The political economy of pensions – Political support and scope for reform

Paul John Bernhard Stenbäck

University of Helsinki
Faculty of Social Sciences
Economics

Master's thesis
May 2012

Tiedekunta/Osasto – Fakultet/Sektion – Faculty Faculty of Social Sciences	Laitos – Institution – Department Department of Political and Economic Studies	
Tekijä – Författare – Author Paul Stenbäck		
Työn nimi – Arbetets titel – Title The political economy of pensions – Political support and scope for reform		
Oppiaine – Läroämne – Subject Economics		
Työn laji – Arbetets art – Level Master's thesis	Aika – Datum – Month and year May 2012	Sivumäärä – Sidoantal – Number of pages 84 + appendix

Tiivistelmä – Referat – Abstract

The ageing population of the developed world questions the financial sustainability of the pay-as-you-go pension system, which redistributes income from the current working population to the retirees. The development has spurred a debate in the public and in academia regarding the future of the system. This thesis analyzes the political support of the pension system and the potential for reform, with a focus on the conflict of interest between young and old generations.

The political support of the pension system is analyzed by surveying the two main theoretical modelling approaches of pension systems, and by analyzing one model of each approach in detail. In the majority voting approach, a pension system arises when at least half of the population has economic reasons to support it. Conde-Ruiz and Galasso (2005) model a pension system as the result of the support of retirees and low-income workers. The latter support the system because of its redistributive component that redistributes from rich to poor.

The interest group approach focuses explicitly on the political competition between groups in society. If the elderly lobby is powerful enough, pension transfers that redistribute from young to old emerge as a result. In the model of Mulligan and Sala-i-Martin (1999), a pension system arises as the result of time-intensive competition. As the young are likely to become old and the old have lower labour productivity, the old have a larger incentive to use their time for political competition. An alternative interpretation is that the elderly are more single-minded, i.e. they have less diverse political goals.

In addition to analyzing the two models above, the paper discusses them from an empirical and methodological point of view. Specifically, the controversial issue of applying economic methodology on political science is discussed. Finally, the thesis surveys political economy models that focus on the political sustainability of pension system reforms. Although the predictions on the potential for reform are pessimistic, recent empirical research suggests that the outlook may not be as gloomy as the models predict.

Sources:

Conde-Ruiz, J.I., Galasso, V., 2005: Positive arithmetic of the welfare state. Journal of Public Economics, 89, 933-955.

Mulligan, C.B., Sala-i-Martin, X., 1999: Gerontocracy, Retirement, and Social Security. NBER working paper 7117, National Bureau of Economic Research. citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.200.5865&rep=rep1&type=pdf. Date: 26.4.2012.

Avainsanat – Nyckelord – Keywords
Pensions, social security, PAYG, pay-as-you-go, ageing population, demographic change, political economy, public choice

Säilytyspaikka – Förvaringställe – Where deposited
Helsinki University Library, HELDA digital database

Muita tietoja – Övriga uppgifter – Additional information

Table of contents

1 Introduction

In an age of rising life expectancy and lower birth rates, rising pension expenditures are putting a strain on public finances. Since the introduction of pay-as-you-go[1] (PAYG) pension systems[2] in the developed world, the support ratio of the working population to the amount of elderly has diminished rapidly, and it is expected to diminish further in the coming decades. The demographic development questions the sustainability of many pension systems in the developed world, most of which are mainly based on PAYG arrangements.

Similarly to most redistributive schemes, PAYG social security systems involve a trade-off between protection and distortion of incentives. On the one hand, pension benefits protect the old from a sharp decline in income after retirement. On the other hand, the benefits are financed by taxing income of the working population. These taxes impose deadweight losses on the economy, because of the changes in behaviour of both the working population and the retired. The optimal social security system must therefore balance between these considerations.

The redistributive design of the PAYG system leads to conflicts of interest between different cohorts of society, as individual preferences for the social security system vary. Furthermore, as political decisions are rarely eternal, there exists an incentive for individuals and groups in society to attempt to reform the system in their favour. It is this conflict of interest that is the focus of the political economy of social security, and the implications the political struggle has on the design of social security systems.

This thesis surveys the two main model families of political theories of social security, namely majority voting models and interest group models. One model of each approach is

---

[1] An unfunded pension system where current pensions are financed by taxing the income of the current working population. In contrast, a funded system is one where current workers save for their own future pensions.

[2] As is common in the academic literature, the pension system is hereafter also referred to with the term social security system.

examined in detail. In doing so, we seek to understand why a redistributive system like the PAYG system exists and why it has grown to become the largest welfare program in the developed world. We also analyze the empirical predictions of the models, and discuss the methodology and model properties. The discussion gives us information on the plausibility of the models, and of the implications their predictions have on real-world policy. Finally, we shall examine the implications of an ageing population and the scope for social security reform.

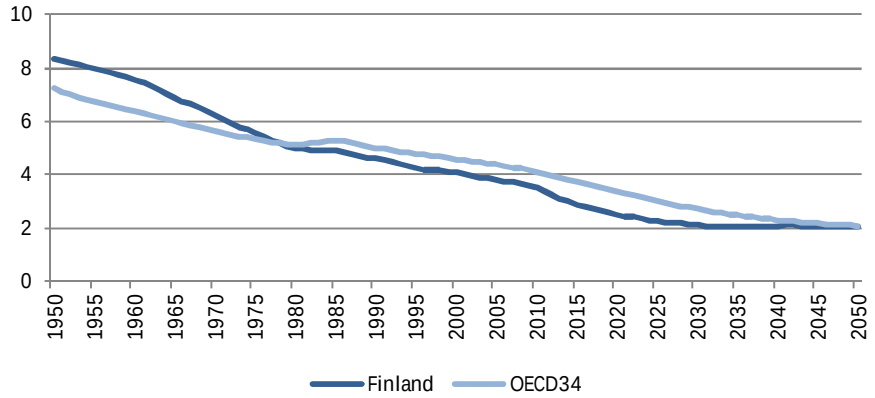
The thesis is organized as follows. The remainder of this chapter provides an overview of the demographic development in the developed world as well as an introduction to the political economy of social security. Chapter 2 analyzes the majority voting model of Conde-Ruiz and Galasso (2005) in detail, and discusses its result from an empirical and methodological point of view. In the following chapter, the interest group model of Mulligan and Sala-i-Martin (1999) is the subject for the same scrutiny. Chapter 4 discusses the implications for reforming social security, and the final chapter concludes.

1.1 Social security and the effects of an ageing population

The rise of the welfare state in the decades after World War 2 was accompanied by a large increase in social security spending. In the US, social security spending rose from 0,3 per cent of GDP in 1950 up to 4,7 per cent of GDP in 1996 (Mulligan and Sala-i-Martin, 1999: p. 2). The development is similar in other countries, and continues to this day. In 34 OECD countries, pension spending increased by 14,5 per cent between 1990 and 2007. Today, social security is the largest single item of public spending, amounting to an average of 7,0 per cent of GDP in 2007. (OECD, 2011)

Coinciding with the continuing rise of social security spending is the demographic transition of the developed world. Fertility rates have remained low for a long time already while life expectancy has increased rapidly, and the development is predicted to continue (OECD, 2011). The trend is well illustrated by the decrease in the support ratio, i.e. the amount of 20-65-year olds divided by the amount of above 65-year olds (picture 1.1). In 1950, the support ratio was 7,2 in 34 OECD countries, whereas it in 2010 had fallen to 4,1. In 2050, the support ratio is predicted to be a mere 2,1. Although the demographic

prediction is subject to uncertainty, the trend is clear: the population of the developed world is ageing. The trend is similar in most individual countries, with some differences in timing and intensity.



Picture 1.1: Support ratio, historical and projected values (OECD, 2011).

The diminishing support ratio has direct implications for the social security system. In most public pension systems in the developed world, a majority of pension revenues are financed through a PAYG system (IMF, 2004: p. 160). For example, in Finland in 2010, 87 per cent of pension payments were financed through payroll taxes and 13 per cent from pension funds (Eläketurvakeskus, 2011: p. 7). Ceteris paribus, the change in support ratio diminishes the working population contributing to pension revenues, and increases the amount of retirees eligible for pension benefits. Therefore, in order to balance revenues and expenditures of the PAYG system, changes must be made. For example, raising the retirement age or increasing immigration of working age immigrants are policies being discussed and implemented in countries throughout the developed world.

The imminent need for reform is obvious to lead to conflict of interests between different groups of society, as the PAYG system is ripe with vested interests. For example, current retirees are unlikely to favour a drastic cut in their pension benefits, after spending their whole working life financing the pensions of the previous generation. Similarly, young individuals entering the workforce are unlikely to favour steep raises in the payroll tax, as the tax is deducted straight from their own wage income. To understand the mechanisms of the PAYG social security system, we must analyze the political support of the system and how parameter changes, such as e.g. an ageing population, affects the system.

1.2 Political theories of social security

The efficiency of pension systems are widely researched topics in the economic literature (see e.g. Mulligan and Sala-i-Martin, 2004). The standard approach of economists is to analyze a theoretical model from the perspective of economic efficiency, and afterwards test this empirically, potentially leading to policy implications and recommendations. However, this approach may in some cases be incomplete when analyzing real-world economic decision-making. There are various reasons for which politicians, when choosing a policy from the set of alternatives, may deviate from economic efficiency and optimality for political purposes; lobbying, contributions to political campaigns and votes of vested interest groups to name a few. This is why social security design and reforms should be analyzed not only from the perspective of economic efficiency theory, but also from the perspective of politics and political economy.

Political science offers various ways by which politics and political decisions can be analyzed. The way in which economics can contribute to this discussion is by applying the tools of rational decision-making on the political process and its agents. This branch of research is known as public choice theory, and was originally developed by economists such as Duncan Black, James Buchanan and Gordon Tullock. It uses concepts from economics, including constrained utility maximization, game theory and decision theory, to analyze political institutions and processes. Contrary to standard economics, it sees not only the individuals but also politicians and government bureaucrats as mostly self-interested agents interacting in a society, and the state not only as a utility-maximizing institution but with potential interests of its own. In addition, it analyzes political institutions, such as representative democracy or constitutional rules, and their implications for political outcomes.

The political landscape of the society has an impact on the pension system in several ways. For example, as the PAYG pension system involves substantial redistribution from young to old, it leads to vested interests and opportunities for strong lobbying groups to affect decision-making. Also, institutional rules and the preferences of citizens affect the level of social security and the involvement of the state. Formalized political economy models of social security attempt to analyze the PAYG social security systems and their rise to

become the most used social insurance instrument in the world. Although the existence of a pension system may seem intuitive at first, the fact that it is unfunded turns it into a major redistributive scheme. The ageing population in the developed world questions the system's sustainability and political support, and puts downward pressures on the level of redistribution. Furthermore, the existence of a PAYG system in such a large part of the world indicates that there is some mechanism or advantage that induces a system to be created and sustained in a society.

It is worth contemplating on the differences between so called efficiency theories of social security and political theories of social security. The former see social security as the outcome of market inefficiencies or myopia. For example, a theory could argue that the market is unable to provide insurance for the old, and that the government is needed to provide social security. Political theories, however, see the PAYG system as an outcome of a political struggle. Groups of citizens compete politically to redistribute resources from each other, and the resulting redistribution from young to old is the social security system. It is the political theories of social security that is the subject of this paper. The theories are generally positive, as opposed to normative, describing how things are and not how they ought to be. However, as the political theories often suggest that social security is economically inefficient, they may be able to suggest efficiency-increasing reforms.

The political economy literature of social security has three main branches. Firstly, the models attempt to answer why programs that transfer substantial resources from the young and middle-aged to the elderly exist, and from where the programs get their political support, taking into account that the elderly constitute only a minority of the population. Secondly, the models try to examine how social security interacts with other redistributive programs of the welfare state as well as among the different characteristics of the social security system. Finally, the literature addresses current heated topics of the pension debate, such as how social security responds to demographic and economic changes, and what possibilities for reform are available.

This thesis will discuss all of the questions mentioned above. Chapters two and three, including the analysis of the models by Conde-Ruiz and Galasso (2005) and Mulligan and

Sala-i-Martin (1999), address the first issue on the support for social security. The second question is touched upon in Conde-Ruiz and Galasso's (2005) model, as it links social security to an income redistribution scheme. Reforming social security will be discussed in chapter four.

The purpose of the political mechanisms of a model is to aggregate individual preferences related to a policy, i.e. social security, into a political outcome. In accordance with the categorization by Galasso and Profeta (2002), the political mechanisms used to formally model social security can be categorized into three groups of models; majority voting, veto power and interest group models. Veto power models are few and less emphasized in the literature, and will therefore receive less attention in this thesis. In majority voting models, citizens vote and social security arises if there are economic reasons for a majority of the population to support the policy. On the other hand, interest group models focus more explicitly on the political struggle. If the elderly as a group are powerful enough, intergenerational redistribution from young to old arises as a result.

In the next two chapters, an overview of majority voting and interest group models is given, and one model of both will be analyzed in further detail. Furthermore, we will provide a discussion of the empirical implications and methodology of the models.

2 Majority voting approach

The most commonly used way to model social security systems is through majority voting models (hereafter also referred to as median voter models). Ever since Browning (1975) applied majority voting on social security, a large range of literature has been published either to expand the original model or to develop new ones. In this section, an overview of majority voting models is provided, after which we will analyze the majority voting model of Conde-Ruiz and Galasso (2005) in detail.

The outline of the majority voting game is simple. The electorate votes on a range of pair-wise votes on social security tax rates, and the option that obtains a majority of the votes against all other tax rate alternatives becomes the policy outcome. The preferences are

assumed to be single-peaked[3], which ensures that a Condorcet winner[4] can be found. This leads to the deciding voter being the median voter and the policy outcome being the option favoured by the median voter. The outcome is a Nash equilibrium, and the result is also known as the median voter theorem (see appendix 1 for a formal expression and proof).

When analyzing social security systems, it is important to extend the equilibrium to several periods. If young and middle-aged agents do not expect the system to be in place when they grow old, they will not support it in the first place. To tackle this problem, several solutions have been used. Browning (1975), among others, used a once-and-for-all election that assumes full commitment over future policies. Full commitment means that no future changes to the system are feasible. In order to avoid this unrealistic assumption, later contributions, including Hu (1982) and Sjoblom (1985) among others, have added recurring elections in every period to the models. If the voting model is multidimensional, i.e. it includes several policy variables, some further restrictions must be made in order to find an equilibrium. This case will be further discussed in section 2.1, where a multidimensional model is analyzed.

As long as the elderly comprise less than half of the population, the system must benefit at least a part of the young and the middle-aged in order for a majority of the population to vote in favour of social security. Otherwise the social security level in equilibrium would be zero. Galasso and Profeta (2002) identify five economic reasons, proposed by the theoretical literature, which could induce younger generations to vote in favour of a positive level of social security: dynamic inefficiency, reduced time horizon, crowding-out effects, within-cohort redistribution and altruism. These categories are briefly presented below.

Dynamic inefficiency of the economy means that the rate of return on capital is smaller than the implicit rate of return of social security. The latter equals the population growth

---

[3] Preferences are single-peaked when agents have a most preferred choice over a set of outcomes, and when the further away from the most preferred choice an outcome is, the less it is preferred.

[4] A Condorcet winner is the option which, compared to all other options, is favoured by a majority of the electorate.

rate times the real wage growth, as this is the rate at which the amount of funds available for PAYG redistribution grows in an unfunded system. If this inequality holds, social security improves welfare for all individuals in the economy, as shown by Aaron (1966). Majority voting models with this assumption include Browning (1975) and Sjoblom (1985), among others. However, the notion of dynamic inefficiency as a source of support for unfunded social security systems has been proven to be unrealistic (Abel, 1989), especially since it has been shown that dynamic inefficiency is not an absolute prerequisite for positive social security levels.

Already Browning (1975) expanded his analysis by showing that, even in a dynamically efficient economy, middle-aged citizens may vote for a positive social security level due to their reduced time horizon. For the middle-aged voters, past contributions are sunk costs and what matters are current and future contributions and benefits. Various models build upon this concept, including Boadway and Wildasin (1989) and Cooley and Soares (1999). Calculations by Galasso (2002) support the suggestion that middle-aged receive a higher internal rate of return from social security compared to other investments.

Cooley and Soares (1999) and Boldrin and Rustichini (2000) extend the so called crowding-out effect to social security. The argument goes that intergenerational redistribution schemes crowds out capital, as less capital is available due to smaller savings compared to a situation without social security. This effect increases the real returns to capital, and in turn decreases the rate of return to labour, i.e. the real wages. Therefore, asset-holders may be inclined to support a social security as it creates redistribution from individuals who focus on labour income to individuals with large capital holdings. The empirical relevance of the crowding-out effect has not been sufficiently tested.

The notion that social security relies on the support of the poor was first proposed by Tabellini (2000). He suggests that the within-cohort redistribution element of social security gives an incentive for less well-off individuals to support social security. Boskin et al. (1987) and Galasso (2002) have shown that low-income individuals yield higher returns from social security compared to high-income individuals in the US. This suggests that in a social security system with a high redistributive component, i.e. in a Beveridgean system

(as opposed to a Bismarckian system), rising income inequality should lead to a higher level of social security. A multidimensional model of social security and inequality will be presented in detail in section 2.1.

Finally, some models including Veall (1986) and Tabellini (2000) describe situations where the working-aged individuals have altruistic preferences towards the elderly. In Veall (1986), young agents save too little leading to an inefficient allocation of resources, as they realize that future young generations will provide them with pension transfers. Tabellini (2000) combines weakly altruistic preferences with the within-cohort redistribution argument discussed above.

In addition to majority voting models with one policy parameter, as in the models discussed above, some majority voting models attempt to analyze the existence of social security combined with other welfare programs. Conde-Ruiz and Galasso (2005) analyze the rise of a social security system jointly with an income redistribution system, a model which shall be presented in detail below. Poutvaara (2006) derives a model where a voting game decides the level of a wage tax, which is used for investment in public education and social security benefits. Similarly, in the model of Conde-Ruiz and Galasso (2003a), voters vote on the size of social security and on the existence of an early retirement provision.

As we have seen, majority voting is a straightforward way to model the support of social security systems in a society. A positive level of social security arises when at least half of the population supports it. For this to happen, at least half of the working population must have economic reasons to support social security. These reasons include dynamic inefficiency, reduced time horizon, crowding-out effects, within-cohort redistribution and altruism. In addition to unidimensional majority voting models, social security has been analyzed in models with two policy variables. In the next section, a multidimensional model of social security and the welfare state will be analyzed in detail.

2.1 A multidimensional model of social security

As mentioned in the previous section, the argument that the popularity of social security is based on the support of low-income young and the elderly was first proposed by Tabellini (2000). He argued that the redistributive component built into many social security systems induces the poor of the currently working population to support pension transfers. Together with the elderly, they constitute a majority of the population. In an overlapping generations model with weakly altruistic preferences and without commitment he arrives in a politico-economic equilibrium with a positive level of social security, which is the larger, the larger the amount of elderly in the population and the larger the inequality in the society. Furthermore, he presents empirical evidence in favour of the theory.

The result of the paper relies on two important assumptions. Firstly, the redistributive component of social security is crucial in order to arrive in an equilibrium with positive levels of social security. This is backed by the stylized fact that almost all pension systems in the world have some kind of redistributive component. Pension contributions are linked to wage income, but benefits are not or to a lesser degree. This assumption is discussed further in the discussion in chapter 2.4.2. Secondly, for the equilibrium to arise in Tabellini (2000), the social security system must be the only redistribution scheme in the society. This is not realistic, as most countries offer social insurance for the poor in other ways besides pensions. Were an additional social insurance scheme introduced into the model, the equilibrium would disappear. The low-income workers would prefer to redistribute through this other scheme instead of through social security.

To counter the problem of a disappearing equilibrium, Conde-Ruiz and Galasso (2005) build a model with a more complete and realistic welfare state, including both a social security system and a separate income distribution scheme. This is the model that is presented in this chapter. The model arrives in an equilibrium with a positive level of social security and income distribution, as long as the economy has sufficiently large income inequality and there are enough retirees in the population.

To complement the theoretical background of social security majority voting models given in the previous section, it should be noted that a similar framework of majority voting has also been applied on income inequality and income redistribution. In fact, the general intuition is the same as in the model of Tabellini (2000). The lower the income of a citizen, the larger the welfare system she supports due to redistribution from poor to rich. The implemented policy is the most preferred policy of the median voter, due to the median voter theorem (see appendix 1). A seminal paper on the subject was written by Meltzer and Richards (1981).

As we have two variables instead of one in the model of Conde-Ruiz and Galasso (2005), the model changes from being a unidimensional model, as in Tabellini (2000) and most other political economy models, into a multidimensional model. This leads to certain difficulties in modelling the equilibrium, as the median voter theorem may not necessarily be applicable. To counter this, the authors use subgame perfect structure-induced equilibrium outcomes, as suggested and defined by Shepsle (1979) and Conde-Ruiz and Galasso (2003a). The method will be presented in more detail below. Also other methods have been used to deal with multidimensional issue spaces in the literature, including legislative bargaining, probabilistic voting and interest group models, the last one which is the subject of chapter 3.

There are several reasons for choosing the model of Conde-Ruiz and Galasso (2005). Firstly, it is a good example of a majority voting model of social security, with the advantages and drawbacks discussed in the end of the chapter. Secondly, the model uses the result of the interest group model of Mulligan and Sala-i-Martin (1999), which is the model analyzed in chapter 3. This makes the models good compliments. Thirdly, the model implies interesting and relevant propositions for the interlinkedness of different welfare programs.

2.1.1 The model economy

The model of Conde-Ruiz and Galasso (2005) consists of an economy with overlapping generations, where two generations are alive every period, defined by the authors as “young” and old”. To make the division more intuitive and clear, I will use the words

"worker" instead of young and "retiree" instead of old. Population growth is constant $\mu > 0$, from which it follows that in any given period t, for every worker there are $1/(1+\mu)$ retirees. The model has no wage growth, and therefore the implicit rate of return on the unfunded social security system equals the population growth μ, as discussed in section 2.

Individuals work in their young age as workers, and retire when becoming old retirees. Consumption takes place only as retired. Workers differ in their working abilities, which are distributed throughout $[e,\bar{e}] \subset \Re_+$, according to a cumulative distribution function $G(.)$. Therefore, an agent born at time t is denoted by $e_t \in [e,\bar{e}]$. The distribution has a mean e_φ and is skewed, $G(e_\varphi) > 1/2$. A skewed labour income distribution means that there are more workers with income below the mean income compared to workers with above-mean income. Such a distribution is observed in most countries throughout the world.

The labour input is transformed into the only consumption good through a production function for every worker. This function is dependent on the worker's ability e_t and the amount of labour supplied by an individual with ability e, $n(e_t)$:

$$y(e_t) = e_t n(e_t). \tag{2.1}$$

A storage technology converts a unit of consumption in period t into $(1+\mathrm{R})$ units of consumption in period $t+1$, i.e. $y_{t+1} = (1+\mathrm{R})y_t$. All private intertemporal transfers of resources occur through this technology, which could be interpreted as e.g. savings in the capital market. Therefore, by assuming $R > \mu$, the economy is dynamically efficient, i.e. the rate of return of private savings is larger than the implicit rate of return of social security.

The utility function of the agents is log-linear and depends on leisure in working age and consumption as retired:

$$U(l_t, c_{t+1}^t) = \ln(l_t) + \beta c_{t+1}^t, \tag{2.2}$$

where l is leisure, c is consumption, β is the individual discount factor of time, superscripts denote the period the individual is born and subscripts denote the current

period. For workers there is a trade-off between the amount of labour $n(e_t)$ and leisure $l(e_t)$, since $n(e_t) = \bar{l} - l(e_t)$. $\bar{l} > 0$ is the total disposable time, which is assumed to be equal across agents. Workers pay a payroll tax, which is divided into an income redistribution tax τ_t and a social security tax σ_t, receive a transfer from the income redistribution scheme and save all their disposable income for future consumption. Retirees make no economic decisions, as they consume all of their saved income as well as their transfer from the pension scheme. The lifetime budget constraint for an individual born at time t with the ability e_t is therefore:

$$c_{t+1}^t = [e_t n(e_t)(1-\tau_t-\sigma_t) + T_t](1+R) + P_{t+1}, \tag{2.3}$$

where T_t is the income redistribution transfer at time t and P_{t+1} is the pension transfer at time $t+1$.

