

## ARTICLES

# INCOME INEQUALITY, MOBILITY, AND THE WELFARE STATE: A POLITICAL ECONOMY MODEL

**LUCA BOSSI**

*University of Pennsylvania*

**GULCIN GUMUS**

*Florida Atlantic University*  
and  
*IZA*

In this paper, we set up a three-period stochastic overlapping-generations model to analyze the implications of income inequality and mobility for demand for redistribution and social insurance. We model the size of two different public programs under the welfare state. We investigate bidimensional voting on the tax rates that determine the allocation of government revenues among transfer payments and old-age pensions. We show that the coalitions formed, the resulting political equilibria, and the demand for redistribution crucially depend on the level of income inequality and mobility.

**Keywords:** Redistribution, Mobility, Inequality, Structure-Induced Equilibrium

## 1. INTRODUCTION

In developed countries, old-age pensions and transfer payments to the working-age make up the core of the welfare state. Tanzi and Schuknecht (2000) provide a detailed analysis of the level and composition of social expenditures in OECD countries as well as a historical summary of the trends. They document that the composition of public spending has changed considerably in industrialized countries since the late 19th century. Until about the 1960s, expenditure growth largely resulted in an extension of government services and the formation of basic social security systems. Direct or real government spending played a central role, whereas cash transfers were of marginal significance. Starting in the early 1960s, however, most public spending growth has been generated by the expansion of social programs, which have often taken the form of cash transfers. They explain that, for the most part, this growth cannot be attributed to technical

We gratefully acknowledge helpful comments and suggestions received from Michele Boldrin, David Easley, David Kelly, Salvador Ortigueira, Pere Gomis-Porqueras, and Karl Shell on earlier versions. We also thank an anonymous referee for constructive comments. The usual disclaimers apply. Address correspondence to: Luca Bossi, Department of Economics, University of Pennsylvania, 160 McNeil Building, 3718 Locust Walk, Philadelphia, PA 19104-6297, USA; e-mail: boluca@econ.upenn.edu.

factors, such as declining government productivity or aging populations, but is brought about by political decisions that extended benefits into universal social programs.

In high-income industrialized countries, besides the traditional old-age pensions, there are a variety of income maintenance programs, such as the provisions targeting disability, sickness, unemployment, and injuries at work. Given this multiplicity of programs under one welfare state, it is important to distinguish among various policy instruments according to the degree to which they provide redistribution and social insurance. Some programs are designed specifically to target inequalities between the rich and the poor, whereas others aim to cope with variations in income over the life cycle. The seminal paper by Meltzer and Richard (1981) relates the overall size of the welfare state to the degree of income inequality across agents in a median voter model. They show that economies that display more unequal incomes will have larger welfare spending without differentiating between different income transfer programs. In this paper, we extend the Meltzer and Richard model in several directions.

First, we acknowledge the variety of programs under the welfare state and model two different programs simultaneously. One is the social security old-age pensions program, which is mainly an intergenerational redistribution tool. The second is a transfer payment program for working-age individuals that redistributes income intragenerationally. We study the sizes of these two programs under the welfare state by employing a political economy model with multidimensional voting. Moene and Wallerstein (2001) claim that income inequality may shape alternative welfare programs differently; they argue that political support for redistribution and social insurance depends critically on the groups to which the benefits are targeted. Simultaneous modeling of the two programs provides the opportunity to capture such dynamics.

Second, in our model, income is determined endogenously through labor supply and we allow for distortionary effects of taxation. Third, for a given income distribution, we study the effects of income mobility, in addition to the effects of income inequality, on the level of redistributive taxation. In a recent paper, Kopczuk et al. (2010) use social security administrative data on earnings to study inequality and mobility in the United States since 1937. They report that, since the 1950s, earnings inequality has been increasing. They show that short-term mobility, defined as year-to-year mobility, has been remarkably stable during the same period, for a variety of measures. However, they find that long-term mobility measures, reflecting mobility over the stages of the working life, have substantially increased. They conclude that a comprehensive analysis of disparity requires studying both inequality and mobility. To the best of our knowledge, our study is the first to simultaneously model the effects of income inequality and mobility on the demand for redistribution in a bidimensional political economy framework. Using a three-period overlapping-generations model, we investigate voting on the tax rates that determine the allocation of tax revenues between these two social programs. Not only does the welfare state play a significant role in

equalizing incomes across groups and over lifetimes, but also income inequality and mobility have crucial implications for redistributive policies.

We are not the first to analyze the relationship between mobility and redistributive taxation in a political economy framework. Lindbeck (1985) shows that, under certain conditions, even some poor individuals may optimally choose not to vote for redistributive policies if they expect to climb up the income ladder sometime in the future. This argument explains why we do not observe full expropriation schemes in reality. Similarly, Piketty (1995) shows that personal mobility experience shapes voters' preferences for redistributive taxation. He claims that differences in perception about social mobility may generate differences in welfare policies across countries. Bénabou and Ok (2001) formalize these insights with their so-called "prospect of upward mobility" (POUM) hypothesis. However, the empirical evidence they provide indicates that the expected income gains might be too small compared to the risks of downward mobility. This suggests that, at least in the United States, demand for social insurance dominates the POUM mechanism.

In a series of papers, Alberto Alesina and his co-authors empirically explore the relationship between the size of the welfare state and individual attitudes toward income inequality and mobility [Alesina et al. (2004); Alesina and Angeletos (2005); Alesina and La Ferrara (2005)]. They claim that, at equilibrium, perceptions of income mobility can actually be sustained as self-fulfilling social beliefs that affect preferences for redistributive policies. These papers use empirical tools to study the observed differences in size between European and U.S.-type welfare states. The intuition goes as follows: if Americans perceive themselves as living in a more mobile society where citizens have more equal access to opportunities, then, at equilibrium, they will display a high effort level and demand less redistribution. Conversely, in Europe, individuals may believe that those at the bottom of the income ladder just happen to be "unlucky," leading to high levels of redistribution and thus to distorted effort level. Therefore, in the absence of altruistic preferences, income mobility may be perceived as a substitute for social insurance.

Alesina and co-authors, among others, have suggested that popular support for redistributive programs may also arise as a result of altruistic motives. If the high-income groups in the economy have a taste for fairness or social justice, they may be willing to accept redistributive taxation. However, it is difficult to argue that popular support for large welfare programs is primarily motivated by altruism [Lindbeck (1985)]. Individuals' attitudes towards redistributive programs may stem uniquely from self-interested concern rather than a pure taste for fairness. There is indeed empirical evidence that the rich do tend to vote for transfers when they are uncertain about their future position in the income distribution, and, once the positions are established, the main motive for supporting a welfare state is self-interest [Ravallion and Lokshin (2000); Beckman et al. (2004)].

In light of this discussion, we pose the following question: if the agents are strictly selfish, how can we explain popular support for inter- and intragenerational redistribution? We suggest that support for redistributive programs may reflect the

self-interest of the poor as well as the rich. In fact, we are able to show that when income mobility is higher, holding everything else constant, the rich demand higher transfer payments whereas the poor demand less. This is because, in our political economy model, the decisive voter for the transfer payments program is a young individual for whom uncertainty about lifetime income is a crucial factor in determining his preferences for redistribution. Among the more productive agents, the younger cohort demand higher transfer payments because they are at risk of moving down the income ladder (prospect of downward mobility). A pension program that provides mostly intergenerational redistribution, on the other hand, does not respond to changes in expected income because the social security pension program is supported by the middle-aged and elderly in our economy, who have no uncertainty regarding their future position in the income distribution.

The existing literature on the effects of income mobility on redistributive taxation considers social spending as a whole and does not differentiate among various types of public spending. Our paper makes an important point by addressing the fact that many different redistributive programs are provided simultaneously. Thus, a more interesting question posed here is what the preferred policy mixture would be. The idea is basically that not only the extent of redistributive taxation but also the optimal policy mix between diverse redistribution instruments may depend on income mobility and other features of the income distribution. Hence, our work contributes to the literature by investigating the effect of income inequality and mobility on the demand for two separate public social programs.

A vast literature analyzes social security and other redistributive programs under the welfare state using political economy models.<sup>1</sup> Although some of these studies do consider multiple issues voted upon by majoritarian rule, their focus is not the effect of income mobility on redistributive taxation.<sup>2</sup> In order to find the equilibria for the multidimensional voting game, we use the so-called “structure induced equilibrium” (SIE) concept. This concept was introduced by Shepsle (1979) and has been applied in the context of social security by Vincenzo Galasso and Ignacio Conde-Ruiz in several papers [e.g., Conde-Ruiz and Galasso (2004, 2005); Galasso (2008)]. The most relevant study for our work is Conde-Ruiz and Galasso (2005), in which they model individual preferences over the generosity of the social security system jointly with individual preferences over a transfer payment program. They show that demand for each of these components depends on the proportion of the elderly and income inequality. Their findings indicate that coalitions may be formed based on the dimension of age as well as income. They conclude that, in order to investigate the effect of income inequality on the welfare system, one needs to study income distribution by age groups.

We extend the framework of Conde-Ruiz and Galasso (2005) with two redistributive instruments to a three-overlapping-generations model in which we incorporate income mobility by introducing individual shocks to labor productivity. We also improve upon their model by endogenizing labor supply and saving decisions. Furthermore, we allow consumption to take place in all three periods as opposed to only during individuals’ working life. The government taxes the labor

income of young and middle-aged individuals and uses these resources to finance lump-sum transfers. Conde-Ruiz and Galasso (2005) define the intragenerational component of the welfare system as a transfer distributed to every working group irrespective of their income. In our model, intergenerational transfers take the form of pensions to the old whereas intragenerational transfer payments are made available to those with low incomes; thus, the latter display a stronger social insurance feature. Consistent with the existing literature, our results indicate that the size of the welfare system depends on both the age and income dimensions. Our analysis offers the additional insight that income mobility plays a crucial role in determining the political outcomes in a multidimensional policy space. Examination of the political equilibria in this economy reveals that coalitions around the policy issues are possibly formed along three dimensions simultaneously: age, income, and future income prospects.

The rest of the paper is organized as follows: in Section 2, we introduce the model and describe the economic equilibrium; in Section 3, we present the voting game and then characterize political equilibria based on the SIE approach; in Section 4, we discuss the main results of our model; and finally, Section 5 concludes. All the proofs and derivations are in the Appendix.

## 2. ECONOMIC MODEL

The underlying economic model is a modification of the Meltzer and Richard (1981) small open economy producing a single commodity, as outlined in Persson and Tabellini (2000). Consider a stochastic overlapping-generations framework where countably many individuals live for three equally long periods. We label periods of life as youth (1), middle-age (2), and old-age (3). For tractability, the focus is on an economy with a linear storage technology. Time is discrete and goes from 0 to  $\infty$ . For the sake of simplicity, we assume no population growth ( $n = 0$ ).<sup>3</sup> Therefore, in any given period  $t$ , there are equal numbers of old, middle-aged, and young individuals. Individuals work in the first two periods of their lives and retire in the last. The labor supplied by the young and the middle-aged is denoted by  $l_1$  and  $l_2$ , respectively. The elderly do not work ( $l_3 = 0$ ) and consume their retirement pension and asset income. A simple storage technology makes it possible to transfer the consumption good from one period to another: a unit of the consumption good today is transformed into  $R = (1 + r)$  units of tomorrow's consumption good.  $R$  is the interest factor and  $r$  is the exogenously given rate of return on capital. Individuals can buy or sell assets in the capital market at this interest rate. We assume that the economy is dynamically efficient, i.e.,  $r > n$ .

