

Young Liberals and Old Conservatives – Inequality, Mobility and Redistribution*

Astri Muren and Sten Nyberg

September 29, 2005

Abstract

The paper examines the impact of income inequality and mobility on income redistribution in a modified median voter model where redistributive conflict takes place both between educational groups and age-groups. The effects of inequality and mobility are not unambiguous but depend on factors such as how mobility changes in different groups and causes of inequality. We also examine the effect of the length of electoral periods on redistribution and welfare for different groups and allow for majority voting on the length of electoral periods. Finally, we extend the model to encompass retirement and baby booms.

JEL Classification: D31, D72, H20, J31, P16.

Keywords: Inequality, mobility, income redistribution, median voter, age-earnings profiles.

*Correspondence: Department of Economics, Stockholm University, 106 91 Stockholm, Sweden. E-mail astri.muren@ne.su.se or sten.nyberg@ne.su.se. The authors gratefully acknowledge helpful comments and suggestions on earlier versions at the EEA meeting in Lausanne, Humboldt University, University of Manchester, University of New South Wales, the National Institute of Economic Research, Stanford University, the Stockholm School of Economics, the Trade Union Institute for Economic Research and Umeå University. We have also benefited from discussions with colleagues at the department, the Institute for International Economic Studies and The Research Institute for Industrial Economics. We also thank Markus Jäntti for providing data, and the Swedish Collegium for Advanced Studies in Social Sciences for offering a stimulating research environment.

*Any man who is under 30, and is not a liberal, has not heart; and any man who is over 30 and is not a conservative, has no brains.*¹

1 Introduction

It is well known that income redistribution is more extensive in Western Europe than in the US. Since the demand for redistribution in democracies is often thought to increase with income inequality, and since the US income distribution is more unequal, this is perceived as something of a puzzle. Similarly, looking at the development over time it is interesting to note that the increase in earnings inequality in the US, as well as in Europe, coincides with tax revolts and retreats in welfare state ambitions.

Economic theory clearly lends some support to the idea that inequality increases the demand for redistribution. However, there are different kinds of inequality, such as inequality between groups in society, and variations in earnings over the individual's life-cycle, and these may have different effects on the support for redistribution.² Some types of inequality may even lead to less redistribution. For example, in a median voter model, an inflow of poor immigrants may yield this result if their incomes are sufficiently low to reduce the mean income more than the median. (See Razin, Sadka & Swagel, 2002). Similarly, high earnings mobility, (i.e., inequality over the life-cycle) may reduce a forward-looking voter's enthusiasm for redistribution, an idea which also has some support in US data. (See, e.g., Bénabou & Ok, 2001, and Alesina & La Ferrara, 2001.)

As for a comparison between the US and Europe, several authors note that differences in income mobility between these seem quite small, e.g. the probabilities of moving from one income quintile, or decile, to another are remarkably similar.³ However, transition probabilities do not fully capture the incentives for redistribution

¹This "orphan quote" exists in many versions. This wording has been attributed to Winston Churchill but as Mark Shirey pointed out to us several earlier versions with other sources exist.

²For a broad discussion of inequality and its development over time in different countries, see Gottschalk & Smeeding [1997].

³See e.g. Bénabou [2000] and Alesina, Glaeser & Sacerdote [2001]. Aaberge et al. [2002] find that the relative reduction in inequality (Gini) from extending the accounting period is also very similar.

since they do not reflect the *value* of transitions – a value that is often *much* higher in the US than in Europe.⁴

In this paper we examine the effects of inequality and mobility in an integrated median voter framework. We allow for both income differences between groups and individual income development over time. Concerning inequality, there is some ambiguity about the effects on redistribution in a median voter model, that depends on how inequality is measured. For instance, there is no apparent correspondence between the Gini-coefficient, and the support for redistribution. We examine a few different aspects of inequality, including the comparison of current mean and median income used in the basic median voter model. We find that, in the presence of mobility, this comparison says very little about the support for redistributive policies. Measures of between-group inequality are more informative. Moreover, the effects of changes in income mobility are not uniform but depend on in which group they take place. These are some of the effects examined in our model, which point to the fruitfulness of examining inequality and mobility simultaneously. Yet, our analysis remains simple enough to allow various extensions, covering e.g. retirement and baby-booms.

There are also other explanations for the difference between Europe and the US. For example, Bénabou [2000] presents a model based on incomplete insurance markets and wealth-bias in voting participation. Explanations relating to social beliefs about the fairness of market outcomes (Alesina & Angeletos, 2005), and trade and the insurance function of a large public sector (Cameron, 1978, and Rodrik, 1998), have also been proposed. Alesina et al. [2001] evaluate several explanations for the US–Europe redistribution puzzle and conclude that racial fragmentation may be amongst the most important reasons for why there is less redistribution in the US. Moreover, Glaeser & Alesina [2004], identify proportional representation as an important factor

⁴An OECD [1996] study reveals that if mobility is measured in terms of median-proportions earnings bands, a measure taking absolute earnings differences into account, the US has considerably higher earnings mobility than the seven European countries included in the study. Another indication that the US earnings distribution is more dispersed is that, in 1983, a transition from the first to the tenth blue-collar wage decile implied an income increase of approximately 30 percent in Sweden, 200 percent in the UK, and 400 percent in the US (Hibbs & Locking, 2000).

contributing to the rise of the welfare state in Europe.⁵

Our basic model contains two groups of agents that differ both in average income and income mobility. Income redistribution is determined by a majority vote and takes the form of proportional taxes and flat transfers. Redistributive conflict can take place both between groups and between generations. For example, old low-skilled workers may earn more than the median income and vote against redistribution while young high-skilled workers may be in favor of it. Furthermore, in the model we abstract from income risk⁶, targeted transfers, interest groups and politicians.

The model shows that the effects of mobility and inequality on the support for redistribution are not uniform but depend on the determinants of mobility and inequality. Higher mobility in low-skilled groups reduces the support for redistribution, while higher mobility among the high-skilled has the opposite effect. Increased inequality between educational groups increases redistribution, but an increase in the number of working poor can lead to less redistribution. We also find that the comparisons between the current mean and median incomes in the population are uninformative about the political support for redistribution.

Mobility and inequality differ across countries and over time and may help to explain differences and changes in redistributive policies.⁷

The support for redistributive policies also depends on their expected duration – long electoral periods lead to less redistribution in the model. Short periods can

⁵Recent research points to the importance of institutions as a link between voter preferences and policy outcomes (see Persson & Tabellini, 2000). Besley & Case [2003] emphasize the empirical shortcomings of the simple median voter model.

⁶Redistribution can serve as income insurance and the demand for this service should increase in income risk. However, income risk seems higher in the US than in Europe which deepens the puzzle.

⁷Returns to experience may be shaped by technological changes. For instance, new CNC machinery or integrated accounting and business management software may increase the return to experience for high-skilled workers while reducing it for low-skilled. When human capital is technology-specific, new technologies may imply steeper age-earnings profiles (Chari & Hopenhayn, 1991, and Thompson, 2003). Such changes can also affect income differences between groups, e.g. technological change may be skill biased. (See Berman, Bound & Griliches, 1994, and Bresnahan, 1999.)

Wage policies and labor market institutions differ across countries and can affect income differences between groups (see Fortin & Lemieux, 1997) as well as age-earnings profiles. There are considerable country differences in within-group wage development. See, e.g., Card & Lemieux [2001] and Brunello & Comi [2004]. For example, the policies of wage-compression in Sweden have tended to flatten age-earnings profiles, while the system of company training in Japan may have caused steeper age-earnings profiles (Klevmarken, 1982, and Hashimoto & Raisian, 1985).

induce “myopia” in young high-skilled agents, who may vote for redistribution even though they are better off without it in a life-time perspective. However, despite efficiency costs of taxation, low-skilled agents as a group benefit from redistribution. In a majority vote on the length of the electoral period, the outcome is a period long enough to remove any myopic incentives, which is not in the low-skilled group’s interest. The median voter in this case is low-skilled and thus betrays her class interest.

We also examine a fully funded retirement system and find that increased retirement savings reduce redistribution but that the effect of changing the retirement age is ambiguous. Finally we consider the effect of a baby-boom. A baby-boom has a positive effect on redistribution as the baby-boomers’ incomes, and thus the tax base, grow. However, as baby-boomers age they become more conservative, and thus the support for redistribution is reduced during the periods when baby-boomers are median voters.

The paper proceeds as follows. In section 2, we present the basic model and examine the effects of mobility and inequality on equilibrium redistribution. We also relate the level of redistribution to the difference between the current mean and median incomes. We then consider the length of electoral period. In sections 3 and 4 we consider the effects of retirement and a baby boom. Section 5 offers some concluding remarks.

2 The model

Let the length of an individual’s life-span be equal to one and assume that new individuals are born continuously at a constant rate so that the age distribution is uniform. The population size is normalized to one. There are two types of agents, H -agents and L -agents. Let p be the proportion of L -agents and $1 - p$ the proportion of H -agents in the population. For each type earnings are assumed to increase linearly

with age, θ .⁸ The age-earnings profile of an individual of type i is

$$y_i(\theta) = a_i + b_i\theta \quad (1)$$

where $i \in \{H, L\}$. H -individuals have steeper age-earnings profiles than L -individuals. The age-earnings profiles of H -agents do, however, have a lower intercept. If the groups are taken to be defined by high and low education this can represent earnings foregone while investing in education. The age-earnings profiles are illustrated in Figure 1a.