The first part of the optimization problem is now to find the labour supply, taking into account the budget constraint. To do that, a maximization problem for the worker can be constructed, where the workers maximize (2.2) with respect to $l(e_t)$, subject to the budget constraint (2.3). The function to be maximized is:

$$\max\{\ln(l(e_t)) + \beta((e_t(\bar{l} - l(e_t))(1-\tau_t-\sigma_t) + T_t)(1+R) + P_{t+1}\}. \tag{2.4}$$

Taking a derivative with respect to $l(e_t)$ gives the first order condition:

$$\frac{1}{l(e_t)} - \beta e_t(1-\tau_t-\sigma_t)(1+R) = 0. \tag{2.5}$$

We now assume that the individual discount factor equals the inverse of the rate of return of the storage technology, $\beta = 1/(1+R)$, so that the labour supply does not depend on the interest rate. Rearranging gives the amount of leisure and labour supply for an agent with ability e_t:

$$l^*(e_t) = \frac{1}{e_t(1-\tau_t-\sigma_t)}, \tag{2.6}$$

$$n^*(e_t) = \bar{l} - \frac{1}{e_t(1-\tau_t-\sigma_t)}. \tag{2.7}$$

As it is impossible for an agent to work less than 0, the labour supply is 0 if $n^*(e_t) \leq 0$. However, we assume that all agents supply a positive amount of labour, which imposes the following restriction on the tax rates:

$$n^*(e_t) = \bar{l} - \frac{1}{\underline{e}_t(1-\tau_t-\sigma_t)} > 0$$

$$\Leftrightarrow \tau_t + \sigma_t < 1 - \frac{1}{\bar{l}\underline{e}_t}. \tag{2.8}$$

We have now established the labour supply in equation (2.7), which depends on the type of the agent e_t. Due to the log-linearity of the utility function (2.2), the labour supply reacts to changes in the tax rate but not to changes in transfer levels. Income effects do not affect the labour supply decision, whereas the substitution effect of taxes does and distorts labour supply decisions. The larger the tax rates, the lower the supply of labour. This simplifies the analysis, as it implies that tomorrow's fiscal policies do not have an effect on today's labour supply.

2.1.1.1 The welfare system

In addition to the labour supply of the agents, we need to know what the level of the income and pension transfers are, given the labour supplied. It should be noted that the welfare system is the only purpose of the government in the model. There is no utilitarian government maximizing the utility of its citizens, but the state is instead purely a device through which taxes are channelled into income redistribution and pension benefits. The budget is assumed to be balanced in every period, and therefore the transfers depend on how much is collected through taxation in every period.

The welfare system of the model consists of two instruments, an income redistribution system and a social security system. The income distribution is an intragenerational redistribution scheme which only affects the workers. All workers benefit from a lump sum transfer T_t, which is financed by a flat payroll tax τ_t on the labour income. This creates redistribution from the rich, i.e. the agents with abilities above mean income, to the poor, i.e. the agents with abilities below mean income, as the amount paid in taxes rises with income whereas the transfer does not. In the real world, this corresponds to e.g. unemployment insurance schemes or other programs that help the less better off.

On the other hand, the social security system is an intergenerational scheme, which redistributes from the workers to the retirees. Workers contribute a payroll tax rate σ_t of their labour income, and retirees benefit from a lump sum transfer P_t, which is independent of the earlier contributions. This also implies redistribution from rich to poor, as the rich pay more contributions but get no more pension benefits than the poor. Finally, it is assumed that both systems balance their budgets every period, which means that total contributions must equal total payments in every period.

From these assumptions we can build budget constraints for the systems. The transfers paid must equal the income from the payroll taxes. The total income of the economy is calculated by adding together the labour income by integrating over the whole population. The budget constraint of the income redistribution system is thus:

$$T_t = \tau_t \int_{\underline{e}}^{\bar{e}} e_t n(e_t) dG(e_t). \tag{2.9}$$

The budget constraint of the social security scheme is similar, although a factor of $1+\mu$ must be added as the old are fewer than the young due to population growth:

$$P_t = \sigma_t (1+\mu) \int_{\underline{e}}^{\bar{e}} e_t n(e_t) dG(e_t). \tag{2.10}$$

As we insert the labour supply (2.7) into the budget constraints (2.9) and (2.10), we obtain new expressions for the budget constraints that are dependent on the tax rates only. For the redistribution system (see appendix 2 for calculations):

$$T_t(\tau_t, \sigma_t) = \tau_t (e_\varphi \bar{l} - \frac{1}{(1-\tau_t-\sigma_t)}), \tag{2.11}$$

and for the pension system:

$$P_t(\tau_t, \sigma_t) = \sigma_t (1+\mu)(e_\varphi \bar{l} - \frac{1}{(1-\tau_t-\sigma_t)}). \tag{2.12}$$

From the constraints it can be seen that the workers' redistribution transfer displays a Laffer curve with respect to the corresponding tax rate τ_t. This means that transfers rise when the tax rate rises up to a certain maximum, and decreases when transfers rise above the maximum. On the other hand, workers' transfers depend negatively on the social security payroll tax rate. As the social security tax decreases the value of (2.3) due to benefit

payments to the currently retired, the tax distorts the labour decision of the workers. This leads to a decrease in the average income in the economy and consequently reduces redistribution benefits of the workers. Similarly, the pension transfer of the retirees shows a Laffer curve with respect to the pension tax rate σ_t and depends negatively on the redistribution tax rate of the workers.

2.1.1.2 The economic equilibrium

So far, we have been able to define the labour supply of the agents as well as the level of transfers, given the labour supplied. The next step is to define the economic equilibrium as well as calculate the indirect utility functions of the young and old agents.

The definition of the economic equilibrium goes as follows: An equilibrium is a sequence of allocations of $\left\{l(e_t), c_{t+1}^t(e_t)\right\}$ at time $t = 0,...,\infty$ and $e_t \in [\underline{e}, \bar{e}]$, given a sequence of tax rates $\{\tau_t, \sigma_t\}$at time $t = 0,...,\infty$, and given a real interest rate R, such that it fulfils the following three conditions. Firstly, the consumer maximizes her utility in every period, i.e. maximizes (2.4). Secondly, the income redistribution and social security budget constraints, (2.9) and (2.10), are satisfied in every period. Finally, the aggregate resource constraint, as defined below, is satisfied in all periods:

$$\int_{\underline{e}}^{\bar{e}} c_t^{t-1}(e_{t-1}) dG(e_{t-1}) = (1+R)\int_{\underline{e}}^{\bar{e}} (1-\sigma_{t-1}) e_{t-1} n(e_{t-1}) dG(e_{t-1}) +$$

$$\sigma_t (1+\mu)\int_{\underline{e}}^{\bar{e}} e_t n(e_t) dG(e_t). \qquad (2.13)$$

The aggregate resource constraint can be interpreted as such that the consumption of the retirees in period t, i.e. the left-hand side of the equation, must equal the current value of the income that was not paid as pensions in period $t-1$, plus the pension benefits of period t, adjusted for population growth.

In order to obtain the utility level of the agents in an economic equilibrium at time t for a worker with ability e_t and a retiree with ability e_{t-1}, we calculate the indirect utility functions by inserting the levels of leisure (2.6) and labour (2.7) as well as the budget constraints given the labour choices (2.11) and (2.12) into the direct utility function (2.2).

For the workers, this indirect utility function depends on the tax rates in periods t and $t+1$ as well as on the ability of the agent:

$$v_t^t(\tau_t,\sigma_t,\tau_{t+1},\sigma_{t+1},e_t) = \ln(\frac{1}{e_t(1-\tau_t-\sigma_t)}) + \beta((e_t(\bar{l} - \frac{1}{e_t(1-\tau_t-\sigma_t)})(1-\tau_t-\sigma_t)$$
$$+ \tau_t(e_\varphi \bar{l} - \frac{1}{(1-\tau_t-\sigma_t)}))(1+R) + \sigma_t(1+\mu)(e_\varphi \bar{l} - \frac{1}{(1-\tau_{t+1}-\sigma_{t+1})})).$$

Simplifying and using the assumption that $\beta = 1/(1+R)$, we arrive at the expression:

$$v_t^t(\tau_t,\sigma_t,\tau_{t+1},\sigma_{t+1},e_t) = -\ln e_t - \ln(1-\tau_t-\sigma_t) - 1 + e_t\bar{l}(1-\tau_t-\sigma_t)$$
$$+ \tau_t(e_\varphi \bar{l} - \frac{1}{(1-\tau_t-\sigma_t)}) + \sigma_t \frac{1+\mu}{1+R}(e_\varphi \bar{l} - \frac{1}{(1-\tau_{t+1}-\sigma_{t+1})}). \qquad (2.14)$$

The right-hand side of the equation can be interpreted as follows. The first two terms depict the utility derived from leisure as worker. The rest of the terms correspond to the consumption in the next period $t+1$ derived from three sources; the income that is earned in period t and not paid as taxes (third and fourth term), the income redistribution in period t (fifth term) as well as the pension transfer in period $t+1$ adjusted for population growth and the discount rate (sixth term).

Analogously, the indirect utility function of the retiree depends the agent's ability as well as on the tax rates in period $t-1$ and t:

$$v_t^{t-1}(\tau_{t-1},\sigma_{t-1},\tau_t,\sigma_t,e_{t-1}) = e_{t-1}n(e_{t-1})(1-\tau_{t-1}-\sigma_{t-1}) + T_{t-1}$$
$$+ \sigma_t \frac{1+\mu}{1+R}(e_\varphi \bar{l} - \frac{1}{(1-\tau_t-\sigma_t)}).$$

As the first two terms are constants at time t, depending on the agent's ability and past taxes but not on current or future taxes, we can simplify the expression to:

$$v_t^{t-1}(\tau_{t-1},\sigma_{t-1},\tau_t,\sigma_t,e_{t-1}) = K(\tau_{t-1},\sigma_{t-1},e_{t-1})$$
$$+ \sigma_t \frac{1+\mu}{1+R}(e_\varphi \bar{l} - \frac{1}{(1-\tau_t-\sigma_t)}), \qquad (2.15)$$

where $K(\tau_{t-1},\sigma_{t-1},e_{t-1}) = e_{t-1}n(e_{t-1})(1-\tau_{t-1}-\sigma_{t-1}) + T_{t-1}$. At current time t, retirees take the constant term K, which depicts decisions made in the past period $t-1$, as given. This leaves the pension transfer of time t as the only relevant variable for the retired.

The indirect utility functions (2.14) and (2.15) show the preference relations of the workers and retirees over current and future taxes. It is worth noting that the retirees' ability affects their utility, but does not have an impact on the preferences over current tax rates. All retired agents share the same preferences over the income distribution and social security tax rates and transfers. Retirees are single-minded, as shall be seen in chapter 3 (Mulligan and Sala-i-Martin, 1999). On the other hand, the tax preferences of the workers do depend on the ability of the agent.

2.1.2 The voting game

So far, standard microeconomic theory has been used to derive the economic equilibrium and the indirect utility functions of the agents in equilibrium. In order to arrive at a policy outcome, a political process which aggregates the preferences of the agents in the economy is needed. Conde-Ruiz and Galasso (2005) consider a majority voting system, where elections take place in every period. All agents alive cast their votes. Every agent has zero mass, which means that no single voter alone can affect the election outcome, and which leads to an assumption of sincere voting[5]. In every election, voters vote on the current tax rates, τ_t and σ_t. This voting game features two important and distinguishable elements; the game is repeated in every period, and the voting game is two-dimensional, i.e. it has two decision variables.

Voters only determine the current policies, which can be changed at zero cost in the next period. As was briefly discussed in section 2, this means that workers may not be inclined to support a social security system if they suspect that their current voting behaviour will not have any relevance for future voting outcomes. If this was the case, workers would vote against social security, or else they would pay a tax without future benefits. However, if current workers can expect that current voting has an impact on future policies, an incentive may arise to support social security. As suggested by Hammond (1975), an implicit contract between generations may arise, in which current workers vote for a positive level

[5] Sincere voting assumes that voters vote in a way that reflects their true preferences, as opposed to strategic voting when voters support a candidate not reflecting their true preferences in order to avoid an undesirable outcome.

of social security and in return expect to be rewarded with a transfer when they retire. Failure to adhere to the implicit contract is punished by removal of pension transfers.

With these assumptions, various sequences of tax policy compose equilibria. In fact, any sequence of tax rates, which in every period makes a majority of voters better off than without social security, are subgame perfect equilibrium[6] outcomes. Different criteria have been used in the literature to select between these equilibria. As the policy issue in this model is two-dimensional, Nash equilibria are typically not found by analyzing a static voting game with commitment. To solve the problem, we follow Shepsle (1979) and Conde-Ruiz and Galasso (2003a), who analyze structure-induced equilibria, in which voters vote simultaneously, but separately, on the issues at hand. This issue-by-issue voting preserves the median voter framework (see appendix 1), which is then used to choose the equilibrium policies. The equilibrium outcomes are referred to as subgame perfect structure-induced equilibrium (SPSIE) outcomes. (See appendix 3 for a formal definition of structure-induced equilibria, and appendix 4 for a formal definition of the voting game and the SPSIE.)

To find the outcomes, we first analyze the structure-induced equilibria of the voting game with commitment over social security. Today's voters decide on the current and future social security tax rates, as well as on the current income redistribution tax rate. An issue-by-issue voting game is used to determine the equilibria. The space of policy alternatives is $(\tau, \sigma) \in [0,1]x[0,1]$ subject to (2.8), which was defined as the condition for a positive labour supply. In order for the median voter framework to be applicable, the agents' preferences over the tax rates must be single-peaked, as discussed in section 2. Derived from the indirect utility functions (2.14) and (2.15), the preferences can be shown to be single-peaked if $N = (1+\mu)/(1+R) > 1/2$ (see appendix 5 for calculations).

In the second phase, we abstract from the assumption of commitment over future social security tax rates. The equilibrium outcomes of the voting game with commitment will then be utilized to find the equilibrium outcomes of the voting game without commitment.

---

[6] A subgame perfect equilibrium is a strategy profile which represents a Nash equilibrium of all subgames of the original game.

2.1.3 Politico-economic equilibria

This section examines the voting game, which determines the size of the welfare state tax rates, (τ, σ). First, we determine the median voters of the two issues in a voting game with commitment. This is done by finding the most preferred income distribution tax rate given the social security tax rates, $\tau(\sigma)$, and vice versa, $\sigma(\tau)$. For each value of σ, the median voter over τ is identified, and vice versa. These median voters' functions cross at (τ^*, σ^*), which is a structure-induced equilibrium of the voting game with commitment. Afterwards, the equilibrium of a repeated game without commitment is considered.

2.1.3.1 Voting on the income redistribution tax rate

The voting on the tax rates depend on the indirect utility functions of the workers, (2.14), and the retirees, (2.15). For the retired, the voting on the income redistribution tax rate is straightforward. As the income redistribution tax reduces the average income in the economy and thus also decreases pension benefits, the indirect utility function depends negatively on the income redistribution tax rate. Therefore, for any positive value of σ, the maximization of (2.15) with respect to τ yields $\tau_{t,old} = 0$.

On the other hand, workers may benefit from income redistribution, depending on their ability. The most preferred income redistribution tax rate for the workers is obtained by maximizing (2.14) with respect to τ_t:

$$\frac{\partial v_t^t(.)}{\partial \tau_t} = (e_\varphi - e_t)\bar{l} - \frac{\tau_t}{(1-\tau_t-\sigma_t)^2} = 0 .$$

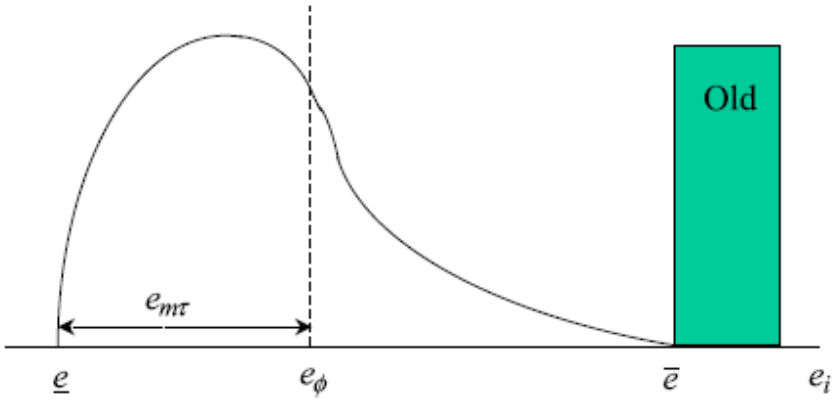
Rearranging and calculating (see appendix 2) gives:

$$\tau_{t,e}(\sigma_t) = \max\left\{0; 1-\sigma_t + \frac{1-\sqrt{1+4(e_\varphi - e_t)\bar{l}(1-\sigma_t)}}{2(e_\varphi - e_t)\bar{l}}\right\}. \qquad (2.16)$$

The equation above depicts the relation between the most preferred income redistribution rate and the ability of the agents. We note that the expression $(e_\varphi - e_t)$ has a negative sign in the numerator and a positive sign in the denominator. This means that increasing (decreasing) the expression causes a decrease (increase) in the equilibrium tax rate. The higher the ability of an agent, the lower is the preferred tax rate. When the expression

$(e_\varphi - e_t)$ is positive, the equilibrium tax rate is positive, and when the expression is negative, the equilibrium rate is 0. In other words, above average income types, $e_t > e_\varphi$, vote for $\tau_t = 0$ together with the retired, whereas below average income types, $e_t < e_\varphi$, vote for a positive tax rate.

The above characteristics of the voting game and the equilibrium tax rates of the workers and retirees allow us to order the agents according to their ability and age. This is done in picture 2.1, where the y-axis is the number of voters. The white area depicts the workers ordered according to their skewed distribution of ability, whereas the green area describes the old population. The population to the left are poor and support a positive income redistribution tax, whereas the population to the right are richer or retired, and vote for a zero tax rate.



Picture 2.1: Voting on income redistribution (Conde-Ruiz and Galasso, 2005).

Since the retired generation, even after adjusting for voting participation rates, represents a minority in the model and in most electorates in the real world, the median voter on the income redistribution tax rate (intragenerational median voter) is a currently working individual. In fact, the intragenerational median voter is the type $m\tau$ worker who divides the population in two. This corresponds to that the workers' cumulative distribution function in the point $e_{m\tau}$, $G(e_{m\tau})$, should equal half of the total population, as the old vote for no tax and are located to the right of the young in picture 2.1. As we know from section 2.1.1 that for every worker there is $1/(1+\mu)$ old, we get:

$$G(e_{m\tau}) = \frac{1}{2} * (1 + \frac{1}{1+\mu}) = \frac{2+\mu}{2(1+\mu)}. \quad (2.17)$$

Therefore, having concluded that the median voter is a worker and following equation (2.16), the equilibrium tax rate is $\tau_{m\tau}(\sigma) > 0$ if the median voter's ability is below the average ability $e_{m\tau} < e_{\varphi}$. If the opposite holds, the tax rate is 0. At this stage, the time index of the tax can be dropped, as we consider steady states.

2.1.3.2 Voting on the social security tax rate

The most preferred social security tax rate can be derived in a similar fashion as the income redistribution tax rate. For the retirees, we maximize their indirect utility function (2.15) with respect to σ_t:

$$\frac{\partial v_t^{t-1}(.)}{\partial \sigma_t} = e_{\varphi}\bar{l} - \frac{1}{1-\tau_t-\sigma_t} - \frac{\sigma_t}{(1-\tau_t-\sigma_t)^2} = 0$$

$$\Leftrightarrow e_{\varphi} n(e_{\varphi}) = \frac{\sigma_t}{(1-\tau_t-\sigma_t)^2} \quad (2.18)$$

$$\Leftrightarrow \sigma_{t,old}(\tau_t) = 1 - \tau_t - \sqrt{\frac{1-\tau_t}{e_{\varphi}\bar{l}}}. \quad (2.19)$$

This corresponds to the retirees' most preferred choice of the social security tax rate, which maximizes their pension transfer (see appendix 2 for calculations).

As we assume commitment over social security tax rates, the voting decision for the workers is derived by maximizing the indirect utility function (2.14) with respect to the current and future social security policy $\sigma_t = \sigma_{t+1} = \sigma$, taking the current and future income redistribution rates, τ_t and τ_{t+1}, as given (see appendix 2 for more calculations):

$$\frac{\partial v_t^t(.)}{\partial \sigma_t} = \frac{1}{1-\tau_t-\sigma_t} - e_t\bar{l} + \frac{\partial T_t}{\partial \sigma_t} + \frac{\frac{\partial P_{t+1}}{\partial \sigma_t}}{1+R} = 0$$

$$\Leftrightarrow e_t n(e_t) - \frac{\partial T_t}{\partial \sigma_t} = \frac{\frac{\partial P_{t+1}}{\partial \sigma_t}}{1+R}. \quad (2.20)$$

This equation gives the intuition of the trade-off faced by the workers when voting for social security. The left-hand side captures the current cost to the workers. They pay a contribution and the pension tax decreases the redistribution transfer, T_t, due to a reduction in the tax base. The right-hand side describes the increase in the pension payment in the next period. This trade-off leads to substitutability of the different welfare programs; a large rate of income redistribution makes the social security system more expensive for the workers.

We denote $N = (1+\mu)/(1+R)$ as the performance of the pension system compared to the saving technology. By setting $\tau_t = \tau_{t+1}$ and restricting our analysis to steady states, we can simplify and rearrange equation (2.20) further (see appendix 2 for calculations):

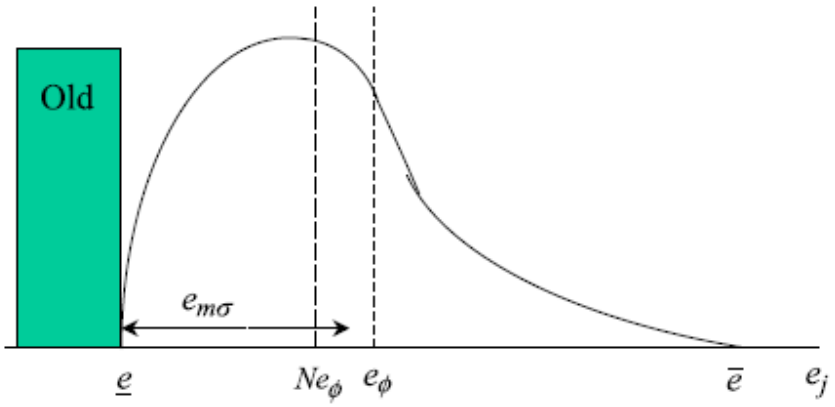
$$e_\varphi n(e_\varphi) - \frac{1}{N}(e_t n(e_t) + \frac{\tau_t}{(1-\tau_t-\sigma_t)^2}) = \frac{\sigma_t}{(1-\tau_t-\sigma_t)^2} \tag{2.21}$$

$$\Leftrightarrow \sigma_{t,e}(\tau_t) = \max\left\{0;1-\tau_t + \frac{1-\sqrt{1+4\bar{l}(N+\tau_t(1-N))(e_\varphi N - e_t)}}{2\bar{l}(e_\varphi N - e_t)}\right\}. \tag{2.22}$$

Substituting values for e shows that the voting decision is decreasing with the ability of the agent, i.e. the lower the ability, the higher the preferred tax rate. This is due to the within-cohort redistribution of the pension system, which combines a proportional income tax with a lump sum pension. Especially, for small enough values for the income redistribution tax, $\tau_t \leq (1-N)/(2-N)$, only the voters whose pre-tax income is below a certain fraction *N* of the pre-tax average income in the economy, $e_t n(e_t) < N e_\varphi n(e_\varphi)$ (with $N < 1$ as the economy is dynamically efficient), will vote for a positive level of social security. Richer workers will vote for a tax level of zero. The calculations of the conditions are not provided in the appendix, as the authors do not provide information on the computations and I was unable to calculate them myself.

Comparing equations (2.18) and (2.21) reveals that all workers will vote for a lower social security tax rate than the retirees. As the left-hand side of (2.21) is smaller than the left-hand side of (2.18), the tax rate σ_t in (2.21) also has to be smaller than in (2.18) in order for the equality to hold. Following these conclusions, the voters' preferences over the social

security tax rate can be ordered according to age and income in picture 2.2, in the same way as above with the income redistribution tax.



Picture 2.2: Voting on social security (Conde-Ruiz and Galasso, 2005).

The median voter on the pension tax rate is now, once again, the type $m\sigma$ worker who divides the electorate in half. However, as the retirees are now voting for a higher tax than the young and are located to the left of the workers in picture 2.2, the corresponding point on the workers' cumulative distribution function should now equal half of the working population less the retired population, i.e.:

$$G(e_{m\sigma}) = \frac{1}{2} * (1 - \frac{1}{1+\mu}) = \frac{\mu}{2(1+\mu)}. \qquad (2.23)$$

This corresponds to an agent with a lower ability compared to the median voter on income redistribution, as the value of equation (2.23) is smaller than the value of (2.17). The value of the median social security tax rate is $\sigma_{m\sigma}(\tau)$ as defined in equation (2.22).

2.1.3.3 Equilibrium outcomes

In the previous two sections we have analyzed the voters' preferences over the income redistribution and the pension schemes. This was done by first determining the median voter over both issues, $e_{m\tau}$ and $e_{m\sigma}$, and then calculating the median voters' most preferred tax rates, $\tau_{m\tau}(\sigma)$ and $\sigma_{m\sigma}(\tau)$. These preferred policies, depicted by equations (2.16) and (2.22), can be interpreted as reaction functions. For each value of the pension (income redistribution) tax rate, equation (2.16) (equation (2.22)) gives the median voter's most preferred tax rate of the income redistribution (pension) system. As defined by Shepsle

(1979), the structure-induced equilibrium outcomes of the voting game correspond to the points where these reaction functions are equal, i.e. cross.