Individuals differ not only in terms of age but also in terms of their productivity. Individual  $i$ , in his first period of life, may be of a high ( $j = H$ ) or a low type ( $j = L$ ). The L-types are less productive than the H-types, which is formalized as observable differences in "effective time" endowments ( $e_H > e_L > 0$ ). Moreover, productivity is subject to an idiosyncratic shock over the life cycle, and the uncertainty about lifetime income is resolved as the agent ages. The productivity shock

is observed at the beginning of the second period of life and follows transitional probabilities according to the Markov transition matrix:

$$\begin{array}{cc}
 & \begin{array}{cc} e_L & e_H \end{array} \\
 \begin{array}{c} e_1 \\ e_2 \end{array} & \begin{array}{cc} \pi_1 & 1 - \pi_1 \\ 1 - \pi_2 & \pi_2 \end{array}
 \end{array}$$

This matrix indicates that the conditional probability of being L-type (H-type) when middle-aged for an agent who was L-type (H-type) when young is  $\pi_1$  ( $\pi_2$ ). Formally,  $\Pr\{e_2 = e_j | e_1 = e_j\} = \pi_d$ , and  $\Pr\{e_2 = e_{-j} | e_1 = e_j\} = 1 - \pi_d$  with  $j = L, H, d = 1$  if  $j = L$ , and  $d = 2$  if  $j = H$ .

We assume that  $\pi_1 > \pi_2 > 1/2$ : the productivity level over time is persistent and the persistence is stronger for the less endowed. Within each young and middle-aged generation, a constant proportion,  $\lambda$ , is assumed to be L-type and the rest  $(1 - \lambda)$  are H-type. The average level of productivity for each generation is given by  $\bar{e} = \lambda e_L + (1 - \lambda)e_H$ . Because we assume a stationary distribution of income (i.e., the probabilities are constant over time), it follows that, at the steady state,

$$\lambda = \frac{1 - \pi_2}{(1 - \pi_1) + (1 - \pi_2)}$$

with  $\lambda > 1/2$ . Therefore, income distribution is right-skewed; i.e., median income is lower than mean income, which is consistent with observed empirical regularities.<sup>4</sup> The ratios  $\frac{e_L}{\bar{e}}$  and  $\frac{e_H}{\bar{e}}$  are both measures of income inequality in this economy. We assume a constant real wage rate for both generations and normalize it to one. Total labor income in each period  $t$  is  $\Psi_t = \Psi_{1,t} + \Psi_{2,t}$ , where  $\Psi_{1,t}$  and  $\Psi_{2,t}$  represent total labor income for the young and the middle-aged, respectively. Note that total labor income is stationary over time and is equal to total labor supply because of the normalization of wages.

The government plays a purely redistributive role. We consider a welfare system that is composed of two policies, both financed through proportional taxes that are levied only on those who supply labor at time  $t$ , i.e., the young and the middle-aged. The first program distributes lump-sum transfer payments,  $B$ , to the working L-types. This component of the welfare state is financed through a tax rate denoted by  $\tau$  and is an instrument to redistribute income intragenerationally, from the H-types to the L-types. Income maintenance programs, such as public assistance in the United States including unemployment insurance, Temporary Assistance for Needy Families (TANF), Supplemental Security Income (SSI), and food stamps, insure individuals against large income reductions due to risks associated with human capital and random events [Hungerford (2008)]. Following Lindbeck (1985) and Bénabou (2000), we assume that private insurance markets are absent. Given the variability of lifetime income, the transfer policy offers social insurance and protects against risks that private insurance markets fail to cover.

As the second component of the welfare state, members of the old generation are entitled to old-age pensions,  $P$ , that are supported via the tax rate  $\theta$ . The pension system plays two roles: it is a means of both intra- and intergenerational redistribution. Pensions take the form of a lump-sum payment, and thus are modeled as a Beveridgean social security system. As Conde-Ruiz and Profeta (2007) point out, in such social security systems with a weak relationship between individuals' contributions and pensions, there is a relatively higher degree of redistribution within cohorts. In contrast, in a Bismarckian social security system, pension benefits are strongly linked to contributions, and thus the degree of within-cohort redistribution is relatively lower.<sup>5</sup> Obviously, these aspects of the welfare state have implications for political equilibria, inducing voters to form coalitions along the dimensions of age and productivity. In every period, the welfare state can be described by a vector  $(B, P)$ : transfers to the young and middle-aged, and pensions for the elderly. The government budgets are balanced every period, and they can be written in per capita terms as

$$B_t = \frac{\tau_t \Psi_t(\tau_t, \theta_t)}{2\lambda}, \tag{1}$$

$$P_t = \theta_t \Psi_t(\tau_t, \theta_t). \tag{2}$$

(1) and (2) imply bidimensional voting on the tax rates  $\tau_t$  and  $\theta_t$ , which in turn determine the transfers  $B_t$  and the pensions  $P_t$ .

Preferences are additively time-separable and agents value both periodic consumption,  $c$ , and leisure,  $x$ . Consumption in the first two periods of life enters into a logarithmic utility function,  $U(c) = \log(c)$ , and linearly as  $U(c_3) = c_3$  when old. As in Persson and Tabellini (2000), the linearity of the utility function in old-age consumption is imposed purely for convenience. This simplifies our computations by allowing all income effects to be absorbed by old-age consumption. The young and the middle-aged allocate their disposable income between current consumption and savings. The budget constraint for the young in period  $t$  is

$$c_{1j,t} = \begin{cases} l_{1j,t}(1 - \tau_t - \theta_t) - s_{1j,t} + B_t & \text{if } j = L \\ l_{1j,t}(1 - \tau_t - \theta_t) - s_{1j,t} & \text{if } j = H, \end{cases} \tag{3}$$

where we denote the consumption, labor supply, and savings of a  $j$ -type young individual in period  $t$  by  $c_{1j,t}$ ,  $l_{1j,t}$ , and  $s_{1j,t}$ , respectively. Similarly, the budget constraint of the middle-aged in period  $t$  for individual consumption can be written as

$$c_{2k,t} = \begin{cases} l_{2k,t}(1 - \tau_t - \theta_t) - s_{2k,t} + (1 + r)s_{1j,t-1} + B_t & \text{if } k = L \\ l_{2k,t}(1 - \tau_t - \theta_t) - s_{2k,t} + (1 + r)s_{1j,t-1} & \text{if } k = H, \end{cases} \tag{4}$$

where we denote the consumption, labor supply, and savings of a  $k$ -type middle-aged individual in period  $t$  by  $c_{2k,t}$ ,  $l_{2k,t}$ , and  $s_{2k,t}$ , respectively. Finally, for the elderly, we have  $c_{3,t} = (1 + r)s_{2k,t-1} + P_t$ . Preferences over leisure are represented by  $W(x) = \alpha \log(x)$ , where  $\alpha$  is a utility parameter. Young and middle-aged

individuals are endowed with  $(1 + e_{j,t})$  unit of effective time to be divided between leisure and labor supply, so the time constraint is given by

$$x_{j,t} = 1 + e_{j,t} - l_{j,t}, \quad j = L, H. \tag{5}$$

The timing of the model is as follows: each period, agents make their labor supply and consumption decisions, knowing their current but not their future types. Next, people vote on the policy vector  $(\tau, \theta)$ . Finally, taxes are paid and consumption takes place. This timing implies that the young and the middle-aged differ not only in terms of age but also in terms of information concerning their position in the future income distribution. Young individuals face uncertainty and thus need to form expectations of future income conditional on their current positions. The elderly have no economic decision to make, as they consume their entire wealth based on pensions together with their lifetime savings and leave no bequests.

The young agents' problem at time  $t$  is choosing the optimum labor supply and consumption to maximize expected utility. Formally, the economic problem for the young L-type is

$$\begin{aligned} \max_{c,l} & U(c_{1L,t}) + W(x_{1L,t}) + \beta\{\pi_1[U(c_{2L,t+1}) \\ & + W(x_{2L,t+1})] + (1 - \pi_1)[U(c_{2H,t+1}) + W(x_{2H,t+1})]\} + \beta^2 U(c_{3,t+2}), \end{aligned} \tag{6}$$

where  $\beta \in (0, 1)$  is the individual discount factor, and we assume, for simplicity, that it is equal to the inverse of the interest factor:  $\beta = 1/R$ . The young H-type agent maximizes the following expected utility:

$$\begin{aligned} \max_{c,l} & U(c_{1H,t}) + W(x_{1H,t}) + \beta\{\pi_2[U(c_{2H,t+1}) \\ & + W(x_{2H,t+1})] + (1 - \pi_2)[U(c_{2L,t+1}) + W(x_{2L,t+1})]\} + \beta^2 U(c_{3,t+2}). \end{aligned} \tag{7}$$

Middle-aged agents face the following problem:

$$\max_{c,l} U(c_{2k,t}) + W(x_{2k,t}) + \beta U(c_{3,t+1}) \quad \text{where } k = L, H. \tag{8}$$

It can be shown that there is a unique solution to agents' problem (see Appendix A.2 for the derivation of economic equilibrium). Utility maximization yields the optimal level of consumption  $c_1^* = c_2^* = 1$  because of the quasi-linear utility specification. Optimal labor supply is given by  $l_{1,t}^* = l_{2,t}^* = 1 + e_{j,t} - \frac{\alpha}{1 - \tau_t - \theta_t}$ , where  $j = L, H$ .<sup>6</sup> Total labor income, in equilibrium, is given by

$$\Psi(\tau_t, \theta_t) = 2(1 + \bar{e}) - \frac{2\alpha}{1 - \tau_t - \theta_t}. \tag{9}$$



Taxes distort labor supply for all groups; thus higher taxes imply a decrease in aggregate labor supply. The aggregate feasibility constraint in this economy is

$$\begin{aligned}
 & (1 - \tau_t - \theta_t)\{\lambda l_{2L,t} + (1 - \lambda)l_{2H,t}\} + [\lambda l_{1L,t} + (1 - \lambda)l_{1H,t}] \\
 & + (\tau_t + \theta_t)\{\lambda l_{2L,t} + (1 - \lambda)l_{2H,t}\} + [\lambda l_{1L,t} + (1 - \lambda)l_{1H,t}] \\
 & = c_{3,t} + [\lambda c_{2L,t} + (1 - \lambda)c_{2H,t}] + [\lambda c_{1L,t} + (1 - \lambda)c_{1H,t}].
 \end{aligned}
 \tag{10}$$

We are now ready to define the economic equilibrium for our economy.

**DEFINITION 1.** *For a given sequence of policies  $\{\tau_t, \theta_t\}_{t=0}^\infty$ , an economic equilibrium is a set of allocations  $\{c_{1j,t}^*, c_{2k,t}^*, x_{1j,t}^*, x_{2k,t}^*, l_{1j,t}^*, l_{2k,t}^*\}_{t=0}^\infty$  such that (i)  $c_{1j,t}^*$  and  $l_{1j,t}^*$  solve the problem of a  $j$ -type young agent in period  $t$ , and  $c_{2k,t}^*$  and  $l_{2k,t}^*$  solve the problem of a  $k$ -type middle-aged agent in period  $t$ ; (ii)  $x_{1j,t}^*$  and  $x_{2k,t}^*$  are determined from (3), (4), and (5); (iii) the government budget constraints in (1) and (2) are satisfied; and finally, (iv) feasibility as in (10) holds.*

In the rest of the paper, we consider the voting outcomes for this model and study how the political equilibria relate to income inequality and mobility.

