UNIVERSITY *of York*

Discussion Papers in Economics

No. 19/02

Generational Bias and Tax Policy

Alan Krause

Department of Economics and Related Studies
University of York
Heslington
York, YO10 5DD

Generational Bias and Tax Policy

Alan Krause*
University of York

28 January 2019

Abstract

We examine the nonlinear taxation of labour income and savings when the government places more weight on the welfare of the elderly than of young people. Our analysis is motivated by the observation that the elderly are more likely to vote. Compared to optimal taxation under a utilitarian social welfare function, we show that savings are subsidised, and young low-skill workers face a higher marginal labour tax rate. We also show that the lifetime utility of low-skill individuals is reduced, and that of high-skill individuals is increased, relative to optimal taxation under utilitarianism. An extension of the model to include generation-specific public spending is also considered.

Keywords: generational policy; nonlinear taxation.

JEL Classifications: H21, H42.

---

*Department of Economics and Related Studies, University of York, Heslington, York, YO10 5DD, U.K. E-mail: alan.krause@york.ac.uk.

1 Introduction

This paper is motivated by the observation that old people are more likely to vote than young people. It then follows that governments, or political parties seeking election, have an incentive to promise and implement policies that lean towards the preferences of the elderly. While much of the literature has examined the government's expenditure policies,[1] in this paper we mainly focus on tax policy. Specifically, how the nonlinear taxation of labour income and savings is affected by a government that places more weight on the welfare of the elderly than of the young. We do, however, consider an extension of the model to include generation-specific public spending.

To this end, we begin by examining nonlinear taxation without public spending in an infinite-horizon overlapping generations (OLG) model. In this case, the tax system is purely redistributive. The OLG setting is well suited to our task, because in each period (except the first) there is a population of old people and young people. Each person works when young, and they may be a high-skill or low-skill worker. When old, individuals are retired and must live off the savings they made when young. The usual assumption in the literature is that the government sets tax policy to maximise a utilitarian social welfare function. However, we assume instead that the government maximises an objective function which is a weighted average of the utilities of the old and young generations. Our main finding is that any departure, however slight, from utilitarianism leads to non-zero savings taxation.[2] In particular, if the objective function is weighted towards the elderly, savings will be subsidised with low-skill individuals receiving a larger subsidy. The qualitative features of the labour income tax schedule are not affected by generational bias. However, the positive marginal labour tax rate faced by low-skill individuals is now higher than under utilitarianism. We also show

[1] In particular, a large literature, beginning with Samuelson (1975), has considered social security and pensions policies. The focus has mainly been on optimal policy, rather than reflecting a generational bias. However, the persistence of pay-as-you-go pensions policies, which by their nature redistribute from the young to the old, may reflect the political power of the elderly. Textbook discussions of the issues involved can be found in Blanchard and Fischer (1989), Myles (1995), and Kaplow (2008), among many others.

[2] The often-cited papers by Chamley (1986) and Judd (1985) show that the optimal savings/capital tax rate is zero. Both Chamley and Judd consider linear taxation, but it follows from Atkinson and Stiglitz (1976) that the same result holds with nonlinear taxation (under standard assumptions).

that low-skill individuals are worse-off over their lifetimes, and high-skill individuals are better-off, relative to utilitarianism. Therefore, generational bias not only redistributes from the young to the old, it also indirectly benefits high-skill individuals at the expense of low-skill individuals. All of these results continue to hold when the model is extended to include generation-specific public spending, on education for the young and on general public spending (e.g., health care) for the old. As intuition would suggest, placing more weight on the welfare of the elderly shifts the composition of public spending away from education for the young and towards general public spending for the old.

The possibility that government policy favours the elderly due to their proclivity to vote has received considerable attention in public discourse. However, to the best of our knowledge, it has not been subjected to a formal analysis. We fill this gap in the literature by examining generational bias in an OLG model with nonlinear (Mirrlees 1971) taxation.[3] There are, however, two strands of literature that examine related issues. First, there is a literature that examines how generational conflict affects linear labour and capital taxation, as well as the government's spending priorities, in a political or voting equilibrium. See, e.g., Razin, et al. (2002), Holtz-Eakin, et al. (2004), and Bassetto (2008). However, as all people (young and old) are assumed to participate in the political process, the resulting equilibrium reflects the will of the majority, rather than a generational bias. Indeed, the median or decisive voter is typically a young person. A second strand of literature examines the generational incidence of specific tax reforms, such as substituting consumption taxes for income taxes. See, e.g., Kotlikoff (2002), Krause (2007), and Conesa and Garriga (2016). Both strands of literature are positive in nature, in that they seek to explain observed policy or the generational incidence of actual policy reforms. By contrast, we take a different approach. We first establish a normative benchmark, by deriving optimal policy under a utilitarian social welfare function. We then examine the effects of deviating from that normative benchmark due

[3] The first paper to examine nonlinear taxation within an OLG setting is Ordover and Phelps (1979). Their focus is on optimal taxation under Bergson-Samuelson and maxi-min social welfare functions, which reflect each generation's lifetime (young and old) utility. Subsequently, a number of papers have examined nonlinear taxation within OLG frameworks, though the specific structure of our model is most closely related to Brett (2012).

to generational bias.

Our paper is also related to the literature on voting over nonlinear income tax schedules (e.g., Roell 2012; Brett and Weymark 2017), which derives the tax schedules preferred by individuals with different skills. The tax schedule chosen via voting, which by the median voter theorem will be that preferred by middle-skill individuals, is then contrasted with the optimal tax schedule. Our focus is similar, in that we examine (albeit indirectly) the nonlinear tax schedules preferred by young versus old people. However, we do not model voting behaviour, as we take it as given that the government will implement a policy which favours the elderly, reflecting the persistent evidence that they are more likely to vote (see, e.g., Gardiner 2016). Current models of voting behaviour, based on the median voter theorem, do not take voter participation into account. The lack of participation by young people — and the resulting generational bias — is the primary motivation for our paper.

The remainder of the paper is organised as follows. Section 2 presents the model and assumptions. Section 3 derives optimal nonlinear utilitarian taxation, as well as nonlinear taxation under a generational bias. Section 4 extends the model to include generation-specific public spending. Section 5 concludes, while a number of mathematical details are contained in an appendix.