The life-time income for an individual of type i is,

$$\bar{y}_i = \int_0^1 (a_i + b_i t) dt = a_i + \frac{b_i}{2}. \quad (2)$$

This is also the individual's average income since the individual's lifespan is normalized to one. The average income in the population \bar{Y} is simply the sum of the average incomes of each type weighted by their frequency in the population,

$$\bar{Y} = p\bar{y}_L + (1 - p)\bar{y}_H. \quad (3)$$

Since the size of the population is one, the average income equals the tax-base of the economy.

Throughout we assume that the life-time income of H -individuals exceeds that of L -individuals. In the educational interpretation the income difference can be seen as the premium required to compensate for non-pecuniary costs of education. Moreover, we assume that the L -agents are in majority. In terms of the model parameters the above assumptions amount to:

$$(A1) \quad 0 < b_L < b_H \quad \text{and} \quad a_H < a_L$$

$$(A2) \quad \Delta\bar{y} = \bar{y}_H - \bar{y}_L > 0$$

$$(A3) \quad p > \frac{1}{2}$$

⁸Actual age-earnings profiles are concave and steeper profiles are associated with higher education (Psacharopoulos & Layard, 1979, Murphy & Welch, 1992, Brunello & Comi, 2004). Profiles measured from cross-section data show lower wages for older individuals. However, it does not necessarily follow that individual earnings decline after some age. See Thornton, Rodgers & Brookshire [1997].

The age-earnings profiles together with the composition of types in the population determine the income distribution. Specifically, the density function is given by

$$f(y) = \begin{cases} (1-p)/b_H & \text{if } y \in [y_H(0), y_L(0)] \cup [y_L(1), y_H(1)] \\ (1-p)/b_H + p/b_L & \text{if } y \in [y_L(0), y_L(1)] \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

Assumptions (A1) and (A2) imply that the income distribution is positively skewed (see Figure 1b). Moreover, it is straightforward to show that the median income \tilde{Y} falls short of the mean income \bar{Y} (Lemma 1). Hence, the income distribution given by expression (4) would result in positive redistribution in a standard median voter model, not taking mobility into account.⁹

Lemma 1 $\bar{Y} > \tilde{Y}$.

Proof: \tilde{Y} is determined by: $\int_{y_H(0)}^{\tilde{Y}} f(t)dt = 0.5$. (A1) - (A3) imply that there is a median income earner in each group. Thus $\tilde{Y} = \frac{pb_H\bar{y}_L + (1-p)b_L\bar{y}_H}{pb_H + (1-p)b_L}$ and $\bar{Y} - \tilde{Y} = \frac{p(1-p)\Delta\bar{y}(b_H - b_L)}{pb_H + (1-p)b_L} > 0$ by (A1) and (A2).

Figure 1 (a and b) about here

Taking earnings mobility into account, individual voting behavior is determined by how the individual's average income over the electoral period, δ , is affected by the tax and transfer policies under consideration. The parameter δ can also be interpreted as a proxy for the duration of redistributive policies. For simplicity, we abstract from discounting. The average period income before tax for a voter of age θ and type i is,

$$\bar{y}_i(\theta) = \begin{cases} \frac{1}{\delta} \int_{\theta}^{\theta+\delta} y_i(t) dt = y_i(\theta) + \frac{b_i}{2}\delta & \text{if } 1 - \theta \geq \delta \\ \frac{1}{1-\theta} \int_{\theta}^1 y_i(t) dt = y_i(\theta) + \frac{b_i}{2}(1 - \theta) & \text{if } 1 - \theta < \delta \end{cases} \quad (5)$$

where $y_i(\theta)$ is the individual's income at the beginning of the electoral period, and the second term is the average increase in income during the period. The upper expression applies if the individual's remaining life-time, $1 - \theta$, exceeds the electoral

⁹See, e.g., Meltzer & Richard [1981].

period and the lower if it does not. In both cases average earnings increase in age but for older agents, $\theta > 1 - \delta$, the rate of increase is only half of that for the young, since the maximum income remains constant for old agents.

Redistribution is assumed to be achieved through fixed transfers financed by proportional taxes. Taxation is costly and we assume that the cost is quadratic in the tax rate and given by $\frac{\rho\tau^2}{2}\bar{Y}$, where ρ measures the efficiency of the tax system and $0 < \rho \leq 2$. Assuming budget balance a tax τ suffices to finance a per capita transfer equal to $\tau(1 - \frac{\rho\tau}{2})\bar{Y}$. For an agent of type i and age θ the average period income after taxes and transfers is

$$\bar{Y}_i(\tau, \theta) = (1 - \tau)\bar{y}_i(\theta) + \tau(1 - \frac{\rho\tau}{2})\bar{Y}. \quad (6)$$

The optimal tax rate for the individual strikes a balance between the marginal increases in benefits and tax payments resulting from a higher tax rate, i.e.

$$(1 - \rho\tau)\bar{Y} - \bar{y}_i(\theta) \leq 0 \quad (7)$$

The optimal tax rate is then

$$\tau_i(\theta) = \max \left\{ 0, \frac{\bar{Y} - \bar{y}_i(\theta)}{\rho\bar{Y}} \right\}. \quad (8)$$

Note that since $\bar{y}_i(\theta)$ is strictly increasing in θ this tax rate decreases with age. Hence, for any given τ we can implicitly define a threshold age for each type, $\theta_i(\tau)$, using (8) and (5), such that all individuals younger than this prefer τ to any lower tax rate and all those older prefer τ to any higher tax. Specifically,

$$\begin{aligned} \theta_L(\tau) &= \max \left\{ 0, \frac{(1-\rho\tau)\bar{Y} - \bar{y}_L(0)}{b_L}, 2\frac{(1-\rho\tau)\bar{Y} - \bar{y}_L}{b_L} \right\} \\ \theta_H(\tau) &= \max \left\{ 0, \frac{(1-\rho\tau)\bar{Y} - \bar{y}_H(0)}{b_H} \right\}. \end{aligned} \quad (9)$$

Very high τ are not even supported by the youngest, and $\theta_i(\tau) = 0$. More moderate tax rates gather wider support and may even attract some older L -types, i.e. with $\theta_L \geq 1 - \delta$. By assumption, old H -types have a period income above \bar{Y} and are not interested in redistribution. Note that the $\theta_i(\tau)$ s are continuous and decreasing in τ and strictly so for τ such that $\theta_i(\tau) > 0$.

The political support for a given tax rate τ can now be expressed in terms of the fraction of the population that prefers τ to any lower tax rate, $\Phi(\tau)$, where

$$\Phi(\tau) = p\theta_L(\tau) + (1-p)\theta_H(\tau). \quad (10)$$

By construction Φ is continuous and decreasing in τ (and ρ) and strictly so for $\Phi > 0$. Moreover, $\Phi(\tau) = 0$ for sufficiently high tax rates and clearly so for $\tau \in [\frac{1}{\rho}, 1]$. Hence,

Proposition 1 *(i) If $\Phi(0) \leq 0.5$ then $\tau^* = 0$. (ii) If $\Phi(0) > 0.5$ there is a unique $\tau^* > 0$, defined by $\Phi(\tau^*) = 0.5$, that wins over every other τ in pairwise comparisons (a Condorcet winner) and $\theta_H(\tau^*) < 0.5 < \theta_L(\tau^*)$. (iii) $\tau^* > 0$ decreases in ρ .*

Proof: In (i) a majority, $1 - \Phi(0)$, prefers a lower τ . Since Φ decreases in τ this holds strictly for any $\tau > 0$. (ii) Since Φ is continuous and decreasing in τ and $\Phi(\tau) = 0$ for τ sufficiently close to 0.5 there exists a τ^* such that $\Phi(\tau^*) = 0.5$. It also follows that if τ is such that $\Phi(\tau) > 0.5$ then $\tau < \tau^*$. Hence, by definition a majority prefers τ^* to τ . A parallel argument holds for τ such that $\Phi(\tau) < 0.5$. Finally, since $\bar{y}_H(0.5) > \bar{Y}$, H -agents older than 0.5 strictly prefer $\tau = 0$ which implies $\theta_H(\tau^*) < 0.5$. Since $\Phi(\tau^*) = 0.5$ it follows that $\theta_L(\tau^*) > 0.5$. (iii) The equilibrium condition only determines the product $\tau^*\rho$. ■

Thus, if there is majority support for a positive tax rate, there exists a unique equilibrium tax rate, which increases in the efficiency of the tax system.

2.1 Mobility and inequality

Let us now consider the effects of income mobility and inequality on equilibrium redistribution. First we examine the effect of increased mobility. To isolate the mobility effect we examine changes in the slope of age-earnings profiles, while adjusting the intercept to keep the average income in each group constant. This removes the inter-group redistribution motive.

Proposition 2 *An increase in b_L , compensated by a decrease in a_L such that \bar{y}_L remains constant, strictly decreases τ^* for $\tau^* > 0$. The reverse is true for a com-*

compensated increase in b_H . A proportional compensated increase in b_L and b_H strictly decreases τ^ .*

Proof: In the Appendix.

A compensated increase in mobility decreases the incomes of young individuals and increases the incomes of the old. Since a threshold L -voter is relatively old, $\theta_L(\tau^*) > 0.5$, she benefits from increased mobility and becomes less interested in redistribution. The opposite is true for a threshold H -voter. A proportional compensated increase in mobility for both groups leads to less redistribution since the negative effect in the L -group dominates the positive effect in the smaller H -group.