### 3. VOTING GAME

The policy to be voted upon is a bidimensional vector. This violates one of the main assumptions of the median voter theorem, and thus it is well known that a majority voting equilibrium may fail to exist.<sup>7</sup> The implausibility of the assumption that the issue space is unidimensional has led to alternative methods of aggregating preferences such as legislative bargaining, probabilistic voting, and interest group models.<sup>8</sup> We assume that public social programs are decided through a majoritarian voting procedure using the SIE concept. There is a process that aggregates agents’ preferences over two issues (tax rates  $\tau$  and  $\theta$ , in our case), where political structures limit voting to only one dimension at a time.

We prefer the SIE framework to alternative approaches because it allows median voters’ identities to differ in each policy dimension. As Lindbeck (1985) points out, the redistribution programs are all fragmented, and each group in the economy has different interests in different components of the welfare state. This fragmentation basically allows politicians to extract votes from each specific group by unbundling the issues: the identity of the median voter can be different for determining separate components of the welfare state.

Elections take place in every period, and all economic agents, regardless of age and income, are involved in the voting process. All agents vote simultaneously and separately on  $\tau$  and  $\theta$ . Given that no individual alone can influence the outcome of the political game, we assume sincere voting. In other words, voters choose their ideal policies according to their true preferences, independent of any strategic consideration. To analyze the political equilibria, we focus on the voting game in a static environment with commitment as in Bénabou and Ok (2001). Current

voters, in this case, are assumed to commit to future policies in a once-and-for-all voting game.<sup>9</sup>

We analyze the political problem of seven voting groups in this economy: young L-type (L), young H-type (H), middle-aged L-type who was a L-type when young (LL), middle-aged H-type who was an L-type when young (LH), middle-aged H-type who was an H-type when young (HH), middle-aged L-type who was an H-type when young (HL), and finally the old (O). Obviously, three overlapping generations, as opposed to two, display more interesting coalition formation possibilities. In what follows, we drop the subscript  $t$  for ease of exposition unless necessary. The indirect utility functions,  $V(\tau, \theta)$ , for each of the groups whose political preferences we analyze are

$$\begin{aligned}
 V_L &= \log\{l_L^*(1 - \tau - \theta) + B\} + \beta\pi_1 \log\{l_{LL}^*(1 - \tau - \theta) + B\} \\
 &\quad + \beta(1 - \pi_1) \log\{l_{LH}^*(1 - \tau - \theta)\} + \beta^2 P + \alpha \log\{1 + e_L - l_L^*\} \\
 &\quad + \beta\pi_1\alpha \log\{1 + e_L - l_{LL}^*\} + \beta(1 - \pi_1)\alpha \log\{1 + e_H - l_{LH}^*\}, \\
 V_H &= \log\{l_H^*(1 - \tau - \theta)\} + \beta\pi_2 \log\{l_{HH}^*(1 - \tau - \theta)\} \\
 &\quad + \beta(1 - \pi_2) \log\{l_{HL}^*(1 - \tau - \theta) + B\} + \beta^2 P + \alpha \log\{1 + e_H - l_H^*\} \\
 &\quad + \beta\pi_2\alpha \log\{1 + e_H - l_{HH}^*\} + \beta(1 - \pi_2)\alpha \log\{1 + e_L - l_{HL}^*\}, \\
 V_{LL} &= \log\{l_{LL}^*(1 - \tau - \theta) + B\} + \beta P + \alpha \log\{1 + e_L - l_{LL}^*\}, \\
 V_{LH} &= \log\{l_{LH}^*(1 - \tau - \theta)\} + \beta P + \alpha \log\{1 + e_H - l_{LH}^*\}, \\
 V_{HH} &= \log\{l_{HH}^*(1 - \tau - \theta)\} + \beta P + \alpha \log\{1 + e_H - l_{HH}^*\}, \\
 V_{HL} &= \log\{l_{HL}^*(1 - \tau - \theta) + B\} + \beta P + \alpha \log\{1 + e_L - l_{HL}^*\}.
 \end{aligned}$$

We examine individual preferences over the space  $(\tau, \theta)$  issue by issue. Following Shepsle (1979), we assume that agents express their preferences over  $\tau$  for a given level of  $\theta$  and then over  $\theta$  for a given level of  $\tau$ . With this Cournot-like procedure, we first rank the preferences and identify the median outcomes along these two policies,  $\tau^m$  and  $\theta^m$ . Then, using the systems of reaction functions, we can identify the SIE of the voting game with commitment. We provide a full and formal characterization of the SIE with commitment in Section 3.3.

**DEFINITION 2 (SIE).** *A structure induced equilibrium with commitment is a vector of policies  $(\tau^m, \theta^m)$  such that  $\tau_{t+1}^m = \tau_t^m = \tau^m, \theta_{t+1}^m = \theta_t^m = \theta^m \forall t$  and*

- (i) *when the pension tax rate expected to be chosen is  $\theta^m$ ,  $\tau^m$  is a Condorcet winner when voting over the transfer taxes, and*
- (ii) *when the transfer tax rate expected to be chosen is  $\tau^m$ ,  $\theta^m$  is a Condorcet winner when voting over the pension component.*

If preferences are single-peaked for each policy, then a SIE exists and is a median in both directions. Thus, before proceeding further with our model, we need to ensure the existence of a SIE.

LEMMA 1 (Single-peaked preferences). *For a given  $\tau$ , preferences over  $\theta$  are single-peaked, and for a given  $\theta$ , preferences over  $\tau$  are single-peaked.*

Proof. See Appendix A.4. ■

The lemma establishes that, given our assumptions, the conditions in Theorem 3.1 of Shepsle (1979) are satisfied for both  $\tau$  and  $\theta$ .

### 3.1. Voting on Social Security Pensions

The old vote to have the tax revenue–maximizing tax rate [i.e.,  $\text{argmax}_\theta \theta\Psi(\tau, \theta)$ ]. They are a homogeneous group as far as voting is concerned: there is no political distinction between old L-types and old H-types because heterogeneity does not affect their voting decisions. By setting the tax rate to the maximum possible, old retirees are just maximizing their consumption, because we assume away any form of altruism. Furthermore, because they do not work, they do not pay taxes, so there is no cost for them in setting a high  $\theta$  to finance the pension system. In other words, retirees want to have this tax rate as high as possible because it does not impact their previously paid contributions, but it does affect their pensions. This reflects the widely known argument of the “single-mindedness” of the elderly when they vote on social security [Mulligan and Sala-i-Martin (1999)].

For young and middle-aged individuals, the choice is not as simple. It depends on several factors: First of all, the tax rates distort their labor supply decisions. Recall from the previous section that  $l_{LL}^* = l_{HL}^* = l_L^*$  and  $l_{HH}^* = l_{LH}^* = l_H^*$  because both pensions and transfers are lump-sum and are not contributory. Thus, the magnitude of tax distortion in each period is the same, regardless of age or type.<sup>10</sup> Their political preferences, instead, differ according to their current and expected positions in the income distribution. Formally, the young will choose the social security tax rate that maximizes their expected intertemporal indirect utility function. The preferred tax rate for the L-type young,  $\theta_L^*$ , is determined by equating the marginal benefit to the marginal cost of taxation (see Appendix A.5),

$$\frac{\partial V_L}{\partial \theta} = \beta^2 (\Psi + \theta\Psi_\theta) + (1 + \beta\pi_1) \frac{\tau\Psi_\theta}{2\lambda} - (1 + \beta\pi_1)l_L^* - \beta(1 - \pi_1)l_H^* = 0, \quad (11)$$

where  $\Psi_\theta$  denotes the derivative of the total labor income with respect to the pension tax. Analogously, for the H-type young,  $\theta_H^*$  is defined by

$$\frac{\partial V_H}{\partial \theta} = \beta^2 (\Psi + \theta\Psi_\theta) + \beta(1 - \pi_2) \frac{\tau\Psi_\theta}{2\lambda} - (1 + \beta\pi_2)l_H^* - \beta(1 - \pi_2)l_L^* = 0. \quad (12)$$

The terms in these two expressions can be interpreted as follows: increasing  $\theta$  brings a future benefit when old [ $\beta^2(\Psi + \theta\Psi_\theta) > 0$ ]. All other terms in (11) and

(12) imply a cost in the first two periods of life due to social security taxation. Note that these costs are different between the young L-type and the young H-type depending on their current income levels as well as their heterogeneous income prospects.

The preferred pension tax rates for the middle-aged types are given by the following equations:

$$\frac{\partial V_{LL}}{\partial \theta} = \frac{\partial V_{HL}}{\partial \theta} = \beta (\Psi + \theta \Psi_{\theta}) + \frac{\tau \Psi_{\theta}}{2\lambda} - l_L^* = 0, \quad (13)$$

$$\frac{\partial V_{HH}}{\partial \theta} = \frac{\partial V_{LH}}{\partial \theta} = \beta (\Psi + \theta \Psi_{\theta}) - l_H^* = 0. \quad (14)$$

The pension benefit outlook is better for the middle-aged than for the young because of the proximity of the retirement period ( $\beta > \beta^2$ ). In addition, the costs associated with a higher tax rate are lower for the middle-aged than for the young because the middle-aged pay taxes for one last period only. Because pensions are not contributory, the preferred  $\theta$  for the middle-aged does not depend on past productivity. This reflects the sunk-cost character of the first period payment into the pension program. Although it is not possible to obtain a tractable closed-form solution for the preferred pension tax, it is still possible to obtain a ranking. The following lemma establishes how we can rank the ideal social security pension tax across all agents in this economy.

LEMMA 2 (Ranking of  $\theta$ ).

$$0 \leq \theta_H^* < \theta_L^* < \theta_{HH}^* = \theta_{LH}^* < \theta_{LL}^* = \theta_{HL}^* < \theta_O^*.$$

Proof. See Appendix A.6. ■

This lemma formally ranks all the pension tax rates. In particular, we find that coalitions are formed along the dimension of age when we consider the pension program. Within both the young and the middle-aged groups, the L-types favor higher tax rates than the H-types because the former are net recipients from the pension program. The H-type individuals are better at saving privately, because we assume that the economy is dynamically efficient and the taxes are distortive. Now we are ready to establish the median among the preferred pension tax rates.

PROPOSITION 1 (Median voter over  $\theta$ ). *The median voter over the issue  $\theta$  is  $\theta^m = \theta_{LL}^* = \theta_{HL}^*$ .*

Proof. See Appendix A.7. ■

Thus, the decisive voter over the pensions is a middle-aged L-type agent. This finding indicates that coalitions based on age and income arise as in Conde-Ruiz and Galasso (2005).