In order to display the equilibrium results in an intuitive way, it is useful to add a measure of the ability of the median voters relative to the mean voters, marked by Δ. For the two systems respectively, we get $\Delta_\tau = (e_\varphi - e_{m\tau})\bar{l}$ and $\Delta_\sigma = (Ne_\varphi - e_{m\sigma})\bar{l}$. It should be noted that Δ_τ only measures the difference between the abilities of the mean and the median voters, whereas Δ_σ is the difference between the median voter's ability and the mean voter's ability weighted by the relative performance of the pension system, N. This is due to the fact that social security is an inferior way to redistribute because of its inefficiency in transferring resources into the future compared to private savings, as was defined early on in section 2.1.1. Additionally, we can introduce two further simplifying notations, $\Delta = \Delta_\tau(1-N) - \Delta_\sigma$ and $\hat{\Delta}_\tau = \Delta_\tau - (1-N)(1+\sqrt{1+4\Delta_\tau})/2$. By inserting the value of $\tau_{t,e}(\sigma_t)$ from (2.16) into the value of $\sigma_{t,e}(\tau_t)$ in (2.22) and vice versa, we get an equation system with two equations and two unknowns. As the equation system is extremely cumbersome to solve by hand without a computer program, the solution is not provided in the appendix. However, the system is solved by Conde-Ruiz and Galasso (2005), and the result is the following proposition, which outlines the structure-induced equilibrium outcome of the voting game with commitment.

Proposition 2.1 The unique structure-induced equilibrium of the voting game with commitment over the social security tax rates exists with the outcome (τ^*, σ^*), such that

(1) if $\Delta_\tau \le 0$ and $\Delta_\sigma \le -(1-N)$, then $\tau^* = 0$ and $\sigma^* = 0$;

(2) if $\Delta_\tau \le 0$ and $\Delta_\sigma > -(1-N)$, then $\tau^* = 0$ and $\sigma^* = 1 + \dfrac{1-\sqrt{1+4N\Delta_\sigma}}{2\Delta_\sigma} > 0$;

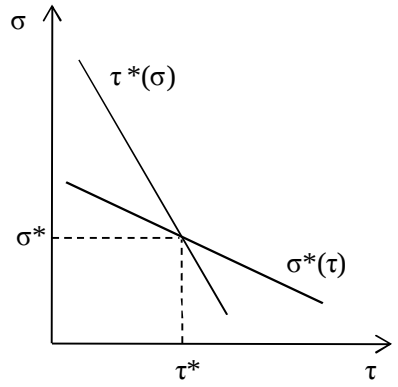
(3) if $\Delta_\tau > 0$ and $\Delta_\sigma \le \hat{\Delta}_\tau$, then $\tau^* = 1 + \dfrac{1-\sqrt{1+4\Delta_\tau}}{2\Delta_\tau}$ and $\sigma^* = 0$;

(4) if $\Delta_\tau > 0$ and $\Delta_\sigma > \hat{\Delta}_\tau$, then $\tau^* = \Delta_\tau \frac{1-2N\Delta-\sqrt{1-4N\Delta}}{2\Delta^2} > 0$ and

$\sigma^* = 1 - N - \tau^*(2 - N - \frac{\Delta_\sigma}{\Delta_\tau}) > 0$.

The proposition links income inequality, i.e. the relative ability of the median voters, with the equilibrium welfare programs. In the first case, income inequality is low and no welfare programs appear in equilibrium. In case 2, the intragenerational median voter's ability is higher than the mean ability, but the median voter on pensions is below the mean, and therefore only the social security system is implemented. This kind of society could arise in an economy with moderate income inequality and a sufficiently large elderly population, or in an economy where the income inequality is mainly due to a large share of elderly citizens. The third case represents a society with large inequality inside the working population, but only a small amount of retirees, resulting in an equilibrium with an income redistribution system only. This corresponds to a society with a young population and large income inequality. Finally, the last case requires a sufficient amount of income inequality and old individuals, and gives rise to a welfare state with both pensions and income redistribution. This is the society that most developed countries comprise today.

Picture 2.3 depicts the reaction functions and the equilibrium in the last case with both welfare programs. The tax rate reaction functions have negative slopes, which mean that an increase in one equilibrium tax rate leads to a decrease in the other equilibrium tax rate.



Picture 2.3: Reaction functions and equilibrium with income redistribution and social security (Conde-Ruiz and Galasso, 2005).

An important insight of the proposition is the fact that, in order to understand the composition of the welfare state and the relation between it and labour income inequality, we must analyze the income redistribution by age groups. Age, instead of income, may be the most important determinant in some agents' voting decision. The income redistribution should first be separated into age groups and only then can the median voter be recomposed to account for income, as was depicted in picture 2.1 and 2.2.

In arriving at the proposition stated above, we have assumed commitment over the social security policy. This assumption will now be dropped. Proposition 2.2 describes the subgame perfect equilibrium outcome of the voting game on welfare tax rates without commitment.

Proposition 2.2 All pairs of outcomes (τ^*, σ^*), which represent structure-induced equilibrium outcomes of the voting game with commitment over the two welfare programs, are subgame perfect structure-induced equilibrium outcomes of the game without commitment.

The proof is provided in appendix 6, but the intuition will be explained here. It is obvious that the retirees support social security whereas the workers support the income redistribution scheme. In addition, workers with low ability, who would vote for a positive level of social security in the game with commitment, would similarly also be willing to enter an implicit contract between successive generations of voters in order to sustain positive social security, as this would bring her similar benefits. As an example, we consider the last case of proposition 2.1, with both tax rates positive. A very low ability worker would now prefer even more income redistribution and less social security, as the latter is an inefficient redistribution scheme. However, we know from section 2.1.2 that all agents have zero mass, and that they can't affect the election outcome. And even if they could, a low ability worker could not change it in the desired direction. Any agent, whose ability is lower than the median voter's ability, can decrease a tax rate by voting for a zero tax rate, thereby reducing the median tax rate. However, the agent is not able to increase the tax rate, as she is already voting for a higher tax than the median voter's tax rate, due to her preferred tax rate being above the median voter's. Therefore, she can't have an impact on

the voting game in the desired direction, and will instead support the implicit contract between generations. The same reasoning can be applied to other voters as well as to other cases of proposition 2.1.

We have now calculated the equilibrium outcomes of the voting game without commitment, as characterized by propositions 2.1 and 2.2. These are the results the model predicts for an economy with a given demography and a given level of income inequality. Before discussing the results, we will investigate what kind of changes the model predicts when the underlying parameters change.

2.1.4 Response to changes in demography and income inequality

To derive interesting real-world implications from the model, it is fruitful to analyze how the model responds to certain parameter changes. Conde-Ruiz and Galasso (2005) evaluate how the equilibrium tax rates change when the income inequality changes. In addition to this, we will analyze how the social security system of the model will change as a result of an ageing population. This is something the authors of the article do not focus on.

The intuition of the effect of a changing age structure is simple. A decrease in the population growth, μ, corresponds to the ageing population of the contemporary developed world. For a given income distribution, an increase of retirees relative to the working population leads to opposite results for the equilibrium tax rates of the median voter: a decrease in the income redistribution tax rate and an increase in the social security tax rate. The effect is the most intuitive by analyzing pictures 2.1 and 2.2. In picture 2.1, an increase in the amount of retirees makes the block of retirees larger and adjusts the median voter for income redistribution rightward. The median voter is now richer than before, preferring a lower income redistribution tax rate. Consequently, in picture 2.2 the adjustment of the median voter is leftward, with the median voter for social security having a lower income than before the parameter change. The result is a higher social security tax rate than before. The same effect can be depicted in picture 2.3: a larger block of retirees shifts $\tau_{m\tau}(\sigma)$ downward and $\sigma_{m\sigma}(\tau)$ upward. For every equilibrium social security tax rate, the equilibrium income redistribution tax rate is lower, and vice versa.

When analyzing the effect of a change in overall inequality, the analysis is less straightforward. Larger inequality means that the difference between the mean ability and the median voter increases. A relative reduction in the income redistribution median voter's ability increases the income redistribution tax rate and decreases the social security tax rate, shifting $\tau_{m\tau}(\sigma)$ up in picture 2.3. For the social security median voter, the effect on the tax rates is the opposite, and the $\sigma_{m\sigma}(\tau)$ also shifts upwards. Thus, the reaction functions move in the same direction. The effect can be shown analytically by taking the total derivative of the change in income inequality on the equilibrium tax rates, thereby decomposing the changes in the equilibrium tax rates into effects related to a change in the income redistribution median voter's relative ability ($d\Delta_\tau$) and in the social security median voter's relative ability ($d\Delta_\sigma$):

$$d\tau^* = \frac{\partial \tau^*}{\partial \Delta_\tau} d\Delta_\tau + \frac{\partial \tau^*}{\partial \Delta_\sigma} d\Delta_\sigma, \tag{2.24}$$

$$d\sigma^* = \frac{\partial \sigma^*}{\partial \Delta_\tau} d\Delta_\tau + \frac{\partial \sigma^*}{\partial \Delta_\sigma} d\Delta_\sigma. \tag{2.25}$$

In 2.24, the first term on the right-hand side is positive whereas the second is negative; the lower ability of the income redistribution median voter relative to the mean ability pushes the income redistribution tax rate up, whereas the lower relative ability of the social security median voter pushes the tax rate down. In 2.25, the first term on the right-hand side is negative whereas the second is positive, which leads to the opposite effect on the tax rates.

From the analytical analysis so far follows that the effect of a change in income inequality is ambiguous. Conde-Ruiz and Galasso (2005) extends the analysis to determine on what conditions the tax rates rise or decrease as a result of a change in income inequality. As the exact analytical calculations to determine this are extremely difficult, I will only explain the intuition here (see original article for calculations). The intuition follows directly from equations (2.24) and (2.25). If the change in the measure Δ_τ of the income redistribution voter's relative ability is sufficiently larger than the change in Δ_σ, then the income redistribution tax increases and the social security tax decreases. Equation (2.24) gets positive values and (2.25) negative values. The opposite effect occurs if the change in Δ_σ

is sufficiently larger than the change in Δ_τ; now (2.24) is negative and (2.25) positive. Finally, if the changes in Δ_τ and Δ_σ are equally large, both taxes increase; both (2.24) and (2.25) are positive. The analysis for decreasing income inequality is the reversed of the above.

We can now summarize the effects of the parameter changes. An increase in the share of retirees, keeping the income inequality constant, leads to a lower amount of income redistribution from rich to poor and to a higher amount of social security from workers to retirees. This is because of the uniformed voting behaviour of the retired, who support social security but not income redistribution. A larger share of retirees is able to affect the policy parameters directly through the median voter. On the other hand, when overall income inequality changes, the resulting changes in the size of the two welfare programs depend on how much the ability of the two median voters' change in relation to the mean ability workers. This is because the change in income inequality leads to opposite effects in the two dimensions of the voting game, and the relative effects of the changes decides if the welfare programs grow or diminish. In order to find a clearer link between income inequality and welfare programs, income inequality should be measured by age group. The retirees, regardless of ability, vote for social security and against income redistribution. However, an increase (decrease) in income inequality among workers only should result in an increase (decrease) in both welfare programs.

2.2 Results

The model of Conde-Ruiz and Galasso (2005) results in a positive level of social security if there are sufficient economic reasons for at least a majority of the population to support it. The population consisting of workers and retirees vote according to their preferences on two issues: the income redistribution tax, which is used to support low-income workers, and the social security tax, which is used to support retirees. In accordance with the median voter theorem, the implemented tax policy is the preferred policy of the median voter. Structure-induced equilibria guarantee that the median voter framework is applicable on the two voting dimensions separately. In a society with a sufficient amount of inequality and enough retirees, a welfare state consisting of an income redistribution scheme and a social security system arises as the result of the voting game.

The intuition of the result is straightforward. Because of the redistributive component in the social security system, low-income workers support both the income redistribution scheme and the social security system. On the other hand, the retirees support only the social security system, because the income redistribution system only reduces the tax base and hence their pensions. Because of the retirees' uniform voting behaviour, they are able to shape the two winning majorities. For social security, the majority consists of the retirees and the low-income workers; for income redistribution, the majority consists of low-income workers, with the decisive median voter being a worker with higher income than the social security median voter's. In other words, the uniform voting behaviour of the retirees creates a wedge between the abilities of the median voters, leading to an equilibrium with both welfare schemes. In the equilibrium, the social security system is larger than the income redistribution system.

Furthermore, in section 2.1.4 we concluded that an ageing population further enlarges this gap. An increase in the amount of retirees combined with the uniform voting of the elderly leads to the median voter of income redistribution to be richer and the median voter of social security to be poorer than before. As high-income types prefer less of both redistribution schemes, the result is a rise in the amount of social security and a drop in the level of income redistribution. On the other hand, the effect of an increase in overall income inequality is ambiguous, and depends on the relative effects of the two dimensions of the voting game. Instead of overall inequality, income inequality should be measured by age group.

The results lead to several interesting testable implications. Firstly, the model predicts that an ageing population should lead to more social security but less income redistribution. Second, the model specifies the relation between income inequality and welfare programs. Especially, the model suggests that larger overall income inequality does not necessarily lead to more social security and more income redistribution, as predicted by Tabellini (2000) and Meltzer and Richards (1981), but the income inequality should be measured by age group. The empirical evidence will be considered in the next section.

2.3 Empirical evidence

The ultimate test of an economic model is how well it explains empirical facts. In this section, we will consider the empirical evidence of the implications derived in the previous section. In particular, we focus on the link between an ageing population and social security as well as the link between income inequality and social security. The effects on the income redistribution program will not be considered.

In order to consider the effect of an ageing population on social security, we must refine our hypothesis to distinguish between a rise in total pension expenditure and a rise in pension expenditure per retiree. This is a crucial distinction; a rise in total pension expenditure without a rise in individual pensions may be seen as a justified consequence of population ageing. On the other hand, a rise in individual pensions points to evidence of the elderly being able to exert larger influence on policy decisions than before, thereby securing higher redistribution benefits per retiree. Therefore, we will consider the empirical evidence of population ageing on total pension expenditure and individual pensions separately.

First consider the effect of an ageing population on total social security expenditure. Various empirical studies have concluded that the total size of social security rises with an older population (Breyer and Craig, 1997; Tabellini, 2000; Disney, 2007; Shelton, 2008; Tepe and Vanhuysse, 2009). The empirical part of Razin, Sadka and Swagel (2002) draws the opposite conclusion. Their study argues that a higher dependency ratio, i.e. a higher ratio of people who are too young or too old to work divided by the working population, has a negative effect on total pension transfers. However, this study has been criticised by Disney (2007) and Tepe and Vanhuysse (2009) among others as incomplete, not taking problems of autocorrelation and non-stationarity into account. Hollanders and Koster (2011) find that pension spending is not significantly associated with a rise in the age of the median voter, but indeed is positively correlated with a rise in the dependency ratio. According to these empirical investigations, it seems to be the case that an older population leads to rising total pension expenditure.

When the dependant variable is pension expenditure per retiree, however, the picture becomes less clear. In earlier research, including Breyer and Craig (1997), the proportion of elderly has no significant effect on the expenditure per retiree. However, recent additions to the literature (Tepe and Vanhuysse, 2009; Hollanders and Koster, 2011), focusing explicitly on this question, find that the relationship is weakly negative: an older population leads to smaller benefits per retiree. What these results imply is that the conventional wisdom of political economy models that an older population leads to larger pension expenditure only holds on the aggregate level. The total amount of resources used for pensions grows with an ageing population, but the absolute level of pensions per retirees decreases. A theoretical explanation for this phenomenon has been presented by Razin, Sadka and Swagel (2002). They use a majority voting model to suggest that a rising dependency ratio leads to lower expected profitability of the PAYG system for the currently working population. An ageing population may cause current workers to expect smaller future pensions compared to the ones paid out today, resulting in current workers favouring smaller pensions already today.

If these empirical analyses are correct, they have some profound implications for some political economy models of social security. Firstly, political economy models of social security succeed in predicting that an ageing population leads to larger spending on retirement benefits. This feature is shared by most models, including the models of Conde-Ruiz and Galasso (2005) and Mulligan and Sala-i-Martin (1999) analyzed in detail in this paper. However, this is not sufficient to explain per retiree changes, as more people share a larger amount of resources. Second, models predicting a rise in the individual level of pensions as a result of an ageing population are likely to be wrong (e.g. Verbon and Verhoeven, 1992; Persson and Tabellini, 2000: p. 123-132). The recent empirical literature indicates that although the total amount of social security increases, the amount of retirees rises even faster, leading to lower pensions per retirees.

The evidence on the effects of income inequality on social security is mixed. Using different measures of income inequality, Perotti (1996) and Tabellini (2000) find that greater inequality in pre-tax income correlates with a larger social security system. On the other hand, Lindert (1996) and Breyer and Craig (1997) find that the relation between their

measures of income inequality and the overall size of social security is non-existent or only weakly statistically significant, respectively.

The empirical evidence does not seem to fully support the original hypothesis of Tabellini (2000), who predicted that larger overall income inequality leads to a larger social security system. However, the analysis of Conde-Ruiz and Galasso (2005) above may bring clarity to the picture. They predict that the impact of a change in overall income inequality on the size of social security is ambiguous. Instead, the income inequality of the working population is the relevant measure. Empirical studies measuring income inequality per age group and social security size has, to my knowledge, not been made.

The model of Conde-Ruiz and Galasso (2005) passes empirical tests quite well. An ageing population does lead to a larger social security system. However, when it comes to pensions per retiree, the effect is negative, as workers seem to retire faster than the amount of pensions rise. On the other hand, the impact of overall income inequality on the size of social security is unclear. An empirical study on the relation between income inequality per age group and the social security system is needed to judge the second hypothesis of the model.

2.4 Discussion

After analyzing the empirical implications of Conde-Ruiz and Galasso (2005), we will now discuss the model and its methodology. In what follows, we will first consider common objections to the use of economic methodology on political concepts. We will not discuss economic modelling in general, but focus on factors relevant for the analysis in this paper, starting from an overall discussion of public choice theory and then considering majority voting models in particular. After that, we will critically assess the model and main assumptions of Conde-Ruiz and Galasso (2005).

2.4.1 On the methodology

Applying economic methodology on issues originally considered outside the realm of economics has been labelled “economic imperialism” (Lazear, 2000). Specifically, public choice theory uses economic concepts including utility-maximizing behaviour, efficiency

and equilibria on the analysis of politics. According to Miller (1997), the introduction of economic constructs into political analysis has been beneficial in several ways. Most importantly, it has compelled political scientists to think in terms of formal, rigorous arguments, introduced political science to mathematical modelling as well as introduced various concepts that are today accepted as mainstream political science. Although it is an exaggeration to consider modern political analysis an extension of economic market analysis, the economic framework has helped to improve political analysis both methodologically and substantively. (Miller, 1997)

Criticism of the public choice approach has also been extensive. A large part of this is similar to the criticism of neoclassical economics in general, and that will not be discussed here. However, we will consider two objections of public choice, which are also relevant for our model of majority voting above. Firstly, it can be argued that a self-motivated, rational and utility-maximizing voter is an unrealistic assumption. One can also argue that the materialist approach of economics is insufficient. Especially in the case of voting and political activity, the framework does not take the consequences of ideology adequately into account. However, these arguments do not dismiss public choice theories. Microeconomic assumptions of rational behaviour do not have to hold for all people at all times. What is required is that agents are motivated by utility-maximization for a sufficient, undetermined fraction of their actions (Buchanan and Tullock, 1963: p. 29). If this is the case, public choice models have a positive value in predicting behaviour and should be evaluated according to standard empirical methods. Therefore, even if it is deemed unrealistic that all voters always are rational and self-interested agents, it is likely that their voting behaviour is determined at least to some extent by a utility function such as the one in the model presented in this chapter.

Second, perhaps the most prominent attack on public choice theory was made by political scientists Donald Green and Ian Shapiro in their book “Pathologies of Rational Choice Theory” (1994). Their main message is that the empirical results of the public choice theories have been weak, and that the branch has done little to promote the understanding of political activity. The debate regarding the fruitfulness of public choice theory continues to this day.

Moving the scope from general public choice theories to majority voting models, we encounter both advantages and disadvantages. The major advantage of majority voting models based on the median voter theorem is that they are simple and intuitive. They allow us to model and develop hypotheses of complicated problems, while maintaining the most important drivers behind the problems. Furthermore, it is often easy to draw conclusions and testable implications from median voter models. These useful characteristics have made the median voter model a benchmark for more complicated models of policy formation in democracies. (Congleton, 2003) However, the median voter models have some potential drawbacks: the way of aggregating preferences is unrealistic, they are applicable only on democracies, they ignore the political process and they often explain only a few variables.

Perhaps the strongest objection towards majority voting models is their unrealistic way of aggregating individual agents' preferences into policy decisions. In the case of a one-dimensional policy space, few societies actually vote on a single policy issue in a pair-wise vote to determine the Condorcet winner, as the median voter theorem projects (see appendix 1). Similarly, in a two-dimensional model like the one of Conde-Ruiz and Galasso (2005), few legislatures work in the way structure-induced equilibria postulates (see appendix 3). Representatives in most countries are seldom elected in general elections with majority rule, and policy issues are rarely voted on one issue at a time. Further, if voting is required for social security to arise, non-democracies should have substantially different social security policies. Mulligan, Gil and Sala-i-Martin (2010) show that this is not the case, as they, holding GDP and age structure constant, find no significant differences in social security policies between democracies and non-democracies.

To counter these arguments, Peltzman (1980) among others has argued that voting in a majority voting model should not be understood literally, but more as a metaphor. Voters do not necessarily explicitly vote, but the voting institution in the model conveys the opinion of the people and this contributes to the viability of certain policies. This view is supported by the conception that the opinions of elected politicians often reflect the opinions of the electorate and that politicians must take the opinion of the public into account. If the view is correct, the problem of unrealistic aggregation of preferences

disappears. Further, if one argues that voting in the majority voting models conveys the opinion of the people in non-democracies, and that non-democratic rulers are to some degree sensitive to popular support, the problem of similar social security programs in democracies and non-democracies disappears as well.

Majority voting models also do not model the political process itself, but assumes that the explicit or implicit voting result is implemented. The lack of a modelled political process makes the model unable to take into account the inefficiency of the political process (Krehbiel, 2004) and the fact that elected politicians also pursue other interests besides the public good or the will of the people (Downs, 1965). Although these are relevant concerns that must be kept in mind when analyzing majority voting models, the problem is double-edged. Including a political process into the framework would diminish the model's main asset, namely its simplicity.

Finally, a good model of social security should be able to explain many facts of social security. Most majority voting models are not able to do that (Mulligan and Sala-i-Martin, 2004). The models often focus and explain only a few of the characteristics present in social security systems throughout the world. This flaw is also present in the analysis of Conde-Ruiz and Galasso (2005), where the focus is on the support of the welfare programs and the implications of the age structure and income inequality on the welfare programs. However, it does not, for example, explain why there exists an induced retirement age or why social security expenditure as a share of GDP is correlated with economic growth.

Despite these theoretical drawbacks, majority voting models have gained empirical support. Poole and Daniels (1985) suggest that the assumption of single-peaked preferences crucial for the median voter theorem is not as unrealistic as it may seem. Further, Congleton and Shughart (1990) and Congleton and Bennett (1995) find that the median voter model explains the support for certain large public projects better than interest group models. In Ahmed and Greene (2000), interest group models are equally good at explaining public spending than the median voter model, although the latter has some advantages over other types of political economy models.

2.4.2 On the model

As the model of Conde-Ruiz and Galasso (2005) is a majority voting model with the properties discussed in the previous section, I will not repeat the advantages and drawbacks of such models here. The subject of this section is a critical discussion of the assumptions that are prerequisites for this model specifically: the single-mindedness of the retirees, the redistributive aspect of social security, the commitment assumption and the implications of differing political participation.

The uniform voting behaviour of the retirees is a crucial assumption that shapes the two winning coalitions on the two welfare programs, and the assumption is one of the main results of the model of Mulligan and Sala-i-Martin (1999) analyzed in detail in chapter 3. The authors argue that elderly voters are single-minded, i.e. they have similar political interests compared to workers whose interests may vary depending on the industry or occupation of individuals (see section 3.1.1 for a more detailed definition). Other researchers emphasizing single-mindedness and its implications for public spending include Peltzman (1980), Lehner and Widmaier (1983), Murrell (1984), Berry (1984) and Mulligan and Sala-i-Martin (2003). However, some researchers argue that this view is over-simplified and does not correspond to the behaviour of retirees in the contemporary environment of ageing populations. Firstly, older voters tend to be relatively loyal to their party identifications, not necessarily voting for the candidate or party that offers them the highest benefits (Goerres, 2009). Second, recent research argues that retirees in practice do not vote as a homogeneous group with a focus on the real value of their pensions (Goerres, 2009). Instead, it is other factors, such as health concerns and intangible assets including social and emotional capital that constitute the largest concern for the elderly (Tepe and Vanhuysse, 2009). These results undermine the assumption of single-mindedness among the elderly voters, and thereby the results of Conde-Ruiz and Galasso (2005). Further research is needed to determine the impact of these attacks on elderly single-mindedness.