Next we consider how income inequality affects redistribution. Inequality may occur in different dimensions, such as income differences between social classes or educational groups, over the life-cycle, or between generations. The political significance of between-group inequality is also likely to be affected by the size of the various groups.

There is an extensive literature on income inequality, containing various measures of inequality. Many measures, such as Gini-coefficients, ratios of decile incomes e.g. the 90/10-ratio, and the mean versus median income comparison common in theoretical median voter models, capture both permanent and transitory income inequality. Short term fluctuations in individual income, e.g. reflecting brief unemployment spells, or low initial incomes reflecting study or apprenticeship periods, have different political implications than lasting income inequality. Some studies try to filter out non-permanent income inequality by extending the accounting period over which income is measured. (See e.g. Aaberge et al., 2002.)

Below, we first consider the effect on redistribution of changes in permanent inequality, keeping mobility constant. Second, we examine the applicability of the mean vs median comparison as a predictor of the existence and extent of redistribution in our model. Finally, we study the effect of changes in the relative size of different income groups in the population.

If we keep mobility constant in our model, changes in inequality can only occur

through changes in the respective group's initial incomes.

Proposition 3 *An increase in a_L strictly decreases τ^* for $\tau^* > 0$. The reverse is true for an increase in a_H .*

Proof: In the Appendix.

Thus, increased inequality between groups unambiguously increases redistribution. As we shall see later, this result may not hold if groups of voters with persistently low or high incomes, who never become median voters, are introduced.

Now, let us turn to the mean vs median measure of inequality. As in all median voter models, the equilibrium tax rate in our model is determined by the voter with the median income. However, the relevant income concept in our model is the average income over the electoral period, $\bar{y}_i(\theta)$, a forward looking concept. Mobility makes the corresponding income distribution differ from the distribution of current incomes. Consequently, a comparison of the current mean versus median income, $\bar{Y} > \tilde{Y}$, may not correctly reflect incentives for redistribution.

Below we examine the extent to which income inequality measured as the difference between the current mean and median incomes can explain the existence of income redistribution. Moreover, we examine whether a more unequal income distribution, as measured by this income difference, is associated with more redistribution in equilibrium.

Proposition 4 *(i) $\bar{Y} - \tilde{Y} > 0 \Rightarrow \Phi(0) > 0.5$ if and only if $\delta = 0$. (ii) Consider two income distributions $f(y)$ and $f(y)'$. Then $\bar{Y} - \tilde{Y} \leq \bar{Y}' - \tilde{Y}'$ does not imply $\tau^* \leq \tau^{*}$.*

Proof: In the Appendix.

In conclusion, the mean versus median comparison is an insufficient condition for the occurrence of redistribution.¹⁰ It is also uninformative about the relative support for redistribution in comparisons between different income distributions, i.e., increased inequality as measured by $\bar{Y} - \tilde{Y}$ is not necessarily associated with more income redistribution.

¹⁰In fact, this criterion is sufficient only when the duration of redistributive policies is infinitely short. Note that if $b_L = b_H = 0$ the income distribution degenerates to two points, at a_L and a_H . Hence, $\bar{Y} > \tilde{Y}$ then implies a strictly positive tax rate.

Finally, we discuss the composition of the population in terms of the proportions of low skilled, p . This factor can differ considerably between countries.¹¹ A change in p has two opposing effects: The direct effect of a higher p is to increase the support for redistribution. However, such a change also erodes the tax base and makes redistribution less attractive to median income voters. The net effect cannot generally be signed. Four cases arise depending on whether $\theta_L \geq 1 - \delta$ and $\theta_H \geq 0$. Below we consider the case where electoral periods are short.

Proposition 5: τ^* is quasi-concave in p for $\tau^* > 0$ and δ such that $\theta_L < 1 - \delta$ and $\theta_H > 0$.

Proof: In the Appendix.

This suggests that an inverse U-shaped pattern could exist between p and the equilibrium tax rate. (Obviously, at the respective end points, $p = 0$ and $p = 1$, the equilibrium tax rate would be zero.)

2.2 The influence of non-median voter groups

So far we have assumed that all individuals start out with below-average incomes but eventually attain incomes above this level. In reality, some groups have bleak prospects of ever attaining the average income, e.g. agents who face long-term unemployment or face little or no wage development over time (“working poor”). Others may have above average incomes throughout their working lives.

Naturally, working poor prefer more redistribution than the median voter while the opposite holds for those with consistently high incomes. Hence, the direct effect of an increased number of poor voters, say due to changing labor market conditions or migration, is to increase the support for redistribution. There is also a counteracting indirect effect caused by the reduction in the mean income, which lowers the feasible transfer level and reduces the median voter’s enthusiasm for redistribution.¹²

¹¹See Blondal, Field & Girouard [2002]. Ultimately the proportion of skilled workers is endogenous and is influenced by the expected degree of redistribution in the future. However, this is beyond the scope of our analysis. See Hassler, Rodríguez-Mora, Storesletten and Zilibotti [2003] for an interesting contribution in this vein.

¹²Razin, Sadka & Swagel [2002] call this effect fiscal leakage. They model the effects of migration

We briefly consider this tradeoff. Suppose a small fraction of the population γ has a constant income, $\bar{y}_P < \bar{y}_L$. The mean income in the population is then $\bar{Y}_{NM} = (1 - \gamma)\bar{Y} + \gamma\bar{y}_P$. The threshold ages are still given by (9), using \bar{Y}_{NM} as defined above instead of \bar{Y} . In analogy with (10) the poverty-adjusted equilibrium tax condition is

$$\Phi(\tau^*) = (1 - \gamma) [p\theta_L(\tau^*) + (1 - p)\theta_H(\tau^*)] + \gamma \equiv 0.5 \quad (11)$$

where $\Phi(\tau)$ is continuous in γ . We illustrate the possibility that a higher γ can lead to less redistribution in the case when $\theta_L(\tau^*) < 1 - \delta$ and $\theta_H(\tau^*) > 0$. Using (11), the first derivative of $\Phi(\tau^*)$ with respect to γ evaluated at $\gamma = 0$, can be expressed as

$$p \frac{(1 - \rho\tau^*)\bar{y}_P - \bar{y}_L(0)}{b_L} + (1 - p) \frac{(1 - \rho\tau^*)\bar{y}_P - \bar{y}_H(0)}{b_H}. \quad (12)$$

The support for redistribution tends to increase in γ if $\rho\tau^*$ is small, which is the case if income differences between groups is small and mobility high, and $\bar{y}_i(0)$ is low, which is the case for small δ . Conversely, the support for redistribution initially decreases in γ if \bar{y}_P is less than $\min\{\bar{y}_L(0), \bar{y}_H(0)\}$, i.e., if the poor are poorer than even the youngest L and H voters. (When electoral periods are long, this is more likely to be the case. If $\delta = 1$ it is always the case.) Note that the incidence of persistent poverty, or low pay, seems to be higher in the US and the UK than in other OECD countries (see OECD, 1997, chapter 2). The former also have less extensive redistribution.

2.3 Empirical illustration

Our model implies that increased income mobility for L -voters (the majority group) should lead to less redistribution. Furthermore, Proposition 5 suggests that an inverted U-shaped relation could exist between transfers and the relative size of educational groups. Below we take a brief look at how this compares with actual patterns of redistribution, education and mobility.

on redistribution and examine the relative importance of the above effects using data on eleven European countries. They find that the fiscal leakage effect is important and may dominate.

We have data on redistribution, mobility and education for 20 OECD countries.¹³ The level of redistribution (τ in the model) is measured as *Public social expenditure as percentage of GDP*, for OECD countries for the year 1997 (Roman et al. [2001]). Our mobility measure is *The percentage increase in income when moving from quintile 3 to quintile 4* in the market income distribution.¹⁴ ¹⁵ The proportion of L -types (p in the model) is measured as the percentages with *Less than upper secondary* and *Upper secondary* education (OECD, 1998).

An OLS regression of social spending on income mobility, the proportion of uneducated, and this variable squared, yields the following results.

$$\tau = - 63.701 - 2.443b_L + 2.787p - 0.019p^2$$

(-1.24)
(-2.17)
(2.19)
(-2.16)

where t-values are in parentheses. The signs of all three coefficients are consistent with the predictions of the model.

3 Electoral periods and “myopia”

In a life-time perspective, H -agents are clearly better off with no redistribution. Since taxation is costly and short electoral periods induce strong redistribution incentives, this could perhaps also be true for L -agents. However, this is not the case.

Proposition 6 *The L -individual’s life-time income is higher with equilibrium redistribution than with no redistribution.*

¹³The countries are Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Ireland, Italy, Luxembourg, Mexico, Netherlands, Norway, Portugal, Spain, Sweden, UK and US.

¹⁴We use decile data from the WIDER World Income Inequality Database, <http://www.wider.unu.edu/wiid/wiid.htm> (the shares are estimated from Luxembourg Income Study data). For some countries data is not available after 1986, and for commensurability we have chosen available years between 1985-92.

¹⁵Ideally, we would like a measure that captures both differences in transition probabilities and the magnitude of income changes. Transition probabilities between equal-width earnings bands (defined as proportions of the median wage) has this feature but is only available for six countries. (See OECD, 1996, chapter 3). Since the quintile transition probabilities are quite similar between countries we have instead chosen to only measure the income differences between the quintiles. Moreover, we would have liked to have group specific mobility measures. There are country studies of education specific age-earnings profiles but cross-country comparisons are problematic. (See, e.g., Brunello & Comi, 2004.) Since the proportion with low education is generally quite large (the average is 79.4), we interpret the measure as mainly reflecting income mobility for individuals with low education.