3.2. Voting on Transfer Payments

From the government budget constraints (1) and (2), it is clear that, for a given  $\theta$ , the transfer payments program will not benefit the old. In fact, this program would reduce their pensions by distorting the labor supply and decreasing the government revenues allocated to the pension program. Therefore, the elderly oppose any transfer payments for the working agents, resulting in  $\tau_0^* = 0$ . Similarly, the middle-aged H-types, both LH and HH, are net contributors to such a transfer program. They would rather have taxes totally allocated to  $P$  rather than to  $B$ , yielding  $\tau_{HH}^* = \tau_{LH}^* = 0$ . The rest of the agents, on the other hand, benefit from this intragenerational transfer scheme, especially the young who also value the social insurance aspect of this program. How much they benefit depends on their current type as well as on the income mobility prospects. By maximizing the expected indirect utility functions of the young, we obtain the following two first-order conditions that define the preferred welfare transfer taxes  $\tau_L^*$  and  $\tau_H^*$ , respectively (see Appendix A.5):

$$\begin{aligned} \frac{\partial V_L}{\partial \tau} &= (1 + \beta\pi_1) \frac{(\Psi + \tau\Psi_\tau)}{\lambda(1+n)(2+n)} + \beta^2\theta\Psi_\tau - (1 + \beta\pi_1)I_L^* \\ &- \beta(1 - \pi_1)I_H^* = 0, \end{aligned} \tag{15}$$

$$\begin{aligned} \frac{\partial V_H}{\partial \tau} &= \beta(1 - \pi_2) \frac{(\Psi + \tau\Psi_\tau)}{\lambda(1+n)(2+n)} + \beta^2\theta\Psi_\tau - (1 + \beta\pi_2)I_H^* \\ &- \beta(1 - \pi_2)I_L^* = 0, \end{aligned} \tag{16}$$

where  $\Psi_\tau$  is the derivative of the total labor income with respect to the welfare transfer tax.<sup>11</sup> The ideal  $\tau$  for the middle-aged is defined by

$$\frac{\partial V_{LL}}{\partial \tau} = \frac{\partial V_{HL}}{\partial \tau} = \frac{(\Psi + \tau\Psi_\tau)}{\lambda(1+n)(2+n)} + \beta\theta\Psi_\tau - I_L^* = 0, \tag{17}$$

$$\frac{\partial V_{HH}}{\partial \tau} = \frac{\partial V_{LH}}{\partial \tau} = \beta\theta\Psi_\tau - I_H^* < 0. \tag{18}$$

The following lemma provides a comparison of preferred transfer taxes across groups.

LEMMA 3 (Ranking of  $\tau$ ).

- (I) If  $e_H > E$ , then  $0 = \tau_0^* = \tau_{HH}^* = \tau_{LH}^* < \tau_H^* < \tau_L^* < \tau_{LL}^* = \tau_{HL}^*$ , and
- (II) if  $e_H < E$ , then  $0 = \tau_0^* = \tau_{HH}^* = \tau_{LH}^* < \tau_H^* < \tau_{LL}^* = \tau_{HL}^* < \tau_L^*$ ,  
where

$$E = \frac{2\alpha\theta [1 - \beta(1 - \pi_1)] + (1 - \pi_1)(1 - \tau - \theta)(\alpha - 1 + \tau + \theta)}{(1 - \pi_1)(1 - \tau - \theta)^2}.$$

Proof. See Appendix A.8. ■

The young agents, particularly the L-types, profit from the risk-sharing component of this program, resulting in  $\tau_L^* > \tau_H^*$ . The prospect of upward income mobility is small for the L-types, so their demand for redistribution is higher. The income mobility prospects induce the risk-averse young to partially smooth their expected income-level differences with the help of the transfers. The groups LL and HL, on the other hand, are net receivers of transfer payments and thus demand a larger transfer payments program. It is not possible to unequivocally rank the relationship between  $\tau_{LL}^*$  and  $\tau_L^*$ . The comparison depends on the level of the initial endowment  $e_H$ : if it is relatively high (low), the upward mobility prospects are relatively better (worse), which would indicate that the young L-types favor a smaller (larger) transfer program than the middle-aged L-types. It is now possible to establish the decisive voter over this issue.

PROPOSITION 2 (Median voter over  $\tau$ ). *The median voter over  $\tau$  is*

- (I)  $\tau^m = \tau_H^*$  if  $\lambda < \frac{3}{4}$ ,
- (II.a)  $\tau^m = \tau_L^*$  if  $\lambda > \frac{3}{4}$  and  $e_H > E$ ,
- (II.b)  $\tau^m = \tau_{LL}^* = \tau_{HL}^*$  if  $\lambda > \frac{3}{4}$  and  $e_H < E$ .

Proof. See Appendix A.9. ■

The intuition behind this result is as follows: middle-aged H-types and the old are always a natural coalition against the intragenerational redistribution program. However, if the fraction of the population with low productivity is relatively small, young H-types join the young and middle-aged L-types, constituting the winner coalition in favor of this program. Most likely, the decisive voter over the size of this program is a young agent. The novelty here is that, in addition to age and income, the young’s uncertainty over future position motivates redistribution.

### 3.3. Political Equilibria

In this section, we describe the political equilibria. We have already characterized the political preferences along each separate dimension in Propositions 1 and 2. Because preferences are single-peaked and the reaction functions have the single-crossing property, we can apply Theorem 3.1 in Shepsle (1979). The next proposition aggregates the preferences over the two dimensions of the policy space and characterizes the SIE of the voting game with commitment.

PROPOSITION 3 (Characterization of the SIE). *A SIE is a vector of policies  $(\tau^m, \theta^m)$  such that*

- (i)  $(\tau^m, \theta^m) = (\tau_H^*, \theta_{LL}^*)$  if  $\lambda < \frac{3}{4}$ ,
  - (ii)  $(\tau^m, \theta^m) = (\tau_L^*, \theta_{LL}^*)$  if  $\lambda > \frac{3}{4}$  and  $e_H > E$ ,
  - (iii)  $(\tau^m, \theta^m) = (\tau_{LL}^*, \theta_{LL}^*)$  if  $\lambda > \frac{3}{4}$  and  $e_H < E$ ,
- where  $E$  is defined in Lemma 3.

Proof. See Appendix A.10. ■

This proposition indicates that multiple political equilibria emerge in this economy.<sup>12</sup> Coalitions around the policy issues are possibly formed along three dimensions simultaneously: age, income, and future income prospects. When the share of L-types is relatively high, a coalition of young and middle-aged L-types will emerge to try to obtain higher transfers. If the fraction of L-types is relatively small, the preferred policies of the coalition of young and middle-aged H-types will prevail. In the case of pension taxes, a coalition based mainly on age emerges. The elderly may have significant influence on political outcomes, because they strongly support pension provision. Nevertheless, the decisive voter in determining the size of the pension program is always a middle-aged L-type. In contrast, the decisive voter for the transfer program is more likely to be a young agent for whom income mobility is a critical issue. The fact that, for both programs, the median voter is either a young or middle-aged citizen is consistent with the empirical evidence [Galasso (1999); Koethenburger et al. (2008)].

#### 4. DISCUSSION

In this section, we provide two comparative statics exercises to investigate the properties of the political equilibria. The first illustrates the effects of a productivity-neutral increase in inequality. In other words, we consider an increase in  $e_H$  that is offset by a decrease in  $e_L$  such that the average productivity  $\bar{e}$  is unchanged. The second comparative statics exercise consists of a productivity-neutral increase in mobility. We consider the scenario in which both  $\pi_1$  and  $\pi_2$  decrease so that the fraction of low-productivity agents in the economy ( $\lambda$ ) remains constant. There are at least two reasons to focus on the effects of productivity-neutral changes in inequality and mobility on the SIE. One is to avoid any changes in overall productivity or labor income such that the tax base stays the same (except for the distortionary effects of taxation). Second, this approach allows us to keep the composition of the electorate unaltered because the population shares for each group remain constant.

The two comparative statics exercises we carry out in the following are linked. In our model, income inequality is defined by the effective time endowment differential. However, inequality may be more or less persistent depending on the level of income mobility in the economy. This is because the variation in labor incomes over an individual's lifetime depends not only on  $e_L$  and  $e_H$  but also on  $\pi_1$  and  $\pi_2$ . Therefore, an increase in mobility may imply a reduction in lifetime income inequality.<sup>13</sup> In what follows, we analyze the effect of changes in mobility and inequality on  $\tau^m$  and  $\theta^m$ . Given our earlier assumptions, the direct effect of changes in taxes on the government revenue dominates the behavioral effect of such changes on the labor supply. Therefore, the total transfer payment program size, as measured by  $\tau\Psi(\theta, \tau)$ , and the total pension program, as measured by  $\theta\Psi(\theta, \tau)$ , move in the same direction as the changes in  $\tau$  and  $\theta$ , respectively (see

note 9). In each case, we also discuss how an increase in inequality or mobility might result in a shift from one SIE to another.

**4.1. Increasing Inequality**

First, we consider an overall widening of inequality and analyze how the preferred tax rates change. We carry out a classical *ceteris paribus* exercise in the form of a mean-preserving spread in the productivity distribution. We introduce a decrease in  $e_L$  and an opposite contemporaneous increase in  $e_H$  such that  $\bar{e}$  is unaffected. For an arbitrarily small  $\varepsilon > 0$ , if  $e_H$  is increased to  $\tilde{e}_H = e_H + \varepsilon$ , then  $e_L$  has to be decreased to  $\tilde{e}_L = e_L - \varepsilon \frac{1-\lambda}{\lambda}$ , which implies that  $de_L = -(\frac{1-\lambda}{\lambda})de_H$ . Both ratios  $e_L/\bar{e}$  and  $e_L/e_H$  suggest an increase in income inequality in this economy. Let us first analyze the transfer payments program. We can write the indirect utilities defining the preferred  $\tau$  as implicit functions of this tax rate and the parameters  $e_L$  and  $e_H$ ,

$$V(\tau; e_L, e_H | \theta) \equiv F(\tau; e_L, e_H | \theta) = 0. \tag{19}$$

These implicit functions are defined by (15), (16), or (17), depending on the identity of the median voter. For any given  $\theta$ , total differentiation of these expressions yields

$$\frac{d\tau}{de_H} = \frac{-F_{e_H} + \left(\frac{1-\lambda}{\lambda}\right) F_{e_L}}{F_\tau}, \tag{20}$$

where the subscripts denote partial derivatives. The sign of this expression depends on the identity of the decisive voter (see Appendix A.11). If the proportion of the L-types ( $\lambda$ ) is relatively high, then the median voter will either be a young or middle-aged L-type who prefers a larger transfer-payments program as a result of increased inequality. If the proportion of the L-types is relatively low, then the median voter will instead be a young H-type who prefers a smaller transfer program. This is because the H-types are net contributors to this program, whereas the L-types are net recipients.

Next, we investigate the effect of widening inequality on the pension program. The indirect utilities defining the preferred  $\theta$  as implicit functions of this tax rate and the parameters  $e_L$  and  $e_H$  can be expressed as

$$V(\theta; e_L, e_H | \tau) \equiv G(\theta; e_L, e_G | \tau) = 0, \tag{21}$$

where the  $G$  function is derived from (13). Conditional on the equilibrium outcome for  $\tau$ , total differentiation of this expression yields

$$\frac{d\theta}{de_H} = \frac{-G_{e_H} + \left(\frac{1-\lambda}{\lambda}\right) G_{e_L}}{G_\theta}. \tag{22}$$

It turns out that higher inequality leads to a larger social security pension program because the LL group now faces even lower productivity, which means they are better off by supporting a larger pension system rather than saving privately (see



Appendix A.12). To summarize, depending on the relative size of the low-skilled population, a mean-preserving spread in the productivity distribution implies a larger pension program, whereas it induces a nonmonotonic effect on the transfer payments program. Given the logarithmic utility assumption, insurance is not a normal good in our model. Therefore, wider income inequality implies that the redistributive aspect of government programs dominates their insurance aspect.

