therefore, the median individual is more prone to setting a high tax rate in the sequential voting game than in the simultaneous one, all else being equal. Additionally, the subsidy rate is higher in the sequential version of the political game since it is not tied anymore to the level of the tax rate. Whereas the two policy rates are equal in the simultaneous voting game, which requires the subsidy rate to stay low in order to prevent too high a degree of income redistribution, the fact that the two rates are disconnected in the sequential game (as shown in (3.15)) allows for a higher degree of education subsidy there.

It has thus been shown that the choice of timing is determinant for the results obtained in equilibrium. The reasons for adopting a sequential timing instead of a simultaneous one, in the context of a model involving only one generation, have already been detailed in subsection 3.2.2 . Appendix 3.B extends the previous analysis by explicitly considering two coexisting generations, with education transfers flowing from an adult towards a child generation. In this appendix, I proceed to show that this setup delivers the same predictions as the sequential time structure above, even in the case where individuals decide simultaneously on the level of the two policy rates, for given values of the model parameters.

In the following section, I will present the latest available data on education subsidy rates and labor income tax rates in various OECD countries, and argue that the proposition that $s > \tau$ in the model presented above is indeed a stylized fact.

3.5 A quantitative assessment of the link between education subsidy and labor income tax rates

A main prediction of the model presented in the previous sections is that the rate of subsidies to education voted at the political equilibrium, s, should be higher than the labor income tax rate τ, in a case where both subsidy and tax policies are constrained to be linear. In order to explore the validity of this proposition, I use data on marginal tax rates taken from OECD (2014b), which I match with data on the share of public expenditures in higher education from OECD (2014a). The rationale for focusing on higher education comes from the fact that lower levels of education are subject to compulsory schooling laws, so that the level of subsidy by the government can be safely assumed to be irrelevant on the decision whether or not to attend these lower levels of school, at least for OECD countries. Concerning τ, I choose to employ the total tax wedge[18] faced by an individual earning 133% of his country's average wage, in order to get rid of the effects of tax breaks put in place for low wage earners in many countries.[19] The latest available figures for both types of data apply to 2011. All the countries included in table 3.1 are members of the OECD; some member countries were not included for data availablility reasons.

[18] The total tax wedge is defined as the sum of the (local- and central-government) income tax, and employer- and employee-based social security contributions, divided by the total labor costs. It is higher than the pure income tax rate, defined as the income tax divided by the gross wage, and even the all-in tax rate, defined as the sum of employee social security contribution and income tax divided by the gross wage. In that respect, using these alternative measures for τ would have only strengthened the result obtained in table 3.1 .

[19] A similar exercise has been performed with individuals earning the average wage: again, because many income tax systems are progressive, this only reinforces the result that $s > \tau$ for most countries

Country	Tax rate τ (%)	Subsidy rate s (%)
Chile	7.0	24.2
Korea	22.9	27.0
United Kingdom	49.0	30.2
Japan	41.7	34.5
United States	41.8	34.8
Australia	42.0	45.6
Israel	38.6	49.0
Canada	34.3	57.4
New Zealand	30.0	64.5
Italy	61.6	66.5
Mexico	25.2	67.1
Portugal	50.7	68.6
Netherlands	49.6	70.8
Poland	36.1	75.5
Slovak Republic	44.4	76.9
Spain	48.1	77.5
Estonia	42.9	80.4
Ireland	56.7	80.5
France	59.7	80.9
Czech Republic	48.6	81.1
Germany	57.4	84.7
Slovenia	51.0	85.2
Austria	60.6	86.9
Sweden	63.1	89.5
Belgium	68.3	90.0
Iceland	43.5	90.6
Denmark	56.1	94.5
Finland	57.2	95.9
Norway	51.2	95.9
Sample mean	46.2	69.2

Tableau 3.1: Labor income tax rates and higher education subsidies (in %) for several OECD countries, in 2011

Judging from table 3.1 , it is undeniable that$s > \tau$ can be established as a stylized fact: indeed, out of the 29 countries in the sample, only three (the USA, the UK, and Japan) have a labor income tax rate that is higher than their rate of subsidy to education. Additionally, the rates of education subsidies and labor income taxation are positively correlated, with a coefficient of correlation between the two series of 0.64: this fact is also consistent with the theory outlined above. It should be noted that the theory does not necessarily predict a perfect correlation between the two variables, since there is no reason *a priori* that two different countries should be characterized by the same wage elasticity of labor supply or wage returns to education, let alone the same distribution of innate abilities (whatever the empirical counterpart of this theoretical variable may be).

in the sample.

3.6 Conclusion and discussion

3.6.1 Summary of the findings

The main message of this paper is that the political process cannot decentralize the optimal income taxation and education subsidy choices of a planner with redistributive concerns, in the case where the two available instruments are linear rates. Such a result hinges on the assumption that the level of the labor income tax rate cannot be committed upon before individuals select their level of education. Once the education decision is sunk, the tax elasticity of income becomes lower than what it would have been if the labor income tax rate had been chosen before the education decision. Additionally, under the specification of the earnings function that was adopted, it is shown that the rate of education subsidy has no impact on the labor income tax rate chosen later, so that it is set at too high a level in the political equilibrium. The prediction that the education subsidy rate is higher than the labor income tax rate at equilibrium, thus generating too much human capital compared to what would be optimal, is indeed verified in the latest available OECD data.

In appendix 3.B , I show that this simple, two-period model with sequential voting yields the same political equilibrium that an overlapping-generation model with intergenerational altruism, in which every generation votes simultaneously on a rate of income taxation that applies to itself and a rate of education subsidy that helps its children acquire an education. This OLG setting, while acknowledging the fact that policy decisions are taken simultaneously on different dimensions, still makes the point that no commitment can be made on the income tax schedule facing a generation at the time when this generation has not yet entered the labor market and is still educating.

3.6.2 Discussion and further research

Several characteristics of the model can be considered to restrict the validity of the message of this chapter, some of which have already been pointed out in the previous literature. Notably, Jacobs and Bovenberg (2011) remark that the efficiency results that serve as a benchmark for my political equilibrium are dependent on the specification of the earnings function employed. In the case of this model, the functional forms of not only the earnings function, but also of the labor disutility of individuals are needed if one is to obtain the crucial technical result that the relative education levels of two individuals are independent from the rate of education subsidy. This result is an essential step towards obtaining all further analytical results, and especially the lack of direct connection between the rate of education subsidy and the income tax rate.

I believe further research should focus on whether political equilibria are able to decentralize efficient allocations of public funds, in other settings involving the two public policies that are studied in the present chapter. This would have the virtue of bringing closer together the normative and political economy analyses on the subject. Additionally, extensions of the present work could include considering the existence of either externalities in education acquisition, or credit constraints to acquire education in the first period of the life of individuals. This last extension was attempted but had not been carried to the end at the time of writing, however. It involves defining a second type of heterogeneity for individuals, this time not only in the ability to benefit from education, but in the ability to pay for it as well.[20] The principal challenges of this approach lie in finding tractable formulations of the problem, as well as identifying the median voter in a setting with a bidimensional policy space, and bidimensional heterogeneity of individuals as well.