2 Model and Assumptions

We consider an infinite-horizon OLG model, in which each individual lives for two periods. In the absence of taxation and public spending, an individual i born in period t solves the following problem. Choose c_i^t, l_i^t, s_i^t, and x_i^{t+1} to maximise their lifetime utility:

$$u(c_i^t) - v(l_i^t) + \delta u(x_i^{t+1}) \tag{2.1}$$

subject to their period t and period $t+1$ budget constraints:

$$c_i^t + s_i^t \leq w^t a_i l_i^t \tag{2.2}$$

$$x_i^{t+1} \leq (1 + r^{t+1}) s_i^t \tag{2.3}$$

where $u(\cdot)$ is increasing and strictly concave, $v(\cdot)$ is increasing and strictly convex, and $\delta \in (0, 1)$ is the discount factor. Consumption and labour supply when young, in period t, are denoted by c_i^t and l_i^t, respectively. Consumption when old, in period $t+1$, is denoted by x_i^{t+1}. Individual i's savings, which are chosen when young, is denoted by s_i^t, and r^{t+1} denotes the rate of return on savings in period $t+1$. Finally, a_i represents individual i's ability level, making $a_i l_i^t$ individual i's effective labour supply, and w^t represents the price (wage rate) of effective labour in period t. We assume that there are two types of individual, type 1 and type 2, such that $a_2 > a_1 > 0$. Therefore, type 1 individuals are low-skill workers and type 2 individuals are high-skill workers. For future reference, we use $y_i^t = w^t a_i l_i^t$ to denote type i's pre-tax labour income.

It is shown in the Appendix that the solution to programme (2.1) – (2.3) yields the marginal conditions:

$$1 = \frac{v'(l_i^t)}{u'(c_i^t) w^t a_i} \tag{2.4}$$

$$1 = \frac{u'(c_i^t)}{\delta(1 + r^{t+1}) u'(x_i^{t+1})} \tag{2.5}$$

Equation (2.4) represents the trade-off between consumption and labour when young, while equation (2.5) represents the trade-off between consuming when young and old.

The production side of the economy consists of a single profit-maximising firm, who produces an aggregate commodity each period according to the production function:

$$Y^t = F(K^t, Z^t) \tag{2.6}$$

where $F(\cdot)$ exhibits constant returns to scale, Y^t is aggregate production in period t, K^t is the economy-wide capital stock in period t, and Z^t is the total supply of effective labour in period t. Total effective labour is the sum of total effective low-skill and high-skill labour: $Z^t = Z_1^t + Z_2^t$ where $Z_1^t = N_1^t a_1 l_1^t$ and $Z_2^t = N_2^t a_2 l_2^t$, with N_i^t denoting the population of type i workers in period t. Therefore, low-skill and high-skill labour are perfect substitutes for one another. For simplicity we assume that $N_1^t = N_2^t$, and that the populations of both types grow at the exogenous rate of n per period.

As the production function exhibits constant returns to scale and has two inputs, it can be rewritten as:

$$Y^t = Z^t f(k^t) \tag{2.7}$$

where $k^t = K^t/Z^t$ is the capital-labour ratio, and $f(\cdot)$ is increasing and strictly concave. Profit maximisation then implies:

$$\frac{\partial Y^t}{\partial Z^t} = f(k^t) - f'(k^t)k^t = w^t \tag{2.8}$$

$$\frac{\partial Y^t}{\partial K^t} = f'(k^t) = r^t \tag{2.9}$$

That is, labour and capital are paid their marginal products.

Equilibrium in each period can be represented by the national accounting identity:

$$F(K^t, Z^t) = C^t + I^t \tag{2.10}$$

where C^t and I^t represent aggregate consumption and investment in period t, respectively. It is shown in the Appendix that equation (2.10) can be rewritten as:

$$f(k^t) = \frac{1}{a_1 l_1^t + a_2 l_2^t}\left[c_1^t + c_2^t + \frac{x_1^t + x_2^t}{1+n}\right] + \frac{(1+n)k^{t+1}(a_1 l_1^{t+1} + a_2 l_2^{t+1})}{a_1 l_1^t + a_2 l_2^t} - k^t \tag{2.11}$$

Following standard practice, we analyse the steady-state equilibrium in which all variables per capita are constant through time. It is shown in the Appendix that the steady-state version of equation (2.11) is:

$$f(k) = \frac{w}{y_1 + y_2}\left[c_1 + c_2 + \frac{x_1 + x_2}{1+n}\right] + nk \tag{2.12}$$

where the absence of time superscripts represents the steady-state (time invariant) value of that variable.

3 Tax Policy

In order to establish a benchmark, we first derive the key characteristics of the optimal nonlinear tax schedule under a utilitarian social welfare function. In this section there is no public spending, so taxation is purely redistributive. The information assumptions are the same as in standard Mirrlees-style models: the government cannot observe any individual's skill type (a_1 or a_2), and therefore cannot implement (first-best) personalised lump-sum taxation. Instead, the government implements second-best taxation, under which individuals are willing to reveal their types. All other variables, including age, are observable.[4] Thus, the young cannot mimic the old, nor vice versa.

We analyse taxation in steady-state. In this case, the government's choice of an optimal nonlinear tax schedule is equivalent to it choosing steady-state allocations for low-skill and high-skill individuals, subject to feasibility and incentive-compatibility constraints. That is, the government chooses c_1, y_1, x_1, c_2, y_2, x_2, and k to maximise:

$$u(c_1) - v\left(\frac{y_1}{wa_1}\right) + \delta u(x_1) + u(c_2) - v\left(\frac{y_2}{wa_2}\right) + \delta u(x_2) \tag{3.1}$$

subject to:

$$f(k) \geq \frac{w}{y_1 + y_2}\left[c_1 + c_2 + \frac{x_1 + x_2}{1+n}\right] + nk \tag{3.2}$$

$$u(c_2) - v\left(\frac{y_2}{wa_2}\right) + \delta u(x_2) \geq u(c_1) - v\left(\frac{y_1}{wa_2}\right) + \delta u(x_1) \tag{3.3}$$

where $w = f(k) - f'(k)k$ by profit maximisation (equation 2.8). The utilitarian social welfare function, equation (3.1), reflects the assumption that there are equal numbers of low-skill and high-skill individuals. Equation (3.2) is the steady-state equilibrium condition, which ensures that the chosen allocation is feasible.[5] Equation (3.3) is the high-skill type's incentive-compatibility constraint. Following common practice, we analyse a 're-

[4] Indeed, there is an interesting literature which shows that basing taxation on both age and income is more efficient than basing taxation on income alone. Efficiency gains are obtained if there is a correlation between (observable) age and (unobservable) ability. See, e.g., Blomquist and Micheletto (2008) and Weinzierl (2011).