The second key requirement for the winning majority of Conde-Ruiz and Galasso (2005) is the redistributive element of the social security system. Without this element, the winning coalition resulting in a positive level of social security would not arise. One measure of the redistributiveness of the pension system is the replacement rate, i.e. the share of a person's

working income she gets as a pension after retiring. If the replacement ratio is the same for everyone, the system is wholly Bismarckian and has no redistribution. On the other hand, if the replacement rate decreases with income, there is redistribution from rich to poor as high-income workers get a lower share of their income as pensions compared to low-income workers. In such a Beveridgean system, benefits are less linked to contributions. Figure 3 in Disney (2004: p. 291) show that all nine developed countries in the figure have differing replacement rates for high- and low-income workers, although to a differing degree. Another way to measure the redistributiveness of a system is by measuring the internal rate of return of investing in the social security system for workers with different income. Boskin et al. (1987) and Galasso (2002) find that the internal rate of return for low-income families is higher than middle- and high-income families. Therefore, the evidence of differing replacement rates and different internal rates of return for low-income workers supports the assumption that a redistributive component in the social security system facilitates the creation of a coalition between retirees and low-income workers.

This discussion raises a follow-up implication for Conde-Ruiz and Galasso (2005). The larger the redistributive component in social security is, the more effective social security is as a redistribution device and the stronger low-income workers support social security. This includes the median voter, which raises the social security level in society according to the mechanisms of the model. Therefore, the model predicts that a Beveridgean system of social security should be larger than Bismarckian ones. However, this is counterfactual, as the latter tend to be larger than the former (Casamatta, Cremer and Pestieau, 2000: p. 504). This is something that Conde-Ruiz and Galasso (2005), as well as other similar models such as Tabellini (2000), fail to explain. They also do not address why some systems have been designed to me more redistributive and others less.

In section 2, we discussed the importance of the notion of commitment in majority voting models. If current workers do not believe they are able to receive similar pension benefits as the current elderly receive, they may not want to support current pensions in the first place, leading to a breakdown of the PAYG social security system. In Conde-Ruiz and Galasso (2005), the final result has no commitment of policy and a condition for social security is that the young expect a pension transfer when they retire, as expressed in their

indirect utility function. An implicit contract between generations in the spirit of Hammond (1975) arises. As a way of modelling social security policy, the feature of no commitment can be criticized. Rose (1990) argues that public expenditure is less a result of the choices of the current government and more a result of inheritance of past government commitments, grounded in institutions and laws. According to Scharpf (2000), this is especially true for social security systems, where past contributions give current pensions a status of vested rights. The criticism implies a stronger form of commitment than the one of Conde-Ruiz and Galasso (2005). The critics argue that even if the implicit contract between generations was broken and the current young would not anymore support social security, actually changing the policy is more difficult than the model predicts.

A final point to be discussed is the political participation of different groups. If one takes the voting in majority voting models literally, as discussed in 2.4.1, the voting participation of cohorts of different age and income must be taken into account in the model of Conde-Ruiz and Galasso (2005). And even if voting is only a metaphor for public support of policies, participation in overall political activities has an impact on the model. Supporting conventional wisdom, there is substantial evidence on that voting participation rises with both higher age and higher income (e.g. Glenn and Grimes, 1968; Frey, 1971; Sigelman et al., 1985). Voting is often considered a proxy of overall political activity, which would imply that older and richer individuals are not only more active voters, but also more active in other political activities. However, Jennings and Markus (1988) find that voting participation of the elderly remains unchanged for above 65 year old individuals, but other more demanding modes of political activities, such as e.g. attending political meetings, declines for the mentioned cohort.

Although these results do not have qualitative implications for the model of Conde-Ruiz and Galasso (2005), they have certain quantitative consequences. For example, focusing on voting participation and not on more demanding political activities, higher relative voting participation of the retirees makes the block of retirees supporting social security and opposing income redistribution stronger, making the median voter of the social security poorer and the median voter of income redistribution richer. Similarly, lower relative voting participation of low-income individuals makes the block of low-income workers supporting

both social security and income redistribution weaker. The median voter of both programs becomes richer. Therefore, the aggregate effect of the voting behaviour described above on the equilibrium policy is negative for the income redistribution system, resulting in less redistribution from rich to poor workers. For the social security system, the aggregate effect is ambiguous and depends on the relative strength of the two effects described above.

2.5 Conclusion

In this chapter, we have described majority voting models of social security, where social security arises as the voting result when at least a majority of the population has economic reasons to support the system. As the elderly do not by themselves constitute a majority of the population, there must be some mechanism to induce the young and middle-aged to support the systems. Mechanisms surveyed here include dynamic inefficiency, reduced time horizon, crowding-out effects, within-cohort redistribution and altruism.

The two-dimensional majority voting model of Conde-Ruiz and Galasso (2005) models the social security system together with another component of the welfare state, an income redistribution system. They find that in equilibrium in a society with sufficient income inequality and a large enough share of elderly, a positive level of both social security and income redistribution arises. Furthermore, a larger share of elderly leads to more social security and less income redistribution, whereas the effect of a change in overall income inequality is ambiguous.

The empirical evidence on social security systems supports these predictions only partially. It seems to be true that an older population leads to more social security. However, models predicting larger individual pensions are likely to be false, as pensions per retiree seem to fall with a larger share of elderly. The contradictory evidence of the effect of income inequality on social security suggests that income inequality should be measured per age group instead of as an aggregate measure.

Finally, the methodology and the model assumptions were discussed. Using economic methodology on political science is still a contradictory theme, although it has been an established branch of economics for some time already. The advantages and drawbacks of

majority voting modelling were also discussed, emphasizing the value of the model as a theoretical benchmark. Furthermore, the two important assumptions of single-mindedness and the redistributive component of social security were scrutinized, as well as the commitment assumption and the impact of the varying political participation of different cohorts.

3 Interest group approach

Besides majority voting models, the other major political economy model family of social security are interest group models. This chapter will present the idea of the models and the major contributions of the literature. It will also analyze the interest group model of Mulligan and Sala-i-Martin (1999) in closer detail, and discuss the model from a theoretical and empirical perspective.

Whereas the premise of majority voting models is a democratic setting where citizens support a policy if they have sufficient economic reasons to do so, the interest group models focus on the political struggle and the relative power of certain groups compared to others. In the case of pensions, the elderly may constitute a group powerful enough to transfer resources from the working population, even if they would not constitute a majority of the population. Following Galasso and Profeta (2002), interest group models can be divided into ones based on influence functions and support functions, depending on if the differences in actions between groups are due to economic or political reasons. These model families will be shortly introduced below.

Becker (1983, 1985) was a seminal contributor to a model family emphasizing the competition between special groups for political influence. Although he did not analyze social security specifically, he argued that political pressure of special groups may interfere with the political equilibrium, impacting on how tax resources are directed. Applications on social security include Becker and Mulligan (2003), Profeta (2004) and Mulligan and Sala-i-Martin (1999), the latter which will be analyzed in more detail in the next section. In these models, young and old compete by undertaking political activities in order to exercise political pressure. The activities can be either time-intensive, as in Mulligan and Sala-i-

Martin (1999), good-intensive, i.e. dependant on resources spent on lobbying and other political contributions as in Becker and Mulligan (2003), or depend on the size of the group, as in Profeta (2004). The policy outcome, which in the case of pensions is an intergenerational transfer, depends on the pressure, size and other characteristics of each group. This is measured through an influence function. Finally, the group exerting more influence receives a transfer from the other group. It is worth to be noted that the influence functions of the groups are symmetric, i.e. all groups have the same fundamental political power. Therefore, in order for social security transfers to arise, there must be economic reasons for the old to exert more pressure than the young population.

On the other hand, in interest group models based on support functions, positive levels of social security arise as a result of a political process, where the government maximizes a political support function that includes the utility of currently living generations. Examples of support function models include Verbon and Verhoeven (1991, 1992) and Grossman and Helpman (1996, 1998). As in the influence function models, social security arises due to the old winning the political competition. However, contrary to the influence function models, the old are successful because of political reasons and not economic ones. For example, in Grossman and Helpman (1998), the fundamental political power is not equal across groups, as the old are assumed to be more efficiently organized in lobbies and therefore more efficient in exerting political pressure. Also, the old make contributions to political campaigns, leading to larger transfers from young to old.

We have now been able to identify the main characteristics of interest group models. The models argue that social security transfers arise from a political competition between the young and the old. The old are more successful in this competition either due to economic reasons, as in the influence function models, or due to political reasons, as in the support function models. In the next section, we will analyze the model by Mulligan and Sala-i-Martin (1999), which is an interest group model with an influence function.

3.1 An interest group model of social security

The offset of Mulligan and Sala-i-Martin's (1999) model is in the fact that the elderly have been politically very successful. The authors argue that the growth in social security spending described in section 1.1 cannot be explained solely by a changing demography. Also, the authors assert that the elderly have politically been more successful than many other minorities that could be argued to be in need of government assistance. A theory explaining the rise of social security should be able to account for why the elderly have been so successful while other minorities with low wages have not.

To explain these factors, Mulligan and Sala-i-Martin (1999) develop a model in which a redistributive program requires political, moral and social pressure to be exercised by the program's beneficiaries. Specifically, this pressure depends on the amount of non-working time enjoyed by the group's members. This is for two reasons: Firstly, it is obvious that groups with plenty of leisure time can devote more time to political activities. Second, it is likely that a group of agents without jobs is likely more homogeneous in its political preferences compared to a group with jobs. The citizens with jobs are probable to be from different industries with different political concerns, having more heterogeneous political preferences. In addition to this single-mindedness of the retired, the fact that most of the people currently working are likely to at some point retire further strengthens the power of the retired in the political competition.

The reasons for choosing the model of Mulligan and Sala-i-Martin (1999) are similar to the ones outlined for the other example model in section 2.1. The model is a suitable example of an interest group model of social security. Also, the single-mindedness result of the elderly in this model gives intuitive support for the voting behaviour of the elderly in the model outlined in chapter 2.

3.1.1 The model setup and the interpretation of time

The setup of the model of Mulligan and Sala-i-Martin (1999) builds upon the simple version of pressure groups with influence functions (Becker, 1983). However, the one important difference is that some of the resources needed to exert political pressure depend

on the time used for the activity, i.e. they are time-intensive. In this section, we will present the basics of the model and introduce interpretations of the concept of time central to the model. After that, we will define the political equilibrium and derive some important features of the equilibrium. Finally, we will discuss the model properties and its empirical predictions. The article is a discussion paper and has probably not gone through the process of peer review related to publishing in an academic journal. Despite this, the article is written by two renowned economists and is widely cited, and can therefore be considered to be of good quality. Furthermore, because of the lack of a peer review process, the text is somewhat long and cumbersome. In what follows, I will attempt to compress the article into a more compact format, without losing the essential parts of the model. Also, I use some differing notation for clarity.

Society consists of an equal number of identical young and identical old agents. Although we used a different notation in chapter 2, in this model it is more intuitive to keep the notations "young" and "old", as the old are not by definition retired. Superscripts describe the agents' age at the given period, indexed with $(i = y, o)$ for young and old individuals, respectively. Subscripts describe the time period of the variable, and a missing subscript is equivalent to the current time period t. Agents use their available time to work and for leisure activities. These activities, denoted by l, include participation in lobbying and other political activities. Of the total leisure activities, the political activities are assumed to constitute a fixed, and perhaps very small, fraction, which is independent of the tax rates. It is also assumed that political activities are normal, as opposed to inferior, goods with respect to the total leisure activities. This means that the level of political activity increases when the level of leisure activities increases, and vice versa. We also assume that agents do not take into consideration the effects of their political activities on other agents. Furthermore, as free-riding is an important element of the analysis, the leisure of an individual group i member is denoted by l^i and the group average by $\bar{l}^i$.

In the tradition of Becker (1983, 1985), the groups exert influence through a political influence function, which allows the two groups to get lump sum transfers from each other. Unlike Becker, in this model the size of the transfer is an increasing function with respect to the relative time used for leisure and political activities. We denote $f(\bar{l}^o, \bar{l}^y)$ as the

influence function, which is the main determinant of the fraction of potential GNP that is transferred from the young to the old (a negative transfer means that the net transfer is from old to young). The marginal product of elderly leisure is concave , i.e. $\partial f / \partial \bar{l}^{o} > 0$ and $\partial f^{2} / \partial^{2} \bar{l}^{o} < 0$. Higher leisure of the old increases the net transfer from young to old, but at a decreasing rate. Consequently, the marginal product of young leisure is convex, $\partial f / \partial \bar{l}^{y} < 0$ and $\partial f^{2} / \partial^{2} \bar{l}^{y} < 0$. Higher leisure of the young decreases the net transfer from young to old, but at a decreasing rate. If the groups use an equal amount of time to lobbying, the outcome of the influence function is 0, i.e. $f(\bar{l},\bar{l}) = 0$. Also, we assume that f is symmetric, meaning that switching the amount of leisure between the groups yields the opposite result for f, i.e. $f(a,b) = -f(b,a)$. If there are no social security taxes and transfers, young and old choose the same amount of leisure and work. This means that $l^{y} = l^{o} = \bar{l}$ and that the outcome of the influence function is 0, $f(\bar{l},\bar{l}) = 0$. From this and the symmetry of f, it follows that the groups have the same fundamental political power, i.e. the means by which the groups compete in the political landscape favour neither the old nor the young.

From the fact that the fundamental political power is the same between groups follows a key hypothesis that the model predicts. This hypothesis is that a social security with a higher political power for the old may arise, even when the underlying fundamental political power of the two groups is the same. In other words, even if the young and the old are identical in every respect except for their time horizon and value of time, a political equilibrium may involve an endogenous creation of a social security system, financed with labour income taxes and including compulsory retirement or retirement-inducing taxes.

Before describing the analytical side of the model, it is worthwhile to contemplate on the interpretation of the influence function, as it plays such an important role in the model. Mulligan and Sala-i-Martin (1999) identify five ways to interpret the function, of which two are the most important. The more important ones are what I will call selfishness and single-mindedness, and the less important ones are segregation, the neo-Marxist interpretation and endogenous political preferences. The different interpretations are shortly introduced below.

The most obvious interpretation of the influence function is that each group cares only about itself, and l^i is related to the resources devoted by each cohort to the political competition between groups. The more time is used for political activities, the more successful the group is in the political competition. This interpretation is what will be called selfishness.

The hypothesis has another less intuitive but, according to the authors, even more relevant and important interpretation. Assume that all agents have a fixed amount of political resources, e.g. time or goods, which they allocate between different political issues. The issue that gets the most aggregate resources is the most successful. In a contemporary specialized economy, workers often tend to focus on issues that affect and facilitate their own industry or occupation in relation to others, and may prefer to have such subsidies instead of a monetary government transfer. Also, political measures of different industries may, to a certain degree, cancel each other out. On the other hand, non-workers do not have such diverse interest and are largely united, or single-minded, in their political activities; their main goals are likely to be monetary transfers and medical care. From this follows that the inclusion of l^i in the influence function represents the amount of political single-mindedness in every group, and the assumption that political single-mindedness facilitates political success.

The three less important interpretations of the influence function are segregation, the neo-Marxist interpretation and endogenous political preferences. The segregation interpretation means that high levels of leisure may enable an age group to segregate itself in time and space, in order to reduce costs of political organization. Pratt (1976, chapter 4) argues that this has been one of the strengths of the retiree movement. The neo-Marxist interpretation asserts that government transfers occur for the purpose of pacifying beneficiaries who are upset because they are not able to get jobs. The argument has been applied on social security, and this literature is surveyed by Pampel and Williamson (1989). Finally, the argument of endogenous political preferences states that l^i should not only be interpreted as time doing politics but also in a more general sense as persuading other people not to oppose certain policies, as has been presented by Becker (1996), among others. This

interpretation of the influence function makes it possible that some young are induced to support policies that favour the elderly, which is certainly a realistic feature.

We have now accounted for the two main interpretations of the influence function relevant to this paper, selfishness and single-mindedness, and these will be discussed from an empirical standpoint in chapter 3.3. The three less important interpretations of the influence function are segregation, the neo-Marxist interpretation and endogenous political preferences. It is useful to maintain these metaphors throughout the presentation of the mathematical model in the next section.

3.1.2 The political equilibrium

The model of Mulligan and Sala-i-Martin (1999) is a standard two-period overlapping generations model. As defined earlier, there are two generations alive at any given time, and young and old agents make decisions on consumption, leisure and savings. Every generation is represented by an interest group, which maximizes its members' aggregate utility by extracting rents from the other group alive at the time through the influence function.

To be able to have an impact on the leisure and political effort of its members, interest groups can tax their members with a fixed-rate labour income tax, which is then distributed to the members as lump sum transfers. This means that the amount of tax paid grows with income, whereas the transfer back to the members does not. The existence of such a tax is not intuitive, and therefore requires further clarification. To assume that this tax is explicit is not a realistic assumption, as lobbies rarely have the ability to tax or gather large membership fees. However, in order to keep members from leaving the group and avoid the taxes, a powerful interest group may be able to use the government to enforce the taxation, making the tax an implicit tax on the members of the interest group. Although it appears as the government levies the taxes and pays the transfers, it may be the case that they do so largely due to the pressure of the interest group. In the model, the only role of the government is to enforce taxation and redistribution within and across cohorts.

It is also important to distinguish between taxes and transfers within the groups and taxes and transfers from one group to another through the influence function. The fixed rate taxes and the resulting tax proceeds used to pay out lump sum transfers mentioned here are assumed to be made only within the groups. On the other hand, the social security transfer resulting from the influence function is the only transfer between groups, and this transfer is financed with a lump sum tax on the other group.

The functionality of the model's tax system is best demonstrated by an example. Consider an old individual, subject to the old group's tax and transfer system (administered by the government) as well as to a transfer from the young group (assuming that the net transfer is from young to old). The agent pays a fixed-rate tax to the own group, receives a lump sum transfer from the own group and receives another lump sum transfer from the young group. As the government administers all taxes and transfers, bidirectional taxes and transfers may cancel each other out. The net transfer for an old agent is likely to be positive, and due to the fixed-rate tax also declining with growing income. In a similar way, a young individual is likely to end up with a negative net transfer, which also is the more negative, the higher the agent's income. The distorting effect of the tax is the interest groups' tool to affect the labour decisions of the agents, and by co-ordinating the members' actions through tax decisions the groups will be able to achieve higher utility for the own group as a whole.

The transfer system corresponds to a pay-as-you-go social security system with limited commitment. The limited commitment concept of the model lies in between full commitment and no commitment discussed in section 2. The old cannot guarantee that the young will be able to tax the next generation when today's young become old. Therefore, the main determinant of the olds' benefits is, in each period, the political pressure they are able to apply. This is supported by the fact that, despite precise benefit formulas of social security, benefit formulas often change over time and the benefits paid out are often only weakly related to taxes paid (Mulligan and Sala-i-Martin, 1999). However, the model assumes that government programs are "persistent". A social security system in place today makes it easier for tomorrow's young to be taxed by tomorrow's elderly. As the unborn generations (including tomorrow's young) are not able to exert political power, all generations currently alive should be in favour of a large level of persistence. This is

because in the case of a social security program with redistribution from old to young, all agents currently alive are at least as well off or better off than in the case with low persistence. The persistence of the government policy is denoted with the parameter $\rho > 0$.

To underline the limited commitment property of the model, the equilibrium concept chosen is a repeated static Nash equilibrium, where the only remnant of previous choices is the policy persistence. In any given period, groups and individuals make their choices, and the young and the old interest groups play a static Nash game to determine the level of redistribution. In the following period, today's old are dead, today's young have become old and a new generation of young has been born. Each period can be divided into three stages:

Stage 1. Interest groups "choose" labour income tax rates for their members. In doing this, they take into account the effect of taxes on political activities, on the utility of their members, and on the interest groups' decisions in the future. Actions of non-members and own past actions are taken as given.
Stage 2. Individuals choose current leisure and consumption. Current and future prices, taxes, and subsidies are taken as given.
Stage 3. Taking the amount of redistribution in the preceding period as given, the period's total leisure by interest group defines the redistribution from one group to another for that period. Holding the behaviour of interest groups constant in every period, redistributive transfers tend to be persistent over time.

We are now prepared to solve the three stages of the model with backward induction, starting from stage 3.

Stage 3
The aggregate leisure choices for both interest groups alive in any period, depicted by the group averages $\bar{l}^{y}$ and $\bar{l}^{o}$[7], determine the taxes and subsidies between the groups

[7] As the groups are of the same size, a higher average amount of leisure is equivalent to a higher aggregate amount of leisure. The words average and aggregate are in this paper used interchangeably, depending on what is more intuitive in the context.

for that period F_t according to the following function:

$$F_t = f(\bar{l}^o, \bar{l}^y) + \rho F_{t-1}, \quad p \in [0,1]. \tag{3.1}$$

Thus, the redistribution today is the sum of the outcome of the political competition today and a share (which depends on the level of persistence) of the redistribution in the previous period. f is assumed to have the properties introduced earlier. Notice the difference between capital F and lower case f; the former denotes the total transfer, whereas the latter depicts the outcome of the influence function in this period. When positive, F_t is a lump sum transfer to the old at time t, financed with a lump sum tax from the income of the young at time t, and vice versa. Furthermore, F is a share of potential GNP, and potential GNP is denoted by $\overline{w}$. We also assume that $\overline{w}$ is a homogeneous function of the labour productivities of the old and young, w^o and w^y. This means that if w^o or w^y is multiplied by a factor, then $\overline{w}$ is multiplied by some power of this factor.

After solving stage 3, we can combine stages 1 and 2. We first assume that the interest groups can dictate the actions of its members, and choose those actions to maximize the utility of the members. After that, we return to the important question of how the members' actions are co-ordinated by the interest groups.

Stages 1 and 2

We solve the stages by age group. The old interest group maximizes its members' utility by choosing aggregate consumption $\bar{c}^o$ and leisure $\bar{l}^o$, taking the wealth of its members a, the choices of the currently young $\bar{l}^y$ and the effect of past redistribution on current policy ρF_{t-1} as given:

$$v(\overline{w}\rho F_{t-1} + \overline{a}; w^o, \bar{l}^y) \equiv \max_{\bar{c}^o, \bar{l}^o} u^o(\bar{c}^o, \bar{l}^o)$$

$$\text{such that } \bar{c}^o = w^o(1 - \bar{l}^o) + \overline{w}\left[f(\bar{l}^o, \bar{l}^y) + \rho F_{t-1}\right] + \overline{a}\,. \tag{3.2}$$

The right-hand side of the identity depicts the utility maximization of the old with respect to their consumption and leisure choices. They are denoted with bars in order to emphasize that it is the interest group making decisions for all its members. The consumption constraint depends on the income from working, on the lump sum transfer from the young as a result of political competition, and on the wealth of the members. The function v

depicts the value of the redistribution program for the old. This is relevant for the young lobby's decision, as it anticipates the aging of its members. The value depends on the legacy left by the lobby representing its members in the previous period as young, i.e. the persisting part of the redistribution system, $\overline{w}\rho F_{t-1}$, and the assets accumulated, $\overline{a}$. Likewise, in stages 1 and 2 the young interest group maximizes its members' utility by choosing aggregate consumption $\bar{c}^{y}$ and leisure $\bar{l}^{y}$, taking the activities of the currently old $\bar{l}^{o}$ and of the future young $\bar{l}^{y}_{t+1}$ as well as the effect of past policy $\overline{w}\rho F_{t-1}$ as given:

$$\max_{\bar{c}^{y},\bar{l}^{y},\bar{a}_{t+1}} \left[u^{y}(\bar{c}^{y},\bar{l}^{y}) + \beta v(\overline{w}_{t+1}\rho F_{t} + \bar{a}_{t+1}; w^{o}_{t+1}, \bar{l}^{y}_{t+1})\right]$$

$$\text{such that } \bar{c}^{y} + R\bar{a}_{t+1} = w^{y}(1-\bar{l}^{y}) - \overline{w}\left[f(\bar{l}^{o},\bar{l}^{y}) + \rho F_{t-1}\right]. \qquad (3.3)$$

$\bar{a}_{t+1}$ depicts the assets the young plan to save for consumption as old. Intuitively, the young interest group maximizes its current utility plus the preference-discounted value of the future redistribution program, βv, subject to a constraint. According to the constraint, current consumption plus the discounted value of assets saved for old age equals the income from working as young less the lump sum taxes paid to the old. The discount factor is $R = 1/(1+r)$, where lending and borrowing occurs at the rate of interest, which is independent of policy choices. This interest rate regime can be interpreted as a small open economy taking r as given, and where lending and borrowing occurs with citizens living longer than two periods. Alternatively, it can be interpreted as a storage technology paying a rate of return r, although this adds the constraint that the amount of stored resources $\bar{a}_{t+1}$ must be non-negative. Non-negative storage corresponds to that agents cannot borrow against their future pension benefits. The constraint may or may not be binding in the model equilibria, depending on the parameter values.