[20] See the review of de Fraja (2002) in subsection 3.1.2 on this topic.

Appendices

3.A Appendix: Simultaneous vote on s and τ

As a benchmark for the sequential-vote politico-economic equilibrium of the two-period model set up in this chapter, I provide hereafter the results of a different model, in which agents vote simultaneously on the education subsidy rate s and the income redistribution rate τ, before agents make any private decision.

Timing

1. ($t = 1$) Parties A and B announce a binding joint policy proposal for both the subsidy rate s and the income tax rate τ. Agents elect the government for both periods based on the policy proposals.
2. Private agents decide on their educational inputs purchase e_i (at a unitary cost of $p_e(1-s)$ in terms of the consumption good)
3. ($t = 2$) Gross hourly wage w_i of individual i is given by education e_i and (heterogeneous) ability γ_i.
4. Agents choose their labor supply l_i. Their gross wage income $z_i = w_i l_i$ is taxed to finance redistribution, recoup subsidies and finance homogenous lump-sum transfers to all individuals; agents pay their education costs and consume c_i.

In this model, the budget constraint, the utility over consumption and leisure and the indirect utility over both policy rates (given by equation (3.12)) are unchanged.

Now, the joint preferred policy over s and τ will be determined for each agent by maximising indirect utility in equation (3.12), with respect to both instruments at the same time, after taking into account the dependency of e_i with respect to both instruments.

Replacing e_i by its expression obtained in equation (3.7) in (3.12), I get:

$$\begin{aligned} v(i;s,\tau) &= [\gamma_i(1-\tau)]^{\frac{1+\epsilon}{\mu}} \left[\frac{\beta}{p_e(1-s)}\right]^{\frac{\beta(1+\epsilon)}{\mu}} \frac{\mu}{\beta(1+\epsilon)} \\ &-sp_e\left[\frac{\beta}{p_e(1-s)}\right]^{1/\mu} \int_{\underline{\gamma}}^{\overline{\gamma}} [\gamma_n(1-\tau)]^{\frac{1+\epsilon}{\mu}} dF(\gamma_n) \\ &+t(1-\tau)^{\frac{1+\epsilon}{\mu}-1}\left[\frac{\beta}{p_e(1-s)}\right]^{\frac{\beta(1+\epsilon)}{\mu}} \int_{\underline{\gamma}}^{\overline{\gamma}} \gamma_n^{\frac{1+\epsilon}{\mu}} dF(\gamma_n) \end{aligned} \tag{3.23}$$

Then taking the partial derivative with respect to τ, one obtains:

$$\frac{\partial v(i;s,\tau)}{\partial t} = 0 \Leftrightarrow \frac{\tau}{1-\tau}\left[\frac{\epsilon+\beta(1+\epsilon)}{\mu}\right] = 1 - \frac{\gamma_i^{\frac{1+\epsilon}{\mu}}}{\int_{\underline{\gamma}}^{\overline{\gamma}} \gamma_n^{\frac{1+\epsilon}{\mu}} dF(\gamma_n)} + \frac{s}{1-s}\frac{\beta(1+\epsilon)}{\mu} \tag{3.24}$$

Similarly, taking the partial derivative with respect to s leads to the following:

$$\frac{\partial v(i;s,\tau)}{\partial s} = 0 \Leftrightarrow \frac{s}{1-s}(-\frac{1}{\mu}) = 1 - \frac{\gamma_i^{\frac{1+\epsilon}{\mu}}}{\int_{\underline{\gamma}}^{\overline{\gamma}} \gamma_n^{\frac{1+\epsilon}{\mu}} dF(\gamma_n)} - \frac{\tau}{1-\tau}\frac{(1+\epsilon)}{\mu} \tag{3.25}$$

It should not come as a surprise to encounter the terms written above, which all correspond to given elasticities in this problem. Namely, we have: $\epsilon_{et} = \frac{-(1+\epsilon)}{\mu}$, $\epsilon_{es} = \frac{1}{\mu}$, $\epsilon_{zt} = \frac{-(\epsilon+\beta(1+\epsilon))}{\mu}$ and $\epsilon_{zs} = \frac{\beta(1+\epsilon)}{\mu}$, which denote respectively the elasticities of education acquisition e and total gross wage z with respect to policy instruments s and τ.

Combining both first-order conditions then yields the following characterisation of an individual i's preferred policy set $(s^*(i);\tau^*(i))$:[21]

$$\frac{s^*(i)}{1-s^*(i)} = \frac{\tau^*(i)}{1-\tau^*(i)} = \frac{\mu}{\epsilon}\left(1 - \frac{\gamma_i^{\frac{1+\epsilon}{\mu}}}{\int_{\underline{\gamma}}^{\overline{\gamma}} \gamma_n^{\frac{1+\epsilon}{\mu}} dF(\gamma_n)}\right) \tag{3.26}$$

as long as individual i is such that $\gamma_i^{\frac{1+\epsilon}{\mu}} > \int_{\underline{\gamma}}^{\overline{\gamma}} \gamma_n^{\frac{1+\epsilon}{\mu}} dF(\gamma_n)$.

[21] The proof of the concavity of v, which involves proving that the Hessian of v is semi-definite negative, is especially tedious and computation-intensive and has therefore not been reported here.

For other individuals, the preferred policy set is instead:

$$\tau^*(i) = 0 \tag{3.27}$$
$$\frac{s^*(i)}{1-s^*(i)} = \mu\left(\frac{\gamma_i^{\frac{1+\epsilon}{\mu}}}{\int_{\underline{\gamma}}^{\overline{\gamma}} \gamma_n^{\frac{1+\epsilon}{\mu}}\, dF(\gamma_n)} - 1\right) \tag{3.28}$$

Inorder to be able to derive the political equilibrium of the voting game, we now proceed to show that the indirect utility function exhibits intermediate preferences à la Grandmont (1978). To do this, it is sufficient to notice that the indirect utility as expressed in equation (3.12) writes in the following fashion:

$$v(i; s, \tau) = J(s, \tau) + K(\gamma_i)H(s, \tau)$$

where:

$$\begin{aligned} J(s,\tau) &= t(1-\tau)^{\frac{1+\epsilon}{\mu}-1}\left[\frac{\beta}{p_e(1-s)}\right]^{\frac{\beta(1+\epsilon)}{\mu}} \int_{\underline{\gamma}}^{\overline{\gamma}} \gamma_n^{\frac{1+\epsilon}{\mu}}\, dF(\gamma_n) \\ &-sp_e\left[\frac{\beta}{p_e(1-s)}\right]^{1/\mu} \int_{\underline{\gamma}}^{\overline{\gamma}} [\gamma_n(1-\tau)]^{\frac{1+\epsilon}{\mu}}\, dF(\gamma_n) \end{aligned}$$

$$K(\gamma_i) = \gamma_i^{\frac{1+\epsilon}{\mu}}$$

$$H(s,\tau) = \frac{\mu}{1+\epsilon}(1-\tau)^{\frac{1+\epsilon}{\mu}}\left[\frac{\beta}{p_e(1-s)}\right]^{\frac{\beta(1+\epsilon)}{\mu}}$$

It can easily be proven that $K' > 0$ and $K'' > 0$ on its whole domain of definition. Then the usual theorem for intermediate preferences applies, and the favorite policy of the median voter (who happens to be the voter with median ability γ_{med}) is the Condorcet winner of the joint voting game on both policy instruments s and τ.