[5] Satisfaction of equation (3.2) also implies that the government's budget constraint is satisfied in the steady-state equilibrium.

distributive equilibrium' in which the redistributive goals of the government create an incentive for high-skill individuals to mimic low-skill individuals, but not vice versa. Accordingly, the low-skill type's incentive-compatibility constraint will be slack, and is therefore omitted.

Differences in the marginal conditions from programme (3.1) $-$ (3.3) versus those that would hold in the absence of taxation (equations 2.4 and 2.5) can be interpreted as implicit marginal tax rates. Let T_i denote type i's implicit marginal labour tax rate, and let τ_i denote type i's implicit marginal savings tax rate.

It is shown in the Appendix that:

Benchmark Results *Optimal taxation under a utilitarian social welfare function yields*: $T_1 > 0$, $T_2 = 0$, *and* $\tau_i = 0$ *for both types.*

In short, under utilitarianism, we obtain the standard results that low-skill individuals face a positive marginal labour tax rate, high-skill individuals face a zero marginal labour tax rate, and both types face zero marginal savings tax rates (Atkinson and Stiglitz 1976).

Although the wage rate is endogenous in our model, we still obtain the standard results because low-skill and high-skill labour are perfectly substitutable, and relative wages cannot change. When low-skill and high-skill labour are not perfect substitutes and relative wages are endogenous, high-skill individuals face a negative marginal labour tax rate (Stiglitz 1982), and non-zero marginal savings taxation may be optimal (Pirttila and Tuomala 2001). It is also worth noting that we are implicitly assuming that the government can commit to its tax policy. Otherwise, after setting income and savings taxes in period t, the government would re-optimise in period $t + 1$. In that case, zero savings taxation is no longer optimal (Farhi, et al. 2012; Brett and Weymark 2019).

3.1 Generational Bias

We now examine the effects of a government that departs from utilitarianism, and places more (or less) weight on the welfare of the elderly. This can be readily captured by replacing equation (3.1) with the objective function:

$$(1-\alpha)\left[u(c_1) - v\left(\frac{y_1}{wa_1}\right) + u(c_2) - v\left(\frac{y_2}{wa_2}\right)\right] + \alpha\delta\left[u(x_1) + u(x_2)\right] \qquad (3.4)$$

where $\alpha \in (0, 1)$ is the weight placed on the utility of the elderly. Equation (3.4) can be viewed as reflecting the end result of a political process or voting behaviour, with the specific value of α reflecting the political power of the elderly. In each period of our OLG model, there is a population of young people, who are high-skill or low-skill workers, as well as a population of old people, who were high-skill or low-skill workers. When $\alpha > 0.5$, the government will implement a steady-state allocation that is relatively more favourable to the elderly than to the young, and vice versa. If $\alpha = 0.5$, the objective function (3.4) becomes a utilitarian social welfare function.

Although we do not model the voting process, equation (3.4) and the assumption that $\alpha > 0.5$ may be justified on the following basis. In each period there is a population of young people and old people. Young people who vote will consider their current welfare, as well as what they expect to obtain when old. On the other hand, old people who vote will only consider their current welfare, as their youth has passed. This again suggests that policy will lean towards the preferences of the elderly.

Other than replacing equation (3.1) with equation (3.4), the government's tax problem remains the same as above. That is, the government implements the feasible (equation 3.2) and incentive-compatible (equation 3.3) allocation which maximises the objective function (3.4).

It is shown in the Appendix that:

Proposition 1 *Nonlinear taxation with a generational bias yields*: (*i*) *if* $\alpha > 0.5$ *then* $T_1 > 0$, $T_2 = 0$, *and* $\tau_1 < \tau_2 < 0$, *and* (*ii*) *if* $\alpha < 0.5$ *then* $T_1 > 0$, $T_2 = 0$, *and* $\tau_1 > \tau_2 > 0$.

The qualitative features of the marginal labour tax rates are unaffected by generational bias, but non-zero marginal savings taxation becomes desirable. In particular, if the government places more weight on the welfare of the elderly ($\alpha > 0.5$), savings will be subsidised, with low-skill individuals receiving a larger subsidy. Savings are subsidised to shift consumption from the young to the old, with low-skill individuals facing a larger subsidy (i.e., larger distortion) to facilitate incentive compatibility. By a reverse argument, when $\alpha < 0.5$, savings are taxed to shift consumption from the old to the young, with low-skill individuals again facing a larger distortion ($\tau_1 > \tau_2$) to facilitate incentive compatibility.

3.2 Further Results

In this subsection, we examine the comparative statics of the tax schedule with respect to the degree of generational bias, using numerical methods.[6] To this end, the utility function (2.1) is assumed to take the form:

$$\ln(c_i) - \frac{(l_i)^{1+\frac{1}{\eta}}}{1+\frac{1}{\eta}} + \delta \ln(x_i) \tag{3.5}$$

where $\eta > 0$ is the labour supply elasticity. We set $\eta = 0.5$, which is consistent with empirical estimates of the labour supply elasticity (e.g., Chetty, et al. 2011). Also, utility is assumed to be logarithmic in consumption, which is consistent with empirical estimates that the coefficient of relative risk aversion is one (e.g., Chetty 2006). Following Kocherlakota (2010), we assume that each period is 20-years in length and that the annual discount rate is 2%, making $\delta = 0.67$.

The production function (2.6) is assumed to take the Cobb-Douglas form:

$$Y = K^{\frac{1}{3}} Z^{\frac{2}{3}} \tag{3.6}$$

with the exponents chosen to reflect the observation that capital's share of national income is approximately one-third in developed economies, while labour's share is approximately two-thirds.

Fang (2006) and Goldin and Katz (2007) find that the college wage premium is about 60%. Assuming that high-skill individuals attend college while low-skill individuals do not, and that the wage premium reflects skill differences, we normalise $a_1 = 1$ and set $a_2 = 1.6$. Finally, as population growth averages around 1% per annum, and each period of our model is 20-years in length, we set $n = 0.22$.

[6] While it would be better to derive the comparative statics results analytically, the literature has shown that it is necessary to assume quasi-linear utility. See, for example, Brett and Weymark (2011) and Simula (2010). Also, some comparative statics results cannot be signed even with the quasi-linear assumption. The advantage of the numerical approach is that quasi-linearity is not necessary, and a full set of comparative statics results can be obtained. The disadvantage is that the results may be dependent upon the numerical calibration, though this disadvantage can be minimised by considering only empirically-plausible calibrations.