Proof: An L -agent's life-time after-tax income plus transfers, $(1-\tau)\bar{y}_L + \tau(1-\frac{\rho\tau^*}{2})\bar{Y} = \frac{1}{2}(\bar{Y} + \bar{y}_L(\theta_L(\tau^*)))$, exceeds her gross life time income, \bar{y}_L , since $\theta_L(\tau^*) > \frac{1}{2}$. ■

The median L -voter is of above-average age and has an income above \bar{y}_L . If $\tau^* > 0$ this voter obviously gains from taxation. Since the net gain from taxation decreases linearly with age, the aggregate gains for the younger L -voters outweigh the aggregate losses for the older. (Concave age-earnings profiles are likely to accentuate this effect.)

The support for redistribution is influenced by the duration of redistributive policies, or the length of the electoral period, δ . As would be expected, a longer time horizon tends to decrease income redistribution.¹⁶ Specifically,

Proposition 7 τ^* decreases in δ and strictly so for $\theta_H(\tau^*) > 0$ and/or $\theta_L(\tau^*) < 1-\delta$.

Proof: In the Appendix.

An increase in δ also affects the composition of the voter support for redistribution. It raises the period income, which reduces θ_i , but the decrease in τ (Proposition 7) has a positive indirect effect on θ_i . The latter effect is stronger for L -types, whose age-earnings profiles are less steep. Since $\Phi(\tau^*) = 0.5$ in equilibrium, changes in θ_L and θ_H must have opposite signs. Thus, θ_L increases and θ_H decreases in δ . (If no H -type supports τ^* then $\Phi(\tau^*) = p\theta_L(\tau^*)$ and δ will not affect $\theta_L(\tau^*)$). Formally,

Corollary 1 $\theta_H(\tau^*)$ strictly decreases and $\theta_L(\tau^*)$ strictly increases in δ for δ such that $\theta_H(\tau^*) > 0$.

Proof: In the Appendix.

We could say that the longer the electoral periods, the more the dividing line in the redistributive conflict goes between skill-groups (classes), while with shorter electoral periods the dividing line goes between age groups (generations).

How, then, does the length of the electoral period affect the welfare of the different groups? Proposition 6 tells us that L -types are better off with equilibrium redistribution than with *no redistribution* but is silent about whether more, or slightly less, redistribution would be even better. However, the results, in Corollary 2 below, are in line with Proposition 6.

¹⁶This result is similar to that in proposition 2 in Bénabou and Ok [2001].

Corollary 2 For $\tau^* > 0$, the life-time income for L -agents decreases in δ . The life-time income for H -agents increases in δ .

Proof: In the Appendix.

Thus, the groups have diametrically opposing interests concerning the length of the electoral period.

Ultimately the length of the electoral period is determined in a political process, where also considerations such as political accountability are likely to be important. Below we examine this choice, but focus exclusively on the redistributive aspects. Consider an extension of the model where the electorate first faces a once-and-for-all vote on δ and then votes on τ in subsequent elections. Once δ is set, τ^* will be the same in all future elections. Hence, when voting on δ agents are concerned with the effect on the *remaining lifetime* income (rather than the period income, equation (6)).

$$\bar{\Upsilon}_i^{RL}(\tau^*(\delta), \theta) = (1 - \tau^*(\delta)) \frac{1}{1 - \theta} \int_{\theta}^1 y_i(t) dt + \tau^*(\delta) \left(1 - \frac{\rho\tau^*(\delta)}{2}\right) \bar{Y}$$

The marginal effect of a higher δ on $\bar{\Upsilon}_i^{RL}(\tau^*(\delta), \theta)$ is a product of two factors, $\partial\bar{\Upsilon}_i^{RL}/\partial\tau$ and $\partial\tau^*/\partial\delta$. Since $\bar{\Upsilon}_i^{RL}(\tau^*(\delta), \theta) \geq \bar{\Upsilon}_i(\tau, \theta)$ the first factor is negative for the median voter. Thus for $\partial\tau^*/\partial\delta < 0$ (Proposition 7) the product is strictly positive. This means that the constitutionally selected δ must be sufficiently long for $\theta_H(\tau^*) = 0$ and $\theta_L(\tau^*) \geq 1 - \delta$. Any such electoral period yields the same degree of redistribution as if the electoral period had covered the remaining life-span for all voters. In sum,

Proposition 8 (i) There exists a $\underline{\delta}$ such that $\tau^*(\delta) = \tau^*(1)$ for $\delta \in [\underline{\delta}, 1]$. (ii) The equilibrium $\delta \geq \underline{\delta}$.

Proof: In the Appendix.

Hence, constitutional voting results in very long electoral periods and a low level of redistribution. This outcome is not favorable for L -types. The reason it still receives the necessary support is, of course, that the median voter is an older L -individual who has no incentives to take care of the welfare of coming generations of L -agents.

4 Retirement

The assumption that earnings increase throughout the voter's life-span does not accord with the fact that people do retire. Below we study a mandatory fully funded retirement system where individual benefits are determined by individual contributions and the retirement age. Retirees may favor redistribution and the support for redistribution depends on the size of benefits. We focus on how changes in contributions and the retirement age affect equilibrium redistribution.

Let retirement benefits be financed by flat deductions, d_i , from agents' taxable income throughout their working lives, i.e. $y_i(\theta) = a_i + b_i\theta - d_i$. Accumulated savings are paid out from the retirement age, r , as a flat taxable retirement income $d_i \frac{r}{1-r}$. For incomes to remain non-negative we require that $a_i > d_i$. For simplicity we assume that $\delta = 1$. Now, three groups are potentially in favor of redistribution, young L -types and agents of both types approaching retirement (when $\delta = 1$, $\theta_H(\tau^*) = 0$).

Figure 2 about here

The tax base depends on the retirement age: $\bar{Y}(r) = rp(a_L + \frac{b_L}{2}r) + r(1-p)(a_H + \frac{b_H}{2}r)$. The average future income of an individual is (c.f. the second case in (5)),

$$\bar{y}_i(\theta) = \begin{cases} \frac{1}{1-\theta} [(a_i + \frac{b_i}{2}(r + \theta))(r - \theta) + d_i\theta] & \text{if } \theta \leq r \\ \frac{d_i r}{1-r} & \text{if } \theta > r \end{cases} \quad (13)$$

which is strictly quasi-concave in θ .¹⁷ Since retirees may also strictly favor redistribution, and by continuity some individuals close to retirement, there may exist two upper threshold ages, $\theta_L^R(\tau)$ and $\theta_H^R(\tau)$ respectively, above which agents favor more redistribution. All thresholds are implicitly defined by

$$(1 - \rho\tau)\bar{Y}(r) = \bar{y}_i(\theta) \quad (14)$$

The support for redistribution is $\Phi(\tau) = p[\theta_L(\tau) + 1 - \theta_L^R(\tau)] + (1-p)[1 - \theta_H^R(\tau)]$. Thus effects that increase $\theta_L(\tau)$ and decrease $\theta_L^R(\tau)$ and $\theta_H^R(\tau)$ increase τ^* .

¹⁷ $\frac{\partial \bar{y}_i}{\partial \theta} = \frac{1}{1-\theta} [\bar{y}_i - a_i - b_i\theta + d_i]$ and $\frac{\partial^2 \bar{y}_i}{\partial \theta^2} = \frac{1}{1-\theta} [2\frac{\partial \bar{y}_i}{\partial \theta} - b_i]$

We now examine the effects of retirement savings, d_i , on redistribution. Higher contributions reduce the disposable income of those who work while making retirees better off. Consequently, θ_L decreases while the upper thresholds increase. Hence,

Proposition 9 *The equilibrium tax decreases in pension contributions.*

Proof: In the Appendix.

Now, let us consider the effects of changes in the retirement age. A higher retirement age raises the period (life-time) incomes, $\bar{y}_i(\theta)$, for all voters, but it also increases the tax base, \bar{Y} . The net effect on the equilibrium tax cannot be determined. Below we consider the effects on the support for redistribution in the different groups.

Proposition 10 *An increase in r strictly increases θ_H^R . The support for redistribution in the L -group, $\theta_L(\tau) + 1 - \theta_L^R(\tau)$, may increase or decrease but both θ_L and θ_L^R cannot decrease.*

Proof: In the Appendix.

The first result follows since for H -types a higher r increases the income more than the tax base. For L -types, a higher r benefits old voters more than young ones.

5 An ageing baby-boom

Demographic changes, such as baby-booms, declining nativity and wars alter the age structure of the population and have economic and political effects that evolve and change over time. One currently much discussed issue is what effect the ageing of the 1940s baby-boom will have on welfare state arrangements. Below we examine how the support for redistributive policies is affected by a baby-boom in our model. We abstract from retirement in the analysis but provide a simple numerical example concerning the timing of welfare state reforms, where we take retirement into account.

A baby-boom gives increased political clout to its age group, and as the group ages, its taste for income redistribution subsides. However, the rising incomes of the baby-boomers increases the tax base and make others more interested in redistribution.

The interaction between these effects determines how the support for redistribution changes over time.