3.B Appendix: an overlapping-generation model with altruism and forward intergenerational public transfers to finance education

In the present appendix, I setup a model with two overlapping generations, in which the oldest generation alive cares for the utility of its children and votes to establish subsidies to the acquisition of education of its children, while deciding at the same time on the level of redistribution to implement within its own generation through a labor income tax. I show that under perfect dynastic altruism, and barring any population growth, the politico-economic equilibirum of this model can be characterized by the same equations as the sequential voting game set up in sections 3.2 through 3.4 .

3.B.1 Time structure and agents

Let us consider the following OLG model with two overlapping generations alive at each date t: we refer to the individuals of the generation born in t as children, while individuals from the other generation (born at $(t-1)$) are referred to as parents (or adults). Let N_t the number of parents alive at date t, and let $n_{t+1} = (N_{t+1} - N_t)/N_t$ the growth rate of adult population between t and $(t+1)$.[22]

A vote concerning a linear tax rate on labor income and a linear subsidy rate to education purchase takes place in every period, before private decisions are taken, with the adult generation being the only one allowed to vote for both taxes. I assume here that the children's generation has not reached electoral majority yet. The two political decisions are this time assumed to be taken simultaneously, recognizing the fact that taxes and expenditures are decided as parts of the same legislative process of establishing the budget. I consider that there are no commitment problems regarding the two policy rates that are voted upon in equilibrium,[23] and that they are not revoted upon once agents have made their private decisions.

3.B.2 Altruism between generations

For the sake of simplicity, it is assumed that a parent from generation t with intrinsic ability γ_i will always produce an offspring with the same ability γ_i, thus barring any consideration of intergenerational mobility.[24] This parent will be interested in the well-being of his child, in the form of a discount δ to the utility of his child, added to the

[22]Note that generations will henceforth be denoted according to the date at which they reach adulthood, and not their birth cohort: hence the subscript t will always refer to individuals born at date $(t-1)$, adult at t.

[23]Indeed, contrary to the model presented in section 3.2 , the collective decision on the two policy rates impacts private decisions right away, so that it is impossible to change the relevant policy rate once agents have made any decision (as opposed to the possibility of changing the labor tax rate once agents have picked their level of education, as explained in subsection 3.2.2).

[24]It follows that the distribution of abilities is stable over time.

parent's utility derived from consumption.[25] I then consider that an individual's utility is:

$$\begin{aligned} V_{i,t} &= c_{i,t} - \frac{l_{i,t}^{1+1/\epsilon}}{1+1/\epsilon} + \delta V_{i,t+1} \\ &= \sum_{k=0}^{+\infty} \delta^k \left(c_{i,t+k} - \frac{l_{i,t+k}^{1+1/\epsilon}}{1+1/\epsilon} \right) \end{aligned} \tag{3.29}$$

with $\delta \in (0,1)$ the measure of downward altruism towards the next generations.

In this OLG model, I make the following assumption concerning the flow of public education subsidies:

Assumption 7 *Parents of generation t vote on education subsidies s_t that apply to their children, and not to themselves.*

This assumption comes from the fact that investment in a child's education has to be made before the child can reach political majority, and that adult parents are too old to benefit from education subsidies once they are eligible to vote. Because the adult's utility includes utility derived from consumption of their own child, and since investment in their child's education entails no personal cost for themselves, the adult will select privately the level of educational acquisition $e_{i,t+1}$ that is optimal for their child, for a set of given tax and subsidy rates. From a political point of view, though, it is unclear so far whether the parent generation wants to optimally subsidize their children's education acquisition through the subsidy rate s_t, as parents might be tempted to increase the lump-sum transfer to themselves g_t at the detriment of the educational subsidy.[26]

3.B.3 The public sector

In this small OLG model, I force the budget of the government to be balanced in any given period, so that the government cannot resort to debt issuance to cover unfunded expenditures.

[25] I follow here the recursive formulation of altruism first introduced in Barro (1974), sometimes called dynastic altruism.

[26] In terms of notation, policy variables are indexed by the time at which the relevant decision is made, in other words by the generation responsible for the underlying political choice: hence τ_t and g_t refer to taxes and transfers from and to the generation adult at t, while s_t is voted by generation t but will help generation e_{t+1} acquire an education.

Let us write the budget constraint of the government in any given period t:

$$\tau_t N_t \int_{\underline{\gamma}}^{\overline{\gamma}} \gamma_{n,t} e_{n,t}^{\beta} l_{n,t} dF(\gamma_n) = g_t N_t + s_t N_{t+1} \int_{\underline{\gamma}}^{\overline{\gamma}} p_{e,d} e_{n,t+1} dF(\gamma_n) \tag{3.30}$$

In per adult terms, one gets:

$$\tau_t \int_{\underline{\gamma}}^{\overline{\gamma}} \gamma_{n,t} e_{n,t}^{\beta} l_{n,t} dF(\gamma_n) = g_t + s_t(1+n_{t+1}) \int_{\underline{\gamma}}^{\overline{\gamma}} p_{e,d} e_{n,t+1} dF(\gamma_n) \tag{3.31}$$

Under equation (3.31), no debt is ever created, which cuts the dynamic link between the different periods coming from the one-period budget constraints: each policy set $\{s_t, \tau_t, g_t\}$ voted in equilibrium needs to equalize resources to transfers. Therefore, the only difference with the two-period model developed before lies in the presence of the population growth rate n_{t+1}.

3.B.4 The voting game

Let us now turn to the policy choice $\{s_t, \tau_t, g_t\}$ made in each period t, by first recognising that it has an impact on the parent generation t's utility only through the utility coming from their own consumption and labor supply decisions, as well as those of their children. Indeed, now that we have ruled out debt creation, the adult generation of t has no means of influencing the welfare of generations $(t+k)$ for $k \geq 2$.[27] An individual from generation t with intrinsic ability γ_i then faces a private budget constraint akin to that of the two-period model I solved before (equation (3.3)), and evaluates his own utility in the same way as before, so that private behavior under given policy rates is left entirely unchanged between the simpler model and the present OLG one.[28]

Therefore, the only possible discrepancies between the two models shall come from the political decisions: let us then examine the decision over each policy variable in turn.

[27] Again, this comes from the quasi-linearity of the utility function: even if generation t acts so as to lower generation $(t+1)$ consumption, this will not change generation $(t+1)$ decisions concerning generation $(t+2)$, since the marginal utility of consumption of generation $(t+1)$ is constant and equal to one.

[28] While it seems obvious for the labor supply decision, the education acquisition one deserves an additional explanation. The decision taken in period t by the adult concerns the education acquisition of the individual born in t, yet the cost of education is borne by the child and gives utility to the parent through the child's utility, so that the adult is a perfect agent of the child and private behavior for given policy variables is unmodified.

Let us start by recognising that the trade-off over the labor tax rate τ_t is left unchanged here. Taxing more labor income will amount to reducing the work incentives, which are otherwise the same as before, while it will provide the same amount of fiscal resources. For a given value of s_t, these fiscal resources will be used to finance more transfers, the marginal value of which is constant and the same as before. Then it is clear that each individual's preferred policy over τ_t remains the same as in the two-period model, for each given value of s_t.