The results are shown in Table 1, where U_i denotes the lifetime utility of type i individuals. Our numerical analysis starts with utilitarianism ($\alpha = 0.5$), and then places slightly more weight on the welfare of the elderly ($\alpha = 0.51$ and $\alpha = 0.52$). In particular, we obtain:

Proposition 2 *Under nonlinear taxation with a generational bias: (i) the marginal savings subsidies are increasing in α, and (ii) the low-skill type's marginal labour tax rate is increasing in α.*

Part (i) of Proposition 2 follows naturally from Proposition 1. In our numerical simulations, $\tau_1 = 0$ under utilitarianism and declines to $\tau_1 = -0.09973$ when $\alpha = 0.52$. Analogously, τ_2 declines from $\tau_2 = 0$ to $\tau_2 = -0.07156$. Part (ii) of Proposition 2 is more interesting: T_1 increases from $T_1 = 0.12261$ under utilitarianism to $T_1 = 0.12948$ when $\alpha = 0.52$. Low-skill individuals are required to work longer, which *ceteris paribus* corresponds to a lower marginal labour tax rate, but they also receive less consumption when young, which *ceteris paribus* corresponds to a higher marginal labour tax rate. On balance, the latter effect dominates, and the low-skill type's marginal labour tax rate increases when more weight is placed on the welfare of the elderly. It is worth noting that high-skill individuals are also required to work longer and receive less consumption when young, but they continue to face a zero marginal labour tax rate.

Proposition 3 *Under nonlinear taxation with a generational bias, low-skill individuals are worse-off and high-skill individuals are better-off over their lifetimes as α increases.*

In our numerical simulations, $U_1 = -1.31405$ under utilitarianism and falls to $U_1 = -1.32162$ when $\alpha = 0.52$. By contrast, U_2 increases from $U_2 = -1.00612$ to $U_2 = -1.00172$. Individuals have the same utility function, but are distinguished by their skills (a_1 and a_2). These skill differences are relevant when individuals are young, and when maximising equation (3.4) the skill differences entice the government to redistribute from high-skill to low-skill individuals. However, as α increases, the objective function (3.4) becomes concentrated on utility when individuals are old, which depends only on consumption. Accordingly, as α increases, the government implements an allocation that is less redistributive than under utilitarianism, thus making high-skill individuals better-off and low-skill individuals worse-off.

4 Extension: Introducing Public Spending

In this section, we introduce public spending into the model. Given our interest in generational bias, we consider generation-specific public spending, i.e., spending on education which benefits the young, and general public spending which benefits the old. Taxation now has two roles: (i) redistribution and (ii) raise revenue to finance public spending. However, the introduction of public spending does not affect the pattern of *marginal* tax rates, as derived in Section 3. Accordingly, in this section we focus on how generational bias affects the distribution of public spending.

With the introduction of public spending, the government chooses c_1, y_1, x_1, c_2, y_2, x_2, k, e_1, e_2, and b to maximise:

$$(1-\alpha)\left[u(c_1) - v\left(\frac{y_1}{wa_1(e_1)}\right) + u(c_2) - v\left(\frac{y_2}{wa_2(e_2)}\right)\right] + \alpha\delta\left[u(x_1) + h(b) + u(x_2) + h(b)\right] \tag{4.1}$$

subject to:

$$f(k) \geq \frac{w}{y_1 + y_2}\left[c_1 + c_2 + \frac{x_1 + x_2}{1+n}\right] + nk + e_1 + e_2 + b \tag{4.2}$$

$$u(c_2) - v\left(\frac{y_2}{wa_2(e_2)}\right) + \delta(u(x_2) + h(b)) \geq u(c_1) - v\left(\frac{y_1}{wa_2(e_2)}\right) + \delta(u(x_1) + h(b)) \tag{4.3}$$

where public spending on education can be directed towards low-skill workers, e_1, or high-skill workers, e_2, which increases their respective skills: $a_1(e_1)$ and $a_2(e_2)$ are both increasing and strictly concave functions. General public spending for the elderly is denoted by b, with utility from b represented by the function $h(b)$ which is increasing and strictly concave. To fix ideas, one can think of b as being public spending on health care for the elderly.[7] Equation (4.2) is the steady-state equilibrium condition, now extended to include public spending (see the Appendix), and equation (4.3) is the high-skill type's incentive-compatibility constraint.

It is shown in the Appendix that:

Proposition 4 *As α increases, there is relatively more public spending for the elderly and relatively less public spending on education for the young.*

[7]For example, medicare in the U.S. is generally only available to persons aged 65 or older.

As intuition would suggest, the distribution of public spending shifts in favour of the elderly as more weight is placed on their welfare. It is worth noting that low-skill education benefits low-skill individuals when young by making it easier for them to supply effective labour. This same benefit is obtained by high-skill individuals when young from high-skill education, but high-skill education has the additional benefit in that it relaxes the incentive-compatibility constraint.[8] However, as α increases, the tax schedule becomes less redistributive, as discussed above. Accordingly, there is also less need to provide high-skill education to relax the incentive-compatibility constraint.

4.1 Further Results

More insights can be obtained by examining the comparative statics with respect to the degree of generational bias, again using numerical methods. With public spending, the utility function (3.5) becomes:

$$\ln(c_i) - \frac{(l_i)^{1+\frac{1}{\eta}}}{1+\frac{1}{\eta}} + \delta\left[\ln(x_i) + \ln(b)\right] \tag{4.4}$$

where for simplicity we are assuming that $h(b) = \ln(b)$. We specify the skill functions as $a_1 = 1 + e_1^{\kappa}$ and $a_2 = 1.6(1 + e_2^{\kappa})$, where $\kappa \in (0, 1)$. These are chosen such that $a_1 = 1.0$ and $a_2 = 1.6$, as above, when $e_1 = 0$ and $e_2 = 0$. In our numerical simulations, we arbitrarily set $\kappa = 0.25$. The remaining parameters are the same as in Subsection 3.2.

The results are shown in Table 2; in particular, we obtain:

Proposition 5 *As α increases, the level of public spending for the elderly increases, while the level of public spending on education for the young decreases.*

While the analytical result, Proposition 4, is concerned with relativities, the numerical result, Proposition 5, is concerned with levels. When α increases from $\alpha = 0.5$ to $\alpha = 0.52$, public spending for the elderly, b, increases from 0.22436 to 0.23215. On the other hand, low-skill education, e_1, falls from 0.01200 to 0.01190 and high-skill education, e_2, falls from 0.02615 to 0.02564. High-skill education falls by proportionately more because an increase in α means that taxation becomes less redistributive, so the motive

[8] This result has been established in the previous literature. See, for example, Krause (2006) for further discussion.

the government has to provide high-skill education to relax the incentive-compatibility constraint is reduced.