To simplify the analysis we assume that the baby-boom occurs at a single point in time and that it constitutes a fraction, γ , of the population. The proportions of L and H -individuals in the baby-boom are the same as in the rest of the population. The income of the baby-boomers is $y_B(\theta_B) = py_L(\theta_B) + (1-p)y_H(\theta_B)$, where $y_i(\theta_B)$ is given by (1). The average income in the population is $\bar{Y} = (1-\gamma)\bar{Y}_{NB} + \gamma y_B(\theta_B)$, where \bar{Y}_{NB} is the average income of non-baby-boomers. Since $\theta_H(\tau) < \theta_L(\tau)$, the support for a tax τ , $\Phi(\tau, \theta_B)$ as a function of θ_B , is determined in one of three regimes, differing in the fraction of baby-boomers who support higher taxes, $\Gamma(\tau, \theta_B)$.

$$\Phi(\tau, \theta_B) = (1-\gamma)[p\theta_L(\tau, \theta_B) + (1-p)\theta_H(\tau, \theta_B)] + \gamma\Gamma(\tau, \theta_B) \quad (15)$$

where

$$\Gamma(\tau, \theta_B) = \begin{cases} 1 & \text{if } \theta_B \leq \theta_H(\tau, \theta_B) \\ p & \text{if } \theta_H(\tau, \theta_B) \leq \theta_B \leq \theta_L(\tau, \theta_B) \\ 0 & \text{otherwise} \end{cases}$$

The function Γ describes three phases in the baby-boom's effect on the support for redistribution. Initially, the baby-boom may increase or decrease the support for redistribution, since $y_B(0) < \bar{Y}$ (cf. the case with a group of permanently poor).¹⁸ Generally, as long as no baby-boomer is a median voter, their rising incomes increase the support for redistribution over time. As θ_B reaches θ_i , some baby-boomers become median voters, and remain so for an interval of time. (An ordinary voter is only a median voter for an instant before being replaced by a slightly younger voter. By contrast baby-boomers remain median voter for as long as it takes to replace the whole group). During this time the support for taxation must decrease. Since there are at most two thresholds, θ_H and θ_L , a baby-boom can create one or two "cycles"

¹⁸This can be seen for the case when $\theta_H(\tau) > 0$ by differentiating (15) with respect to γ . We then have

$$\frac{\partial \Phi}{\partial \gamma} = \frac{1}{2(1-\gamma)} - (1-\gamma)(\bar{Y}_{NB} - y_B(\theta_B)) \left[p \frac{\partial \theta_L}{\partial \bar{Y}} + (1-p) \frac{\partial \theta_H}{\partial \bar{Y}} \right]$$

which is positive if $\bar{Y}_{NB} - y_B(\theta_B)$ is not too large but may otherwise be negative.

with increasing and then decreasing support for redistribution. Note that the extent to which this translates into corresponding variations in the tax rate depends on the frequency of elections, i.e., δ . Figure 3 illustrates a case with $\delta = 0$.

Figure 3 about here

Now, let us consider a simple numerical example concerning the timing of the (second) peak of welfare state spending, where we take retirement into account. Suppose agents vote between the ages of 20 and 80 years and retire at 60. Moreover, suppose (the parameters are such that) all H -types, including retirees, vote for lower taxes, while retired L -types vote for higher taxes. If we abstract from forward-looking behavior of L -types approaching retirement, we can approximate the condition determining the age of the median L -voter by

$$\frac{60 - 20}{60}p \theta_L + \frac{80 - 60}{60}p = 0.5$$

where the first term reflects the support for higher taxes from individuals in working age and the second term that from retired L -types. Suppose finally that $p = 0.8$, implying that $\theta_L \simeq 0.44$. This translates into a physical age of $20 + 60 \cdot 0.44 = 46.4$ years. If the baby-boom began in 1945, then welfare state spending would have peaked in 1991. While this lags somewhat behind the tax reforms and the retreats in European welfare states, there are obviously many factors not accounted for in this example. For instance, the presence of working poor would tend to shift the peak year back in time. Taking forward-looking behavior into account would also have this effect. By contrast, a higher p , or a later retirement age, would have the opposite effect.

6 Concluding remarks

The median voter model in its most narrow form has not been entirely successful in explaining redistribution. Extended versions of the model seem to perform better

(see e.g. Bénabou & Ok, 2001, who consider mobility, and Razin, Sadka & Swagel, 2002, who consider between-group inequality).

In this paper we have examined how income mobility and income inequality interact in determining income redistribution in a modified median voter model. Mobility and inequality can arise in different ways, e.g., mobility may differ between groups, and inequality can exist between generations as well as between groups. The model shows that the effects of mobility and inequality on the support for redistribution are not uniform but depend on the determinants of mobility and inequality. For example, higher mobility in low-skilled groups reduces the support for redistribution, while higher mobility among the high-skilled has the opposite effect. Also, a wider income gap between educational groups tends to increase redistribution but an increase in the number of working poor can lead to less redistribution.

We also examine the traditional comparison between the current mean and median incomes in the population in the context of our model, and find that such comparisons are uninformative about the political support for redistribution. Moreover, we consider the effect of the length of electoral periods. Short periods accentuate the redistributive conflict between generations while long periods make the equilibrium tax rate more sensitive to between-group inequality. A majority vote on the period length minimizes the equilibrium tax rate, to the detriment of low-skilled voters.

We believe that it would be valuable to study these issues empirically, as more comparable cross-country data on mobility and earnings in different educational groups becomes available. Our admittedly very rough empirical illustration is somewhat encouraging in this respect.

Appendix

Proof of Proposition 2: The effect on θ_i of increased mobility keeping \bar{y}_i constant is $\left. \frac{\partial \theta_i}{\partial b_i} \right|_{d\bar{Y}=0} = \frac{\partial \theta_i}{\partial b_i} - \frac{1}{2} \frac{\partial \theta_i}{\partial a_i}$, which is derived by differentiating the threshold $\theta_i(\tau)$ s with respect to b_i while imposing $d\bar{y}_i = da_i + \frac{1}{2}db_i = 0$. Since \bar{Y} remains constant, the threshold age of the other type is not affected and $\left. \frac{\partial \Phi}{\partial b_i} \right|_{d\bar{Y}=0}$ has the same sign as $\left. \frac{\partial \theta_i}{\partial b_i} \right|_{d\bar{Y}=0}$. This also gives us the sign of the effect on τ^* ($\left. \frac{\partial \tau^*}{\partial b_i} \right|_{d\bar{Y}=0} = -\left. \frac{\partial \Phi}{\partial b_i} \right|_{d\bar{Y}=0} / \left. \frac{\partial \Phi}{\partial \tau} \right)$, since Φ is strictly decreasing in τ .

The partial derivatives of the $\theta_i(\tau)$ s (expression (9)) depend on whether the age of the threshold agent is above or below $1 - \delta$. Excluding $\theta_L = 0$ since $\theta_L(\tau^*) > 0$ for $\tau^* \geq 0$ the partial derivatives are

$$\begin{aligned} \frac{\partial \theta_L}{\partial a_L} &= \begin{cases} \frac{(1-\rho\tau)p}{b_L} - \frac{1}{b_L} \\ \frac{2(1-\rho\tau)p}{b_L} - \frac{2}{b_L} \end{cases} & \frac{\partial \theta_H}{\partial a_H} &= \begin{cases} 0 \\ \frac{(1-\rho\tau)(1-p)}{b_H} - \frac{1}{b_H} \end{cases} \\ \frac{\partial \theta_L}{\partial b_L} &= \begin{cases} \frac{1}{b_L} \left[\frac{(1-\rho\tau)p}{2} - \frac{\delta}{2} \right] - \frac{\theta_L(\tau^*)}{b_L} \\ \frac{2}{b_L} \left[\frac{(1-\rho\tau)p}{2} - \frac{1}{2} \right] - \frac{\theta_L(\tau^*)}{b_L} \end{cases} & \frac{\partial \theta_H}{\partial b_H} &= \begin{cases} 0 \\ \frac{1}{b_H} \left[\frac{(1-\rho\tau)(1-p)}{2} - \frac{\delta}{2} \right] - \frac{\theta_H(\tau^*)}{b_H} \end{cases} \end{aligned}$$

Hence,

$$\left. \frac{\partial \theta_L}{\partial b_L} \right|_{d\bar{Y}=0} = \begin{cases} \frac{1}{b_L} \left(\frac{1}{2} - \frac{\delta}{2} - \theta_L(\tau^*) \right) < 0 \\ \frac{2}{b_L} \left(\frac{1}{2} - \frac{1}{2} - \theta_L(\tau^*) \right) < 0 \end{cases} \quad \left. \frac{\partial \theta_H}{\partial b_H} \right|_{d\bar{Y}=0} = \begin{cases} 0 \\ \frac{1}{b_H} \left(\frac{1}{2} - \frac{\delta}{2} - \theta_H(\tau^*) \right) > 0 \end{cases}$$

which gives the result.