### 4.2. Increasing Mobility

The second comparative statics exercise we consider is a productivity-neutral increase in income mobility. For an arbitrarily small  $\varepsilon > 0$ , if  $\pi_1$  is reduced to  $\tilde{\pi}_1 = \pi_1 - \varepsilon$ , then  $\pi_2$  has to be reduced to  $\tilde{\pi}_2 = \pi_2 - \varepsilon \frac{1-\pi_2}{1-\pi_1}$  so that  $\lambda$  remains unchanged. Therefore, we have  $d\pi_2 = \frac{1-\pi_2}{1-\pi_1} d\pi_1$ . For the purpose of the comparative statics exercises, we can write any indirect utility defining the preferred  $\tau$  as an implicit function of this tax rate and the parameters  $\pi_1$  and  $\pi_2$ :

$$V(\tau; \pi_1, \pi_2 | \theta) \equiv Q(\tau; \pi_1, \pi_2 | \theta) = 0. \tag{23}$$

These implicit functions are again defined by (15), (16), or (17). For a given  $\theta$ , total differentiation yields

$$\frac{d\tau}{d\pi_1} = \frac{-Q_{\pi_1} - \left(\frac{1-\pi_2}{1-\pi_1}\right) Q_{\pi_2}}{Q_{\tau}}, \tag{24}$$

where the subscripts denote partial derivatives. It is possible to show that this expression is negative if the median voter is a young H-type and positive if the median voter is a young L-type (see Appendix A.13). Thus, an increase in mobility would decrease the preferred level of the transfer payments program of the young L-type because higher mobility would encourage the L-type to demand less social insurance. The young H-type, on the other hand, would prefer a larger risk-sharing transfer payments program as a result of reduced persistence. This finding is in line with Bénabou and Ok (2001). In this case, neither the median nor the mean income is affected, yet the demand for social insurance changes simply because of a change in expected incomes.

To analyze how an increase in mobility affects the pension program, we express the indirect utility defining the preferred  $\theta$  as an implicit function of this tax rate and the parameters  $\pi_1$  and  $\pi_2$ :

$$V(\theta; \pi_1, \pi_2 | \tau) \equiv Z(\theta; \pi_1, \pi_2 | \tau) = 0, \tag{25}$$

where the  $Z$  function is derived from (13). Conditional on  $\tau$ , total differentiation of this expression yields

$$\frac{d\theta}{d\pi_1} = \frac{-Z_{\pi_1} - \left(\frac{1-\pi_2}{1-\pi_1}\right) Z_{\pi_2}}{Z_{\theta}}. \tag{26}$$

As expected, a change in mobility does not affect the size of the preferred social security pension program as determined by the LL group. In fact, changes in mobility would not affect the size of the preferred transfer payments program of this group either (see Appendix A.14). The intuition relies on the absence of complete information regarding their lifetime income. In sum, rising mobility affects the political support for redistribution to the extent that it reduces the inequality of lifetime incomes. Therefore, the results in this section are consistent with those reported in the previous one.

Finally, a few additional remarks are in order. Our comparative statics exercises abstract from any changes in the identity of the median voter due to changes in income inequality and mobility. Neither the changes in inequality nor those in mobility can affect the identity of the median voter over the pension program. The identity of the decisive voter over the transfer payment program, on the other hand, might change if the increase in  $e_H$  or the decrease in  $\pi_1$  were large enough. In particular, with increasing income mobility, the young L-type agent is more likely to be the median voter than the middle-aged L-type agent. This shift from one SIE to another is due to the change in the ordering of the preferred transfer payment tax rates for these two groups (see Lemma 3). The intuition is again based on future prospects: higher  $e_H$  and lower  $\pi_1$  both increase expected utility for the young L-type, but not for the middle-aged L-type.

### 4.3. Voting without Commitment

In the preceding analysis, the commitment assumption allows us to abstract from the complications created by the dynamic nature of the voting game. However, the political decisions of the young, the middle-aged, and the retirees may have reciprocal effects on each other. In particular, without a device to commit to future policies, young members may refuse to politically support a transfer of resources to older individuals because they are not guaranteed that this transfer policy will be kept in place in the future. Boldrin and Rustichini (2000), among others, make this point, which represents an issue to most voting models of intergenerational transfers. However, when the model includes uncertainties regarding future income positions, one needs to focus on voting with commitment to maintain the political outcomes. Bénabou and Ok (2001, p. 454) point this out, explaining that “mobility considerations can enter into voter preferences only if current policy has lasting effects.” They also provide a description of institutional motivations behind the policy commitment assumption.

The repeated voting framework would be the same voting game described earlier, but each period, voters would determine the current size of the redistributive programs while perceiving that their vote may shape future policy choices as well. Combining the game-theoretic insights offered by Boldrin and Rustichini (2000) on intergenerational implicit contracts with the equilibrium concept introduced by Shepsle (1979), Conde-Ruiz and Galasso develop the concept of stationary subgame perfect structure induced equilibrium in the context of social security models

(2003, 2004, 2005). The aim is not to characterize the full set of possible equilibria of the repeated voting game, but rather to show that the political equilibria of the static voting game would still be possible outcomes of the dynamic game without commitment. As Conde-Ruiz and Galasso (2005) show, trigger strategies can be defined to sustain the equilibrium outcome of the once-and-for-all voting game as a subgame-perfect Nash equilibrium in the repeated game. Extending our results to the case of repeated voting, following this approach, is beyond the scope of this paper.

## 5. CONCLUSIONS

In this paper, we pose the following question: how do income inequality and mobility affect the demand for different redistribution programs under one welfare state? In particular, we consider the allocation of tax revenues between pensions to the elderly and transfer payments to the working individuals. We allow for productivity differences across members of the same generation. In addition, we introduce individual shocks to productivity over lifetime which follow exogenously given transition probabilities. We study the size of each component of the welfare state and the implications of income inequality and mobility for demand for redistribution and social insurance. We analyze the outcomes of the voting game using the SIE approach introduced by Shepsle (1979).

Our results indicate that the coalitions formed, and thus the resulting political equilibria, crucially depend on the level of income inequality in the economy. As in Meltzer and Richard (1981), a more unequal income distribution will very likely lead to higher levels of redistributive taxation. Not only the differences in current levels of income but also the expected lifetime income plays an essential role in determining the demand for redistribution and social insurance. We show that an increase in income mobility may influence the demand for each component differently depending on the redistribution mechanism. In the case of transfers to the poor, the decisive voter is a young individual who has incomplete information concerning his position in the future income distribution. Thus, higher mobility leads rich individuals to support higher transfer payments, whereas the poor demand less, because of their improved prospects of moving up the income ladder. The pension program, on the other hand, does not respond to variations in lifetime income. This is due to the fact that the social security pension program is supported by the middle-aged and elderly, who have no uncertainty about their future positions in the income distribution. These findings suggest that although changes in income mobility may not necessarily affect the size of the welfare state, they may lead to a change in the relative allocation of tax revenues across programs.

The model in this paper provides important insights into the phenomenon of the median voter in the presence of multiple redistributive programs. Consistent with the previous literature, we find that both age and income are critical determinants of political preferences regarding each redistributive component of the welfare state.

The pension program providing intergenerational redistribution may be sustained by a coalition of the old and the middle-aged, as in the seminal paper by Browning (1975). The transfer payments program, on the other hand, is supported by a coalition of low-income types, as in Tabellini (2000) and Casamatta et al. (2000). The present analysis offers further intuition that income mobility plays a crucial role in determining the political outcomes in a multidimensional policy space. We show that heterogeneity in future income prospects play a key role, in that the demand for social insurance responds to changes in expected income even when median and mean income remain the same.

The model presented here could be extended in several ways. First, we assume that the income mobility process is exogenous, and we do not model its sources. It would be interesting to analyze to what extent income mobility is explained by job displacements or destructions due to structural changes in the economy. Another possible extension would allow the labor supply to respond to tax rates at the extensive margin as well as the intensive margin, so that distortions in the form of participation decisions could also be modeled. Similarly, one could study a setting where individuals can alter their future positions through their educational choices to enhance their prospects of upward mobility. It would also make the analysis richer to allow a more realistic income distribution by modeling income as a continuous variable. Other potentially useful extensions could be the introduction of contributory benefits rather than lump-sum ones and a more comprehensive tax system that would also tax the interest income. These modifications would induce additional layers of heterogeneity in the voting behavior of the middle-aged.

It would be a useful exercise to gauge the relative efficiency of the political process; however, constructing the social welfare function in overlapping-generations models is rather arbitrary. This is because one would have to assign weights to each generation to reflect their relative importance, which is a subjective call. Because of such difficulties, our analysis remains silent on the comparison of the equilibrium tax rates to the socially optimal ones. Finally, in this paper, we consider income mobility as the income variability over an individual's lifetime. Analyzing intergenerational income mobility would be an alternative approach. However, our main focus here is on social security and redistribution programs, most of which are designed to shift incomes between different periods of a lifetime or between contingencies [Osberg et al. (2004)]. Therefore, most of these programs target temporary income fluctuations over the life cycle rather than income inequality in lifetime incomes across individuals. It would be a natural extension to study how redistributive programs may equalize long-term or lifetime incomes. We leave these possible refinements to future research.

## NOTES

1. Persson and Tabellini (2000) and Galasso and Profeta (2002) provide comprehensive overviews of this literature.

2. For example, Boldrin and Montes (2005) and Poutvaara (2006) investigate the joint determination of public education and pension spending. Bethencourt and Galasso (2008) study health care and social security and show that there might be political complementarities between these two programs.

3. The Appendix presents the analysis under a more relaxed assumption of constant population growth rate  $n \geq 0$  until the comparative statics exercises, where we return to the assumption of no population growth.

4. Creedy et al. (2011) report that the ratio of median income to mean income among 25 democracies around the world in the early 2000s varied between 0.75 and 0.91.

5. The within- and across-cohorts redistribution of income and risk by the welfare system has been studied extensively by several others. See, for example, Persson and Tabellini (2000) and Galasso and Profeta (2002).

6. A sufficient condition to ensure an interior solution to the problem of the agent is that  $\alpha \leq 1 - \tau - \theta$ . Holding everything else constant, individuals with high productivity choose to supply more labor than those with low productivity because  $e_H > e_L$ . This is true for both the young and the middle-aged regardless of the tax rates. Also note that the labor supply of the middle-aged depends only on their current productivity and not the past.

7. See, for example, Chapter 5 in Mueller (2003).

8. See Persson and Tabellini (2000) for a survey of these methods.

9. We focus only on interior equilibria of the voting game and assume that we are operating on the left-hand side of the Laffer curve, such that a tax cut would always lower government revenue [as in Tabellini and Persson (2000, Chapter 6) and in Bethencourt and Galasso (2008)]. Formally, we have (i) for any given  $\theta$ ,  $0 < \tau < \arg\max_{\tau} \Psi(\tau, \theta)$ ; and (ii) for any given  $\tau$ ,  $0 < \theta < \arg\max_{\theta} \Psi(\tau, \theta)$ .

$$10. \partial I_1^* / \partial \theta = \partial I_2^* / \partial \theta = -\frac{\tau - \alpha}{(1 - \tau - \theta)^2} < 0.$$

11. From (16), it is clear that young H-types would not demand any transfers if they were certain that their position in the income distribution would be maintained ( $\pi_2 = 1$ ).

12. Note that case (iii) in Proposition 3 by itself presents a situation in which multiple equilibria arise because the two reaction functions overlap. Thus, this is not a full characterization of the SIE of the game [see also Proposition 5.5 of Bethencourt and Galasso (2008)].