We can now define the political equilibrium. For given values of $a, \bar{l}^{y}_{t+1}, w^{o}, w^{y}, w^{y}_{t+1}, F_{t-1}$, the vector $Y = (\bar{c}^{o}, \bar{l}^{o}, \bar{c}^{y}, \bar{l}^{y}, a_{t+1}, \bar{c}^{o}_{t+1}, \bar{l}^{o}_{t+1})$ constitutes a political equilibrium, such that:

$\bar{c}^{o}$ and $\bar{l}^{o}$ solve (3.2), given $\bar{l}^{y}, a, w^{o}, F_{t-1}$;

$\bar{c}^{y}$, $\bar{l}^{y}$ and a_{t+1} solve (3.3), given $\bar{l}^{o}, \bar{l}^{y}_{t+1}, w^{y}, w^{o}_{t+1}, F_{t-1}$;

$\bar{c}^{o}_{t+1}$ and $\bar{l}^{o}_{t+1}$ solve (3.2), given $\bar{l}^{y}_{t+1}, a_{t+1}, w^{o}_{t+1}, F_{t}$.

The equilibrium consists of the currently old solving their optimization problem (3.2), the currently young solving their optimization problem (3.3) and the next period's old solving their optimization problem (3.2). We assume that a political equilibrium as defined here exists.

Furthermore, it is useful to define a steady state equilibrium, where $\bar{l}^o$ and $\bar{l}^y$ are constant over time. For given values of g, β, R, ρ, the vector $Z = (\bar{c}^o, \bar{l}^o, \bar{c}^y, \bar{l}^y, a, F_{t-1})$ constitutes a steady state political equilibrium such that the following two conditions are fulfilled:

$$F_{t-1} = \frac{f(\bar{l}^o, \bar{l}^y)}{1-\rho};$$

for any $w^o > 0$ and given $a, \bar{l}^y, w^o, w^y = w^o(1+g), w^y = w^o(1+g), F_{t-1}$,

$(\bar{c}^o, \bar{l}^o, \bar{c}^y, \bar{l}^y, a(1+g), \bar{c}^o(1+g), \bar{l}^o)$ is a political equilibrium.

The first condition arises from simple algebra, by setting the result of the political competition to be equal in all periods, $F_t = F_{t-1}$, and inserting this into (3.1):

$$F_t = F_{t-1} = f(\bar{l}^o, \bar{l}^o) + \rho F_{t-1} \Leftrightarrow F_t = F_{t-1} = \frac{f(\bar{l}^o, \bar{l}^y)}{1-\rho}.$$

Thus, the steady state equilibrium level of government transfers, $F_t = F_{t-1}$, is larger than the outcome of the political struggle in the current period, $f(\bar{l}^o, \bar{l}^y)$, due to the persistence of the policy from the previous period. The second condition states that if the labour productivity of the old agents w^o is positive and the labour productivity of the young agents is $(1+g)w^o$, then, for the young, the corresponding equilibrium level of assets accumulated when growing old and the level of consumption as old is also larger by a factor of $(1+g)$.

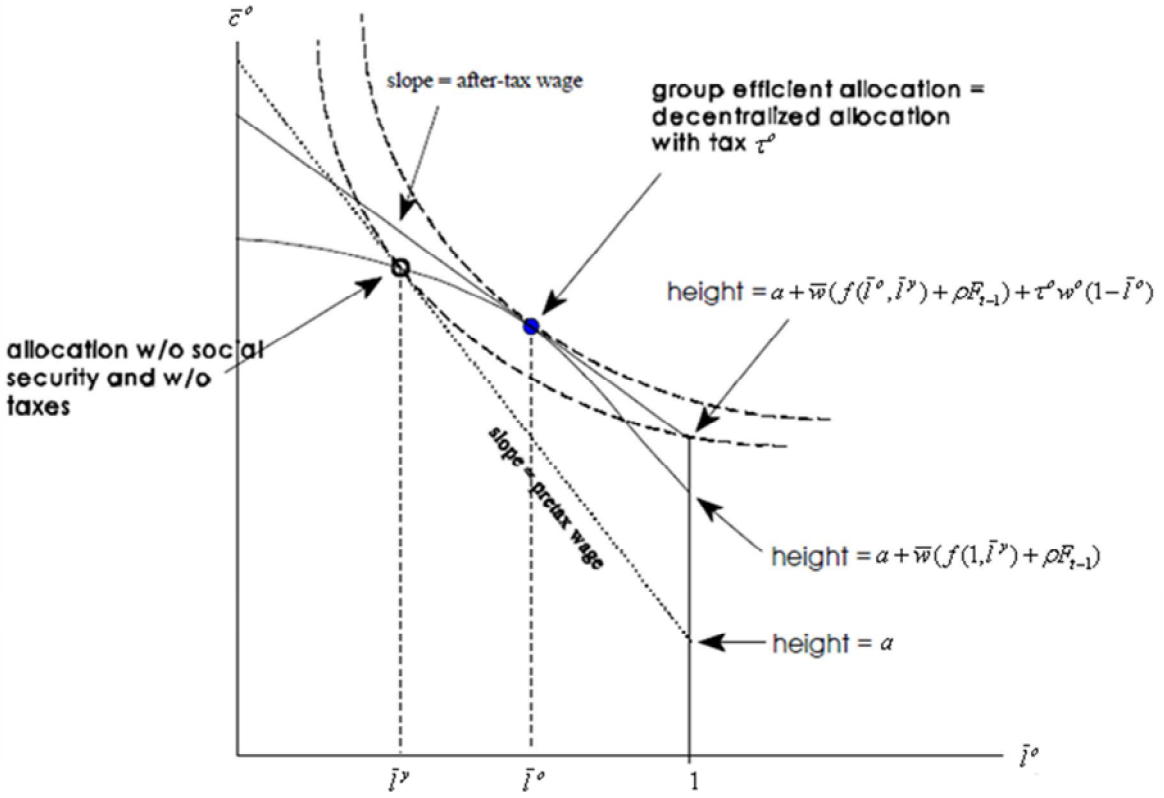
It should be noted that a steady state political equilibrium does not necessarily exist for all utility functions. The existence of this equilibrium will be verified in a special case in section 3.1.5.

3.1.3 Fiscal policy co-ordinating group activities

We will now consider the political equilibrium in combination with a set of taxes and transfers, used by the interest groups to co-ordinate the actions of their members. First, we derive the optimal choice of leisure and labour for the individuals in the absence of intra-group taxes and transfers. We then compare this outcome with the outcome that the groups would choose for its members in order to maximize the aggregate utility of the groups. Finally, we will analyze how the interest groups are able to achieve this result for the groups by introducing explicit or implicit taxes and transfers on their members.

3.1.3.1 The old interest group

First consider the old interest group, depicted in picture 3.1 below.



Picture 3.1: Old group consumption-leisure choices (Mulligan and Sala-i-Martin, 1999).

The solid curve represents the old agent's budget constraint $\bar{c}^o$ in equation (3.2). If leisure equals one, i.e. no time is used to work, the old group consumes only their accumulated

wealth a and the pension transfer, $\overline{w}\left[f(\bar{l}^{o},\bar{l}^{y})+\rho F_{t-1}\right]$, they are able to extract from the young group. The budget constraint is either concave, as the influence function was in section 3.1.1 defined to be concave with respect to $\bar{l}^{o}$, or linear, if the influence function has a value of zero. The dashed curves represent the old agent's indifference curves.

Now suppose that all old group members choose their amount of work freely. The budget constraint that the agents face is similar to the group's constraint in equation (3.2), except for that c and *l* are now the individual agent's decisions and not aggregates, and that they take the actions of all other agents, $(\bar{l}^{o},\bar{l}^{y})$, as given. This is equivalent to taking also the function $f(\bar{l}^{o},\bar{l}^{y})$ as given. The constraint is now:

$$c^{o}=w^{o}(1-l^{o})+\overline{w}\left[f(\bar{l}^{o},\bar{l}^{y})+\rho F_{t-1}\right]+a\,.$$

This is a linear constraint, depicted by the dotted line in picture 3.1. Assuming that the optimal choice is an interior solution (and not a corner solution), the individual agent's labour supply is given by the point that equals the marginal rate of substitution with the wage, $MRS_{\bar{l}^{o},\bar{c}^{o}}=w^{o}$, given by standard labour theory. This is the empty dot in picture 3.1.

We now move from the old individual to the old interest group, whose goal is to maximize the utility of the group as a whole. Still assuming an interior solution, the optimal point is the tangency point between the budget constraint and the indifference curve, denoted with the solid dot in picture 3.1. This is the optimal point from the point of view of the old group as a whole. From standard microeconomics and the implicit function theorem, we know that in this optimal point, the marginal rate of substitution (MRS) equals the ratio of the marginal utilities, which equals minus the slope of the indifference curve:

$$MRS_{\bar{l}^{o},\bar{c}^{o}}=\frac{\partial u^{o}/\partial\bar{l}^{o}}{\partial u^{o}/\partial\bar{c}^{o}}=-\frac{d\bar{c}^{o}}{d\bar{l}^{o}}=w^{o}-\overline{w}\frac{\partial f}{\partial\bar{l}^{o}}<w^{o}\,. \qquad (3.4)$$

In other words, the MRS equals the productivity of the elderly minus the change in redistribution related to a change in leisure.

The important point to note is that in the group optimal point (3.4), the marginal rate of substitution is not equal to the wage, but in fact lower than the wage. Holding constant the work levels of other groups, an old individual would prefer to substitute leisure in exchange

for his wage, as this would allow for a higher personal level of consumption and higher personal utility. This would continue until the marginal rate of substitution equals the wage. In other words, taking the other group members' amount of work as given, it is unrealistic to think that an old agent would accept the amount of work corresponding to the group optimum – he prefers to work more.

The reason for the different optimal choices is that whereas the individual agent takes the transfer from the young group as given, a group choosing for its members can influence this variable. By choosing less work and more leisure, the group earns less income but gets a larger redistribution transfer from the other group, resulting in higher aggregate utility and in a point on a higher indifference curve in picture 3.1. In other words, an agent's leisure has externalities on the other members of the group. However, without a co-ordinating device, there is no way the group can prevent its members from free-riding, with members choosing to work more and the final outcome corresponding to the one with individuals making their own decisions. This is a version of the classical game theoretical problem of the prisoner's dilemma; a co-ordinated choice would maximize aggregate utility, but all agents have a personal incentive to deviate from the co-ordinated choice, resulting in lower utility for all agents.

However, the old interest group is able to avoid the free-riding problem and achieve the group's optimal point in a decentralised way by introducing a flat-rate labour tax and paying out the tax proceeds as lump-sum transfers to its members. Denoting the rate of labour income taxation levied on the old group members with τ^o, the revenue associated with the tax is $\tau^o(1-\bar{l}^o)w^o$. Furthermore, the budget constraint for the old individual is now:

$$c^o = (1-\tau^o)w^o(1-l^o) + \tau^o w^o(1-\bar{l}^o) + \overline{w}\left[f(\bar{l}^o,\bar{l}^y) + \rho F_{t-1}\right] + a, \qquad (3.5)$$

where the first term is the labour tax paid, the second term is the group transfer, the third term is the social security transfer and the fourth term is the wealth of the agent. It is worth noting that the labour tax paid depends on the agent's own labour supply, whereas the transfers depend on the whole group's labour supply. Therefore, an increase (decrease) in the agent's own amount of work implies a decrease (increase) in the net amount of benefits achieved through the intra-group tax system. Equation (3.5) also assumes that, taking the

lump sum taxes and transfers between groups as given, no labour taxes paid by the elderly are enjoyed by the young. One example of a budget constraint of the form of (3.5) is depicted by the solid line in picture 3.1.

To calculate the optimal tax rate, we must equate the slope of the individual's budget constraint with a labour tax in equation (3.5) with the slope of the group budget constraint in the group optimal allocation of equation (3.2). This is equivalent to the former being tangent to the latter in the optimal point derived from the group's indifference curve, exemplified by the solid line in picture 3.1. We equate the slopes by setting the derivatives of the constraints with respect to leisure equal, and performing some simple arithmetic:

$$\frac{\partial c^o}{\partial l^o} = (1-\tau^o)(-w^o) = -w^o + \overline{w}\frac{\partial f}{\partial \bar{l}^o} = \frac{\partial \bar{c}^o}{\partial \bar{l}^o}$$

$$\Leftrightarrow \tau^o = \frac{\overline{w}}{w^o}\frac{\partial f}{\partial \bar{l}^o} > 0\,. \qquad (3.6)$$

The tax rate is positive in order to induce the members to work less. Thereby the group is able to prevent free-riding, internalize the externalities of leisure and co-ordinate the members' decisions to be equivalent to the group's optimal decision. It should be noted that the utility function u^o is not an important determinant of τ^o. The issue at hand is not how the labour supply responds to taxes, but how the individual incentives are aligned with the group's. This leads us to the model's first proposition.

Proposition 3.1 The political equilibrium includes an implicit or explicit positive rate of labour income tax for the elderly, τ^o, with the value given by equation (3.6).

From proposition 3.1 we can draw the important conclusion that the political competition between young and old leads to a distortion in the old group's labour supply. The old work less than they would in a world without political competition. This is the case even when cross-group taxes and transfers are of lump-sum fashion, as in this model, and not progressive. According to standard labour economics, progressive taxes would distort choices further as the marginal tax rate would rise with income.

3.1.3.2 The young interest group

In a similar way as for the old, we will analyze the choices of the young individuals and the fiscal policy of the young lobby. We will begin by deriving the actions of the young in the absence of fiscal policies implied by the interest group. However, in contrast to the old individuals, the young must make two decisions instead of one; the optimal choice between consumption in this period and savings used for consumption in the next period, as well as the optimal choice between consumption and leisure in this period. In the calculations below, we once again assume interior solutions.

The first condition determining the marginal rate of substitution between consumption in this period and in the next is calculated in appendix 7 to be:

$$\frac{\partial u^y / \partial \bar{c}^y}{\partial u^o_{t+1} / \partial \bar{c}^o_{t+1}} = \frac{\beta}{R}. \tag{3.7}$$

In the optimal point, the marginal utility of consumption in this period must equal the marginal utility of consumption in the next period multiplied by a factor of β / R. This means that the MRS is the higher, the higher the interest group values utility in the following period, β, and the lower the discount rate, R.

The intertemporal budget constraint facing a young agent is the constraints of (3.2) and (3.3) combined, taking the actions of others, $(\bar{l}^o, \bar{l}^y, \bar{l}^o_{t+1}, \bar{l}^y_{t+1})$, and hence also (F_t, F_{t+1}), as given (see appendix 7 for calculations):

$$c^y + Rc^o_{t+1} = w^y(1-l^y) - \overline{w}F_t + R\left[w^o_{t+1}(1-l^o_{t+1}) + \overline{w}_{t+1}F_{t+1}\right]. \tag{3.8}$$

Consumption today plus the discounted value of consumption tomorrow equals labour income today minus the redistribution payment today, plus the discounted value of the labour and redistribution income tomorrow. As for an old agent, the young agent's choice of labour supply equates his marginal rate of substitution between leisure and consumption in this period with his wage, $MRS_{\bar{l}^y, \bar{c}^y} = w^y$.

Moving from the individual to the group, the condition for optimality between leisure and consumption in this period is calculated as in (3.4) (see appendix 7 for further calculations):

$$MRS_{\bar{l}^y,\bar{c}^y} = \frac{\partial u^y / \partial \bar{l}^y}{\partial u^y / \partial \bar{c}^y} = -\frac{d\bar{c}^y}{d\bar{l}^y} = w^y + (1 - \rho R \frac{\overline{w}_{t+1}}{\overline{w}})\overline{w}\frac{\partial f}{\partial \bar{l}^y} < w^y . \quad (3.9)$$

Intuitively, this means that the MRS equals the productivity of the young plus the change in redistribution in this period related to a change in leisure minus the discounted and persistent value of the change in redistribution in the next period related to a change in leisure in this period. We will here assume that:

$$\rho R(\overline{w}_{t+1} / \overline{w}) < 1, \quad (3.10)$$

which is true when ρ is much less then 1, $\rho << 1$, or when the interest rate is greater than the economic growth rate, $R(\overline{w}_{t+1} / \overline{w}) < 1 \Leftrightarrow (\overline{w}_{t+1} / \overline{w}) < (1 + r)$. Under this assumption, the MRS in (3.9) is less than w^y, as $\partial f / \partial \bar{l}^y < 0$.

We note from equation (3.3) that the young interest group must take two considerations into account that the old must not. Firstly, the effect of its choices on future redistribution policy, and secondly, the effect of its choices on the actions of the interest group representing its members when they are old. The effects are in equation (3.3) embodied by the terms $\overline{w}_{t+1}\rho F_t$ and a_{t+1}, respectively, in the value function v. Accounting for the first effect means that the young interest group takes into account that fighting against redistribution today, i.e. increasing $\bar{l}^y$, decreases the redistribution in the future. In the optimality condition (3.9), this appears as the term $\rho R\overline{w}_{t+1}(\partial f / \partial \bar{l}^y)$. Accounting for the second effect means that increasing leisure, $\bar{l}^y$, in this period decreases savings used for consumption as old, $\bar{a}_{t+1}$. Ceteris paribus, this decreases old age leisure due to the income effect (assuming that leisure is a normal, as opposed to inferior, good), and consequently decreases old age redistribution. As seen in the young's maximization problem (3.3) and the definition of the value function v in (3.2), the young lobby anticipates that there will exist a lobby driving their interests as old in the following period, and that this lobby will discourage its members from working in order to drive the lobby's interests.

In the exact equivalent way as for the old group, we can draw a key conclusion. From the condition (3.9) we know that in the group's optimal point, MRS is lower than the wage. This means that in this point, the individual agent has an incentive to work more, and it is therefore not realistic to assume that the agent will accept the amount of work imposed by

the group. We arrive at a similar free-riding problem as for the old group, where all individuals have an incentive to deviate from the group's optimal decision, and we end up in the point that the individuals would have chosen without co-ordination.

Still following the same reasoning as before, the young group is able to co-ordinate and implement the optimal allocation in a decentralized manner by introducing an implicit or explicit fiscal policy. However, in contrast to the old group, the young group levies two taxes instead of one; a flat-rate tax on labour income and a flat-rate tax on savings. The total tax proceeds are then rebated back to the young group's members in a lump-sum fashion. The tax revenue associated with the young group's payroll tax rate, τ^y, is $\tau^y(1-\bar{l}^y)w^y$, and with the savings tax rate, σ, is $\sigma R\bar{a}_{t+1}$. By adding the within-group taxes and transfers together with the cross-group taxes and transfers, we get the budget constraint for a young individual subject to the young group's fiscal policy:

$$c^y + (1+\sigma)Rc^o_{t+1} = (1-\tau^y)w^y(1-l^y) - \sigma R\bar{a}_{t+1} + \tau^y w^y(1-\bar{l}^y) - \overline{w}F_t$$
$$+ (1+\sigma)R\left[(1-\tau^o_{t+1})w^o_{t+1}(1-l^o_{t+1}) + \tau^o_{t+1}w^o_{t+1}(1-\bar{l}^o_{t+1}) + \overline{w}_{t+1}F_{t+1}\right]. \qquad (3.11)$$

This equation is similar to (3.8), except for that taxes and transfers have been added. The left-hand side consists of consumption today plus the discounted value of tomorrow's consumption, which includes the effect of the savings tax in two ways; c^o_{t+1} is smaller because the tax decreases future assets, but the tax is transferred back to the members through the factor $(1+\sigma)$. Similarly, the right-hand side includes tax payments and transfers for both taxes in this period and the next. We assume that, besides the cross-group taxes and transfers, no labour or savings taxes paid by the young agents are enjoyed by the elderly. Also, as with the old group, bidirectional taxes and transfers may cancel each other out.

We calculate the optimal labour tax rate in the same way as for the old in equation (3.6), by equating the slope of the individual's budget constraint with a labour tax, (3.11), with the slope of the group budget constraint in the group optimal allocation, (3.3). This is made by setting the derivatives of the constraints with respect to leisure equal (see appendix 7 for detailed calculations):

$$\frac{\partial c^y}{\partial l^y} = (1-\tau^y)(-w^y) = -w^y - (1-\rho R\frac{\overline{w}_{t+1}}{\overline{w}})\overline{w}\frac{\partial f}{\partial \bar{l}^y} = \frac{\partial \bar{c}^y}{\partial \bar{l}^y}$$

$$\Leftrightarrow \tau^y = -\frac{\overline{w}}{w^y}(1-\rho R\frac{\overline{w}_{t+1}}{\overline{w}})\frac{\partial f}{\partial \bar{l}^y} > 0, \qquad (3.12)$$

as $\partial f / \partial \bar{l}^y < 0$ and assumption (3.10) holds. Once again, with a positive tax rate the group is able to prevent free-riding, internalize the externalities of leisure and co-ordinate the members' decisions to be equal to the group's optimal decision. The tax rate is different from the old group's in that it includes the effect of the tax on the redistribution policy in the next period. This is the second important result of the model, formally stated in the proposition below; political competition between generations discourages work by the young.

Proposition 3.2 If assumption (3.10) holds, the political equilibrium includes an implicit or explicit positive rate of labour income tax for the young, τ^y, with the value given by equation (3.12).

To infer the group's optimal level of savings tax, we must once again calculate the group's decision on consumption between this and the next period, and compare it to the result of the individual's optimal decision in (3.7). However, as calculations in appendix 7 show, the optimality condition for the group is the same as for the individual, i.e. equation (3.7). Therefore, the individual's decision coincides with the group's decision. The group does not have to impose a tax for the individual to behave as the group wants. From this follows the next proposition.

Proposition 3.3 The political equilibrium does not include a savings tax or subsidy.

Intuitively, there are two ways in which the young group's incentives may differ from the individuals'. Firstly, as has been shown above, without young labour income taxes, the young work too much. This can be corrected by introducing an optimally chosen labour income tax. Second, without old labour income taxes, a young person will plan to work too much as old. Similarly, this can be corrected by expecting that there exists an interest group when the young grow old, and that they impose an optimally chosen labour income tax as

well. Given the current and future optimally chosen labour taxes, group and individual incentives to save coincide and no savings tax or subsidy is needed.

Another way to interpret proposition 3.3 is that the transfers to the elderly are not means-tested, i.e. all members of the old group regardless of income or assets are eligible for a pension transfer. As earnings and savings are the only private sources of old age income, a savings tax in addition to a labour income tax is essentially the same as including a means test on old age benefits. The poor, who have smaller savings, pay less tax but get the same lump-sum transfer as individuals who pay larger savings taxes. The tax would encourage young age leisure due to the mentioned redistribution, and discourage old age leisure because individuals would rather work as old than young. This is an interesting prediction, as stronger means-testing is one proposed reform in the international pension debate.

To sum up the section, proposition 3.1 and 3.2 have concluded that in order for the interest groups to align the individual members' incentives with their own, both the young and the old interest groups levy fixed-rate taxes on their members and pay out the proceeds as lump-sum transfers. The tax discourages the members from working too much, and the taxes can be thought of as either explicit or implicit. Also, proposition 3.3 states that the young will not impose a savings tax, as the group's incentives to save coincides with the members'. In the next section, we will perform comparative statics on the tax rates in order to evaluate their properties.

3.1.4 Tax rate comparative statics

So far, we have concluded that in a political equilibrium, both interest group levy a labour income tax on their members. By making some simple further assumptions, we will be able to derive some of the properties of the tax rates, especially in relation to each other.

To be able to do this, we will assume that the influence function takes the following linear form:

$$f(\bar{l}^o, \bar{l}^y) = \phi \overline{w}(\bar{l}^o + \bar{l}^y), \tag{3.13}$$

where $\overline{w}$ is a homogeneous function of w^o and w^y, and $\phi < 1$ is a constant. By inserting this function into the optimal tax rates defined in equations (3.6) and (3.12), we can

determine the following relationship between the two tax rates (see appendix 7 for calculations):

$$\frac{\tau^o}{\tau^y} = \frac{w^y}{w^o} \frac{1}{(1 - \rho R \frac{\overline{w}_{t+1}}{\overline{w}})}. \tag{3.14}$$

From this relationship we are able to draw a number of conclusions, namely propositions 3.4, 3.5 and 3.6. The gap between the taxes τ^o / τ^y, i.e. how much higher the old group's tax is compared to the young group's, is the larger, the larger the difference in wages w^y / w^o, the larger the policy persistence ρ and the larger the growth in GNP $\overline{w}_{t+1} / \overline{w}$.

Proposition 3.4 If assumption (3.13) holds, a larger policy persistence ρ increases τ^o / τ^y.

Proposition 3.5 If assumptions (3.10) and (3.13) hold, higher economic growth, i.e. higher w^y / w^o and $\overline{w}_{t+1} / \overline{w}$, increases τ^o / τ^y.

In order to infer the relative size of the tax rates, we introduce two more assumptions:

$$\rho > 0, \; w^y \geq w^o. \tag{3.15}$$

Proposition 3.6 If assumptions (3.13) and (3.15) hold, then $\tau^o > \tau^y$.

Proposition 3.6 is a key result of the model. When policy is persistent and the young have a higher value of time, i.e. a higher wage, the old interest group will encourage its members to choose more leisure compared to the young lobby. This is consistent with the empirical notion that the elderly have less incentive to work than the young, for example due to retirement legislation.

To intuitively explain the fact that tax rates are inversely related to the pretax wage, we note that the cost of taxing group members and discouraging them from working depends on the members' wages. On the other hand, benefits are only weakly proportional to the members' wages, as they also depend on the redistribution transfer from the other group. In optimum, the groups balance marginal cost and marginal benefit of taxation. At any given

tax rate, the old have a smaller wage than the young, and therefore relatively low costs and high benefits compared to the young. This means that the old can set a higher tax rate to get the same marginal loss of labour income as the young.