A proportional compensated increase in b_L and b_H is defined as an increase in β and a reduction in α_L and α_H such that $\bar{y}_i = \alpha_i a_i + \beta b_i$, and consequently also \bar{Y} , remain constant. Thus, $\frac{\partial \alpha_i}{\partial \beta} = -\frac{1}{2} \frac{b_i}{a_i}$. We evaluate $\frac{\partial \theta_i}{\partial \beta}$ at $\alpha_i = \beta = 1$. Again excluding $\theta_L = 0$, $\frac{\partial \theta_L}{\partial \beta} = \left\{ \frac{1-\delta}{2} - \theta_L, -\theta_L \right\}$, which is negative in both cases, and $\frac{\partial \theta_H}{\partial \beta} = \left\{ 0, \frac{1-\delta}{2} - \theta_H \right\}$. Then $\frac{\partial \Phi}{\partial \beta} = p \frac{\partial \theta_L}{\partial \beta} + (1-p) \frac{\partial \theta_H}{\partial \beta} < 0$ in all cases since $p\theta_L + (1-p)\theta_H = \frac{1}{2}$. ■

Proof of Proposition 3: The effects on τ are derived implicitly from $\Phi(\tau^*) \equiv 0.5$. Since Φ is strictly decreasing in τ , the effect on τ^* of a change in a_i has the same sign as $\frac{\partial \Phi}{\partial a_i} = p \frac{\partial \theta_L}{\partial a_i} + (1-p) \frac{\partial \theta_H}{\partial a_i}$ evaluated at τ^* . In addition to $\frac{\partial \theta_L}{\partial a_L}$ and $\frac{\partial \theta_H}{\partial a_H}$, given in the above proof, we also need,

$$\frac{\partial \theta_H}{\partial a_L} = \begin{cases} 0 \\ \frac{(1-\rho\tau)p}{b_H} - \frac{1}{b_H} \end{cases} \quad \frac{\partial \theta_L}{\partial a_H} = \begin{cases} \frac{(1-\rho\tau)(1-p)}{b_L} \\ \frac{2(1-\rho\tau)(1-p)}{b_L} \end{cases}$$

since \bar{Y} is not constant. For each of the comparative static effects there are four possible cases: $\theta_L(\tau^*) \leq 1 - \delta$ and $\theta_H(\tau^*) = 0$; $\theta_L(\tau^*) \leq 1 - \delta$ and $\theta_H(\tau^*) > 0$; $\theta_L(\tau^*) > 1 - \delta$ and $\theta_H(\tau^*) = 0$; $\theta_L(\tau^*) > 1 - \delta$ and $\theta_H(\tau^*) > 0$. We compute $\frac{\partial \Phi}{\partial a_i}$ for each case, using $\Phi(\tau^*) \equiv 0.5$, assumptions (A1) - (A3) and Proposition 1. The effects of changes in a_i for each case are listed below.

$$\begin{aligned}
\text{(L1)} \quad & \frac{\partial \Phi}{\partial a_L} = \frac{p}{b_L} [(1 - \rho\tau)p - 1] < 0 \\
\text{(L2)} \quad & \frac{\partial \Phi}{\partial a_L} = \frac{p}{b_L} \left[(1 - \rho\tau) \left(p + \frac{b_L}{b_H} (1 - p) \right) - 1 \right] < 0 \\
\text{(L3)} \quad & \frac{\partial \Phi}{\partial a_L} = \frac{2p}{b_L} [(1 - \rho\tau)p - 1] < 0 \\
\text{(L4)} \quad & \frac{\partial \Phi}{\partial a_L} = \frac{p}{b_L} \left[(1 - \rho\tau) \left(2p + \frac{b_L}{b_H} (1 - p) \right) - 2 \right] < 0 \\
\text{(H1)} \quad & \frac{\partial \Phi}{\partial a_H} = \frac{p}{b_L} (1 - \rho\tau)(1 - p) > 0 \\
\text{(H2)} \quad & \frac{\partial \Phi}{\partial a_H} = \frac{1-p}{b_H} \left[(1 - \rho\tau) \left(\frac{b_H}{b_L} p + (1 - p) \right) - 1 \right] > 0 \\
\text{(H3)} \quad & \frac{\partial \Phi}{\partial a_H} = \frac{2p}{b_L} (1 - \rho\tau)(1 - p) > 0 \\
\text{(H4)} \quad & \frac{\partial \Phi}{\partial a_H} = \frac{1-p}{b_H} \left[(1 - \rho\tau) \left(2\frac{b_H}{b_L} p + (1 - p) \right) - 1 \right] > 0
\end{aligned}$$

where H2 and H4 follow since Proposition 1 (ii), $\theta_L(\tau^*) > 0.5$, implies that

$$(1 - \rho\tau) \left[\frac{pb_H}{b_L} + (1 - p) \right] > \frac{\bar{y}_L}{\bar{Y}} \frac{[pb_H - (1-p)b_L]}{b_L} = \frac{\bar{y}_L}{b_L} \frac{[p(b_H - b_L) + b_L]}{[\bar{y}_L + (1-p)\Delta\bar{y}]} > 1$$

since by (A1) and (A2) $p(b_H - b_L) > \Delta\bar{y}$. ■

Proof of Proposition 4: (i) For $\delta > 0$ and $\tau = 0$, (9) is given by $\theta_L(0) = \max \left\{ 0, \frac{1-p}{b_L} \Delta\bar{y} + \frac{1-\delta}{2}, 2\frac{1-p}{b_L} \Delta\bar{y} \right\}$ and $\theta_H(0) = \max \left\{ 0, -\frac{p}{b_H} \Delta\bar{y} + \frac{1-\delta}{2} \right\}$. Since (A2) only requires $\Delta\bar{y} > 0$, $\Phi(0) = p\theta_L(0) + (1-p)\theta_H(0) < 0.5$ for any sufficiently small $\Delta\bar{y}$. But if $\delta = 0$ then $\Phi(0) = \frac{1}{2} + p(1-p)\Delta\bar{y} \left(\frac{1}{b_L} - \frac{1}{b_H} \right) > 0.5$. (ii) Suppose $f(y)$ is such that $\tau^* > 0$ and $\theta_H(\tau^*) = 0$. We construct $f(y)'$ from $f(y)$ by compensated changes in b_i . A small compensated increase in b_L (b_H) strictly increases (decreases) \tilde{Y} , while \bar{Y} remains unchanged, i.e. $\bar{Y}' = \bar{Y}$. We can balance compensated increases in b_L and b_H so that $\tilde{Y}' = \tilde{Y}$. However, Proposition 2 implies $\tau^* > \tau^{*'}$, i.e. a contradiction. Moreover, an additional slight increase in b_H yields $\bar{Y} - \tilde{Y} < \bar{Y}' - \tilde{Y}'$ and $\tau^* > \tau^{*'}$. ■

Proof of Proposition 5: Differentiating implicitly gives

$$\frac{\partial \tau^*}{\partial p} = -\frac{\frac{\partial \Phi}{\partial p}}{\frac{\partial \Phi}{\partial \tau}} = -\frac{p\frac{\partial \theta_L}{\partial p} + (1-p)\frac{\partial \theta_H}{\partial p} + \theta_L - \theta_H}{p\frac{\partial \theta_L}{\partial \tau} + (1-p)\frac{\partial \theta_H}{\partial \tau}}.$$

The second derivative can be written

$$\frac{\partial^2 \tau^*}{\partial p^2} = -\frac{\frac{\partial^2 \Phi}{\partial p^2} \frac{\partial \Phi}{\partial \tau} - \frac{\partial \Phi}{\partial p} \frac{\partial^2 \Phi}{\partial \tau \partial p}}{\left(\frac{\partial \Phi}{\partial \tau}\right)^2}$$

and is negative for $\frac{\partial \Phi}{\partial p} = 0$ since $\frac{\partial \Phi}{\partial \tau} < 0$ and $\frac{\partial^2 \Phi}{\partial p^2} = 2\left(\frac{\partial \theta_L}{\partial p} - \frac{\partial \theta_H}{\partial p}\right) < 0$ (see expression (9)). ■

Proof of Proposition 7: The effect of δ on τ^* can be derived implicitly from $\Phi(\tau^*) \equiv 0.5$: $\frac{\partial \tau}{\partial \delta} = -\frac{\partial \Phi / \partial \delta}{\partial \Phi / \partial \tau}$. It is straightforward to show, using (9), that the denominator is strictly negative for $\theta_L(\tau^*, \delta) > 0$, which is true for any $\tau^* \geq 0$. Similarly, it is easy to see that $\frac{\partial \Phi}{\partial \delta} \leq 0$ with equality only if both $\theta_L(\tau^*) > 1 - \delta$ and $\theta_L(\tau^*) = 0$, when $\frac{\partial \theta_L}{\partial \delta} = \frac{\partial \theta_H}{\partial \delta} = 0$. ■

Proof of Corollary 1: First, the effect of δ on the $\theta_i(\tau^*)$ is $\frac{d\theta_i}{d\delta} = \frac{\partial \theta_i}{\partial \delta} + \frac{\partial \theta_i}{\partial \tau} \frac{\partial \tau}{\partial \delta}$. Second, note that $\frac{\partial \Phi}{\partial \delta} = p \frac{\partial \theta_L}{\partial \delta} + (1-p) \frac{\partial \theta_H}{\partial \delta} = 0$ and thus the partial derivatives for the different types must have opposite signs. We now evaluate $\frac{d\theta_i}{d\delta}$ for the different cases that can arise in equation (9).