At this point, the only relevant tradeoff to compare the two models concerns the choice of the education subsidy rate s_t: for a given level of τ_t, an increase in s_t will translate into higher educational investment by the children generation, thus leading to higher utility to the child (which is itself valued at a discounted rate δ in the parent's utility). This comes at the expense of net transfers g_t to the parent generation. Formally, if we write $v_{i,t}$ the indirect utility of individual i of generation t over the policy set, the following holds:

$$\frac{\partial v_{i,t}}{\partial s_t} = \frac{\partial g_t}{\partial s_t} + \delta \frac{\partial}{\partial s_t}\left[c_{i,t+1} - \frac{(l_{i,t+1})^{1+\epsilon}}{1+\epsilon}\right] \tag{3.32}$$

Notice that the utility coming from the individual's own labor income and leisure is unaffected by changes in s_t in that case, since $e_{i,t}$ is already fixed at the time of the vote and thus s_t has no impact on the individual's own labor productivity. Therefore, the level of subsidies will only have an impact on the voting generation's utilities through the transfers each individual receives g_t, and through the increase in educational attainment, and hence on utility, that it brings to the voting generation's children.

Using (3.31), the transfers g_t can be rewritten in the following form, in the absence of any governmental debt:

$$g_t = \tau_t \int \gamma_{n,t} e_{n,t}^{\beta} l_{n,t} dF(\gamma_n) - s_t(1+n_{t+1}) \int p_e e_{n,t+1} dF(\gamma_n) \tag{3.33}$$

There are quite a few differences between (3.33) and (3.8), the two governmental budget constraint equations, respectively in the OLG and 2-period cases. Indeed, in both cases the tax base depends on the average productivity of the currently voting generation, which itself depends on the level of education acquired by all members of this generation. Yet, in (3.33) this level of education $e_{n,t}$ is fixed for all individuals n at the time of the vote, while in (3.8) it will increase following an increase in the subsidy rate s. The second term of the right-hand side of both equations, representing the total cost of the subsidies, is of course increased when individuals decide to acquire more education, be it the currently voting population (in the 2-period case) or its children (in the OLG case). Therefore, when a higher subsidy rate s is voted upon, the total outlays for education subsidies increase, not only because the subsidy is higher for a given level of education, but also because individuals wish to educate more. This mechanism is true in both cases, since the subsidy rate applies in both cases to the generation currently acquiring education. However, a slight difference remains in equation (3.33) with respect to the OLG case, with the presence of the growth rate in the numbers of individuals in a generation n_{t+1}. This is to be expected: if there is population growth, educational subsidies will be diluted as the recipient generation is larger than the contributing one.

Let us now examine the second term on the right-hand side of equation (3.32):

$$\begin{aligned}\frac{\partial}{\partial s_t}\left[c_{i,t+1} - \frac{(l_{i,t+1})^{1+\epsilon}}{1+\epsilon}\right] &= \frac{\partial}{\partial s_t}\left[(1-\tau_{t+1})w_{i,t+1}l_{i,t+1} - \frac{(l_{i,t+1})^{1+\epsilon}}{1+\epsilon}\right] \\ &- \frac{\partial}{\partial s_t}\left[(1-s_t)p_e e_{i,t+1}\right] + \frac{\partial g_{t+1}}{\partial s_t} \qquad (3.34)\end{aligned}$$

When s_t increases, the individual's child will increase his investment in education $e_{i,t+1}$, thereby increasing his labor productivity $w_{i,t+1} = \gamma_{i,t+1}e^{\beta}_{i,t+1}$ once adult. He will therefore supply more labor and earn more labor income, in the same exact manner as the individual himself would have done in the 2-period model.[29] Therefore all the terms but the last one are proved to be unchanged at this point.

Concerning g_{t+1}, it can be seen that the tax base next period is itself changed, just as the tax base available in the second period would be in the 2-period model. But in the OLG case, the amount of government expenditures in the next period is not affected by the current level of the tax subsidy. This can be seen by writing (3.33) in period

[29] It has been shown that the labor income tax rate τ does not vary between the two models. Equation (3.15) even shows that it is independent from the distribution of total (i.e. education-augmented) labor productivity, so that the disincentives to supply labor and acquire education coming from expected taxation of income are exactly the same under both models. Therefore, for any given value of the subsidy rate s, the level of education attainment, and then of labor income, will be the same under both models as well.

$(t+1)$ and taking the derivative with respect to s_t:

$$\begin{aligned}\frac{\partial g_{t+1}}{\partial s_t} &= \frac{\partial}{\partial s_t}\left[\tau_{t+1}\int \gamma_{n,t+1}e^{\beta}_{n,t+1}l_{n,t+1}dF(\gamma_n) - s_{t+1}(1+n_{t+2})\int p_e e_{n,t+2}dF(\gamma_n)\right] \\ &= \frac{\partial}{\partial s_t}\left[\tau_{t+1}\int \gamma_{n,t+1}e^{\beta}_{n,t+1}l_{n,t+1}dF(\gamma_n)\right] \qquad (3.35)\end{aligned}$$

The first term on the right-hand side of (3.35), representing the amount of taxes collected in $(t+1)$, clearly depends on s_t, since the subsidies increase the labor productivity of the workforce in the next period (hence the tax base as well). On the other hand, the amount of expenditures to subsidize education next period is independent from the current level of subsidy. To see why, it is helpful to remember that the utility function exhibits no income effect in our setup. Therefore, the decisions made in period $(t+1)$ by individuals belonging to the generation born in period t will not depend on the income they have, but only on relative prices. When individuals born in t receive a higher subsidy to educate, they become more productive and thus supply more labor, thereby receiving more income: yet, this bears no incidence on their decision to invest on their children's own education, since this decision is all about the trade-off between the transfer they receive (which has a constant marginal utility) and the utility they get from their children's own utility. It should be noted that all along, the quasi-linearity of the utility function is the one reason why the decisions about redistribution and educational investment in a given period can be disentangled, and also why there is essentially no dynamic link between the policies voted upon in successive periods.