The numerical results in Table 2 also show that introducing public spending has no qualitative effect on how an increase in α affects the marginal tax rates. It can also be seen that the lifetime utility of low-skill individuals is still decreasing in α, while that for high-skill individuals is still increasing.

5 Concluding Comments

The usual assumption in the literature is that the optimal tax system is that which maximises a social welfare function. To the extent that the social welfare function represents society's preferences, the optimal tax system is the tax system that is most preferred by society. In this paper, we consider nonlinear taxation when the government places more weight on the preferences of the elderly. Such an assumption seems reasonable, given that governments are more likely to be responsive to the preferences of voters, rather than the population at large. Our main results are intuitive: a tax schedule which favours the elderly subsidises savings, in order to shift consumption from the young to the old, and young people are required to work longer. Moreover, low-skill individuals are worse-off and high-skill individuals are better-off than under utilitarianism, because the tax schedule is less redistributive. All of these results continue to hold when the model is extended to include generation-specific public spending.

Our analysis is motivated by the fact that old people are more likely to vote than young people, which has seemingly held in a number of countries and over many years. Indeed, a recent study of voting behaviour in the UK (Gardiner 2016) presents evidence that the generational voting gap is widening. Nevertheless, given that generational bias redistributes from the young to the old, and indirectly harms low-skill individuals, one would expect the young (in particular, young low-skill workers) to become motivated to vote. The persistent reluctance of young people to vote remains a puzzle.

6 Appendix

A.1 Derivation of Equations (2.4) and (2.5)

The first-order conditions corresponding to programme (2.1) – (2.3) are:

$$u'(c_i^t) - \lambda^t = 0 \tag{A.1}$$

$$-v'(l_i^t) + \lambda^t w^t a_i = 0 \tag{A.2}$$

$$-\lambda^t + \lambda^{t+1}(1 + r^{t+1}) = 0 \tag{A.3}$$

$$\delta u'(x_i^{t+1}) - \lambda^{t+1} = 0 \tag{A.4}$$

where $\lambda^t > 0$ and $\lambda^{t+1} > 0$ are the multipliers on constraints (2.2) and (2.3), respectively. Algebraic manipulation of (A.1) and (A.2) yields equation (2.4), while algebraic manipulation of (A.1), (A.3) and (A.4) yields equation (2.5).

A.2 Derivation of Equations (2.11) and (2.12)

The national accounting identity, equation (2.10), can be written as:

$$Z^t f(k^t) = N_1^t c_1^t + N_2^t c_2^t + N_1^{t-1} x_1^t + N_2^{t-1} x_2^t + K^{t+1} - K^t \tag{A.5}$$

where $N_1^t c_1^t + N_2^t c_2^t$ is total consumption by young people in period t, $N_1^{t-1} x_1^t + N_2^{t-1} x_2^t$ is total consumption by old people in period t, and $K^{t+1} - K^t$ is investment in period t (i.e., change in the capital stock, assuming no capital depreciation). Dividing (A.5) by Z^t yields:

$$f(k^t) = \frac{N_1^t c_1^t}{Z^t} + \frac{N_2^t c_2^t}{Z^t} + \frac{N_1^{t-1} x_1^t}{Z^t} + \frac{N_2^{t-1} x_2^t}{Z^t} + \frac{Z^{t+1}}{Z^t} k^{t+1} - k^t \tag{A.6}$$

Recalling that $Z^t = N_1^t a_1 l_1^t + N_2^t a_2 l_2^t$ and the assumption that $N_1^t = N_2^t$, equation (A.6) reduces to:

$$f(k^t) = \frac{1}{a_1 l_1^t + a_2 l_2^t}\left[c_1^t + c_2^t + \frac{x_1^t + x_2^t}{1+n}\right] + \frac{(1+n)k^{t+1}(a_1 l_1^{t+1} + a_2 l_2^{t+1})}{a_1 l_1^t + a_2 l_2^t} - k^t \tag{A.7}$$

which is equation (2.11).

In steady-state, all variables per capita are time invariant, and $y_i = w a_i l_i$. Thus, the steady-state version of (A.7) is:

$$f(k) = \frac{w}{y_1 + y_2}\left[c_1 + c_2 + \frac{x_1 + x_2}{1 + n}\right] + nk \tag{A.8}$$

which is equation (2.12).

A.3 Derivation of the Benchmark Results

The first-order conditions corresponding to programme (3.1) – (3.3) are:

$$(1 - \theta_2)u'(c_1) - \frac{\lambda w}{y_1 + y_2} = 0 \tag{A.9}$$

$$-v'\left(\frac{y_1}{wa_1}\right)\frac{1}{wa_1} + \theta_2 v'\left(\frac{y_1}{wa_2}\right)\frac{1}{wa_2} + \frac{\lambda w}{(y_1 + y_2)^2}\left[c_1 + c_2 + \frac{x_1 + x_2}{1 + n}\right] = 0 \tag{A.10}$$

$$(1 - \theta_2)\delta u'(x_1) - \frac{\lambda w}{(1 + n)(y_1 + y_2)} = 0 \tag{A.11}$$

$$(1 + \theta_2)u'(c_2) - \frac{\lambda w}{y_1 + y_2} = 0 \tag{A.12}$$

$$-(1 + \theta_2)v'\left(\frac{y_2}{wa_2}\right)\frac{1}{wa_2} + \frac{\lambda w}{(y_1 + y_2)^2}\left[c_1 + c_2 + \frac{x_1 + x_2}{1 + n}\right] = 0 \tag{A.13}$$

$$(1 + \theta_2)\delta u'(x_2) - \frac{\lambda w}{(1 + n)(y_1 + y_2)} = 0 \tag{A.14}$$

$$\begin{aligned}\lambda\left(f'(k) - n\right) + \frac{\partial w}{\partial k}\left[v'\left(\frac{y_1}{wa_1}\right)\frac{y_1}{w^2 a_1} - \theta_2 v'\left(\frac{y_1}{wa_2}\right)\frac{y_1}{w^2 a_2} + (1 + \theta_2)v'\left(\frac{y_2}{wa_2}\right)\frac{y_2}{w^2 a_2}\right] \\ - \frac{\partial w}{\partial k}\frac{\lambda}{(y_1 + y_2)}\left(c_1 + c_2 + \frac{x_1 + x_2}{1 + n}\right) = 0\end{aligned} \tag{A.15}$$

where $\lambda > 0$ and $\theta_2 > 0$ are the multipliers on constraints (3.2) and (3.3), respectively.