Since $\theta_L(\tau^*, \delta) > 0$ we only need to consider whether $\theta_L(\tau^*) > 1 - \delta$ or not, and whether $\theta_H(\tau^*) = 0$ or not. If $\theta_H(\tau^*) = 0$, then $\frac{d\theta_H}{d\delta} = 0$ which implies that also $\frac{d\theta_L}{d\delta} = 0$. If $\theta_H(\tau^*) > 0$ and $\theta_L(\tau^*) > 1 - \delta$ then $\frac{\partial \theta_L}{\partial \delta} = 0$ and thus $\frac{d\theta_L}{d\delta} > 0$ and therefore $\frac{d\theta_H}{d\delta} < 0$. Finally, if $\theta_H(\tau^*) > 0$ and $\theta_L(\tau^*) \leq 1 - \delta$, then $\frac{\partial \theta_L}{\partial \delta} = \frac{\partial \theta_H}{\partial \delta} = -\frac{\delta}{2}$ but since $\frac{\partial \theta_L}{\partial \tau} < \frac{\partial \theta_H}{\partial \tau} < 0$ we again have that $\frac{d\theta_L}{d\delta} > 0$ and $\frac{d\theta_H}{d\delta} < 0$. ■

Proof of Corollary 2: The life-time income as a function of δ is $\bar{Y}_i(\tau^*(\delta), 0) = (1 - \tau^*(\delta))\bar{y}_i + \tau^*(\delta)(1 - \frac{\rho\tau^*(\delta)}{2})\bar{Y}$. The effect of δ goes only via τ^* . For $\tau^* > 0$, condition (7) holds with equality, i.e., $(1 - \rho\tau^*)\bar{Y} - \bar{y}_L(\theta_L) = 0$ implying that $\partial \bar{Y}_L / \partial \tau > 0$. By Proposition 7, $\partial \tau / \partial \delta \leq 0$ and for sufficiently low δ this holds strictly. For H -types, $\bar{y}_H > \bar{y}_H(\theta_H)$ and thus we have that $\partial \bar{Y}_H / \partial \tau < 0$, unless $\theta_H(\tau) = 0$ when $\partial \bar{Y}_H / \partial \tau = 0$, which gives the result. ■

Proof of Proposition 8:

(i) For sufficiently large $\delta \in [0, 1)$, $\theta_H(\tau^*, \delta) = 0$ and $\theta_L(\tau^*, \delta) \geq 1 - \delta$. The reason is that both $\bar{Y} - \bar{y}_H(0, \delta)$ and $1 - \delta - \theta_L(\tau^*, \delta)$ decrease monotonically in δ and equal

zero for a sufficiently large δ . Hence there exists a lowest δ , $\underline{\delta}$, such that $\theta_H(\tau^*, \delta) = 0$ and $\theta_L(\tau^*, \delta) \geq 1 - \delta$. By Proposition 7, τ^* is unaffected by further increases in δ .

(ii) Note that $\partial \bar{\Upsilon}_i^{RL} / \partial \tau \leq \partial \bar{\Upsilon}_i / \partial \tau$, where strict inequality holds for $\theta_i < 1 - \delta$ since the remaining lifetime then exceeds δ . Furthermore, for the median voter $\partial \bar{\Upsilon}_i / \partial \tau = 0$ at τ^* . Hence, if $\theta_H(\tau^*) > 0$ or $\theta_L(\tau^*) < 1 - \delta$ we have that $\partial \bar{\Upsilon}_i^{RL} / \partial \tau < 0$ for the median voter, who then prefers higher δ (Proposition 7). Consequently, the equilibrium δ must exceed $\underline{\delta}$. ■

Proof of Proposition 9: We derive the effect on θ_i implicitly from $(1 - \rho\tau)\bar{Y} = \bar{y}_i(\theta)$:

$$\frac{d\theta_i}{dd_i} = - \frac{\partial \bar{y}_i}{\partial d_i} \bigg/ \frac{\partial \bar{y}_i}{\partial \theta}$$

Note that $\partial \bar{y}_i / \partial d_i > 0$ and $\partial \bar{y}_i / \partial \theta = (\bar{y}_i(\theta) - y_i(\theta)) / (1 - \theta)$, where $\frac{\partial \bar{y}_L}{\partial \theta} > 0$ for $\theta_L(\tau)$; $\frac{\partial \bar{y}_L}{\partial \theta} < 0$ for $\theta_L^R(\tau)$ and $\frac{\partial \bar{y}_H}{\partial \theta} < 0$ for $\theta_H^R(\tau)$, since the latter thresholds are located to the right of the maximal $\bar{y}_i(\theta)$. Thus higher d_L reduces support for higher taxes among L -voters while not affecting H -voters, and higher d_H reduces support for higher taxes among H -voters and does not affect L -voters. This implies that τ^* decreases in d_i . ■

Proof of Proposition 10: The effect of r on θ_i can be derived implicitly from $(1 - \rho\tau)\bar{Y} = \bar{y}_i(\theta)$:

$$\frac{d\theta_i}{dr} = - \left((1 - \rho\tau)\bar{Y} \frac{\partial \bar{Y}}{\partial dr} - \frac{\partial \bar{y}_i}{\partial r} \right) \bigg/ \frac{\partial \bar{y}_i}{\partial \theta}.$$

Note that $\partial \bar{y}_i / \partial \theta = (\bar{y}_i(\theta) - y_i(\theta)) / (1 - \theta)$ is positive at low thresholds and negative at high, since the thresholds must be located on either side of the maximal $\bar{y}_i(\theta)$. ($\theta_H(0) = 0$ for $\delta = 1$). Now, if we expand the numerator we get

$$(1 - \rho\tau)(p(a_L + b_L r) + (1 - p)(a_H + b_H r)) - \frac{1}{1 - \theta}(a_i + b_i r).$$

This is strictly positive for H -types proving the first claim. Since $\frac{1}{1 - \theta}(a_i + b_i r)$ is greater at a higher threshold than at a lower one it follows that if the numerator is negative at the lower threshold then this is also the case at the higher threshold. Hence, we can rule out that $d\theta_L/dr < 0$ and $d\theta_L^R/dr < 0$. The other cases can occur. ■

References

- Aaberge, Rolf, Anders Björklund, Markus Jäntti, Mårten Palme, Peder J. Pedersen, Nina Smith, Tom Wennemo, (2002), "Income Inequality and Income Mobility in the Scandinavian Countries Compared to the United States". *Review of Income & Wealth*, 48(4), 443-470.
- Alesina, Alberto and Edward L. Glaeser, (2004), *Fighting poverty in the US and Europe: A world of difference*. Rodolfo Debenedetti Lectures. Oxford and New York: Oxford University Press.
- Alesina, Alberto, Edward Glaeser and Bruce Sacerdote, (2001), "Why Doesn't the United States Have a European-Style Welfare State?". *Brookings Papers on Economic Activity*, 2:2001, 187-277.
- Alesina Alberto and George-Marios Angeletos, (2005), "Fairness and Redistribution". *The American Economic Review*, 95(4).
- Besley, Timothy and Anne Case, (2003), "Political Institutions and Policy Choices: Evidence from the United States". *Journal of Economic Literature*, 41(1), 7-73.
- Bénabou, Roland, (2000), "Unequal Societies: Income Distribution and the Social Contract". *American Economic Review*, 90, 96-129.
- Bénabou, Roland and Effe Ok, (2001), "Social Mobility and the Demand for Redistribution: The POUM Hypothesis". *Quarterly Journal of Economics*, 116, 447-487.
- Berman, Eli, John Bound and Zvi Griliches, (1994), "Changes in the Demand for Skilled Labor within U.S. Manufacturing: Evidence from the Annual Survey of Manufacturers". *Quarterly Journal of Economics*, 109(2), 367-397.
- Bresnahan, Timothy, (1999), "Computerization and Wage Dispersion: An Analytical Reinterpretation". *The Economic Journal*, 109, F390-F415.
- Brunello, Giorgio and Simona Comi, (2004), "Education and Earnings Growth: Evidence from 11 European Countries". *Economics of Education Review*, 23(1), 75-83.
- Cameron, David, (1978), "The Expansion of the Public Economy: A Comparative Analysis". *American Political Science Review*, 72, 1243-1261.
- Card, David and Thomas Lemieux, (2001), "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis". *Quarterly Journal of Economics*, 116, 705-46.
- Chari, V. V. and Hugo Hopenhayn, (1991), "Vintage Human Capital, Growth, and the Diffusion of New Technology". *Journal of Political Economy*, 99(6), 1142-65.
- Fortin, Nicole M. and Thomas Lemieux, (1997), "Institutional Changes and Rising Wage Inequality: Is there a Linkage?". *Journal of Economic Perspectives*, 11(2), 75-96.