I then turn to the most important result of this appendix, stating under which conditions the OLG and the 2-period equilibria are characterized by the same equations:

Proposition 5 *The OLG model is characterized by the same set of equations as the two-period model when $\delta \to 1$ and $n_{t+1} \to 0$, $\forall t$. In this case, the policies voted in equilibrium in every period of the OLG model, as well as the private decisions taken by all agents, are the same as those of the two-period model.*

Proof Let us start by rewriting how the transfers g_t vary following a change in the subsidy rate s_t when $n_{t+1} = 0$, using (3.33):

$$\frac{\partial g_t}{\partial s_t} = -\frac{\partial}{\partial s_t}\left[s_t \int p_e e_{n,t+1}dF(\gamma_n)\right] \qquad (3.36)$$

Then equation (3.32) can be rewritten in the following way, in the case where $\delta = 1$, using (3.34) and (3.35):

$$\begin{aligned}\frac{\partial v_{i,t}}{\partial s_t} &= \frac{\partial}{\partial s_t}\left[(1-\tau_{t+1})w_{i,t+1}l_{i,t+1} - \frac{(l_{i,t+1})^{1+\epsilon}}{1+\epsilon} - (1-s_t)p_e e_{i,t+1}\right] \\ &+ \frac{\partial}{\partial s_t}\left[\tau_{t+1}\int \gamma_{n,t+1}e^{\beta}_{n,t+1}l_{n,t+1}dF(\gamma_n)\right] - \frac{\partial}{\partial s_t}\left[s_t\int p_e e_{n,t+1}dF(\gamma_n)\right] \quad (3.37)\end{aligned}$$

Equation (3.37) shows that, in this OLG model, the indirect utility of an individual of generation t depends on s_t only through variables that relate to the individuals in the next generation, be it their educational attainment or their level of labor supply. The first term on the right-hand side of (3.37) is basically the indirect utility coming from the child's own labor income and labor supply, reduced by his expenditures on education. This utility coming from the child's utility is undiscounted when $\delta \to 1$. Let us now examine the other two terms, which come respectively from the expressions of the transfers received by the current generation in power at t and the transfers accruing to the next one. However, they both relate to the costs and benefits of subsidising education acquisition for the whole generation born in t (and adult in $(t+1)$). Indeed, the second-to-last term represents the benefits of increasing education subsidies for the transfers received by the next generation: when s_t increases, children educate more, so that the tax base next period is bigger and the transfers to any adult of $(t+1)$ increase, which increases the utility of the adults of t. Then the last term represents the costs of educating the children of period t, borne by the adult of t: increasing these subsidies reduces the transfers received by the adults of t. Then, under the specific case where $\delta \to 1$, these two variations in transfers for different generations are valued the same in the utility of generation t.

Interpreting the result As stated in subsection 3.4.3 , the timing of collective decisions with respect to the individual ones is critical to the results that are obtained, in a setup where only one generation is considered. By introducing two overlapping generations with linkages due to altruism as in the present model, however, I show that it is perfectly possible to reconcile the fact that votes on several policy dimensions happen simultaneously, and that they impact one single given generation in a sequential way. Of course, characterising the equilibrium of the OLG model in the exact same fashion as the sequential 2-period voting game requires some stringent assumptions on the degree of altruism δ and the growth rate of population. Nevertheless, the main message of this section concerning the importance of identifying which generation is concerned by each policy at any given time remains true for other values of these parameters, even if it becomes too difficult to solve analytically for the equilibrium in these cases.

Bibliography

H. Aaron. The social insurance paradox. *Canadian Journal of Economics and Political Science*, 32:371–74, 1966.

A. B. Abel, N. G. Mankiw, L. H. Summers, and R. J. Zeckhauser. Assessing dynamic efficiency: Theory and evidence. *The Review of Economic Studies*, 56(1):1–19, Jan. 1989.

D. Acemoglu and J. Angrist. How Large are Human-Capital Externalities? Evidence from Compulsory-Schooling Laws. In *NBER Macroeconomics Annual 2000, Volume 15*, NBER Chapters, pages 9–74. National Bureau of Economic Research, Inc, October 2001.

F. Andersson and K. A. Konrad. Human capital investment and globalization in extortionary states. *Journal of Public Economics*, 87(7-8):1539–1555, August 2003.

R. Arawatari and T. Ono. Inequality, mobility and redistributive politics. *Journal of Economic Theory*, 148(1):353–375, 2013.

K. J. Arrow. A Utilitarian Approach to the Concept of Equality in Public Expenditure. *The Quarterly Journal of Economics*, 85(3):409–15, August 1971.

A. Atkinson and J. Stiglitz. The design of tax structure: Direct versus indirect taxation. *Journal of Public Economics*, 6:55 – 75, 1976.

A. B. Atkinson and J. E. Stiglitz. *Lectures on Public Economics*. McGraw-Hill, 1980.

A. J. Auerbach, J. Gokhale, and L. J. Kotlikoff. Generational accounting: A meaningful way to evaluate fiscal policy. *The Journal of Economic Perspectives*, 8(1):pp. 73–94, 1994.

V. Barham, R. Boadway, M. Marchand, and P. Pestieau. Education and the poverty trap. *European Economic Review*, 39(7):1257–1275, August 1995.

R. J. Barro. Are government bonds net wealth? *Journal of Political Economy*, 82: 1095–117, 1974.

M. Bassetto. Political economy of taxation in an overlapping-generations economy. *Review of Economic Dynamics*, 11(1):18–43, Jan 2008.

G. S. Becker and K. M. Murphy. The family and the state. *Journal of Law and Economics*, 31(1):pp. 1–18, 1988.

G. Bellettini and C. Berti Ceroni. Is social security really bad for growth? *Review of Economic Dynamics*, 2(4):796–819, Oct 1999.

R. Benabou. Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency? *Econometrica*, 70(2): 481–517, March 2002.

M. B. Berkman and E. Plutzer. Gray peril or loyal support? the effects of the elderly on educational expenditures*. *Social Science Quarterly*, 85(5):1178–1192, Dec 2004.

M. Bernasconi and P. Profeta. Redistribution or Education? The Political Economy of the Social Race. CESifo Working Paper Series 1934, CESifo Group Munich, 2007.

C. Bethencourt and V. Galasso. Political complements in the welfare state: Health care and social security. *Journal of Public Economics*, 92(3–4):609–632, Apr. 2008.

D. Black. On the rationale of group decision-making. *Journal of Political Economy*, 56 (1):23–34, Feb. 1948.

O. J. Blanchard. Debt, Deficits, and Finite Horizons. *Journal of Political Economy*, 93 (2):223–47, April 1985.

S. Blondal and S. Scarpetta. The retirement decision in oecd countries. Technical Report AWP 1.4, OECD, 1998.

D. E. Bloom, D. Canning, and G. Fink. Implications of population ageing for economic growth. *Oxford Review of Economic Policy*, 26(4):583–612, Dec 2010.

T. Blumkin and E. Sadka. A case for taxing education. *International Tax and Public Finance*, 15(2):145–163, April 2008.

R. Boadway, N. Marceau, and M. Marchand. Investment in Education and the Time Inconsistency of Redistributive Tax Policy. *Economica*, 63(250):171–89, May 1996.

R. W. Boadway and D. E. Wildasin. A median voter model of social security. *International Economic Review*, 30(2):307–328, May 1989.

H. Bohn. *Social Security and Demographic Uncertainty: The Risk-Sharing Properties of Alternative Policies*, pages 203–246. University of Chicago Press, January 2001.

M. Boldrin and A. Montes. The intergenerational state: Education and pensions. *Review of Economic Studies*, 72(3):651–664, 07 2005.

M. Boldrin and A. Rustichini. Political equilibria with social security. *Review of Economic Dynamics*, 3(1):41–78, 2000.

L.-E. Borge and J. Rattsø. Young and old competing for public welfare services. CESifo Working Paper Series 2223, CESifo Group Munich, 2008.

A. L. Bovenberg and B. Jacobs. Redistribution and education subsidies are Siamese twins. *Journal of Public Economics*, 89(11-12):2005–2035, December 2005.