As the wage of the young is larger than or equal to that of the elderly, $w^y \geq w^o$, the tax rate in political equilibrium for the old is larger than that of the young, $\tau^o > \tau^y$. This leads to the elderly choosing more leisure than the young, $\bar{l}^o > \bar{l}^y$, and this in turn results in an outcome of the political competition that redistributes from young to old, $f(\bar{l}^o, \bar{l}^y) > 0$. In other words, the political competition leads to the creation of a social security system, with relatively higher labour income taxes for the old compared to the young, providing an incentive for retirement. By restricting the form of the utility function, we will in the next section prove that this constitutes a steady state political equilibrium as defined in section 3.1.2.

3.1.5 Gerontocracy and retirement as an outcome of political competition

After determining some important properties of the labour tax rates, we will now show that a social security system with retirement and redistribution from young to old constitutes a political equilibrium. Mulligan and Sala-i-Martin (1999) mention a number of reasons why retirement, i.e. $l^y < l^o_{t+1}$, may be consistent with an individual's choice of savings and labour supply over a lifetime (see Weiss, 1972, for an analysis). Firstly, it may be that work is in present value terms more profitable as young, $w^y \geq Rw^o_{t+1}$. Second, even if work is equally profitable throughout life, agents may prefer to consume and work less during old age, i.e. $\beta > 1$. Third, agents may find work more painful or leisure more desirable as old. And finally, even without retirement, $l^y \geq l^o_{t+1}$, it may be that old choose to work less than the young because they have higher lifetime wealth.

To keep the focus on how time-intensive political competition between the age groups induces retirement, we abstract from the mentioned effects by introducing the following assumptions:

$$u^i(c,l) = \log c + \gamma \log l\,,\ \beta = R\,,\ w^y = w^o_{t+1},\ w^y \geq w^o\,. \tag{3.16}$$

Intuitively, the assumptions mean that the utility function is logarithmic and not dependent on age, the rate of time preference equals rate of interest, wages are constant throughout an agent's life, and wages are at least as high as the previous generation's wages, respectively. From this assumption, we can infer the following lemma:

Lemma If assumption (3.16) holds, a steady state political equilibrium exists, with a constant amount of leisure and a constant fraction of potential GNP transferred between the young and old. The leisure of the elderly relative to the young depends only on the relative tax rates, and the relation can be calculated according to:

$$\frac{l^o}{l^y} = \frac{1-\tau^y}{1-\tau^o},$$

if $l^o < 1$. As the proof of the lemma is technically very difficult, it is not presented in the appendix and we instead refer to the original article of Mulligan and Sala-i-Martin (1999).

In order to be able to derive the next proposition, we must impose a weak restriction on the second derivatives of the influence function:

$$\text{If } x < y, \text{ then } \frac{\partial f(x,y)}{\partial x} > \left[\frac{1}{1+g} - \rho R\right] \left|\frac{\partial f(x,y)}{\partial y}\right|. \tag{3.17}$$

The restriction is fulfilled e.g. when $\rho R(1+g) = 1$, as the right-hand side is 0 and $\partial f / \partial x > 0$. The restriction is used in the proof of proposition 3.7, which states the following:

Proposition 3.7 If assumptions (3.15), (3.16) and (3.17) hold, a steady state political equilibrium with retirement and social security redistribution from young to old exists:

$$\tau^o > \tau^y,\ l^y < l^o_{t+1} = l^o,\ f > 0.$$

A proof can be found in appendix 8. This is the most important result of the model; political competition between interest groups lead to a steady state political equilibrium with the characteristics of higher taxes and more leisure for the old compared to the young, resulting in induced retirement and redistribution from young to old. This holds even when the young and old have the same fundamental political power, and only differ in their time

horizon and value of time. Although the authors extend the analytical model further, the main result is the one above and our discussion will focus on this.

3.2 Results

We have now derived how the interest-group model of Mulligan and Sala-i-Martin (1999) results in a social security system. The model is an overlapping generations model, with two identical cohorts of young and old agents, each cohort represented by an interest group. Political competition is measured by the influence function, which decides how large transfers are made between the two groups. The influence function depends on the amount of leisure of the two groups. The more leisure the own group chooses in relation to the other group, the more time is available for political competition and the more resources are transferred to the own group from the other. An alternative, and possibly more relevant, metaphor for the influence function is that it measures the level of single-mindedness of the two groups. The more single-minded the group, the more efficient it is in its political activities and the more resources it is able to secure through the political competition.

In each period, three stages take place. In stage one, interest groups choose implicit tax rates for its members. In the second stage, individuals choose their amount of work and leisure. Finally, stage three determines the redistribution from one group to the other according to the influence function. As proposition 3.7 outlines, the direction of the redistributive transfer is from young to old, corresponding to a social security system seen in most parts of the contemporary world.

The key assumptions behind the social security system are that the elderly have lower labour productivity, and that the currently young expect to become old in the future. Further, the ability of interest groups to levy implicit taxes on its members distorts the members' labour choices, resulting in less work, more leisure and larger transfers from the other group. Because the increase in transfers from the other group offsets the fall in labour income, the taxes are welfare-enhancing for the own group, although not necessarily for the economy as a whole. Without distorting taxes as a coordinating device, interest groups would not be able to keep members from free-riding and from choosing to work more.

The main result of steady state social security from young to old, $F > 0$, relies on the result that the old have more leisure than the young, $l^y < l^o$. The empirical evidence of this result will be examined in the upcoming section.

3.3 Empirical evidence

Mulligan and Sala-i-Martin (1999) draws a large number of conclusions from their model. In this section, I have attempted to choose the most important empirical implications, with a focus on the main result and the most important interpretations of this. Recall from section 3.1.1 that there were various metaphors behind the interpretation of l^i. The two most important were that l^i measures the amount of time used for political activities, and that l^i is a measure of political single-mindedness.

The literal interpretation, labelled selfishness in section 3.1.1, stated that l^i is the amount of time devoted to political activities to further the goals of the own group. The empirical evidence on the political activity on the old and young seems to only partly agree with the result of the model. In the discussion of political participation in section 2.4.2, we noted that there is a broad empirical literature determining that political participation rises with age. The hypothesis is further supported by American time diary studies that find that the old use more time for reading, participating in organizations and other political activities (Robinson and Godbey, 1997: p. 192, 214). However, in section 2.4.2 we also noted that some studies have found that more demanding modes of political participation actually decline for above 65 year old individuals. Therefore, although it is likely that the elderly vote more often than younger cohorts, Mulligan and Sala-i-Martin (1999) focus on more time-consuming political activities, such as e.g. lobbying or campaigning. For such activities, it is unclear if the activity level actually rises or decreases with age, especially for retirees.

As retirement is closely related to the amount of leisure of the elderly, it should be included in the analysis. In Mulligan and Sala-i-Martin (1999), induced retirement is represented by a larger implicit tax rate of the elderly. The larger the implicit tax rate, the stronger the retirement incentives and the more leisure the elderly have available for political activities, which through the influence function leads to a higher share of GNP transferred as social

security from old to young. Therefore, a higher ratio of social security to GNP should correlate with higher retirement incentives. Incentives are in practice provided by, for example, a retirement age determining social security eligibility and early retirement schemes. The correlation is supported by empirical evidence (Samwick, 1997; Gruber and Wise, 1999: p. 1-35). Therefore, the model successfully predicts that more social security is associated with stronger retirement incentives.

When it comes to political single-mindedness, we noticed in section 2.4.2 that the evidence of elderly single-mindedness is ambiguous, as some researchers argue that the elderly have more diverse political goals than what Mulligan and Sala-i-Martin (1999) project. Although the focus in section 2.4.2 was on voting behaviour, it is logical that the voting preferences of the elderly are similar to the ones regarding other political activities. Supporting the theory of lower single-mindedness among the working population compared to the retirees, Robinson and Godbey (1997: p. 215) find that a person's employment is at least an as important factor of time use related to political activities as a person's age. Despite this, as concluded in section 2.4.2, the prediction of elderly single-mindedness remains only partially supported by empirical findings.

When comparing the two most important interpretations of the main result of Mulligan and Sala-i-Martin (1999) with empirical findings, the outcome is only partly satisfactory. Although the model correctly predicts that social security is correlated with strong retirement incentives, the two main metaphors of selfishness and single-mindedness are not fully supported. After analyzing the main results of the model, we will in the following section discuss the model and its most important assumptions.

3.4 Discussion

Following the outline of section 2.4, we will begin this section by providing a brief discussion of Mulligan and Sala-i-Martin (1999) from the perspective of methodology. Afterwards, we will assess the assumptions and properties of the model itself.

3.4.1 On the methodology

In the beginning of section 2.4.1, we discussed the methodology of public choice extensively. Regarding the advantages and disadvantages of using economic methodology on political processes, we refer to the discussion there. In this section, we will focus on interest-group modelling. Noting that the public choice approach is the same as for majority voting models, the model families also differ in various ways, especially related to their relation to democracy and their explanatory power.

Interest-group models with influence functions, including Mulligan and Sala-i-Martin (1999), emphasize the pressure on the political system derived from economic differences (in e.g. wages), and de-emphasizes the mechanisms by which this pressure evolves into policy outcomes. This can be seen as a weakness, as the models do not have much to say about the explicit political process. However, it can also be regarded as one of the main strengths of the model, as it does not regard a certain political system as a prerequisite for social security. For example, this is consistent with the evidence on that there are no systematic differences in social security policy between democracies and non-democracies (Mulligan, Gil and Sala-i-Martin, 2010).

Furthermore, another advantage of interest-group models is that they are often able to explain more properties of existing social security systems compared to majority voting models (Mulligan and Sala-i-Martin, 2004). This is also true for our example model analyzed above. For example, the model is able to explain why there exists an induced retirement age and why economic growth is correlated with higher social security as a share of GDP (Mulligan and Sala-i-Martin, 1999: p. 43-44).

As with majority voting models, it can be argued that the setup of interest-group models itself is unrealistic. There are sometimes, but far from always, explicit groups driving the interest of similar-minded individuals in society. In the case of social security, whereas there may be e.g. retirement associations and trade unions defending the rights of retirees, there are few explicit organizations working for the rights of current workers towards current pension liabilities. However, when one considers implicit interest-groups instead of explicit organizations, the model regains its realism. Current workers do take into account

their pension liabilities in their political activities, and thereby form an implicit interest-group promoting similar interests, even if they are not explicitly organized. Following this reasoning, explicit or implicit interest-groups competing for resources may not seem that unrealistic after all.

3.4.2 On the model

Changing focus from overall interest-group models to the model of Mulligan and Sala-i-Martin (1999) examined in detail in this chapter, we will next discuss the assumptions and properties of this model. These include the two main assumptions of the model, i.e. that the young become old and the lower productivity of the elderly, as well as the strength, size and taxation ability of the interest groups and the commitment assumption.

The two main assumptions behind the result of Mulligan and Sala-i-Martin (1999) are that young will be old in the future, and that the elderly have lower productivity than the young. The former is needed to induce the young to support social security to some extent, as the persistence of the program secures a part of their future pension benefits. The assumption is quite obvious to be true, and is one of the reasons that the elderly are so successful compared to other interest groups and minorities. It is less likely for a person to, for example, change gender or race compared to the person growing old in the future.

The second main assumption states that the old have lower productivity and thereby lower wage, leading to their opportunity cost of not working being smaller for the elderly. This results in an equilibrium with a higher implicit tax, higher leisure and larger transfer from the young. The plausibility of the argument of lower productivity of the elderly has been surveyed by Skirbekk (2008), who concludes that most studies predict that productivity peaks as middle-aged and falls thereafter. Therefore, it seems that the assumptions of Mulligan and Sala-i-Martin (1999) are supported by the empirical evidence.

The interest groups of the old and young have a key role in the model. In the previous section 3.4.1 we noted that the interest groups do not necessarily have to be thought of as explicit organizations. However, as Mulligan and Sala-i-Martin (1999) assume that each generation is represented by an interest group, the relative strength of the interest group

should have an impact on social security policy in the real world. Pratt (1976) and Costa (1998) have suggested that pressure groups of the elderly have been important for the emergence of social security in the US. Furthermore, Clague, Balraj and Kramer (1971) argue that American trade unions have been important lobbyists for social security benefits. In Europe where unions are often more powerful than in the US, the unions are likely to have an even stronger impact on pension benefits. In contrast, there are no organized groups fighting explicitly for the working population against their pension liabilities. This imbalance is likely to strengthen the power of the elderly lobby further.

The relative strength of the interest groups is closely related to the taxing ability of the interest groups. Recall from section 3.1.2 that interest groups use the government to tax its members in order to incentivize them to work less and thereby get larger transfers from the other group. The more powerful group, the elderly, is likely to have better means for accomplishing this compared to the less powerful group, the young.

Mulligan and Sala-i-Martin (1999) also do not incorporate interest group size in their model. This is done in an interest group model by Profeta (2004), who arrives at the result that an increase in the proportion of elderly increases total social security. However, if the pressure of the group has decreasing marginal returns to size, the per retiree benefits may decrease. The empirical evidence of a change in demography on total and individual social security was described in section 2.3.

Finally, we shall briefly comment on the commitment assumption of policies. In section 2.4.2, we discussed the assumption of no commitment and its limitations. Mulligan and Sala-i-Martin (1999) use a concept of limited commitment, described in section 3.1.2. Limited commitment with a persistence parameter seems to be a good compromise between no commitment and full commitment. It respects each generation’s ability to form its own policies, but realizes that current policies are often partly the result of previous policies.

3.5 Conclusion

Whereas in the majority voting models social security arises when more than half of the population has economic reasons to support it, interest group models focus explicitly on the

economic or political resources available for different groups in society. These resources are then used to influence the political decisions. The models can be divided into ones based on influence functions, where the groups have economic differences, and ones based on support functions, where the old group is more powerful due to differences in political power.

In our influence function-based example model of Mulligan and Sala-i-Martin (1999), the old interest group is able to implement an intergenerational redistribution policy as the result of time-intensive political competition. Due to the high likelihood of the young growing old in the future and the assumption of lower productivity of the old, the old interest group induces the elderly to work less, uses more resources for political competition and thereby gets a transfer from the young generation through the influence function. However, the literal interpretation of the result that the elderly use more time for political activities gets only partial support from empirical investigations. The same is true for the interpretation of single-mindedness, as some research suggests that the elderly actually have more diverse political goals than what the model predicts.

Finally, the most important finding in the discussion of the methodology of interest group models was that the models are able to predict the existence of social security systems regardless of the political institutions in place in society. Furthermore, the assumptions and the mechanism of the model of Mulligan and Sala-i-Martin (1999) were discussed critically.

4 Implications for social security reform

As presented in the introduction, the population of the developed world is ageing rapidly. The normative message of the research on the financial sustainability of the social security system has been that reform is needed to keep revenues and expenditures in balance. This has spurred a debate in both the public and academia regarding the future of the pension system. In this section, we shall review the implications of majority voting and interest group models on reforming social security, as well as the implications of the specific models reviewed in detail above. Maintaining the perspective of the thesis, we analyze

these questions from the viewpoint of political support and political sustainability and de-emphasize the economic and financial aspects of the system.

Reforms of PAYG social security can be done in two ways. A structural change would change the structure of the system, either to a fully funded system, where current pension payments are saved to be paid out later, or to a mixed system, with both a PAYG component and saved pension payments. In practice, many systems are already today mixed to some degree, although the PAYG component is often dominant. The other alternative is parametric reform, where the parameters related to the pension system are changed. For example, this could mean raising the retirement age or reducing current pension entitlements.

Majority voting models have been especially gloomy on the political feasibility of social security reform. This is because in majority voting models there is a direct link between demographic change and the most preferred policy of the median voter. As the median voter grows older due to population ageing, the policy result of the median voter theorem is closer to the preferences of the old cohort. This development makes the opposition towards social security reform even stronger, as reform often includes transition costs imposed on the currently voting generations. Discussing Germany, Sinn and Uebelmesser (2003) conclude that the year 2016 will be "Germany's last chance for a partial transition to a funded pension system. Thereafter, the country will effectively be a gerontocracy". After that year, the median age is higher than the age of indifference for a gradual structural change to a mixed system, meaning that more than half of the population supports the status quo compared to the reform. According to the IMF (2004: p. 165), more than half of the population of a range of developed countries will be 50 years of age or older by 2019, adjusted for voting participation. After this, reform will be difficult. In Galasso and Profeta (2004), microeconomic simulations based on a majority voting model predicts that the political influence of the elderly will increase with the ageing population, and that the most effective way to mitigate the increase in the size of the system is by increasing the effective retirement age. Bütler (2000) uses a majority voting model and simulations on the Swiss economy to assess the relative political feasibility of different parametric reforms of the social security system. The author concludes that although reforming social security may be

difficult, a sufficiently low internal rate of return of social security compared to the real interest rate and a sufficiently high distortion from social security taxation could induce the population to favour parametric reform. This would imply that although most majority voting models argue that reform is very difficult, reform is nevertheless possible when the system becomes sufficiently unprofitable and distortive.

Conde-Ruiz and Galasso (2005) are able to contribute to the debate on reform by emphasizing the redistributive component of social security. If, for example, structural or parametric reform is proposed where the redistributive component is reduced, the reform may be opposed by low-income workers. This could be the case even if the total size of social security would rise.

The literature applying interest group models on social security reform is relatively scarce. However, the emphasis of interest group models on the political power of the elderly would suggest that potential for reform would decrease with an ageing population. In particular, reforms that the elderly oppose are less likely to be carried out when the elderly lobby increases in power. Meijdam and Verbon (1996) arrives at the result that an anticipated decrease in the population growth rate increases the social security tax rate, in accordance with the elderly power argument. Similarly as was reviewed in section 3.4, Profeta (2004) predicts that a larger amount of retirees increases the size of social security, although decreasing marginal returns to elderly group size may decrease individual pension size.

The model of Mulligan and Sala-i-Martin (1999) also has implications for social security reform. Firstly, the model argues that the higher the persistence of a program, the more likely it is to be successful. Higher persistence means that the young are more likely to achieve similar benefits in the next period as they have to pay in this period. Second, increased life expectancy and a reduction of the retirement age should also increase the likelihood of success for a reform. A longer life expectancy or lower retirement age is similar to raising the probability of a young individual becoming a member of the old group in the future (captured by the persistence parameter, ρ). This induces young agents to support the program. Finally, the authors argue, a forced savings program, where the young are forced to save through a payroll tax and the savings are paid out at old age, may

increase intergenerational redistribution, i.e. the PAYG system, further. The savings payment at old age increases the leisure of the old because of the extra income, and decreases the leisure of the young because of the additional tax on income. This relative increase in leisure in favour of the old increases the intergenerational redistribution through the influence function. The argument assumes that the young are not able to borrow against old age pensions, or otherwise the forced savings program could be negated. For a complete discussion of this final argument, see Mulligan and Sala-i-Martin (1999).

Political economy models have a pessimistic view on the sustainability of the social security system and on the potential for reforming the system. However, as we concluded in the empirical part in section 2.3, recent research suggests that only the aggregate size of the pension system has risen, whereas individual pensions have diminished. These results imply that the actual development may be less determined by growing elderly power, and that the gloomy predictions of political economy models surveyed here may be exaggerated.

5 Conclusion

The political economy of social security analyzes the political support and feasibility of the social security system. By analyzing the two main modelling families, majority voting and interest group models, and examining one model of each in detail, we have attempted to understand why a redistributive system like the PAYG system is so popular and why it has grown to be the largest welfare program in developed economies.

Majority voting models link the rise of social security with demographic development. In these models, social security arises when a majority of the population has economic reason to support the system. Preferences of the individuals are aggregated through a voting game, and the policy of the median voter becomes the implemented policy. As the retirees are not able to form a majority by themselves, various other mechanisms have been proposed for social security to arise. These mechanisms are surveyed in section 2.

Our example model of Conde-Ruiz and Galasso (2005) analyzes a two-dimensional majority voting model with an income redistribution system and a social security system. Using the concept of structure-induced equilibria, a welfare state with both systems arise in a society with a sufficient amount of income inequality and a large enough share of elderly. The social security system is supported by a coalition of retirees and low-income workers; the latter support the system due to the redistributive component in the social security system. The main asset of the model is in how it links the size of social security not only with demographic considerations, but also with the level of income inequality. The authors argue that income inequality in the working population should have an impact on social security, and not the overall income inequality as previous models have suggested. Furthermore, income inequality has implications for social security reform, as reforms reducing the redistributive component of the system may be opposed by low-income workers.

In the discussion of majority voting models, we noted that using economic methodology and modelling on issues traditionally in the scope of political science is still a controversial concept, especially among political scientists. Majority voting models are a prime example of the advantages and drawbacks of such models. They are simple and intuitive models and serve as theoretical benchmarks for further analysis. On the other hand, as with most economic models, they are simplified versions of reality and often contain unrealistic assumptions of the political process.

The median voter and demographics alone, however, is not by itself able to explain the whole magnitude of the increase in social security expenditure. Interest group models de-emphasize voting and focus on the relative power of different groups in society. Social security arises as the result of political competition between the old and the young. In models based on influence functions, the underlying difference between the cohorts leading to elderly power is economic, and in models based on support functions the difference is political.

In our second example model of Mulligan and Sala-i-Martin (1999), social security arises as the result of time-intensive competition. Assuming that the young eventually grow old and the elderly have lower productivity than the young, the elderly have a smaller

opportunity cost of leisure and is able to use more time for political activities, resulting in social security transfers from young to old. A potentially more relevant interpretation of the result of more time used for political activities is that the elderly are more single-minded, i.e. they have less diverse political goals compared to the working population. However, the empirical section of the chapter challenges this interpretation, as there is some evidence on that the old generation's political goals may be more diverse than what the model suggests. In addition to the political support of pensions, a relevant topic for contemporary social security policy is what consequences the ageing population will have on the social security system and what options are available for reform. The political economy models have identified two opposite effects of ageing on social security. On the one hand, the decreasing support ratio reduces the profitability of the system making social security less appealing, especially for younger generations. On the other hand, the rising age of the median voter as well as the increasing power of the elderly creates more support for a large social security system.

Theoretical and simulated models of social security have mainly emphasized the latter effect, linking the historical rise in social security expenditure with the rise in political support. Consequently, most models of social security reform predict that reforming social security will be difficult. As reform often includes reducing the benefits of the growing old generation, it is challenging to obtain the majority's support for such a reform. Nevertheless, some research suggests that when the system is unprofitable enough and the social security taxes are sufficiently distortive, reform will be possible.

The empirical literature on the size of social security is slightly more optimistic, suggesting that reform and change are plausible options. Although the evidence states that the size of total social security has grown as a result of the ageing population, recent empirical research suggests that average individual, per retiree pensions have diminished. This implies that the elderly have not been powerful enough to raise individual pensions, and that there is a limit to the argument of elderly power as a result of an ageing population. The working population has had to carry the burden of financing a larger PAYG system, but the retirees have had to settle for smaller individual pensions.

The political economy of social security is a highly relevant topic, both as a subject for policy recommendations and for academic research. There are various ways in which the literature could be further complemented, of which I shall list a few here. First, in order to test the predictions of Conde-Ruiz and Galasso (2005), one would need an empirical study on how the income inequality of the working population affects social security. To date, empirical papers have focused on overall inequality of the population. Second, in chapter 2.4.2 the assumption of single-mindedness, which is closely related to both models analyzed in detail, was questioned. A fruitful research topic would be to include other factors besides pensions that the elderly evidently value and use, such as health care provision, into the analysis. This could potentially be analyzed in the framework of a multidimensional majority voting model. Third, as interest group power is likely to be dependent on the size of the group, endogenizing group size into the model of Mulligan and Sala-i-Martin (1999) could reveal interesting effects of group size on the other variables and vice versa. Finally, if the empirical result of larger total social security and smaller individual pensions is accurate, many models of social security reform have been excessively pessimistic on the scope for reform. More research is needed to get more accurate predictions in this area.

References

Aaron, H., 1966: The social insurance paradox. *Canadian Journal of Economics and Political Science*, 32, 371-374.

Abel, A.B., 1989: Assessing dynamic efficiency: theory and evidence. *Review of Economic Studies*, 56, 1-19.

Ahmed, S., Greene, K.V., 2000: Is the median voter a clear-cut winner?: Comparing the median voter theory and competing theories in explaining local government spending. *Public Choice*, 105, 207-230.

Becker, G.S., 1983: A theory of competition among pressure groups for political influence. *The Quarterly Journal of Economics*, 98, 371-400.

Becker, G.S., 1985: Public policies, pressure groups and deadweight costs. *Journal of Public Economics*, 28, 329-347.

Becker, G.S., 1996: Accounting for Tastes. 1st ed. Harvard University Press, Cambridge, Massachusetts, US.

Becker, G.S., Mulligan, C., 2003: Deadweight costs and the size of government. *Journal of Law and Economics*, 46, 293-340.

Berry, J.M., 1984: The Interest Group Society. 1st ed. Little, Brown, Boston, US.

Black, D., 1948: On the rationale of group decision making. *Journal of Political Economy*, 56, 23-34.