We first show that $f'(k) = n$, which implies by profit maximisation that $r = n$. Equations (A.10) and (A.13) imply that:

$$v'\left(\frac{y_1}{wa_1}\right)\frac{y_1}{w^2 a_1} - \theta_2 v'\left(\frac{y_1}{wa_2}\right)\frac{y_1}{w^2 a_2} = \frac{\lambda y_1}{(y_1 + y_2)^2}\left[c_1 + c_2 + \frac{x_1 + x_2}{1 + n}\right] \tag{A.16}$$

$$(1 + \theta_2)v'\left(\frac{y_2}{wa_2}\right)\frac{y_2}{w^2 a_2} = \frac{\lambda y_2}{(y_1 + y_2)^2}\left[c_1 + c_2 + \frac{x_1 + x_2}{1 + n}\right] \tag{A.17}$$

which together imply that:

$$v'\left(\frac{y_1}{wa_1}\right)\frac{y_1}{w^2a_1}-\theta_2 v'\left(\frac{y_1}{wa_2}\right)\frac{y_1}{w^2a_2}+(1+\theta_2)v'\left(\frac{y_2}{wa_2}\right)\frac{y_2}{w^2a_2}=\frac{\lambda}{(y_1+y_2)}\left(c_1+c_2+\frac{x_1+x_2}{1+n}\right) \tag{A.18}$$

Therefore, (A.15) reduces to $\lambda\left(f'(k)-n\right)=0$, which establishes that $f'(k)=n$.

From (A.9) and (A.11) we obtain:

$$(1-\theta_2)\delta u'(x_1)(1+n)=(1-\theta_2)u'(c_1) \tag{A.19}$$

Since $r=n$, and (A.9) implies that $1-\theta_2>0$, equation (A.19) can be manipulated to yield:

$$1=\frac{u'(c_1)}{\delta(1+r)u'(x_1)} \tag{A.20}$$

which using equation (2.5) establishes that $\tau_1=0$.

Likewise, from (A.12) and (A.14) we obtain:

$$(1+\theta_2)\delta u'(x_2)(1+n)=(1+\theta_2)u'(c_2) \tag{A.21}$$

which yields:

$$1=\frac{u'(c_2)}{\delta(1+r)u'(x_2)} \tag{A.22}$$

and using (2.5) establishes that $\tau_2=0$.

From (A.9) and (A.10) we obtain:

$$(1-\theta_2)u'(c_1)=(1-\theta_2)v'\left(\frac{y_1}{wa_1}\right)\frac{1}{wa_1}+\theta_2\left[v'\left(\frac{y_1}{wa_1}\right)\frac{1}{wa_1}-v'\left(\frac{y_1}{wa_2}\right)\frac{1}{wa_2}\right]$$
$$+\frac{\lambda w}{(y_1+y_2)}\left[1-\frac{c_1+c_2}{y_1+y_2}-\frac{x_1+x_2}{(1+n)(y_1+y_2)}\right] \tag{A.23}$$

From equation (3.2) we obtain:

$$\left[f(k)-nk\right](y_1+y_2)=w\left[c_1+c_2+\frac{x_1+x_2}{1+n}\right] \tag{A.24}$$

Total output is equal to labour income and capital income, that is:

$$Zf(k) = wZ + rK \qquad \Longrightarrow \qquad f(k) = w + rk \tag{A.25}$$

Since $r = n$, equations (A.24) and (A.25) imply that the last term in equation (A.23) equals zero. Thus, (A.23) can be manipulated to yield:

$$1 - \frac{v'(l_1)}{u'(c_1)wa_1} = \frac{\theta_2}{(1-\theta_2)u'(c_1)}\left[v'\left(\frac{y_1}{wa_1}\right)\frac{1}{wa_1} - v'\left(\frac{y_1}{wa_2}\right)\frac{1}{wa_2}\right] > 0 \tag{A.26}$$

which using (2.4) establishes that $T_1 > 0$.

From (A.12) and (A.13) we obtain:

$$(1+\theta_2)u'(c_2) = (1+\theta_2)v'\left(\frac{y_2}{wa_2}\right)\frac{1}{wa_2} + \frac{\lambda w}{(y_1+y_2)}\left[1 - \frac{c_1+c_2}{y_1+y_2} - \frac{x_1+x_2}{(1+n)(y_1+y_2)}\right] \tag{A.27}$$

where feasibility again implies that the last term in equation (A.27) equals zero. Thus, (A.27) can be manipulated to yield:

$$1 - \frac{v'(l_2)}{u'(c_2)wa_2} = 0 \tag{A.28}$$

which using (2.4) establishes that $T_2 = 0$. ■

A.4 Proof of Proposition 1

Maximisation of equation (3.4) subject to constraints (3.2) and (3.3) yields the first-order conditions:

$$(1-\alpha-\theta_2)u'(c_1) - \frac{\lambda w}{y_1+y_2} = 0 \tag{A.29}$$

$$-(1-\alpha)v'\left(\frac{y_1}{wa_1}\right)\frac{1}{wa_1} + \theta_2 v'\left(\frac{y_1}{wa_2}\right)\frac{1}{wa_2} + \frac{\lambda w}{(y_1+y_2)^2}\left[c_1+c_2+\frac{x_1+x_2}{1+n}\right] = 0 \tag{A.30}$$

$$(\alpha-\theta_2)\delta u'(x_1) - \frac{\lambda w}{(1+n)(y_1+y_2)} = 0 \tag{A.31}$$

$$(1-\alpha+\theta_2)u'(c_2) - \frac{\lambda w}{y_1+y_2} = 0 \tag{A.32}$$

$$-(1-\alpha+\theta_2)v'\left(\frac{y_2}{wa_2}\right)\frac{1}{wa_2} + \frac{\lambda w}{(y_1+y_2)^2}\left[c_1+c_2+\frac{x_1+x_2}{1+n}\right] = 0 \tag{A.33}$$

$$(\alpha+\theta_2)\delta u'(x_2)-\frac{\lambda w}{(1+n)(y_1+y_2)}=0 \tag{A.34}$$

$$\lambda\left(f'(k)-n\right)+\frac{\partial w}{\partial k}\left[(1-\alpha)v'\left(\frac{y_1}{wa_1}\right)\frac{y_1}{w^2a_1}-\theta_2 v'\left(\frac{y_1}{wa_2}\right)\frac{y_1}{w^2a_2}+(1-\alpha+\theta_2)v'\left(\frac{y_2}{wa_2}\right)\frac{y_2}{w^2a_2}\right]$$
$$-\frac{\partial w}{\partial k}\frac{\lambda}{(y_1+y_2)}\left(c_1+c_2+\frac{x_1+x_2}{1+n}\right)=0 \tag{A.35}$$

As above, using (A.30), (A.33) and (A.35) we obtain $f'(k)=n$, which implies that $r=n$ by profit maximisation.