F. Breyer. On the intergenerational pareto efficiency of pay-as-you-go financed pension systems. *Journal of institutional and theoretical economics*, 145:643–58, 1989.

F. Breyer. Voting on social security when labor supply is endogenous. *Economics & Politics*, 6(2):119–130, 1994.

F. Breyer and B. Craig. Voting on social security: Evidence from oecd countries. *European Journal of Political Economy*, 13(4):705–724, December 1997.

F. Breyer and M. Straub. Welfare effects of unfunded pension systems when labor supply is endogenous. *Journal of Public Economics*, 50(1):77–91, Jan. 1993.

E. K. Browning. Why the social insurance budget is too large in a democracy. *Economic Inquiry*, 13(3):373–88, September 1975.

E. Brunner and E. Balsdon. Intergenerational conflict and the political economy of school spending. *Journal of Urban Economics*, 56(2):369–388, Sep 2004.

J. Bryant. Modelling the effect of population ageing on government social expenditures. Treasury Working Paper Series 03/15, New Zealand Treasury, Mar. 2003.

M. R. Busemeyer. Education funding and individual preferences for redistribution. *European Sociological Review*, 29(6):1122–1133, Jan 2013.

M. R. Busemeyer and T. Iversen. The politics of opting out: explaining educational financing and popular support for public spending. *Socio-Economic Review*, 12(2): 299–328, Mar 2014.

M. R. Busemeyer, A. Goerres, and S. Weschle. Attitudes towards redistributive spending in an era of demographic ageing : the rival pressures from age and income in 14 OECD countries. *Journal of European Social Policy*, 19(3):195–212, 2009.

P. Carneiro and J. Heckman. Human Capital Policy. In B. M. Friedman, editor, *Inequality in America: What Role for Human Capital Policies?* MIT Press, 2004.

G. Casamatta, H. Cremer, and P. Pestieau. The political economy of social security. *The Scandinavian Journal of Economics*, 102(3):503–522, 2000.

G. Casamatta, H. Cremer, and P. Pestieau. Voting on pensions with endogenous retirement age. *International Tax and Public Finance*, 12(1):7–28, Jan. 2005.

G. Casamatta, H. Cremer, and P. Pestieau. Is there a political support for the double burden on prolonged activity? *Economics of Governance*, 7(2):143–154, May 2006.

M. A. Cattaneo and S. C. Wolter. Are the elderly a threat to educational expenditures? *European Journal of Political Economy*, 25(2):225–236, Jun 2009.

J. Conde-Ruiz and V. Galasso. Early retirement. *Review of Economic Dynamics*, 6(1): 12–36, jan 2003.

J. Conde-Ruiz and V. Galasso. The macroeconomics of early retirement. *Journal of Public Economics*, 88(9–10):1849–1869, Aug. 2004.

J. I. Conde-Ruiz, V. Galasso, and P. Profeta. The evolution of retirement. CEPR Discussion Papers 4863, C.E.P.R. Discussion Papers, Jan. 2005.

J. C. Conesa and C. Garriga. Optimal fiscal policy in the design of social security reforms. *International Economic Review*, 49(1):291–318, Feb. 2008.

J. C. Conesa and C. Garriga. Optimal response to a transitory demographic shock in social security financing. *Federal Reserve Bank of St. Louis Review*, 33-48(Jan):33–48, 2009.

J. C. Conesa and D. Krueger. Social security reform with heterogeneous agents. *Review of Economic Dynamics*, 2(4):757–795, October 1999.

T. F. Cooley and J. Soares. A positive theory of social security based on reputation. *Journal of Political Economy*, 107(1):135–160, Feb. 1999.

V. P. Crawford and D. M. Lilien. Social security and the retirement decision. *The Quarterly Journal of Economics*, 96(3):505–529, Aug. 1981.

H. Cremer and P. Pestieau. Intergenerational Transfer of Human Capital and Optimal Education Policy. *Journal of Public Economic Theory*, 8(4):529–545, October 2006.

H. Cremer, F. Gahvari, and P. Pestieau. Fertility, human capital accumulation, and the pension system. *Journal of Public Economics*, 95(11):1272–1279, 2011.

P. De Donder and J. Hindriks. Equilibrium social insurance with policy-motivated parties. *European Journal of Political Economy*, 23(3):624–640, September 2007.

G. de Fraja. The Design of Optimal Education Policies. *Review of Economic Studies*, 69(2):437–466, 2002.

A. de Tocqueville. *De la démocratie en Amérique*, volume II. Robert Laffont, 1986 edition, 1840.

W. J. den Haan and A. Marcet. Solving the Stochastic Growth Model by Parameterizing Expectations. *Journal of Business & Economic Statistics*, 8(1):31–34, January 1990.

P. A. Diamond. National debt in a neoclassical growth model. *The American Economic Review*, 55(5):1126–1150, Dec. 1965.

R. Disney. Population ageing and the size of the welfare state: Is there a puzzle to explain? *European Journal of Political Economy*, 23(2):542–553, June 2007.

A. Downs. *An economic theory of democracy*. Harper and Row, New York, 1957.

A. Drazen. *Political Economy in Macroeconomics*. Princeton University Press, 2000.

D. Durevall and M. Henrekson. The futile quest for a grand explanation of long-run government expenditure. *Journal of Public Economics*, 95(7-8):708–722, August 2011.

D. Epple and R. E. Romano. Public provision of private goods. *Journal of Political Economy*, 104(1):57–84, Feb. 1996.

European Commission. Progress and key challenges in the delivery of adequate and sustainable pensions in europe (a joint report on pensions). Occasional Papers 71, European Commission, Directorate-General for Economic and Financiall Affairs, November 2010.

M. S. Feldstein. Distributional Equity and the Optimal Structure of Public Prices. *American Economic Review*, 62(1):32–36, March 1972.

R. Fernandez and R. Rogerson. On the Political Economy of Education Subsidies. *Review of Economic Studies*, 62(2):249–62, April 1995.

R. Fernandez and R. Rogerson. The determinants of public education expenditures: Evidence from the states, 1950-1990. Working Papers 97-16, C.V. Starr Center for Applied Economics, New York University, 1997.

S. Findeisen and D. Sachs. Education Policies and Taxation without Commitment. Working Papers 14-16, University of Mannheim, Department of Economics, 2014.

L. Forni. Social security as markov equilibrium in olg models. *Review of Economic Dynamics*, 8(1):178–194, January 2005.

D. Fudenberg and J. Tirole. *Game Theory*, volume 1 of *MIT Press Books*. The MIT Press, 1991.

V. Galasso. *The Political Future of Social Security in Aging Societies*. MIT Press Books. The MIT Press, 2006.

V. Galasso. Postponing retirement: the political effect of aging. *Journal of Public Economics*, 92(10–11):2157–2169, Oct. 2008.

V. Galasso and P. Profeta. Lessons for an aging society: the political sustainability of social security systems. *Economic Policy*, 19:63–115, 2004.

V. Galasso and P. Profeta. How does ageing affect the welfare state? *European Journal of Political Economy*, 23:554–63, 2007.