- Gottschalk, Peter and Timothy Smeeding, (1997), "Cross-National Comparisons of Earnings and Income Inequality". *Journal of Economic Literature*, 35, 633-687.
- Hassler, John, José V. Rodríguez Mora, Kjetil Storesletten and Fabrizio Zilibotti, (2003), "The Survival of the Welfare State". *American Economic Review*, 93(1), 87-112.
- Hibbs, Douglas A, Jr and Håkan Locking, (2000), "Wage Dispersion and Productive Efficiency: Evidence for Sweden", *Journal of Labor Economics*, 18(4), 755-82.
- Klevmarcken, N. Anders, (1982), "On the Stability of Age-Earnings Profiles". *Scandinavian Journal of Economics*, 84(4), 531-54.
- Hashimoto, Masanori and John Raisian, (1985), "Employment Tenure and Earnings Profiles in Japan and the United States". *American Economic Review*, 75(4), 721-35.
- Meltzer, Allan H. and Scott F. Richard, (1981), "A Rational Theory of the Size of Government". *Journal of Political Economy*, 5, 914-927.
- Murphy, Kevin M. and Finis Welch, (1992) "The Structure of Wages". *The Quarterly Journal of Economics*, 107(1), 285-326.
- Neumark, David and Paul Taubman, (1995), "Why Do Wage Profiles Slope Upward?". *Journal of Labor Economics*, 13(4), 736-761.
- OECD, (1996), Employment Outlook.
- OECD, (1997), Employment Outlook.
- OECD, (1998), Work-force ageing in OECD countries, in *Work-force ageing: Consequences and policy responses*, Working paper AWP 4.1.
- Persson, Torsten and Guido Tabellini, 2000, *Political economics: Explaining economic policy*. Zeuthen Lecture Book Series. Cambridge and London: MIT Press.
- Psacharopoulos, G. and R. Layard, (1979), "Human Capital and Earnings: British Evidence and a Critique". *Review of Economic Studies*, 46(3), 485-503.
- Rodriguez, Francisco, (1999), "Does Distributional Skewness Lead to Redistribution? Evidence from the United States". *Economics & Politics*, 11(2), 171-199.
- Roberts, Kevin, (1977), "Voting over Income Tax Schedules." *Journal of Public Economics*, 8, 329-340.
- Rodrik, Dani, (1998), "Why Do More Open Economies Have Bigger Governments?" *Journal of Political Economy*, 106, 997-1032.
- Razin, Assaf, Efraim Sadka and Phillip Swagel, (2002), "Tax burden and migration: a political economy theory and evidence", *Journal of Public Economics*, 85(2), 167-190.
- Roman, Arjona, Maxime Ladaique and Mark Pearson, (2001), Growth, Inequality and Social Protection, OECD.

Thompson, P., (2003), "Technological Change and the Age-Earnings Profile: Evidence from the International Merchant Marine, 1861-1912". *Review of Economic Dynamics*, 6(3), 578-601.

Thornton, Robert J., James D. Rodgers and Michael L. Brookshire, (1997), "On the interpretation of age-earnings profiles". *Journal of Labor Research*, 18(2), 351-366.

WIDER World Income Inequality Database, <http://www.wider.unu.edu/wiid/wiid.htm>.

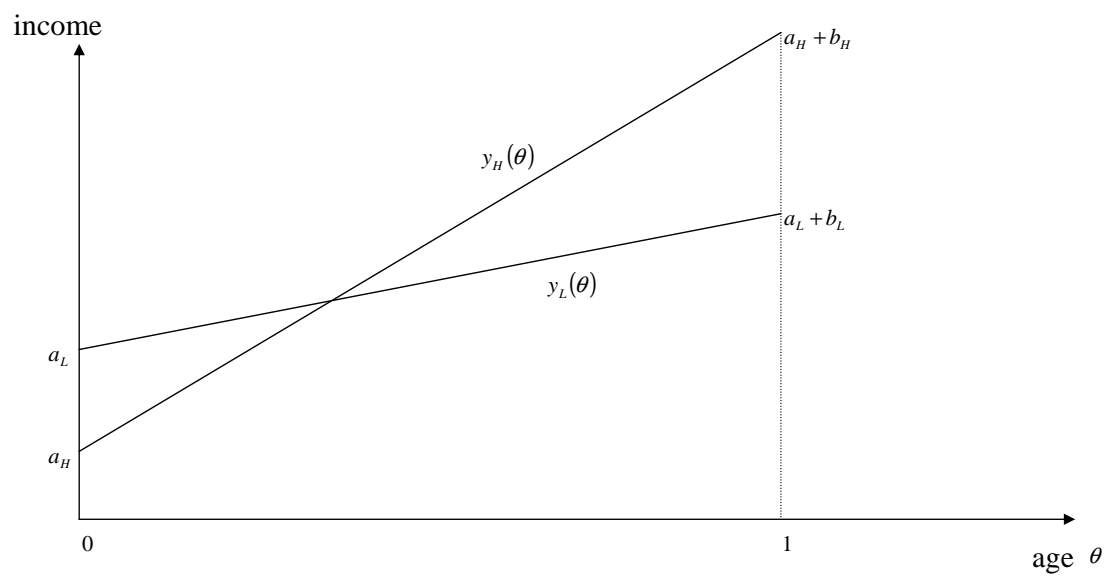


Figure 1a: Age earnings profiles

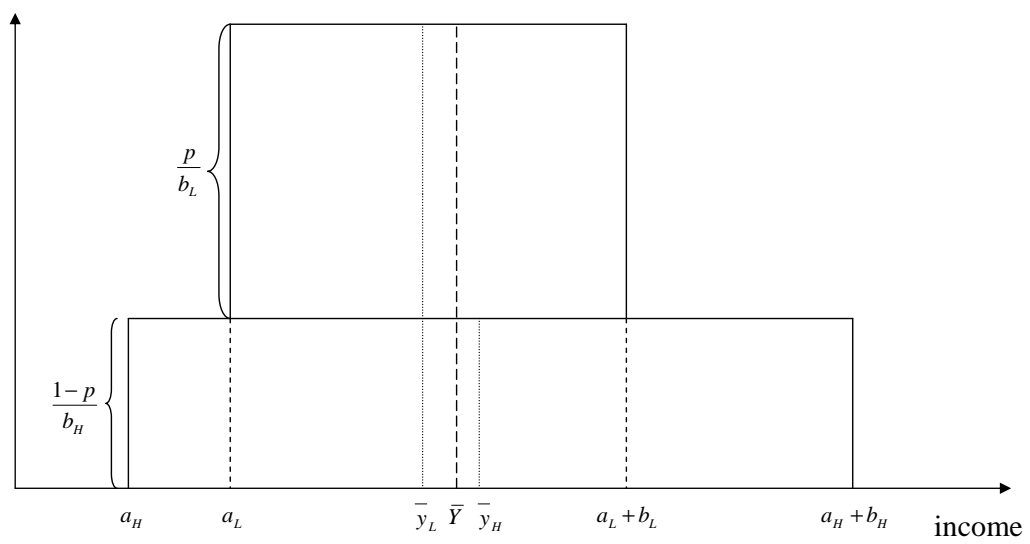


Figure 1b: The income distribution

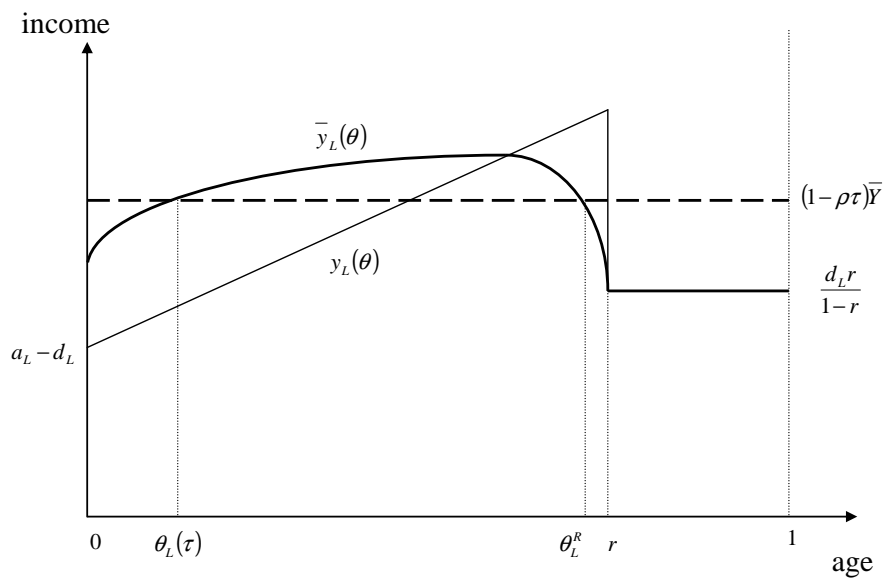


Figure 2: Threshold ages, current income and period income for L-agents retiring at r .

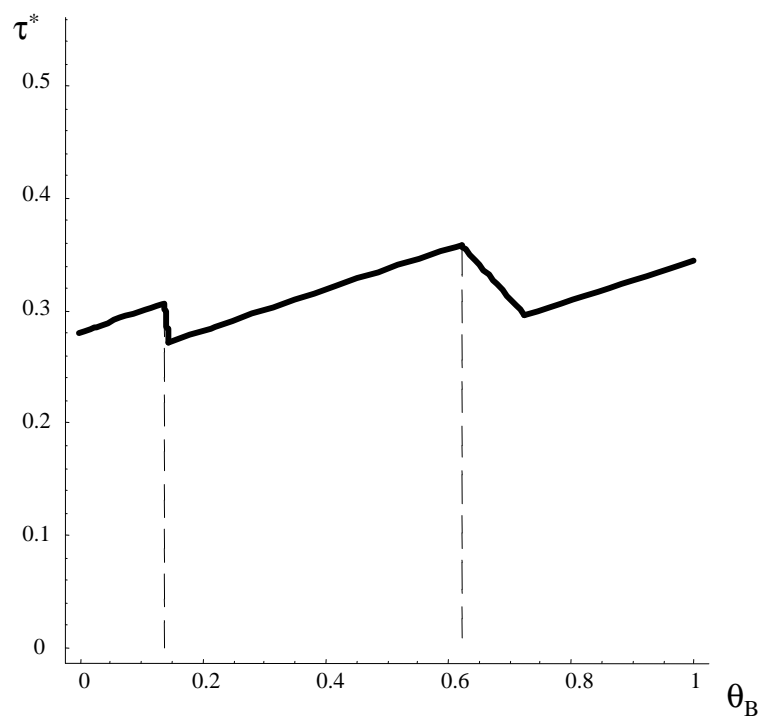


Figure 3: The tax rate as a function of the age of the baby-boom