From (A.29) and (A.31) we obtain:

$$(\alpha-\theta_2)\delta u'(x_1)(1+n)=(\alpha-\theta_2)u'(c_1)+(1-2\alpha)u'(c_1) \tag{A.36}$$

Since $r=n$, and (A.31) implies that $\alpha-\theta_2>0$, equation (A.36) can be manipulated to yield:

$$1-\frac{u'(c_1)}{\delta(1+r)u'(x_1)}=\frac{(1-2\alpha)u'(c_1)}{(\alpha-\theta_2)\delta(1+r)u'(x_1)} \tag{A.37}$$

which using equation (2.5) establishes that $\tau_1 \gtreqless 0$ if and only if $\alpha \lesseqgtr 0.5$.

Likewise, using (A.32) and (A.34) we obtain:

$$(\alpha+\theta_2)\delta u'(x_2)(1+n)=(\alpha+\theta_2)u'(c_2)+(1-2\alpha)u'(c_2) \tag{A.38}$$

which can be manipulated to yield:

$$1-\frac{u'(c_2)}{\delta(1+r)u'(x_2)}=\frac{(1-2\alpha)u'(c_2)}{(\alpha+\theta_2)\delta(1+r)u'(x_2)} \tag{A.39}$$

which using equation (2.5) establishes that $\tau_2 \gtreqless 0$ if and only if $\alpha \lesseqgtr 0.5$.

From (A.37) and (A.39) we obtain:

$$\tau_2-\tau_1=(1-2\alpha)\left[\frac{u'(c_2)}{(\alpha+\theta_2)\delta(1+r)u'(x_2)}-\frac{u'(c_1)}{(\alpha-\theta_2)\delta(1+r)u'(x_1)}\right] \tag{A.40}$$

which using (A.29), (A.31), (A.32) and (A.34) becomes:

$$\tau_2 - \tau_1 = (1-2\alpha)\left[\frac{1}{1-\alpha+\theta_2} - \frac{1}{1-\alpha-\theta_2}\right] \tag{A.41}$$

Therefore, when $\alpha > 0.5$ we have $\tau_1 < \tau_2$, and when $\alpha < 0.5$ we have $\tau_1 > \tau_2$.

From (A.29) and (A.30) we obtain:

$$(1-\alpha-\theta_2)u'(c_1) = (1-\alpha-\theta_2)v'\left(\frac{y_1}{wa_1}\right)\frac{1}{wa_1} + \theta_2\left[v'\left(\frac{y_1}{wa_1}\right)\frac{1}{wa_1} - v'\left(\frac{y_1}{wa_2}\right)\frac{1}{wa_2}\right]$$

$$+\frac{\lambda w}{(y_1+y_2)}\left[1 - \frac{c_1+c_2}{y_1+y_2} - \frac{x_1+x_2}{(1+n)(y_1+y_2)}\right] \tag{A.42}$$

where (A.29) implies that $1-\alpha-\theta_2 > 0$, and feasibility (as above) implies that the last term in equation (A.42) equals zero. Thus, (A.42) can be manipulated to yield:

$$1 - \frac{v'(l_1)}{u'(c_1)wa_1} = \frac{\theta_2}{(1-\alpha-\theta_2)u'(c_1)}\left[v'\left(\frac{y_1}{wa_1}\right)\frac{1}{wa_1} - v'\left(\frac{y_1}{wa_2}\right)\frac{1}{wa_2}\right] > 0 \tag{A.43}$$

which using (2.4) establishes that $T_1 > 0$ for all $\alpha \in (0,1)$.

From (A.32) and (A.33) we obtain:

$$(1-\alpha+\theta_2)u'(c_2) = (1-\alpha+\theta_2)v'\left(\frac{y_2}{wa_2}\right)\frac{1}{wa_2} + \frac{\lambda w}{(y_1+y_2)}\left[1 - \frac{c_1+c_2}{y_1+y_2} - \frac{x_1+x_2}{(1+n)(y_1+y_2)}\right] \tag{A.44}$$

where feasibility again implies that the last term in equation (A.44) equals zero. Thus, (A.44) can be manipulated to yield:

$$1 - \frac{v'(l_2)}{u'(c_2)wa_2} = 0 \tag{A.45}$$

which using (2.4) establishes that $T_2 = 0$ for all $\alpha \in (0,1)$. ■

A.5 Derivation of Equation (4.2)

With public spending, the national accounting identity becomes:

$$F(K^t, Z^t) = C^t + I^t + G^t \tag{A.46}$$

where G^t is total public spending in period t. Equation (A.5) becomes:

$$Z^t f(k^t) = N_1^t c_1^t + N_2^t c_2^t + N_1^{t-1} x_1^t + N_2^{t-1} x_2^t + K^{t+1} - K^t + E_1^t + E_2^t + B^t \qquad \text{(A.47)}$$

where $G^t = E_1^t + E_2^t + B^t$, with E_1^t and E_2^t representing public spending in period t on low-skill and high-skill education, respectively, and B^t representing public spending in period t for the elderly.

Dividing (A.47) by total effective labour Z^t yields:

$$f(k^t) = \frac{N_1^t c_1^t}{Z^t} + \frac{N_2^t c_2^t}{Z^t} + \frac{N_1^{t-1} x_1^t}{Z^t} + \frac{N_2^{t-1} x_2^t}{Z^t} + \frac{Z^{t+1}}{Z^t} k^{t+1} - k^t + e_1^t + e_2^t + b^t \qquad \text{(A.48)}$$

with the lower-case e_1^t, e_2^t, and b^t denoting spending per unit of effective labour. Equation (A.48) can be reduced to:

$$f(k^t) = \frac{1}{a_1 l_1^t + a_2 l_2^t} \left[c_1^t + c_2^t + \frac{x_1^t + x_2^t}{1+n} \right] + \frac{(1+n) k^{t+1} (a_1 l_1^{t+1} + a_2 l_2^{t+1})}{a_1 l_1^t + a_2 l_2^t} - k^t + e_1^t + e_2^t + b^t \qquad \text{(A.49)}$$

The steady-state version of (A.49) is:

$$f(k) = \frac{w}{y_1 + y_2} \left[c_1 + c_2 + \frac{x_1 + x_2}{1+n} \right] + nk + e_1 + e_2 + b \qquad \text{(A.50)}$$

which is equation (4.2).