Boadway, R.W., Wildasin, D.E., 1989: A median voter model of social security. *International Economic Review*, 30, 307-328.

Boldrin, M., Rustichini, A., 2000: Political equilibria with social security. *Review of Economic Dynamics*, 3, 41-78.

Boskin, M.J., Kotlikoff, L.J., Puffert, D.J., Shoven, J.B., 1987: Social security: a financial appraisal across and within generations. *National Tax Journal*, 40, 19-34.

Breyer, F., Craig, B., 1997: Voting on social security: evidence from OECD countries. *European Journal of Political Economy*, 13, 705-724.

Browning, E.K., 1975: Why the social insurance budget is too large in a democracy. *Economic Inquiry*, 13, 373-388.

Buchanan, J.M., Tullock, G., 1962: The Calculus of Consent: Logical Foundations of Constitutional Democracy. 1st ed. University of Michigan Press, Ann Arbor, Michigan, US.

Bütler, M., 2000: The political feasibility of pension reform options: the case of Switzerland. *Journal of Public Economics*, 75, 389-416.

Casamatta, G., Cremer, H., Pestieau, P., 2000: The political economy of social security. *Scandinavian Journal of Economics*, 102, 503-522.

Clague, E., Balraj, P., Kramer, L., 1971: The Aging Worker and the Union. 1st ed. Praeger Publishers, New York, US.

Conde-Ruiz, J.I., Galasso, V., 2003a: Early retirement. *Review of Economic Dynamics*, 6, 12-36.

Conde-Ruiz, J.I., Galasso, V., 2003b: Positive arithmetic of the welfare state. FEDEA working paper, Documento de trabajo 2003-04. http://www.fedea.es/pub/Papers/2003/dt2003-04.pdf. Date: 26.4.2012.

Conde-Ruiz, J.I., Galasso, V., 2005: Positive arithmetic of the welfare state. *Journal of Public Economics,* 89, 933-955.

Congleton, R.D., 2003: The Median Voter Model. In C.K. Rowley, F. Schneider: The Encyclopedia of Public Choice 2, p. 707-712. 1st ed. Springer, US.

Congleton, R.D., Bennett, R.W., 1995: On the Political Economy of State Highway Expenditures: Some Evidence of the Relative Performance of Alternative Public Choice Models. *Public Choice*, 84, 1-24.

Congleton, R.D., Shughart, W.F., 1990: The Growth of Social Security: Electoral Push or Political Pull?. *Economic Inquiry*, 28, 109-132.

Cooley, T.F., Soares, J., 1999: A positive theory of social security based on reputation. *Journal of Political Economy*, 107, 135-160.

Costa, D.L., 1998: The Evolution of Retirement: An American Economic History, 1880-1990. 1st ed. University of Chicago Press, Chicago, US.

Disney, R., 2007: Population ageing and the size of the welfare state: Is there a puzzle to explain? *European Journal of Political Economy*, 23, 542-553.

Downs, A., 1965: A Theory of Bureaucracy. *The American Economic Review,* 55, 439-446.

Eläketurvakeskus, 2011: Työeläkkeiden rahoitus vuonna 2010. Eläketurvakeskuksen tilastoraportteja 06/2011. http://www.etk.fi/fi/gateway/PTARGS_0_2712_459_440_3034_43/http%3B/content.etk.fi%3B7087/publishedcontent/publish/etkfi/fi/julkaisut/tilastojulkaisut/tilastoraportit/tyoelakkeiden_rahoitus_vuonna_2010_7.pdf. Date: 19.4.2012.

Frey, B.S., 1971: Why do high income people participate more in politics? *Public Choice*, 11, 101-105.

Galasso, V., 2002: The US social security: a financial appraisal for the median voter. *Social Security Bulletin*, 64, 57-65.

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

Galasso, V., Profeta, P., 2004: Lessons for an ageing society: the political sustainability of social security systems. *Economic Policy*, 19, 63-115.

Glenn, N.D., Grimes, M., 1968: Aging, Voting, and Political Interest. *American Sociological Review*, 33, 563-575.

Goerres, A., 2009: The Political Participation of Older People in Ageing Europe: The Greying of our Democracies. 1st ed. Palgrave Macmillan, Basingstoke, UK.

Green, P.G., Shapiro, I., 1994: Pathologies of Rational Choice Theory: A Critique of Applications in Political Science. 1st ed. Yale University Press, New Haven, US.

Grossman, G.M., Helpman, E., 1996: Electoral competition and special interest politics. *Review of Economic Studies*, 63, 265-286.

Grossman, G.M., Helpman, E., 1998: Intergenerational redistribution with short-lived governments. *The Economic Journal*, 108, 1299-1329.

Gruber, J., Wise, D.A., 1999: Social Security and Retirement around the World. 1st ed. University of Chicago Press, Chicago, US.

Hammond, P., 1975: Charity: Altruism or cooperative egoism. In E.S. Phelps: Altruism, Morality and Economic Theory, p. 115-131. 1st ed. Russell Sage Foundation, New York, US.

Hollanders, D., Kosters, F., 2011: The Graying of the Median Voter. Aging and the Politics of the Welfare State in OECD Countries. Discussion paper 1, Netspar. http://arno.uvt.nl/show.cgi?fid=113622. Date: 26.4.2012.

Hu, S.C., 1982: Social security, majority-voting equilibrium and dynamic efficiency. *International Economic Review*, 23, 269-287.

IMF, 2004: World Economic Outlook, September 2004: The Global Demographic Transition.

Jennings, M.K., Markus, G.B., 1988: Political Involvement in the Later Years: A Longitudinal Survey. *American Journal of Political Science*, 32, 302-316.

Krehbiel, K., 2004: Legislative Organization. *Journal of Economic Perspectives*, 118, 113-128.

Lazear, E.P., 2000: Economic Imperialism. *Quarterly Journal of Economics*, 115, 99-146.

Lehner, F., Widmaier, U., 1983: Market Failure and Growth of Government: A Sociological Explanation. In C.L. Taylor: Why Governments Grow. 1st ed. Sage, Beverly Hills, Canada.

Lindert, P.H., 1996: What limits social spending? *Explorations in Economic History*, 33, 1-34.

Meijdam, L., Verbon, H.A.A., 1996: Aging and political decision making on public pensions. *Journal of Population Economics*, 9, 141-158.

Meltzer, A.H., Richard, S.F., 1981: A rational theory of the size of government. *Journal of Political Economy*, 89, 914-927.

Miller, G.J., 1997: The Impact of Economics on Contemporary Political Science. *Journal of Economic Literature*, 35, 1173-1204.

Mulligan, C.B., Gil, R., Sala-i-Martin, X., 2010: Social Security and Democracy. *The B.E. Journal of Economic Analysis & Policy*, 10, article 18, 1-44.

Mulligan, C.B., Sala-i-Martin, X., 1999: Gerontocracy, Retirement, and Social Security. NBER working paper 7117, National Bureau of Economic Research. citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.200.5865&rep=rep1&type=pdf. Date: 26.4.2012.

Mulligan, C.B., Sala-i-Martin, X., 2003: Social Security, Retirement and the Single-Mindedness of the Electorate. NBER working paper 9691, National Bureau of Economic Research. citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.160.3898&rep=rep1&type=pdf. Date: 26.4.2012.

Mulligan, C.B., Sala-i-Martin, X., 2004: Political and Economic Forces Sustaining Social Security. *Advances in Economic Analysis & Policy*, 4, article 5, 1-54.

Murrell, P., 1984: An Examination of the Factors Affecting the Formation of Interest Groups in OECD Countries. *Public Choice*, 43, 151-171.

OECD, 2011: Pensions at a Glance 2011: Retirement-Income Systems in OECD and G20 Countries. Including statistics database. www.oecd.org/els/social/pensions/PAG. Date 19.4.2012.

Pampel, F.C., Williamson, J.B., 1989: Age, Class, Politics and the Welfare State. 1st ed. Cambridge University Press, Cambridge, UK.

Peltzman, S., 1980: The Growth of Government. *Journal of Law and Economics*, 23, 209-287.

Perotti, R., 1996: Growth, income distribution, and democracy: what the data say. *Jounal of Economic Growth*, 1, 149-188.

Persson, T., Tabellini, G., 2000: Political Economics: Explaining Economic Policy. 1st ed. Massachusetts Insitute of Technology, Massachusetts, US.

Poole, K.T., Daniels, R.S., 1985: Ideology, Party and Voting in the US Congress, 1959-1980. *American Political Science Review*, 79, 373-399.

Poutvaara, P., 2006: On the political economy of social security and public education. *Journal of Population Economics*, 19, 345-365.

Pratt, H.J., 1976: The Gray Lobby. 1st ed. The University of Chicago Press, Chicago, US.

Profeta, P., 2004: Aging, Retirement and Social Security in a Model of Interest Groups. *Mathematical Population Studies*, 11, 93-120.

Razin, A., Sadka, E., Swagel, P., 2002: The aging population and the size of the welfare state. *Journal of Political Economy*, 110, 900-918.

Robinson, J.P., Godbey, G., 1997: Time for Life. 1st ed. Pennsylvania State University Press, University Park, Pennsylvania, US.

Rose, R., 1990: Inheritance before Choice in Public Policy. *Journal of Theoretical Politics*, 2, 263-291.

Samwick, A.A., 1997: New evidence on pensions, social security, and the timing of retirement. *Journal of Public Economics*, 70, 207-236.

Scharpf, F., 2000: The Viability of Advanced Welfare States in the International Economy: Vulnerabilities and Options. *Journal of European Public Policy*, 7, 190-228.

Shelton, C., 2008: The ageing population and the size of the welfare state: Is there a puzzle? *Journal of Public Economics*, 92, 647-651.

Shepsle, K.A., 1979: Institutional arrangements and equilibrium in multidimensional voting models. *American Journal of Political Science*, 66, 467-482.

Sigelman, L., Roeder, P.W., Jewell, M.E., Baer, M.A., 1985: Voting and Nonvoting: A Multi-Election Perspective. *American Journal of Political Science*, 29, 749-765.

Sinn, H-W., Uebelmesser, S., 2002: Pensions and the path to gerontocracy in Germany. *European Journal of Political Economy*, 19, 153-158.

Sjoblom, K., 1985: Voting for social security. *Public Choice*, 45, 225-240.

Skirbekk, V., 2008: Age and Productivity Capacity: Descriptions, Causes and Policy Options. *Ageing Horizons*, 8, 4-12.

Tabellini, G., 2000: A Positive Theory of Social Security. *Scandinavian Journal of Economics*, 102, 523-545.

Tepe, M., Vanhuysse, P., 2009: Are Aging OECD Welfare States on the Path to the Politics of Gerontocracy? Evidence from 18 Democracies, 1980-2002. *Journal of Public Policy*, 29, 1-28.

Veall, M.R., 1986: Public pensions as optimal social contracts. *Journal of Public Economics*, 31, 237-251.

Verbon, H.A.A., Verhoeven, M.J.M., 1991: Expectations on pension schemes under non-stationary conditions. *Economic Letters*, 36, 99-103.

Verbon, H.A.A., Verhoeven, M.J.M., 1992: Decision making on pension schemes under rational expectations. *Journal of Economics*, 56, 71-97.

Weiss, Y, 1972: On the Optimal Lifetime Pattern of Labour Supply. *Economic Journal*, 82, 1293-1315.

Appendix 1: The median voter theorem and single-peaked preferences

Following Black (1948) and Persson and Tabellini (2000), the median voter theorem is a modelling tool for aggregation of preferences by pure majority rule. This rule is defined by three assumptions: direct democracy, i.e. the citizens make the policy choices, sincere voting, i.e. each citizen votes for the option that gives her the highest utility, and open agenda, which means that the population votes on pairs of policy alternatives, in such a way that the winning policy in one round is posed against a new policy alternative. The option that wins the vote over all other policy options is called a Condorcet winner.

Now consider a unidimensional policy space, i.e. there is one policy variable, denoted by q. In the case of models of social security, the policy variable is often a feature of the social security system, e.g. its size. Preferences are defined as single-peaked for voter i if his preference ordering for different policies depends on the relative distance from his most preferred policy, $q(\alpha^i)$, where α^i denotes specific features of voter i. In other words, a policy closer to $q(\alpha^i)$ is preferred over more distant options. Formally, preferences are single-peaked if the following holds:

> If $q'' \leq q' \leq q(i)$ or $q'' \geq q' \geq q(i)$, it follows that $W(q''; \alpha^i) \leq W(q'; \alpha^i)$,

where W denotes the indirect utility function of the voters. This results in the median voter theorem:

> If all voters have single-peaked preferences over a range of policy alternatives, a Condorcet winner exists and is equal to the most preferred policy of the median voter.

To prove this, we fix the parameter α and order the voters according to their most preferred policies, and denote the median voter's preferred policy with q^m. Suppose that this policy is voted against another policy $q' < q^m$. Due to single-peakedness of preferences, all voters whose most preferred policy satisfies $q^m \leq q(\alpha^i)$ prefer q^m over q', as it is

closer to the preferred policy. This coalition constitutes a majority of the voters. The analogous argument can be applied on $q' > q^m$, which leads to the result that q^m is the Condorcet winner. Therefore, we can conclude that under the mentioned assumptions of direct democracy, sincere voting and open agenda, the policy of the median voter q^m is the unique equilibrium policy. □

Appendix 2: Calculations related to Conde-Ruiz and Galasso (2005)

Most of the stages of the calculations in this appendix are not found in the original article by Conde-Ruiz and Galasso (2005), and I have calculated them myself.

Calculations of the welfare system budget constraints

Inserting the labour supply (2.7) into the budget constraint of the income redistribution system (2.9) gives:

$$T_t = \tau_t \int_{\underline{e}}^{\bar{e}} e_t (\bar{l} - \frac{1}{e_t(1-\tau_t-\sigma_t)}) dG(e_t) = \tau_t (\bar{l} \int_{\underline{e}}^{\bar{e}} e_t dG(e_t) - \frac{1}{(1-\tau_t-\sigma_t)} \int_{\underline{e}}^{\bar{e}} dG(e_t)).$$

On the right-hand side, we see that the first integral equals the definition of the expected value of e_t, i.e. e_φ, and the second integral equals the whole cumulative distribution function, i.e. 1. Therefore, the budget constraint depending on the tax rates is:

$$T_t(\tau_t, \sigma_t) = \tau_t (e_\varphi \bar{l} - \frac{1}{(1-\tau_t-\sigma_t)}).$$

Analogous calculations give the budget constraint for the pension system:

$$P_t(\tau_t, \sigma_t) = \sigma_t (1+\mu)(e_\varphi \bar{l} - \frac{1}{(1-\tau_t-\sigma_t)}).$$

Calculations of the workers' most preferred income redistribution tax rate

Maximizing the indirect utility (2.14) with respect to the income redistribution tax rate gives:

$$\frac{\partial v_t^t(.)}{\partial \tau_t} = (e_\varphi - e_t)\bar{l} - \frac{\tau_t}{(1-\tau_t-\sigma_t)^2} = 0.$$

Simplifying and rearranging gives a second-order equation, which is then solved in the usual way:

$$\tau_t^2 + (2\sigma_t - 2 - \frac{1}{(e_\varphi - e_t)\bar{l}})\tau_t + (1-\sigma_t)^2 = 0$$

$$\Leftrightarrow \tau_t = 1 - \sigma_t + \frac{1 \pm \sqrt{1 + 4(e_\varphi - e_t)\bar{l}(1-\sigma_t)}}{2(e_\varphi - e_t)\bar{l}}.$$

As we only consider the solution that is between the values 0 and 1, we only consider the minus sign and we get the most preferred income redistribution tax rate to be:

$$\tau_{t,e}(\sigma_t) = \max\left\{0; 1 - \sigma_t + \frac{1 - \sqrt{1 + 4(e_\varphi - e_t)\bar{l}(1-\sigma_t)}}{2(e_\varphi - e_t)\bar{l}}\right\}.$$

Calculations of the most preferred social security tax rate

Beginning with the retirees, we maximize their indirect utility function (2.15) with respect to the social security tax rate:

$$\frac{\partial v_t^{t-1}(.)}{\partial \sigma_t} = e_\varphi \bar{l} - \frac{1}{1-\tau_t-\sigma_t} - \frac{\sigma_t}{(1-\tau_t-\sigma_t)^2} = 0.$$

Noting that $e_\varphi \bar{l} - 1/(1-\tau_t-\sigma_t) = e_\varphi(\bar{l} - 1/(e_\varphi(1-\tau_t-\sigma_t))) = e_\varphi n(e_\varphi)$, the above can also be written as the equation (2.18):

$$e_\varphi n(e_\varphi) = \frac{\sigma_t}{(1-\tau_t-\sigma_t)^2}.$$

Simplifying and rearranging again gives a second order equation, which is solved below.

$$\sigma_t^2 + 2(\tau_t - 1)\sigma_t + (\tau - 1)^2 - \frac{1-\tau_t}{e_\varphi \bar{l}} = 0$$

$$\Leftrightarrow \sigma_t = 1 - \tau_t \pm \sqrt{\frac{1-\tau_t}{e_\varphi \bar{l}}}.$$

We once again consider the solution between 0 and 1, and get the following expression for the retirees' most preferred social security tax rate:

$$\Leftrightarrow \sigma_{t,old}(\tau_t) = 1 - \tau_t - \sqrt{\frac{1-\tau_t}{e_\varphi \bar{l}}}.$$

For the workers, we employ similar calculations. We start by maximizing the indirect utility function (2.14) with respect to the current and future social security policy $\sigma_t = \sigma_{t+1} = \sigma$ and taking the current and future income redistribution rates, τ_t and τ_{t+1}, as given. Taking partial derivatives and rearranging gives the equation (2.20):

$$\frac{\partial v_t^t(.)}{\partial \sigma_t} = \frac{1}{1-\tau_t-\sigma_t} - e_t\bar{l} + \frac{\partial T_t}{\partial \sigma_t} + \frac{\frac{\partial P_{t+1}}{\partial \sigma_t}}{1+R} = 0$$

$$\Leftrightarrow e_t n(e_t) - \frac{\partial T_t}{\partial \sigma_t} = \frac{\frac{\partial P_{t+1}}{\partial \sigma_t}}{1+R}.$$

Restricting our analysis to steady states by imposing $\tau_t = \tau_{t+1} = \tau$, further calculations gives the equation (2.21):

$$e_t n(e_t) + \frac{\tau_t}{(1-\tau_t-\sigma_t)^2} = N(e_\varphi \bar{l} - \frac{1}{1-\tau_t-\sigma_t} - \frac{\sigma_t}{(1-\tau_t-\sigma_t)^2})$$

$$\Leftrightarrow e_\varphi n(e_\varphi) - \frac{1}{N}(e_t n(e_t) + \frac{\tau_t}{(1-\tau_t-\sigma_t)^2}) = \frac{\sigma_t}{(1-\tau_t-\sigma_t)^2}.$$

Rearranging further, we get a second-order equation, which is solved below:

$$\sigma_t^2 + (2\tau_t - 2 - \frac{1}{\bar{l}(e_\varphi N - e_t)})\sigma_t + (\tau-1)^2 + \frac{(N-2)\tau_t + 1 - N}{\bar{l}(e_\varphi N - e_t)} = 0$$

$$\Leftrightarrow \sigma_t = 1 - \tau_t + \frac{1 \pm \sqrt{1 + 4\bar{l}(N + \tau_t(1-N))(e_\varphi N - e_t)}}{2\bar{l}(e_\varphi N - e_t)}.$$

As in the case of the retirees, we only consider values between 0 and 1, and we get the expression for the most preferred social security tax rate of the workers:

$$\sigma_{t,e}(\tau_t) = \max\left\{0; 1 - \tau_t + \frac{1 - \sqrt{1 + 4\bar{l}(N + \tau_t(1-N))(e_\varphi N - e_t)}}{2\bar{l}(e_\varphi N - e_t)}\right\}.$$

Appendix 3: Structure-induced equilibria

This appendix provides a formal definition of structure-induced equilibria, which later will be applied in the model of chapter 2. The definition is based on Conde-Ruiz and Galasso (2003b), which is an earlier version of the paper presented in chapter 2. This discussion

paper in turn follows the structure-induced equilibrium concept of Shepsle (1979). Some additions are provided for clarification.

A certain institutional arrangement, which is described below, characterizes how the political system aggregates the preferences of the individuals into a policy outcome (τ^*, σ^*). The electorate, E, has $1+1/(1+\mu)$ members. The space of policy alternatives, or issues, is $(\tau, \sigma) \in \Re^2$ subject to equation (2.8), which was the condition for a positive labour supply of the workers. Individual preferences over the issue space are defined by the indirect utility functions (2.14) and (2.15). At this point, we assume perfect commitment over social security policy and no commitment over the income redistribution tax policy.

The institutional arrangement consists of a committee system, a jurisdictional arrangement, an assignment rule, as well as of an amendment control rule. The committee system divides the electorate into committees $\{C_j\}$. The jurisdictional arrangement, J, separates the issues (τ, σ) into jurisdictions $\{J_k\}$. These jurisdictions are associated to committees according to an assignment rule, $f : C_j \to J_k$. In this way, each issue is assigned to a particular committee. Each committee is then able to make proposals to change the current value of the issue (the status quo) that is part of its jurisdiction. Finally, the amendment control rule determines how proposals can be amended before reaching the final stage, and the possibly amended proposal is then voted in a majority rule, pairwise vote against the status quo by the electorate.

The system adopted in this model is characterized by the following arrangements: (i) Committee of the Whole: there exists a unique committee, which is equal to the whole electorate, i.e. $C = \{E\}$; (ii) Simple Jurisdictions: every jurisdiction is one dimension of the issue space, i.e. $J = \{\{\tau\}, \{\sigma\}\}$. This means that one jurisdiction decides on the income redistribution tax rate and another on the social security tax rate; (iii) All jurisdictions are assigned to the committee of the whole, i.e. $f : E \to \{\{\tau\}, \{\sigma\}\}$; (iv) Germaneness Amendment Control Rule: changes to the proposals are allowed only on the issue that falls in the jurisdiction of the committee. In other words, if a proposal regards σ, only amendments regarding σ are permitted and vice versa.

This political system leads to a situation where the whole electorate is entitled to make proposals on the two issues, but only separately issue by issue. Following Shepsle (1979), simple jurisdictions and the germaneness amendment rule changes our two-dimensional game into a simultaneous issue-by-issue voting game. This helps to overcome the potential lack of a Condorcet winner of the original two-dimensional majority voting game. Furthermore, according to Shepsle (1979), a sufficient condition of an equilibrium to exist in the voting game played under the described institutional arrangement is that the voter's preferences are single-peaked over each issue. The proof of single-peaked preferences in our model is provided in appendix 5. Therefore, the issue-by-issue voting game is able to preserve the framework of the median voter theorem (presented in appendix 1), applied for each issue separately.

Next, we refer to the most preferred point of a voter i in the j-th direction of the issue space (e.g. τ) as the tax rate that maximizes i's indirect utility function along the j-th dimension (τ), while the other issue (σ) is fixed. We are now prepared to define the structure induced equilibrium of a voting game with commitment over social security:

> Let X_j^* refer to the set of j-th components derived from the most preferred points of the electorate in the direction j from the status quo x^o. In the case of a committee of the whole, simple jurisdictions, a germaneness rule for amendments, and single-peaked preferences, x^o is a structure-induced equilibrium outcome if and only if, for all j, x_j^* equals the median of X_j^*.

The proof is defined as follows. As preferences are single-peaked over all single issues, and if x_j^* is the median of X_j^*, x^o defeats all other points along the j-th dimension. This is in accordance with the median voter theorem (see appendix 1). With simple jurisdictions and germaneness rule of amendments, issues are voted on one by one. Now, since x_j^* can't be defeated by any points on any dimension j, then x^o is a structure induced equilibrium outcome. To further illustrate, suppose that x' is an outcome, where

$x' \neq x_j^o = median(X_j^*) \forall j$ along some dimension i. As we have a committee of the whole, x' would always be outvoted by x^o along the i-th dimension. □

Appendix 4: Subgame perfection with structure-induced equilibria

After Shepsle (1979) and Conde-Ruiz and Galasso (2005), this appendix provides a formal definition of the voting game and of the subgame perfect structure-induced equilibrium. The definition is used in the model in chapter 2.

The public history of the game at time t consists of a sequence of income redistribution and social security taxes until $t-1$, $h_t = (\tau_0, \sigma_0), ..., (\tau_{t-1}, \sigma_{t-1}) \in H_t$, where H_t is the set of all potential histories at time t. An action for a worker with type e at time t is a pair of tax rates, $a_{t,e}^w = (\tau, \sigma) \in \mathrm{Y}$, where $\mathrm{Y} = \left\{(\tau, \sigma) : \tau \in [0,1], \sigma \in [0,1], \tau + \sigma \leq 1 - 1/(\bar{l}\underline{e})\right\}$. Similarly, an action for a type e retiree at time t is $a_{t,e}^r = (\tau, \sigma) \in \mathrm{Y}$. We define a_t as the action profile of all workers and retirees at time t: $a_t = (a_t^w \cup a_t^r)$, where $a_t^w = \bigcup_{e \in [\underline{e}, \bar{e}]} a_{t,e}^w$ and $a_t^r = \bigcup_{e \in [\underline{e}, \bar{e}]} a_{t,e}^r$. Furthermore, a strategy at time t for a type e worker is a mapping from the historical development of the game into the action space: $s_{t,e}^w : h_t \to \mathrm{Y}$, and similarly for a type e retiree at time t: $s_{t,e}^r : h_t \to \mathrm{Y}$. A set of strategies, i.e. a strategy profile, played by all individuals at time t is $s_t = (s_t^w \cup s_t^r)$, where $s_t^w = \bigcup_{e \in [\underline{e}, \bar{e}]} s_{t,e}^w$ and $s_t^r = \bigcup_{e \in [\underline{e}, \bar{e}]} s_{t,e}^r$.