13. As eloquently stated by Atkinson et al. (1992, p. 6): "One of the reasons why mobility is of interest is that it reduces inequality in the lifetime sum of earnings relative to that in a single period."

## REFERENCES

- Alesina, Alberto and George-Marios Angeletos (2005) Fairness and redistribution. *American Economic Review* 95(4), 960–980.
- Alesina, Alberto, Rafael Di Tella, and Robert MacCulloch (2004) Inequality and happiness: are Europeans and Americans different? *Journal of Public Economics* 88(9–10), 2009–2042.
- Alesina, Alberto and Eliana La Ferrara (2005) Preferences for redistribution in the land of opportunities. *Journal of Public Economics* 89(5–6), 897–931.
- Atkinson, Anthony Barnes, François Bourguignon, and Christian Morrisson (1992) *Empirical Studies of Earnings Mobility*. Chur, Switzerland: Harwood Academic.
- Beckman, Steven R., John P. Formby, and W. James Smith (2004) Efficiency, equity and democracy: Experimental evidence on Okun's leaky bucket. *Research on Economic Inequality* 11, 17–42.
- Bénabou, Roland (2000) Unequal societies: Income distribution and the social contract. *American Economic Review* 90(1), 96–129.
- Bénabou, Roland and Efe A. Ok (2001) Social mobility and the demand for redistribution: The POUM hypothesis. *Quarterly Journal of Economics* 116(2), 447–487.
- Bethencourt, Carlos and Vincenzo Galasso (2008) Political complements in the welfare state: Health care and social security. *Journal of Public Economics* 92(3–4), 609–632.

- Boldrin, Michele and Ana Montes (2005) The intergenerational state education and pensions. *Review of Economic Studies* 72(3), 651–664.
- Boldrin, Michele and Aldo Rustichini (2000) Political equilibria with social security. *Review of Economic Dynamics* 3(1), 41–78.
- Browning, Edgar K. (1975) Why the social insurance budget is too large in a democracy. *Economic Inquiry* 13(3), 373–388.
- Casamatta, Georges, Helmuth Cremer, and Pierre Pestieau (2000) The political economy of social security. *Scandinavian Journal of Economics* 102(3), 503–522.
- Conde-Ruiz, J. Ignacio and Vincenzo Galasso (2003) Early retirement. *Review of Economic Dynamics* 6, 12–36.
- Conde-Ruiz, J. Ignacio and Vincenzo Galasso (2004) The macroeconomics of early retirement. *Journal of Public Economics* 88(9–10), 1849–1869.
- Conde-Ruiz, J. Ignacio and Vincenzo Galasso (2005) Positive arithmetic of the welfare state. *Journal of Public Economics* 89(5–6), 933–955.
- Conde-Ruiz, J. Ignacio and Paola Profeta (2007) The redistributive design of social security systems. *Economic Journal* 117(520), 686–712.
- Creedy, John, Shuyun May Li, and Solmaz Moslehi (2011) The composition of government expenditure: Economic conditions and preferences. *Economic Inquiry* 49(1), 94–107.
- Galasso, Vincenzo (1999) The U.S. social security system: What does political sustainability imply? *Review of Economic Dynamics* 2(3), 698–730.
- Galasso, Vincenzo (2008) Postponing retirement: The political effect of aging. *Journal of Public Economics* 92(10–11), 2157–2169.
- Galasso, Vincenzo and Paola Profeta (2002) The political economy of social security: A survey. *European Journal of Political Economy* 18(1), 1–29.
- Hungerford, Thomas L. (2008) Income Inequality, Income Mobility, and Economic Policy: U.S. Trends in the 1980s and 1990s. Report to Congress RL34434. Washington, DC: Congressional Research Service.
- Koethenbueger, Marko, Panu Poutvaara, and Paola Profeta (2008) Why are more redistributive social security systems smaller? A median voter approach. *Oxford Economic Papers* 60(2), 275–292.
- Kopczuk, Wojciech, Emmanuel Saez, and Jae Song (2010) Earnings inequality and mobility in the United States: Evidence from social security data since 1937. *Quarterly Journal of Economics* 125(1), 91–128.
- Lindbeck, Assar (1985) Redistribution policy and the expansion of the public sector. *Journal of Public Economics* 28(3), 309–328.
- Meltzer, Allan H. and Scott F. Richard (1981) A rational theory of the size of government. *Journal of Political Economy* 89(5), 914–927.
- Moene, Karl O. and Michael Wallerstein (2001) Inequality, social insurance and redistribution. *American Political Science Review* 95(4), 859–874.
- Mueller, Dennis C. (2003) *Public Choice III*. New York: Cambridge University Press.
- Mulligan, Casey B. and Xavier Sala-i-Martin (1999) Social Security in Theory and Practice (I): Facts and Political Theories. NBER Working Paper 7118.
- Osberg, Lars, Timothy M. Smeeding, and Jonathan Schwabish (2004) Income distribution and public social expenditure: Theories, effects, and evidence. In Kathryn M. Neckerman (ed.), *Social Inequality*, pp. 821–859. New York: Russell Sage Foundation.
- Persson, Torsten and Guido Tabellini (2000) *Political Economics: Explaining Economic Policy*. Cambridge, MA: MIT Press.
- Piketty, Thomas (1995) Social mobility and redistributive politics. *Quarterly Journal of Economics* 110(3), 551–584.
- Poutvaara, Panu (2006) On the political economy of social security and public education. *Journal of Population Economics* 19(2), 345–365.
- Ravallion, Martin and Michael Lokshin (2000) Who wants to redistribute? The tunnel effect in 1990s Russia. *Journal of Public Economics* 76(1), 87–104.

Shepsle, Kenneth A. (1979) Institutional arrangements and equilibrium in multidimensional voting models. *American Journal of Political Science* 23(1), 27–59.  
 Tabellini, Guido (2000) A positive theory of social security. *Scandinavian Journal of Economics* 102(3), 523–545.  
 Tanzi, Vito and Ludger Schuknecht (2000) *Public Spending in the 20th Century: A Global Perspective*. Cambridge, MA: Cambridge University Press.

## APPENDIX

### A.1. CONSTANT POPULATION GROWTH RATE

If population is assumed to grow at a constant rate of  $n \geq 0$ , then the size of the newborn population at time  $t$  is given by

$$N_t = (1 + n)N_{t-1} = (1 + n)^t N_0, \quad N_0 \equiv 1. \tag{A.1}$$

It follows directly from (A.1) that in any period  $t$ , for each old individual there are  $(1 + n)$  middle-aged and  $(1 + n)^2$  young. Total labor income in every period  $t$  is given by

$$\Psi(\tau_t, \theta_t) = \Psi_t = (1 + n)^2 \Psi_{1,t} + (1 + n) \Psi_{2,t} (1 + n)(2 + n)(1 + \bar{e}) - \frac{(1 + n)(2 + n)\alpha}{1 - \tau_t - \theta_t}. \tag{A.2}$$

Aggregate income is stationary over time if we exclude the growth rate of the population. The government budget for the transfer payments program can be written in per capita terms as

$$B_t = \frac{\tau_t \Psi_t(\tau_t, \theta_t)}{\lambda(1 + n)(2 + n)}, \tag{A.3}$$

whereas the one for the pension program remains the same as in the text (see (2)). Finally, the aggregate feasibility constraint would be modified as

$$\begin{aligned} & (1 - \tau_t - \theta_t)\{(1 + n)[\lambda l_{2L,t} + (1 - \lambda)l_{2H,t}] + (1 + n)^2[\lambda l_{1L,t} + (1 - \lambda)l_{1H,t}]\} \\ & + (\tau_t + \theta_t)\{(1 + n)[\lambda l_{2L,t} + (1 - \lambda)l_{2H,t}] + (1 + n)^2[\lambda l_{1L,t} + (1 - \lambda)l_{1H,t}]\} \\ & = c_{3,t} + (1 + n)[\lambda c_{2L,t} + (1 - \lambda)c_{2H,t}] + (1 + n)^2[\lambda c_{1L,t} + (1 - \lambda)c_{1H,t}] \end{aligned} \tag{A.4}$$

### A.2. ECONOMIC EQUILIBRIUM

To solve the economic problem of the young L-type, first we need to write down his intertemporal budget constraint:

$$\begin{aligned} c_{1L,t} + \frac{\pi_1 c_{2L,t+1} + (1 - \pi_1)c_{2H,t+1}}{R} + \frac{c_{3,t+2}}{R^2} &= \{l_{1L,t}(1 - \tau_t - \theta_t) + B_t\} \\ + \frac{\pi_1 \{l_{2L,t+1}(1 - \tau_{t+1} - \theta_{t+1}) + B_{t+1}\} + (1 - \pi_1)\{l_{2H,t+1}(1 - \tau_{t+1} - \theta_{t+1})\}}{R} &+ \frac{P_{t+2}}{R^2}. \end{aligned}$$

Using this, we solve for  $c_{3,t+2}$  and plug it back into (6). From the first-order conditions, we obtain  $U'(c_{1L,t}^*) = U'(c_{2k,t+2}^*) = 1$  which follows directly from our assumption about linearity of utility in old-age consumption. Again from the first-order conditions, we have  $W'(l^*) = U'(c^*)(1 - \tau_t - \theta_t)$ . Using the specific functional forms for  $U$  and  $W$ , one can obtain  $c_{1L,t}^* = c_{2k,t+2}^* = 1$ . His optimal labor supply is  $l_{1L,t}^* = 1 + e_L - \frac{\alpha}{1-\tau_t-\theta_t}$  when young, and  $l_{2L,t+1}^* = 1 + e_L - \frac{\alpha}{1-\tau_{t+1}-\theta_{t+1}}$  when middle-aged. The economic problems in (7) and (8) can be solved in a similar fashion.

**A.3. CONDITIONS IMPOSED ON THE PARAMETERS**

First, recall that in note 6, we impose a sufficient condition to ensure an interior solution to the problem of the agent. We assume that, for any given tax rates  $\tau$  and  $\theta$ , we have  $\alpha \leq 1 - \tau - \theta$ . To ensure  $W' > 0$  and  $W'' < 0$ , we need  $\alpha > 0$ .

Second, for any agent, we want to have  $c_3^* > 0, \forall \tau, \theta$ . Recall that the agents choose  $c_{1j}^* = c_{2k}^* = 1$ , regardless of their type. We consider an individual in the most unfavorable situation, i.e., the L-type young without any redistribution program ( $B = P = 0$ ). His lifetime budget constraint is given by

$$1 + \frac{1}{R} + \frac{c_3^*}{R^2} = l_L^* + \frac{\pi_1 l_L^* + (1 - \pi_1) l_H^*}{R},$$

which yields  $c_3^* = R^2(l_L^* - 1) + R[\pi_1 l_L^* + (1 - \pi_1) l_H^* - 1]$ . A sufficient condition for this expression to be positive is  $e_L > 1$ .

Third, we want to ensure that, for any given tax rates  $\tau$  and  $\theta$ , the net income of H-type is higher than the net income of L-type for both the young and the middle-aged; i.e.,  $l_H^*(1 - \tau - \theta) > l_L^*(1 - \tau - \theta) + B$ . This requires the following condition to be met for any pair of  $\tau$  and  $\theta$ :  $B < (e_H - e_L)(1 - \tau - \theta)$  which can also be expressed as  $e_H - e_L > \frac{\tau l_H^*}{\lambda(1-\theta)}$ . This means that income inequality has to be large enough.