O. Galor and J. Zeira. Income Distribution and Macroeconomics. *Review of Economic Studies*, 60(1):35–52, January 1993.

G. Glomm and B. Ravikumar. Public versus Private Investment in Human Capital Endogenous Growth and Income Inequality. *Journal of Political Economy*, 100(4): 818–34, August 1992.

J. A. Goldstone, E. P. Kaufmann, and M. Duffy Toft, editors. *Political Demography How Population Changes Are Reshaping International Security and National Politics*. Oxford University Press, 2011.

M. Gonzalez-Eiras and D. Niepelt. Population ageing, government budgets, and productivity growth in politico-economic equilibrium. CEPR Discussion Papers 6581, C.E.P.R. Discussion Papers, 2007.

M. Gonzalez-Eiras and D. Niepelt. The future of social security. *Journal of Monetary Economics*, 55(2):197–218, Mar. 2008.

M. Gonzalez-Eiras and D. Niepelt. Ageing, government budgets, retirement, and growth. *European Economic Review*, 56(1):97–115, Jan 2012.

M. Gradstein. An economic rationale for public education: The value of commitment. *Journal of Monetary Economics*, 45(2):463–474, April 2000.

M. Gradstein and M. Kaganovich. Aging population and education finance. *Journal of Public Economics*, 88(12):2469–2485, Dec. 2004.

M. Gradstein, M. Justman, and V. Meier. *The Political Economy of Education: Implications for Growth and Inequality*. The MIT Press, Jul. 2004.

J.-M. Grandmont. Intermediate Preferences and the Majority Rule. *Econometrica*, 46 (2):317–30, March 1978.

U. Grob and S. C. Wolter. Demographic change and public education spending: A conflict between young and old? *Education Economics*, 15(3):277–292, 2007.

J. Gruber and D. Wise. *Social Security and Retirement Around theWorld*. University of Chicago Press, Chicago, 1999.

P. Hammond. Charity: Altruism or cooperative egoism. In E. S. Phelps, editor, *Altruism, Morality and Economic Theory*, pages 115–31. New York: Russel Sage Foundation, 1975.

S. Hanley. Explaining the success of pensioner parties: a qualitative comparative study of 31 polities. In P. Vanhuysse and A. Goerres, editors, *Ageing Populations in Post-industrial Democracies: Comparative studies of policies and politics*, chapter 2, pages 23–53. Routledge, 2012.

P. G. Hare and D. T. Ulph. On Education and Distribution. *Journal of Political Economy*, 87(5):S193–212, October 1979.

A. R. Harris, W. N. Evans, and R. M. Schwab. Education spending in an aging america. *Journal of Public Economics*, 81(3):449–472, September 2001.

J. Heckman and P. Klenow. Human capital policy. In M. Boskin, editor, *Policies to Promote Human Capital Formation*. Hoover Institution, 1998.

R. Holzmann. Global pension systems and their reform : worldwide drivers, trends, and challenges. Social Protection Discussion Papers 68934, The World Bank, May 2012.

S. Homburg. The efficiency of unfunded pension schemes. *Journal of institutional and theoretical economics*, 146:640–47, 1990.

S. Homburg. Interest and growth in an economy with land. *The Canadian Journal of Economics / Revue canadienne d'Economique*, 24(2):450–459, May 1991.

N. Howe and R. Jackson. *The 2003 Aging Vulnerability Index: An Assessment of the Capacity of Twelve Developed Countries to Meet the Aging Challenge*. CSIS and Watson Wyatt Worldwide, 2003.

S. C. Hu. Social security, Majority-Voting equilibrium and dynamic efficiency. *International Economic Review*, 23(2):269–287, June 1982.

A. Ichino, L. Karabarbounis, and E. Moretti. The Political Economy Of Intergenerational Income Mobility. *Economic Inquiry*, 49(1):47–69, Jan. 2011.

I. Iturbe-Ormaetxe and G. Valera. Social security reform and the support for public education. *Journal of Population Economics*, 25(2):609–634, January 2012.

B. Jacobs and A. L. Bovenberg. Optimal Taxation of Human Capital and the Earnings Function. *Journal of Public Economic Theory*, 13(6):957–971, December 2011.

K. L. Judd. *Numerical Methods in Economics*, volume 1 of *MIT Press Books*. The MIT Press, December 1998.

M. Kaganovich and V. Meier. Social security systems, human capital, and growth in a small open economy. *Journal of Public Economic Theory*, 14(4):573–600, Aug 2012.

M. Kaganovich and I. Zilcha. Pay-as-you-go or funded social security? a general equilibrium comparison. *Journal of Economic Dynamics and Control*, 36(4):455–467, Apr 2012.

A. Kemnitz. Demographic structure and the political economy of education subsidies. *Public Choice*, 101(3-4):235–49, December 1999.

A. Kemnitz. Social security, public education, and growth in a representative democracy. *Journal of Population Economics*, 13(3):443–462, 2000.

D. Krueger and F. Kubler. Pareto-Improving social security reform when financial markets are incomplete!? *The American Economic Review*, 96(3):737–755, June 2006.

P. Krusell, V. Quadrini, and J.-V. Rios-Rull. Politico-economic equilibrium and economic growth. *Journal of Economic Dynamics and Control*, 21(1):243–272, Jan. 1997.

J. Lacomba and F. Lagos. Population aging and legal retirement age. *Journal of Population Economics*, 19:507–519, 2006.

J. Lacomba and F. Lagos. Political election on legal retirement age. *Social Choice and Welfare*, 29:1–17, 2007.

H. F. Ladd and S. E. Murray. Intergenerational conflict reconsidered: county demographic structure and the demand for public education. *Economics of Education Review*, 20(4):343–357, August 2001.

F. Lancia and A. Russo. A dynamic politico-economic model of intergenerational contracts. Vienna Economics Papers 1304, University of Vienna, Department of Economics, Mar. 2013.

G. Levy. The politics of public provision of education. *The Quarterly Journal of Economics*, 120(4):1507–1534, November 2005.

A. Lindbeck and J. Weibull. Balanced-budget redistribution as the outcome of political competition. *Public Choice*, 52:195–209, 1987.

P. H. Lindert. The rise of social spending, 1880-1930. *Explorations in Economic History*, 31(1):1–37, January 1994.

P. H. Lindert. What limits social spending? *Explorations in Economic History*, 33(1): 1–34, January 1996.

R. E. Lucas. On the mechanics of economic development. *Journal of Monetary Economics*, 22(1):3–42, Jul 1988.

W. Lutz, W. Sanderson, and S. Scherbov. The coming acceleration of global population ageing. *Nature*, 451(7179):716–719, Jan 2008.

D. Maldonado. Education policies and optimal taxation. *International Tax and Public Finance*, 15(2):131–143, April 2008.

X. Mateos-Planas. Demographics and the Politics of Capital Taxation in a Life-Cycle Economy. *American Economic Review*, 100(1):337–63, March 2010.

V. Meier and M. Werding. Ageing and the welfare state: securing sustainability. *Oxford Review of Economic Policy*, 26(4):655–673, Dec 2010.

A. H. Meltzer and S. F. Richard. A Rational Theory of the Size of Government. *Journal of Political Economy*, 89(5):914–27, October 1981.