For a given action profile a_t, the pair of tax rates (τ_t^m, σ_t^m) represents the medians of the distributions of the voters' most preferred tax rates at time t. Therefore, we take the outcome function of the voting game to be (τ_t^m, σ_t^m) at time t (see appendix 3). Now, the history of the voting game is updated accordingly; at time $t+1$, $h_{t+1} = (\tau_0, \sigma_0), ..., (\tau_{t-1}, \sigma_{t-1}), (\tau_t^m, \sigma_t^m) \in H_{t+1}$. Each agent has a payoff function, which corresponds to the indirect utility function. Expressed formally, for a certain sequence of action profiles $(a_0, ..., a_t, a_{t+1}, ...)$ and of the corresponding realizations, $((\tau_0, \sigma_0), ..., (\tau_t, \sigma_t), (\tau_{t+1}, \sigma_{t+1}) ...)$, the payoff function for a worker with type e at time t is

$v_t^t(\tau_t, \sigma_t, \tau_{t+1}, \sigma_{t+1}, e)$, according to (2.14), and for a type e retiree is $v_t^t(\tau_t, \sigma_t, e)$, as defined in (2.15).

We let $s_{t|\hat{e}}^w = s_t^w / s_{t,\hat{e}}^w$ be the strategy profile for all workers at time t except for type $\hat{e}$, and $s_{t|\hat{e}}^r = s_t^r / s_{t,\hat{e}}^r$ be the strategy profile for all retirees except for the type $\hat{e}$. From this follows that, at time t, a type $\hat{e}$ worker maximizes

$$V_{t,\hat{e}}^t(s_0,...,(s_{t|\hat{e}}^w, s_{t,\hat{e}}^w), s_t^r, s_{t+1},...) = v_t^t(\tau_t^m, \sigma_t^m, \tau_{t+1}^m, \sigma_{t+1}^m, \hat{e}),$$

and a type $\hat{e}$ retiree maximizes

$$V_{t,\hat{e}}^{t-1}(s_0,...,(s_{t|\hat{e}}^r, s_{t,\hat{e}}^r), s_t^w, s_{t+1},...) = v_t^{t-1}(\tau_t^m, \sigma_t^m, \hat{e}),$$

where in accordance with our earlier definition of the outcome function, (τ_t^m, σ_t^m) and $(\tau_{t+1}^m, \sigma_{t+1}^m)$ are the medians among the actions over the welfare program tax rates at time t and $t+1$, respectively. These maximization strategies are subgame perfect equilibria, i.e. they are the optimal strategies given the other players' strategies.

Following the previous, we are now able to define a subgame perfect structure-induced equilibrium (SPSIE) of the voting game: The voting strategy profile $s = \left\{(s_t^w \cup s_t^r)\right\}_{t=0}^{\infty}$ is a SPSIE if the following two conditions are satisfied. Firstly, s is a subgame perfect equilibrium (as defined above); and secondly, the equilibrium outcome associated with s is a structure-induced equilibrium of the voting game with commitment over social security in all periods t (as defined in appendix 3).

Appendix 5: Proof of single-peakedness

Based on Conde-Ruiz and Galasso (2003b) with some clarifying comments, this appendix proves under what conditions the preferences of the electorate are single-peaked. This is one of the conditions for a structure-induced equilibrium to arise, as defined in appendix 3.

Single-peakedness of the workers' preferences

For the workers, the preferences are represented by the utility function (2.14). To infer single-peakedness, we can use the second partial derivative test. For this test we need the Hessian matrix, which we calculate by taking the second derivatives and cross derivatives. We impose $\sigma = \sigma_t = \sigma_{t+1}$, due to commitment over social security, and the stationarity condition $\tau = \tau_t = \tau_{t+1}$.

$$\frac{\partial^2 v_t^t(.)}{\partial \tau^2} = -\frac{1}{(1-\tau-\sigma)^2} - \frac{2\tau}{(1-\tau-\sigma)^3} = -\frac{1+\tau-\sigma}{(1-\tau_t-\sigma_t)^3},$$

$$\frac{\partial^2 v_t^t(.)}{\partial \sigma^2} = \frac{1-N}{(1-\tau-\sigma)^2} - \frac{N}{(1-\tau-\sigma)^2} + \frac{2(\tau+N\sigma)}{(1-\tau-\sigma)^3} = \frac{1-3\tau-\sigma-2N+2N\tau}{(1-\tau_t-\sigma_t)^3},$$

$$\frac{\partial^2 v_t^t(.)}{\partial \tau \partial \sigma} = \frac{\partial^2 v_t^t(.)}{\partial \sigma \partial \tau} = -\frac{2\tau}{(1-\tau-\sigma)^3}.$$

Consequently, our Hessian matrix is the following:

$$\begin{bmatrix} \frac{\partial^2 v_t^t(.)}{\partial \tau^2} & \frac{\partial^2 v_t^t(.)}{\partial \tau \partial \sigma} \\ \frac{\partial^2 v_t^t(.)}{\partial \sigma \partial \tau} & \frac{\partial^2 v_t^t(.)}{\partial \sigma^2} \end{bmatrix} = \begin{bmatrix} -\frac{1+\tau-\sigma}{(1-\tau_t-\sigma_t)^3} & -\frac{2\tau}{(1-\tau-\sigma)^3} \\ -\frac{2\tau}{(1-\tau-\sigma)^3} & \frac{1-3\tau-\sigma-2N+2N\tau}{(1-\tau_t-\sigma_t)^3} \end{bmatrix}.$$

Now, the Hessian is said to be semi definite negative if the determinant of the matrix is equal to or larger than zero, and if the indirect utility function is concave with respect to the tax rates, i.e. its second derivatives are negative. To evaluate the first condition, we calculate the determinant of the Hessian:

$$(-\frac{1+\tau-\sigma}{(1-\tau_t-\sigma_t)^3}) * (\frac{1-3\tau-\sigma-2N+2N\tau}{(1-\tau_t-\sigma_t)^3}) - (-\frac{2\tau}{(1-\tau-\sigma)^3})^2 \geq 0$$

$$\Leftrightarrow \frac{2(1-\sigma+\tau\sigma-\tau^2)N-(1-\tau-\sigma)^2}{(1-\tau_t-\sigma_t)^6} \geq 0.$$

As the denominator is larger than 0, the above term is equal to or larger than zero if its numerator is equal to or larger than zero:

$$2(1-\sigma+\tau\sigma-\tau^2)N-(1-\tau-\sigma)^2 \geq 0$$

$$\Leftrightarrow N \geq \frac{(1-\tau-\sigma)^2}{2(1-\sigma-\tau(\tau-\sigma))} \leq \frac{1}{2}.$$

The last inequality holds as 1/2 is the largest value the term can get given the restrictions on the tax rates. Therefore, the determinant of the Hessian matrix is equal to or larger than zero when $N \geq 1/2$.

To determine concavity with respect to the tax rates, we first note that the second derivative of the indirect utility function with respect to τ is negative, as $(\tau, \sigma) \in [0,1]$ and $\tau + \sigma \leq 1$. Similarly, the second derivative with respect to σ is also negative, as long as $1/2 \leq N < 1$. This restriction on N was imposed from the calculations above and from the assumption of dynamic efficiency in section 2.1.1.

Therefore, as the two conditions for negative semi definiteness hold, we can conclude that the Hessian is semi definite negative if $N \geq 1/2$. From this follows that the workers' preferences are quasi concave and, by Shepsle's lemma 3.1 (Shepsle, 1979), also single-peaked. □

Single-peakedness of the retirees' preferences

The preferences of the retirees in equation (2.15) are not concave with respect to τ, as the most preferred income redistribution tax rate is always zero. Therefore, we cannot use the second partial derivative test to infer single-peakedness. Instead, the authors proceed by introducing a formal definition of single-peaked preferences over a line in the issue space, and use this to prove single-peakedness. As these calculations are extremely difficult, we do not present them here but refer to the original article of Conde-Ruiz and Galasso (2003b).

Appendix 6: Proof of proposition 2.2

The proof follows that of Conde-Ruiz and Galasso (2005), with some clarifying comments.

From the result of proposition 2.1, we build on the notion that (τ^*, σ^*) is the structure-induced equilibrium outcome of the voting game with commitment over social security tax rates. We can then define the realization of the public history of the game as follows:

$$H_t^0 = \{h_t \in H_t | \sigma_k = 0, k = 0,..., t-1\}$$

and

$$H_t^{\sigma} = \left\{ h_t \in H_t \middle| t_0 \in \{0,1,...,t-1\} : \sigma_t = 0 \forall t < t_0 \wedge \sigma_t = \sigma^* \forall t \geq t_0 \right\}.$$

The first equation consists of a set of histories where no social security has ever been imposed, and the second equation consists of a set of histories where the equilibrium level of social security has been imposed at time t_0, and sustained ever since. The key insight is that these histories cover all histories except for the ones where a social security system has been imposed but later changed or abandoned, or the ones where the level of social security has been imposed with a level other than that of the equilibrium. We also note that $H_t^0 \cap H_t^{\sigma} = \emptyset$, i.e. a history can't be a part of both history sets.

Next, we consider the following strategy $s = (s_{t,e}^w, s_{t,e}^r)$ for a worker with type $e \leq e_{m\sigma}$:

$$s_{t,e}^w = \begin{cases} (\tau_{t,e}(\sigma^*), \sigma^*) \\ (\tau_{t,e}(0), 0) \end{cases} \begin{matrix} if \quad h_t \in H_t^0 \cup H_t^{\sigma} \\ h_t \in H_t / \left\{ H_t^0 \cup H_t^{\sigma} \right\} \end{matrix},$$

for a worker with type $e > e_{m\sigma}$:

$$s_{t,e}^w = \begin{cases} (\tau_{t,e}(\sigma^*), \sigma_{t,e}(\tau^*)) \\ (\tau_{t,e}(0), 0) \end{cases} \begin{matrix} if \quad h_t \in H_t^0 \cup H_t^{\sigma} \\ h_t \in H_t / \left\{ H_t^0 \cup H_t^{\sigma} \right\} \end{matrix},$$

and for a retiree:

$$s_{t,e}^r = (0, \sigma_{t,old}(\tau^*)) \ if \ h_t \in H_t,$$

where $\tau_{t,e}(\sigma^*)$ is defined in equation (2.16), $\sigma_{t,e}(\tau^*)$ in (2.22) and $\sigma_{t,old}(\tau^*)$ in (2.19). As we know from appendix 4 that $\tau^*(\sigma^*) = \tau_{t,e_{m\tau}}(\sigma^*)$ and $\sigma^*(\tau^*) = \sigma_{t,e_{m\sigma}}(\tau^*)$, it follows that:

$$\tau_{t,e}(\sigma^*) \geq \tau^* \forall e \leq e_{m\tau},$$

$$\sigma_{t,e}(\tau^*) \geq \sigma^* \forall e \leq e_{m\sigma}.$$

Low-ability workers' most preferred tax rates are higher than the equilibrium outcome of the median voters.

In appendix 3, we defined the outcome function of the voting game at time t as the median outcomes, (τ_t^m, σ_t^m). Now no agent wants to deviate from the strategy above. To prove this, first consider a retiree. She would like to have a higher social security tax rate above σ^*

and an income redistribution tax lower than τ^*. However, deviating from the above strategy $s^r_{t,e}$ will not change the outcome in the desired direction. As $\sigma_{t,old}(\tau^*)$ is already larger than $\tau^*(\sigma^*) = \tau_{t,e_{m\tau}}(\sigma^*)$, an increase in $\sigma_{t,old}(\tau^*)$ would not change the equilibrium median voter policy σ^*; similarly, $\tau_{t,old}(\sigma^*)$ already equals zero and is below $\tau^*(\sigma^*) = \tau_{t,e_{m\tau}}(\sigma^*)$. A similar reasoning applies to a type e worker with $e < e_{m\sigma} < e_{m\tau}$. Although she would prefer both a higher social security and income redistribution tax rate than the outcome, voting for higher rates does not change the median outcome as she already votes for higher tax rates. Finally, a type e worker with $e > e_{m\sigma}$ would like to reduce the social security tax, but her vote is already below the median one and therefore she can't reduce the outcome. In a similar way, both a type $e < e_{m\tau}$ worker, who would like to reduce τ^*, and a type $e > e_{m\tau}$, who would like to decrease it, can't have an impact on the equilibrium outcome.

It should be noted that the voting behaviour of the workers is conditional on the existence of an implicit contract between generations, as explained in section 2.1.2. Otherwise, workers will not support the social security system at all, as future generations would not support it either. This corresponds to that the system has not been abandoned before, i.e. $h_t \in H_t^0 \cup H_t^\sigma$. If this holds, the previous strategy profile $(s^w_{t,e}, s^r_{t,e})$ is a subgame perfect equilibrium of the voting game without commitment, with the equilibrium outcome (τ^*, σ^*). □

Appendix 7: Calculations related to Mulligan and Sala-i-Martin (1999)

Most of the stages of the calculations in this appendix are not found in the article of Mulligan and Sala-i-Martin (1999) and are of my own doing.

<u>Calculations of the MRS between consumption in this period and the next</u>

From equation (3.2), we can derive the maximization problem of the current young group in the next period, i.e. the future old, by changing the time indices:

$$v(\overline{w}_{t+1}\rho F_t + \overline{a}_{t+1}; w^o_{t+1}, \bar{l}^y_{t+1}) \equiv \max_{\bar{c}^o_{t+1}, \bar{l}^o_{t+1}} u^o_{t+1}(\bar{c}^o_{t+1}, \bar{l}^o_{t+1})$$

such that $\bar{c}^o_{t+1} = w^o_{t+1}(1-\bar{l}^o_{t+1}) + \overline{w}_{t+1}\left[f_{t+1}(\bar{l}^o_{t+1}, \bar{l}^y_{t+1}) + \rho F_t\right] + \bar{a}_{t+1}$. (A7.1)

Consequently, equation (3.3) gives the maximization problem of the young group in this period:

$$\max_{\bar{c}^y, \bar{l}^y, \bar{a}_{t+1}} \left[u^y(\bar{c}^y, \bar{l}^y) + \beta v(\overline{w}_{t+1}\rho F_t + \bar{a}_{t+1}; w^o_{t+1}, \bar{l}^y_{t+1})\right]$$

such that $\bar{c}^y + R\bar{a}_{t+1} = w^y(1-\bar{l}^y) - \overline{w}\left[f(\bar{l}^o, \bar{l}^y) + \rho F_{t-1}\right]$. (A7.2)

Note that $\bar{a}_{t+1}$ is the amount saved as young to be consumed as old. By maximizing (A7.2) with respect to $\bar{a}_{t+1}$ subject to the constraint, we get the optimal amount for the young to save. Rearranging the constraint in (A7.2), we get:

$$\bar{c}^y = -R\bar{a}_{t+1} w^y(1-\bar{l}^y) - \overline{w}\left[f(\bar{l}^o, \bar{l}^y) + \rho F_{t-1}\right].$$

Now, we can calculate the partial derivative of the first term in (A7.2) by using the chain rule:

$$\frac{du^y}{d\bar{a}_{t+1}} = \frac{du^y}{d\bar{c}^y}\frac{d\bar{c}^y}{d\bar{a}_{t+1}} = \frac{\partial u^y}{\partial \bar{c}^y}(-R). \quad \text{(A7.3)}$$

To get the partial derivative of the second term v in (A7.2), we again use the chain rule but on equation (A7.1):

$$\frac{du^o_{t+1}}{d\bar{a}_{t+1}} = \frac{du^o_{t+1}}{d\bar{c}^o_{t+1}}\frac{d\bar{c}^o_{t+1}}{d\bar{a}_{t+1}} = \frac{\partial u^o_{t+1}}{\partial \bar{c}^o_{t+1}}. \quad \text{(A7.4)}$$

According to the envelope theorem, this is equal to the partial derivative of v with respect to $\bar{a}_{t+1}$ in (A7.2). Combining (A7.3) and (A7.4) gives the optimality condition:

$$\frac{\partial u^y}{\partial \bar{c}^y}(-R) + \beta\frac{\partial u^o_{t+1}}{\partial \bar{c}^o_{t+1}} = 0 \Leftrightarrow \frac{\partial u^y / \partial \bar{c}^y}{\partial u^o_{t+1} / \partial \bar{c}^o_{t+1}} = \frac{\beta}{R}.$$

This is the MRS between the young group's consumption in this period and in the next period.

When calculating the MRS between the young individual's (and not the group's) consumption in this period and in the next, the only difference is that the individuals take the actions of others, $(\bar{l}^o, \bar{l}^y, \bar{l}^o_{t+1}, \bar{l}^y_{t+1})$, as given. This difference does not change the calculations and result above.

Calculations of the young group's MRS between consumption and leisure

We know from standard microeconomics and the implicit function theorem, explained before calculating equation (3.4), that the following holds:

$$MRS_{\bar{l}^y,\bar{c}^y} = \frac{\partial u^y / \partial \bar{l}^y}{\partial u^y / \partial \bar{c}^y} = -\frac{d\bar{c}^y}{d\bar{l}^y}.$$

From (3.3), we know that the young group's constraint is the following:

$$\bar{c}^y + R\bar{a}_{t+1} = w^y(1-\bar{l}^y) - \overline{w}\left[f(\bar{l}^o, \bar{l}^y) + \rho F_{t-1}\right]. \tag{A7.5}$$

To be able to take a derivative with respect to current leisure, we must derive the expression for $\bar{a}_{t+1}$ from equation (3.2), with all variables changed to represent the next period:

$$\bar{a}_{t+1} = \bar{c}^o_{t+1} - w^o_{t+1}(1-\bar{l}^o_{t+1}) - \overline{w}_{t+1}\left[f(\bar{l}^o_{t+1}, \bar{l}^y_{t+1}) + \rho F_t\right]. \tag{A7.6}$$

Inserting this into (A7.5) gives the following expression:

$$\begin{aligned}\bar{c}^y &= w^y(1-\bar{l}^y) - \overline{w}\left[f(\bar{l}^o, \bar{l}^y) + \rho F_{t-1}\right] \\ &- R\left[\bar{c}^o_{t+1} - w^o_{t+1}(1-\bar{l}^o_{t+1}) - \overline{w}_{t+1}\left[f(\bar{l}^o_{t+1}, \bar{l}^y_{t+1}) + \rho F_t\right]\right].\end{aligned} \tag{A7.7}$$

We can now derive the wanted expression:

$$\begin{aligned}&\frac{d\bar{c}^y}{d\bar{l}^y} = -w^y - \overline{w}\frac{\partial f}{\partial \bar{l}^y} - \rho R\overline{w}_{t+1}\frac{\partial f}{\partial \bar{l}^y} = -w^y - (1-\rho R\frac{\overline{w}_{t+1}}{\overline{w}})\overline{w}\frac{\partial f}{\partial \bar{l}^y} \\ &\Leftrightarrow MRS_{\bar{l}^y,\bar{c}^y} = \frac{\partial u^y / \partial \bar{l}^y}{\partial u^y / \partial \bar{c}^y} = -\frac{d\bar{c}^y}{d\bar{l}^y} = w^y + (1-\rho R\frac{\overline{w}_{t+1}}{\overline{w}})\overline{w}\frac{\partial f}{\partial \bar{l}^y}.\end{aligned} \tag{A7.8}$$

Calculations of the young individual's intertemporal budget constraint

We follow the previous section and insert (A7.6) into (A7.5) to end up with (A7.7), except that we calculate the constraint for the individual and not the group. Afterwards, some simple rearrangements can be made to finish at the desired outcome:

$$\begin{aligned}c^y &= w^y(1-l^y) - \overline{w}\left[f(\bar{l}^o, \bar{l}^y) + \rho F_{t-1}\right] \\ &- R\left[c^o_{t+1} - w^o_{t+1}(1-l^o_{t+1}) - \overline{w}_{t+1}\left[f(\bar{l}^o_{t+1}, \bar{l}^y_{t+1}) + \rho F_t\right]\right] \\ &\Leftrightarrow c^y + Rc^o_{t+1} = w^y(1-l^y) - \overline{w}F_t + R\left[w^o_{t+1}(1-l^o_{t+1}) + \overline{w}_{t+1}F_{t+1}\right].\end{aligned} \tag{A7.9}$$

Calculations of the young group's optimal labour income tax

From (A7.9), we can calculate the derivative of the individual's budget constraint with respect to leisure. This is equated with the group's derivative of the budget constraint with respect to leisure, calculated in (A7.8). Rearranging gives the optimal tax.

$$\frac{\partial c^y}{\partial l^y} = (1-\tau^y)(-w^y) = -w^y - (1-\rho R\frac{\overline{w}_{t+1}}{\overline{w}})\overline{w}\frac{\partial f}{\partial \bar{l}^y} = \frac{\partial \bar{c}^y}{\partial \bar{l}^y}$$

$$\Leftrightarrow \tau^y = -\frac{\overline{w}}{w^y}(1-\rho R\frac{\overline{w}_{t+1}}{\overline{w}})\frac{\partial f}{\partial \bar{l}^y}.$$

Calculations of the relationship between the labour income tax rates

From equation (3.13) we get that $(\partial f / \partial \bar{l}^o) = \phi\overline{w}$ and $(\partial f / \partial \bar{l}_y) = -\phi\overline{w}$. We insert the former into the old group's optimal tax rate (3.6):

$$\tau_o = \frac{\overline{w}}{w^o}\frac{\partial f}{\partial \bar{l}^o} = \frac{\overline{w}}{w^o}\phi\overline{w} \Leftrightarrow \phi\overline{w}^2 = w^o\tau^o.$$

We now insert this equation into the young group's optimal tax rate (3.12) and rearrange to arrive at the desired equation:

$$\tau^y = -\frac{\overline{w}}{w^y}(1-\rho R\frac{\overline{w}_{t+1}}{\overline{w}})\frac{\partial f}{\partial \bar{l}^y} = \frac{\phi\overline{w}^2}{w^y}(1-\rho R\frac{\overline{w}_{t+1}}{\overline{w}}) = \frac{w^o\tau^o}{w^y}(1-\rho R\frac{\overline{w}_{t+1}}{\overline{w}})$$

$$\Leftrightarrow \frac{\tau^y}{\tau^o} = \frac{w^o}{w^y}(1-\rho R\frac{\overline{w}_{t+1}}{\overline{w}}) \Leftrightarrow \frac{\tau^o}{\tau^y} = \frac{w^y}{w^o}\frac{1}{(1-\rho R\frac{\overline{w}_{t+1}}{\overline{w}})}.$$

Appendix 8: Proof of proposition 3.7

The proof is based on the one in Mulligan and Sala-i-Martin (1999), with clarifying extensions.

We assume an interior solution for l^o and suppose the old enjoy less leisure than the young, $l^o < l^y$, which is the opposite of proposition 3.6. By the lemma and the properties of f, $l^o < l^y$ is true if and only if $\tau^o < \tau^y$ and if and only if $f < 0$. Now, we insert assumption (3.17) into the equation for the old group's optimal tax (3.6) and compare it to the young group's optimal tax (3.12):

$$\tau^o = \frac{\overline{w}}{w^o}\frac{\partial f}{\partial \bar{l}^o} > \frac{\overline{w}}{w^o}\left[\frac{1}{1+g} - \rho R\right]\left|\frac{\partial f}{\partial \bar{l}^y}\right| \geq -\frac{\overline{w}}{w^y}(1 - \rho R\frac{\overline{w}_{t+1}}{\overline{w}})\frac{\partial f}{\partial \bar{l}^y} = \tau^y. \qquad \text{(A8.1)}$$

Since $\left|\partial f / \partial \bar{l}^y\right| = -\partial f / \partial \bar{l}^y > 0$, $\overline{w}_{t+1} / \overline{w} = (1+g)$ and $w^y / w^o \geq (1+g)$ by assumption (3.16) and the definition of the steady state political equilibrium in section 3.1.2, we can simplify the second inequality to:

$$\frac{w^y}{w^o}\left[\frac{1}{1+g} - \rho R\right] \geq (1+g)\left[\frac{1}{1+g} - \rho R\right] = (1 - \rho R(1+g)). \qquad \text{(A8.2)}$$

As (A8.2) is true, it follows that (A8.1) is also true, i.e. $\tau^o > \tau^y$. However, this contradicts the condition set by the lemma that $\tau^o < \tau^y$.

Therefore, as the above was showed to contradict itself, it must be that under the relevant assumptions (3.15), (3.16) and (3.17), there is retirement and social security, i.e. $\tau^o > \tau^y$, $l^y < l^o_{t+1} = l^o$, $f > 0$. □