Note that these conditions imposed on parameters do not contradict each other. To sum up, we assume that for any  $\tau$  and  $\theta$ ,

- (i)  $0 < \alpha \leq 1 - \tau - \theta$  to ensure  $W' > 0$  and  $W'' < 0$  as well as an interior solution for optimal labor decisions;
- (ii)  $e_H > e_L > 1 > \alpha$  to have an interior solution for optimal consumption decisions;
- (iii)  $e_H - e_L > \frac{\tau l_H^*}{\lambda(1-\theta)}$ , to ensure that the net income of H-type is higher than the net income of L-type, for both the young and the middle-aged.

**A.4. PROOF OF LEMMA 1: SINGLE-PEAKEDNESS AND SINGLE-CROSSING PROPERTIES OF THE REACTION FUNCTIONS**

First, note that the preferred policies take the following form when  $n > 0$ :

$$\frac{\partial V_O}{\partial \theta} = \frac{\partial(\theta \Psi)}{\partial \theta} = (\Psi + \theta \Psi_\theta) > 0, \tag{A.5}$$

$$\frac{\partial V_L}{\partial \theta} = \beta^2 (\Psi + \theta \Psi_\theta) + (1 + \beta \pi_1) \frac{\tau \Psi_\theta}{\lambda(1+n)(2+n)} - (1 + \beta \pi_1) l_L^* - \beta(1 - \pi_1) l_H^* = 0, \tag{A.6}$$



$$\frac{\partial V_H}{\partial \theta} = \beta^2 (\Psi + \theta \Psi_\theta) + \beta(1 - \pi_2) \frac{\tau \Psi_\theta}{\lambda(1+n)(2+n)} - (1 + \beta \pi_2) l_H^* - \beta(1 - \pi_2) l_L^* = 0, \tag{A.7}$$

$$\frac{\partial V_{LL}}{\partial \theta} = \frac{\partial V_{HL}}{\partial \theta} = \beta (\Psi + \theta \Psi_\theta) + \frac{\tau \Psi_\theta}{\lambda(1+n)(2+n)} - l_L^* = 0. \tag{A.8}$$

However, (14) remains the same as in the text. A higher population growth rate  $n$  would make public pensions more attractive because the same tax rate would then bring higher benefits. To show that voters have a unique preferred  $\tau$  and a unique preferred  $\theta$ , it is enough to show that the preferences are strictly concave in  $\theta$  for a given  $\tau$ , and vice versa.

Note that (A.2) implies  $\Psi_\theta = \partial \Psi / \partial \theta = -\frac{\alpha(1+n)(2+n)}{(1-\tau-\theta)^2} = \Psi_\tau = \partial \Psi / \partial \tau$ , which is always negative. Taking the derivative of this expression with respect to  $\theta$  ( $\tau$ ) yields:  $\Psi_{\theta\theta} = \Psi_{\tau\tau} = -\frac{2\alpha(1+n)(2+n)}{(1-\tau-\theta)^3}$ , which is also always negative. In the following, we present the single-peakedness arguments for the young L-type. The other cases can be derived in a similar fashion.

- (i) Taking the second derivative of the indirect utility function with respect to  $\theta$ , we get

$$\frac{\partial^2 V_L}{\partial \theta^2} = \beta^2 (2\Psi_\theta + \theta \Psi_{\theta\theta}) + \frac{(1 + \beta \pi_1)}{\lambda(1+n)(2+n)} (\tau \Psi_{\theta\theta}) < 0.$$

Thus, we have shown that  $V_L$  is concave in  $\theta$  for a given  $\tau$  and can conclude that the preferences are single-peaked in  $\theta$ .

- (ii) To show single-peakedness in  $\tau$ , we need to take the second derivative of the indirect utility function with respect to  $\tau$ :

$$\frac{\partial^2 V_L}{\partial \tau^2} = \frac{(1 + \beta \pi_1)}{\lambda(1+n)(2+n)} (2\Psi_\tau + \tau \Psi_{\tau\tau}) < 0.$$

We have shown that  $V_L$  is concave in  $\tau$  for a given  $\theta$  and can conclude that the preferences are single-peaked in  $\tau$ .

This concludes the proof of Lemma 1. The single-crossing property holds, because for each possible median voter,  $\partial V / \partial \theta$  is monotone in  $\tau$ , and  $\partial V / \partial \tau$  is monotone in  $\theta$ . ■

**A.5. PREFERRED POLICIES**

Once again, we present the case for the young L-type only because it is straightforward to apply the same procedure to other cases. To find his preferred tax rate  $\theta_L$ , we maximize his indirect utility function with respect to  $\theta$ :

$$\begin{aligned} \frac{\partial V_L}{\partial \theta} &= U'(c_L^*) \frac{\partial c_L^*}{\partial \theta} + W'(l_L^*) \frac{\partial l_L^*}{\partial \theta} + \beta \pi_1 \left[ U'(c_{LL}^*) \frac{\partial c_{LL}^*}{\partial \theta} + W'(l_{LL}^*) \frac{\partial l_{LL}^*}{\partial \theta} \right] \\ &+ \beta(1 - \pi_1) \left[ U'(c_{LH}^*) \frac{\partial c_{LH}^*}{\partial \theta} + W'(l_{LH}^*) \frac{\partial l_{LH}^*}{\partial \theta} \right] + \beta^2 \frac{\partial P}{\partial \theta} = 0. \end{aligned}$$

Applying the envelope theorem yields

$$\frac{\partial V_L}{\partial \theta} = \beta^2 (\Psi + \theta \Psi_\theta) + (1 + \beta \pi_1) \frac{\tau \Psi_\theta}{\lambda(1+n)(2+n)} - (1 + \beta \pi_1) l_L^* - \beta(1 - \pi_1) l_H^* = 0.$$

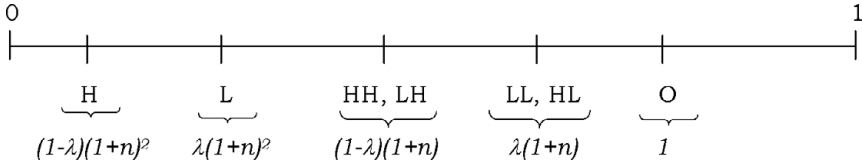


FIGURE A.1. Ranking of the preferred  $\theta$ 's and population shares.

This equation defines  $\theta_L^*$ . Analogously,  $\tau_L^*$  can be found by maximizing his indirect utility function with respect to  $\tau$ :

$$\begin{aligned} \frac{\partial V_L}{\partial \tau} = & U'(c_L^*) \frac{\partial c_L^*}{\partial \tau} + W'(I_L^*) \frac{\partial I_L^*}{\partial \tau} + \beta \pi_1 \left[ U'(c_{LL}^*) \frac{\partial c_{LL}^*}{\partial \tau} + W'(I_{LL}^*) \frac{\partial I_{LL}^*}{\partial \tau} \right] \\ & + \beta(1 - \pi_1) \left[ U'(c_{LH}^*) \frac{\partial c_{LH}^*}{\partial \tau} + W'(I_{LH}^*) \frac{\partial I_{LH}^*}{\partial \tau} \right] + \beta^2 \frac{\partial P}{\partial \tau} = 0. \end{aligned}$$

By the envelope theorem, this can be rewritten as

$$\frac{\partial V_L}{\partial \tau} = (1 + \beta \pi_1) \frac{\Psi + \tau \Psi_\tau}{\lambda(1+n)(2+n)} - (1 + \beta \pi_1) I_L^* - \beta(1 - \pi_1) I_H^* + \beta^2 \theta \Psi_\tau = 0.$$

**A.6. PROOF OF LEMMA 2: RANKING OF  $\theta$**

To determine the ranking of preferred  $\theta$ 's, we will simply manipulate equations (A.5)–(A.8) and (14). As expressed in (A.8) and (14),  $\theta_{LL} = \theta_{HL}$  and  $\theta_{HH} = \theta_{LH}$ . Given the assumptions listed in Section A.3, we have  $e_H - e_L > \frac{B}{1-\tau-\theta}$ . Simple algebra yields  $e_H - e_L > \frac{\tau\alpha}{(1-\tau-\theta)^2\lambda}$ . This implies that the H-types prefer a lower social security tax rate than the L-types, and this is true for both the young and the middle-aged. Thus, we have  $\theta_{HH} = \theta_{LH} < \theta_{LL} = \theta_{HL}$  and  $\theta_H < \theta_L$ . Next, let us compare preferred pension tax between the young and the middle-aged individuals. Note that each of the four equations (A.6)–(A.8) and (14) includes one positive term ( $\Psi + \theta \Psi_\theta$ ), whereas the remaining terms are negative. To determine the comparison across the terms in these equations, we start by noting that  $\theta_{HH} = \theta_{LH} > \theta_H$ . This is because the net benefit from the pension program is greater for the middle-aged H-type than for the young H-type. Similarly,  $\theta_{LL} = \theta_{HL} > \theta_L$ . Furthermore,  $\theta_L < \theta_{HH}$  as long as  $e_H - e_L < I_L^*/\beta\pi_1$ , which is a reasonable restriction. Finally, note that the elderly are net recipients of this program ((A.5) includes only a positive term), which is why they prefer the highest possible  $\theta$ . ■

Figure A.1 illustrates the ranking of preferred  $\theta$ 's.

**A.7. PROOF OF PROPOSITION 1: MEDIAN VOTER OVER  $\theta$**

When  $n > 0$ , Proposition 1 can be restated as follows.

The median voter over the issue  $\theta$  is

- (1)  $\theta^m = \theta_{HH}^* = \theta_{LH}^*$  if  $\lambda < \Gamma(n)$ , and
  - (2)  $\theta^m = \theta_{LL}^* = \theta_{HL}^*$  if  $\lambda > \Gamma(n)$ ,
- where  $\Gamma(n) = \frac{n^2+3n+1}{2(n+1)}$ .