G. M. Milesi-Ferretti and N. Roubini. On the taxation of human and physical capital in models of endogenous growth. *Journal of Public Economics*, 70(2):237–254, November 1998.

P. Milgrom and C. Shannon. Monotone Comparative Statics. *Econometrica*, 62(1): 157–80, January 1994.

C. Miller. Demographics and spending for public education: a test of interest group influence. *Economics of Education Review*, 15(2):175–185, Apr 1996.

J. A. Mincer. *Schooling, Experience, and Earnings*. Number 74-1 in NBER Books. National Bureau of Economic Research, 1974.

J. D. Moreno-Ternero and J. E. Roemer. The political economy of health care finance. CORE Discussion Papers 2007031, Université catholique de Louvain, Center for Operations Research and Econometrics (CORE), May 2007.

C. B. Mulligan and X. Sala-i Martin. Social security in theory and practice (i): Facts and political theories. Economics Working Papers 384, Department of Economics and Business, Universitat Pompeu Fabra, Apr. 1999.

K. Naito. Two-sided intergenerational transfer policy and economic development: A politico-economic approach. *Journal of Economic Dynamics and Control*, 36(9):1340–1348, Sep 2012.

S. B. Nielsen and P. B. Sorensen. On the optimality of the Nordic system of dual income taxation. *Journal of Public Economics*, 63(3):311–329, February 1997.

S. Nishiyama and K. Smetters. Does social security privatization produce efficiency gains? *The Quarterly Journal of Economics*, 122(4):1677–1719, November 2007.

OECD. *OECD Pensions Outlook 2012*. OECD Publishing, 2012.

OECD. *Pensions at a Glance 2013: OECD and G20 Indicators*. OECD Publishing, 2013.

OECD. *Education at a Glance 2014: OECD Indicators*. OECD Publishing, 2014a.

OECD. Tax data base 2014. Technical report, OECD, Paris, 2014b.

T. Ono. The political economy of environmental taxes with an aging population. *Environmental & Resource Economics*, 30(2):165–194, Feb. 2005.

T. Ono. Public education and social security: A political economy approach. Discussion Papers in Economics and Business 13-06, Osaka University, Graduate School of Economics and Osaka School of International Public Policy (OSIPP), May 2013.

F. C. Pampel and J. B. Williamson. *Age, Class, Politics, and the Welfare State*. ASA Rose Monograph Series. Cambridge University Press, 1989.

T. Persson and G. Tabellini. *Political Economics: Explaining Economic Policy*. MIT Press, 2000.

P. G. Peterson. Gray dawn: The global aging crisis. *Foreign Policy*, 78(1):42–55, 1999.

T. J. Philipson and G. S. Becker. Old-age longevity and mortality-contingent claims. *Journal of Political Economy*, 106(3):551–573, June 1998.

C. Pierson. *Beyond the welfare state? : the new political economy of welfare*. Penn State University Press, University Park, Pennsylvania, 2007.

C. R. Plott. A notion of equilibrium and its possibility under majority rule. *American Economic Review*, 57:787–806, 1967.

J. M. Poterba. Demographic structure and the political economy of public education. *Journal of Policy Analysis and Management*, 16(1):48–66, 1997.

P. Poutvaara. On the political economy of social security and public education. *Journal of Population Economics*, 19(2):345–365, June 2006.

P. Profeta. Aging and retirement: Evidence across countries. *International Tax and Public Finance*, 9(6):651–672, November 2002.

A. Rangel. Social security reform: Efficiency gains or intergenerational redistribution. mimeo, Harvard university, 1997.

A. Rangel. Forward and backward intergenerational goods: Why is social security good for the environment? *American Economic Review*, 93(3):813–834, Jun 2003.

A. Razin, E. Sadka, and P. Swagel. The aging population and the size of the welfare state. *Journal of Political Economy*, 110(4):900–918, August 2002.

J. E. Roemer. *Political competition: Theory and applications*. Harvard University Press, Cambridge, 2001.

G. Saint-Paul and T. Verdier. Education, democracy and growth. *Journal of Development Economics*, 42(2):399–407, December 1993.

P. A. Samuelson. An exact Consumption-Loan model of interest with or without the social contrivance of money. *Journal of Political Economy*, 66(6):467–482, Dec. 1958.

I. Sanz and F. J. Velazquez. The role of ageing in the growth of government and social welfare spending in the oecd. *European Journal of Political Economy*, 23(4):917–931, December 2007.

C. A. Shelton. The aging population and the size of the welfare state: Is there a puzzle? *Journal of Public Economics*, 92(3-4):647–651, April 2008.

K. Shepsle. Institutional arrangements and equilibrium in multidimensional voting models. *American Journal of Political Science*, 23(1):27–59, 1979.

E. Sheshinski. A model of social security and retirement decisions. *Journal of Public Economics*, 10(3):337–360, Dec. 1978.

A. Simonovits. Can population ageing imply a smaller welfare state? *European Journal of Political Economy*, 23(2):534 – 541, 2007.

H.-W. Sinn and S. Uebelmesser. Pensions and the path to gerontocracy in Germany. *European Journal of Political Economy*, 19(1):153–158, March 2003.

K. Sjoblom. Voting for social security. *Public Choice*, 45:225–240, 1985.

J. Soares. A dynamic general equilibrium analysis of the political economy of public education. *Journal of Population Economics*, 19(2):367–389, Jun 2006.

R. J. Sørensen. Does aging affect preferences for welfare spending? a study of peoples' spending preferences in 22 countries, 1985–2006. *European Journal of Political Economy*, 29(C):259–271, 2013.

N. L. Stokey, J. R. E. Lucas, and E. C. Prescott. *Recursive Methods in Economic Dynamics*. Harvard University Press, 1989.

G. Tabellini. A positive theory of social security. *The Scandinavian Journal of Economics*, 102(3):523–545, 2000.

M. Tepe and P. Vanhuysse. Accelerating smaller cutbacks to delay larger ones? the politics of timing and alarm bells in oecd pension generosity retrenchment. In P. Vanhuysse and A. Goerres, editors, *Ageing Populations in Post-industrial Democracies: Comparative studies of policies and politics*, chapter 6, pages 127–144. Routledge, 2012.

United Nations. World population prospects: The 2012 revision, volume ii, demographic profiles. Technical report, U.N. Department of Economic and Social Affairs, Population Division, 2013.

U.S. Bureau of Labor Statistics. Current population survey. Technical report, U.S. Department of Labor, 2013.

U.S. Social Security Administration. Social security programs throughout the world: Europe, 2014. Technical Report 13-11801, U.S. Social Security Administration, 2014.

S. Valdes-Prieto. The latin american experience with pension reform. *Ann Public & Coop Econ*, 69(4):483–516, Dec 1998.

B. van Groezen, T. Leers, and L. Meijdam. Social security and endogenous fertility: pensions and child allowances as siamese twins. *Journal of Public Economics*, 87(2): 233–251, February 2003.

P. Vanhuysse and A. Goerres, editors. *Ageing Populations in Post-industrial Democracies: Comparative studies of policies and politics*. Routledge, 2012.

M. E. Yaari. Uncertain lifetime, life insurance, and the theory of the consumer. *The Review of Economic Studies*, 32(2):pp. 137–150, 1965.