Thus, the dominant coalition and the identity of the median voter depend on the proportion of L-types in the population. Given the rankings for  $\theta$  as stated in Lemma 2, we can analyze possible coalitions. The population share for each group is as follows: 1 for the old,  $\lambda\pi_1(1+n)$  for the LL,  $\lambda(1-\pi_1)(1+n)$  for the LH,  $(1-\lambda)\pi_2(1+n)$  for the HH,  $\lambda(1-\pi_2)(1+n)$  for the HL,  $(1-\lambda)(1+n)^2$  for the H, and finally,  $\lambda(1+n)^2$  for the L. Recall that we have  $\lambda = \frac{1-\pi_2}{(1-\pi_1)+(1-\pi_2)} > \frac{1}{2}$ . Simple derivations yield that the median voter for  $\theta$  depends on how  $\frac{n^2+3n+1}{2(n+1)} \equiv \Gamma(n)$  compares to  $\lambda$ . ■

**A.8. PROOF OF LEMMA 3: RANKING OF  $\tau$**

When  $n > 0$ , Lemma 3 can be restated as follows:

- (I) If  $e_H > E$ , then  $0 = \tau_O^* = \tau_{HH}^* = \tau_{LH}^* < \tau_H^* < \tau_L^* < \tau_{LL}^* = \tau_{HL}^*$ , and
  - (II) if  $e_H < E$ , then  $0 = \tau_O^* = \tau_{HH}^* = \tau_{LH}^* < \tau_H^* < \tau_{LL}^* = \tau_{HL}^* < \tau_L^*$ ,
- where

$$E = \frac{\alpha(1+n)(2+n)\theta [1 - \beta(1 - \pi_1)] + (1 - \pi_1)(1 - \tau - \theta)(\alpha - 1 + \tau + \theta)}{(1 - \pi_1)(1 - \tau - \theta)^2}.$$

We can obtain the ranking of preferred  $\tau$ 's by manipulating the expressions given in (15)–(18). First, we start by noting that  $l_{LL}^* = l_{HL}^*$ , which implies that the preferred  $\tau$  will be identical for LL and HL. The same argument holds for HH and LH, because we have  $l_{HH}^* = l_{LH}^*$ . Second, because  $l_H^* < l_{H'}^*$ , the preferred  $\tau$  will be lower for HH and LH groups than for the LL and HL groups. Similarly, for young individuals, we know that the preferred  $\tau$  will be lower for H-types than for the L-types. It is also straightforward to show that LL and HL groups prefer a higher  $\tau$  than the H group. The ranking between the young L-types and the middle-aged L-types is not that trivial. The comparison between (15) and (17) reduces to the comparison between  $e_H$  and

$$E = \frac{\alpha(1+n)(2+n)\theta [1 - \beta(1 - \pi_1)] + (1 - \pi_1)(1 - \tau - \theta)(\alpha - 1 + \tau + \theta)}{(1 - \pi_1)(1 - \tau - \theta)^2}.$$

Figure A.2 illustrates the ranking of preferred  $\tau$ 's. ■

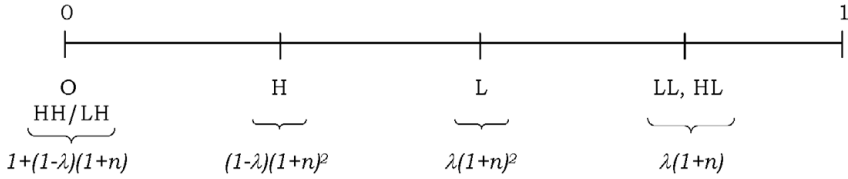
**A.9. PROOF OF PROPOSITION 2: MEDIAN VOTER OVER  $\tau$**

Under  $n > 0$ , Proposition 2 can be restated as follows:

- (I)  $\tau^m = \tau_H^*$  if  $\lambda < \Omega(n)$ ,
  - (II.a)  $\tau^m = \tau_L^*$  if  $\lambda > \Omega(n)$  and  $e_H > E$ ,
  - (II.b)  $\tau^m = \tau_{LL}^* = \tau_{HL}^*$  if  $\lambda > \Omega(n)$  and  $e_H < E$ ,
- where  $\Omega(n) = \frac{1+(n+1)(n+2)}{2(n+1)(n+2)}$ .

Given the population shares, the median voter for  $\tau$  will be the young H-type if the coalition formed by young and middle-aged H-types together with the elderly constitutes the majority, i.e., if  $(1+n)^2(1-\lambda) + (1+n)(1-\lambda) + 1 > (1+n)^2\lambda + (1+n)\lambda$ . This boils down to the condition that  $\lambda < \frac{1+(1+n)(2+n)}{2(1+n)(2+n)} \equiv \Omega(n)$ . If, instead,  $\lambda > \Omega(n)$ , then the median voter will be either the young L-type or the middle-aged L-type, according to the cases listed in Lemma 3. ■

CASE I:  $e_H > E$



CASE II:  $e_H < E$

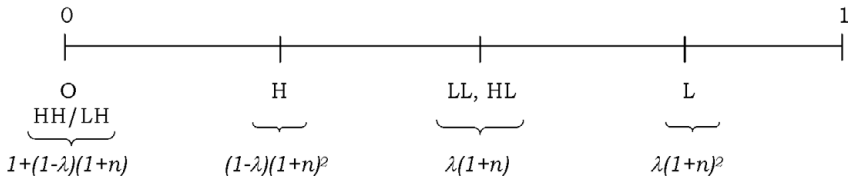


FIGURE A.2. Ranking of the preferred  $\tau$ 's and population shares.

**A.10. PROOF OF PROPOSITION 3: CHARACTERIZATION OF THE SIE**

Proposition 3 can be rewritten as follows under  $n > 0$ :

- (a)  $(\tau^m, \theta^m) = (\tau_H^*, \theta_{HH}^*)$  if  $\lambda < \Gamma(n)$ ,
- (b)  $(\tau^m, \theta^m) = (\tau_H^*, \theta_{LL}^*)$  if  $\Gamma(n) < \lambda < \Omega(n)$ ,
- (c)  $(\tau^m, \theta^m) = (\tau_L^*, \theta_{LL}^*)$  if  $\lambda > \Gamma(n)$  and  $e_H > E$ ,
- (d)  $(\tau^m, \theta^m) = (\tau_{LL}^*, \theta_{LL}^*)$  if  $\lambda > \Omega(n)$  and  $e_H < E$ ,

where

$$\Gamma(n) = \frac{n^2 + 3n + 1}{2(n + 1)}, \quad \Omega(n) = \frac{1 + (n + 1)(n + 2)}{2(n + 1)(n + 2)},$$

and 
$$E = \frac{\alpha(1 + n)(2 + n)\theta [1 - \beta(1 - \pi_1)] + (1 - \pi_1)(1 - \tau - \theta)(\alpha - 1 + \tau + \theta)}{(1 - \pi_1)(1 - \tau - \theta)^2}.$$

The proof for this proposition is a simple combination of the median voters along both dimensions together with the properties of the reaction functions presented in Appendix A.4. In addition, we have  $\Gamma(n) < \Omega(n)$ , as long as  $n < 0.2$ , which is a reasonable restriction. ■

**A.11. EFFECT OF RISING INEQUALITY ON  $\tau$**

The median voter over  $\tau$  can either be L, H, or LL. We will consider each of these one at a time. For L, we have  $F_\tau = (1 + \beta\pi_1)(\frac{2\Psi_\tau + \tau\Psi_{\tau\tau}}{2\lambda}) + \beta^2\theta\Psi_{\tau\tau} < 0$ . Now, we need to check the sign of the numerator in (20).  $F_{e_H} = (1 + \beta\pi_1)\frac{\Psi_{e_H}}{2\lambda} - \beta(1 - \pi_1) = (1 + \beta\pi_1)\frac{1 - \lambda}{\lambda} - \beta(1 - \pi_1) > 0$  because  $\lambda = \frac{1 - \pi_2}{2 - \pi_1 - \pi_2}$ .  $F_{e_L} = (1 + \beta\pi_1)\frac{\Psi_{e_L}}{2\lambda} - (1 + \beta\pi_1) = 0$ . Therefore, the sign of (20) when L is the median voter is positive.

For H, we have  $F_\tau = \beta(1 - \pi_2) \left( \frac{2\Psi_\tau + \tau\Psi_{\tau\tau}}{2\lambda} \right) + \beta^2\theta\Psi_{\tau\tau} < 0$ . Now, we need to check the sign of the numerator in (20).  $F_{e_H} = \beta(1 - \pi_2) \frac{\Psi_{e_H}}{2\lambda} - (1 + \beta\pi_2) = \beta(1 - \pi_2) \frac{1-\lambda}{\lambda} - (1 + \beta\pi_2) < 0$  once we substitute for  $\lambda$  as before.  $F_{e_L} = \beta(1 - \pi_2)\Psi_{e_L}/2\lambda - \beta(1 - \pi_2) = 0$ . Hence, the sign of (20) when H is the median voter is negative.

Last, for LL,  $F_\tau = \frac{2\Psi_\tau + \tau\Psi_{\tau\tau}}{2\lambda} + \beta\theta\Psi_{\tau\tau} < 0$ . Now, we need to check the sign of the numerator in (20).  $F_{e_H} = \Psi_{e_H}/2\lambda = \frac{1-\lambda}{\lambda} > 0$ .  $F_{e_L} = \Psi_{e_L}/2\lambda - 1 = 0$ . Therefore, the sign of (20) when LL is the median voter is positive.

**A.12. EFFECT OF RISING INEQUALITY ON  $\theta$**

The median voter over  $\theta$  is LL. From (A.8),  $G_\theta = \beta(2\Psi_\theta + \theta\Psi_{\theta\theta}) + \tau\Psi_{\theta\theta}/2\lambda < 0$ . Now we check the sign of the numerator in (22). We have  $G_{e_H} = \beta\Psi_{e_H} = 2\beta(1 - \lambda)$  and  $G_{e_L} = \beta\Psi_{e_L} - 1 = 2\beta\lambda - 1$ . Thus,  $-G_{e_H} + \frac{(1-\lambda)}{\lambda}G_{e_L} = \frac{(\lambda-1)}{\lambda} < 0$ . Hence,  $d\theta/d e_H > 0$ .

**A.13. EFFECT OF RISING MOBILITY ON  $\tau$**

The median voter over  $\tau$  can either be L, H, or LL. We will consider each of these one at a time. In the case of L, the denominator of (24) is  $Q_\tau = F_\tau < 0$ . For the numerator, we have  $Q_{\pi_1} = \beta/2\lambda(\Psi + \tau\Psi_\tau) + (1 + \beta\pi_1)\Psi_{\pi_1}/2\lambda - \beta(l_L^* - l_H^*)$  and  $Q_{\pi_2} = (1 + \beta\pi_1)\Psi_{\pi_2}/2\lambda$ . Therefore,  $-Q_{\pi_1} - \left(\frac{1-\pi_2}{1-\pi_1}\right)Q_{\pi_2} = -\frac{\beta}{2\lambda}(\Psi + \tau\Psi_\tau) - \beta(e_H - e_L) < 0$ . Hence, the sign of  $\frac{d\tau}{d\pi_1}$  in this case is positive.

For H, the denominator of (24) is  $Q_\tau = F_\tau < 0$ . For the numerator, we have  $Q_{\pi_1} = \beta(1 - \pi_2)\Psi_{\pi_1}/2\lambda$  and  $Q_{\pi_2} = -\frac{\beta}{2\lambda}(\Psi + \tau\Psi_\tau) + \beta(1 - \pi_2)\Psi_{\pi_2}/2\lambda - \beta(l_H^* - l_L^*)$ . Therefore,  $-Q_{\pi_1} - \left(\frac{1-\pi_2}{1-\pi_1}\right)Q_{\pi_2} = \frac{\beta\lambda}{1-\lambda} \left( \frac{\Psi + \tau\Psi_\tau}{2\lambda} + e_H - e_L \right) > 0$ . Hence, the sign of  $d\tau/d\pi_1$  in this case is negative.

Finally, for LL, the denominator of (24) is  $Q_\tau = F_\tau < 0$ . For the numerator, we have  $Q_{\pi_1} = \Psi_{\pi_1}/2\lambda$  and  $Q_{\pi_2} = \Psi_{\pi_2}/2\lambda$ . Thus,  $-Q_{\pi_1} - \left(\frac{1-\pi_2}{1-\pi_1}\right)Q_{\pi_2} = 0$ . Hence, we have  $d\tau/d\pi_1 = 0$  in this case.

**A.14. EFFECT OF RISING MOBILITY ON  $\theta$**

The median voter over  $\theta$  is LL. From (A.8),  $Z_\theta = G_\theta < 0$ . Now we check the sign of the numerator in (26). We have  $Z_{\pi_1} = Z_{\pi_2} = 0$  because  $\lambda$  is kept constant. Therefore,  $d\theta/d\pi_1 = 0$ .