A.6 Proof of Proposition 4

The first-order conditions on e_1, e_2, and b in programme (4.1) – (4.3) can be simplified to, respectively:

$$(1-\alpha) v' \left(\frac{y_1}{w a_1} \right) \frac{a_1'(e_1) y_1}{w a_1^2} - \lambda = 0 \qquad \text{(A.51)}$$

$$(1-\alpha) v' \left(\frac{y_2}{w a_2} \right) \frac{a_2'(e_2) y_2}{w a_2^2} + \frac{\theta_2 a_2'(e_2)}{w a_2^2} \left[v' \left(\frac{y_2}{w a_2} \right) y_2 - v' \left(\frac{y_1}{w a_2} \right) y_1 \right] - \lambda = 0 \qquad \text{(A.52)}$$

$$\alpha \delta 2 h'(b) - \lambda = 0 \qquad \text{(A.53)}$$

where $\lambda > 0$ and $\theta_2 > 0$ are the multipliers on constraints (4.2) and (4.3), respectively.

Manipulation of (A.51) and (A.53) yields:

$$\frac{v'\left(\frac{y_1}{wa_1}\right)\frac{a_1'(e_1)y_1}{wa_1^2}}{\delta 2h'(b)} = \frac{\alpha}{1-\alpha} \tag{A.54}$$

The numerator on the left-hand side of (A.54) represents the marginal benefit of e_1. Under our assumptions regarding the curvature of the functions, this marginal benefit is decreasing in e_1. Likewise, the denominator represents the marginal benefit of b, which is decreasing in b. An increase in α implies that the left-hand side of (A.54) increases, which means that e_1 must decline relative to b.

Manipulation of (A.52) and (A.53) yields:

$$\begin{aligned}(1-\alpha)&\left\{v'\left(\frac{y_2}{wa_2}\right)\frac{a_2'(e_2)y_2}{wa_2^2} + \frac{\theta_2 a_2'(e_2)}{wa_2^2}\left[v'\left(\frac{y_2}{wa_2}\right)y_2 - v'\left(\frac{y_1}{wa_2}\right)y_1\right]\right\} \\ &= \alpha\left\{\delta 2h'(b) - \frac{\theta_2 a_2'(e_2)}{wa_2^2}\left[v'\left(\frac{y_2}{wa_2}\right)y_2 - v'\left(\frac{y_1}{wa_2}\right)y_1\right]\right\}\end{aligned} \tag{A.55}$$

Incentive compatibility implies that the steady-state allocation is monotonic, in particular $y_2 > y_1$. Thus, the left-hand side of (A.55) is positive, implying that the right-hand side must also be positive. From (A.55) we obtain:

$$\frac{v'\left(\frac{y_2}{wa_2}\right)\frac{a_2'(e_2)y_2}{wa_2^2} + \frac{\theta_2 a_2'(e_2)}{wa_2^2}\left[v'\left(\frac{y_2}{wa_2}\right)y_2 - v'\left(\frac{y_1}{wa_2}\right)y_1\right]}{\delta 2h'(b) - \frac{\theta_2 a_2'(e_2)}{wa_2^2}\left[v'\left(\frac{y_2}{wa_2}\right)y_2 - v'\left(\frac{y_1}{wa_2}\right)y_1\right]} = \frac{\alpha}{1-\alpha} \tag{A.56}$$

The numerator on the left-hand side of (A.56) represents the marginal benefit of e_2, which includes the benefit of relaxing the incentive-compatibility constraint. Again, under our assumptions, this marginal benefit is decreasing in e_2. The denominator represents the marginal benefit of b, net of the effect of e_2 on the incentive-compatibility constraint. An increase in α implies that the left-hand side of (A.56) must increase, which means that e_2 declines relative to b. ■

References

[1] Atkinson, A. and J. Stiglitz (1976), "The Design of Tax Structure: Direct Versus Indirect Taxation", *Journal of Public Economics*, 6, 55-75.

[2] Bassetto, M. (2008), "Political Economy of Taxation in an Overlapping-Generations Economy", *Review of Economic Dynamics*, 11, 18-43.

[3] Blanchard, O. and S. Fischer (1989), *Lectures on Macroeconomics*, MIT Press.

[4] Blomquist, S. and L. Micheletto (2008), "Age-Related Optimal Income Taxation", *Scandinavian Journal of Economics,* 110, 45-71.

[5] Brett, C. (2012), "The Effects of Population Aging on Optimal Redistributive Taxes in an Overlapping Generations Model", *International Tax and Public Finance*, 19, 777-799.

[6] Brett, C. and J. Weymark (2011), "How Optimal Nonlinear Income Taxes Change when the Distribution of the Population Changes", *Journal of Public Economics*, 95, 1239-1247.

[7] Brett, C. and J. Weymark (2017), "Voting over Selfishly Optimal Nonlinear Income Tax Schedules", *Games and Economic Behavior*, 101, 172-188.

[8] Brett, C. and J. Weymark (2019), "Optimal Nonlinear Taxation of Income and Savings without Commitment", *Journal of Public Economic Theory*, 21, 5-43.

[9] Chamley, C. (1986), "Optimal Taxation of Capital Income in General Equilibrium with Infinite lives", *Econometrica*, 54, 607-622.

[10] Chetty, R. (2006), "A New Method of Estimating Risk Aversion", *American Economic Review*, 96, 1821-1834.

[11] Chetty, R., A. Guren, D. Manoli and A. Weber (2011), "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins", *American Economic Review*, 101, 471-475.

[12] Conesa, J. and C. Garriga (2016), "Intergenerational Policy and the Measurement of Tax Incidence", *European Economic Review*, 83, 1-18.

[13] Fang, H. (2006), "Disentangling the College Wage Premium: Estimating a Model with Endogenous Education Choices", *International Economic Review*, 47, 1151-1185.

[14] Farhi, E., C. Sleet, I. Werning and S. Yeltekin (2012), "Nonlinear Capital Taxation Without Commitment", *Review of Economic Studies*, 79, 1469-1493.

[15] Gardiner, L. (2016), "Votey McVoteface: Understanding the Growing Turnout Gap between the Generations", *Resolution Foundation Discussion Paper.*

[16] Goldin, C. and L. Katz (2007), "The Race between Education and Technology: The Evolution of U.S. Educational Wage Differentials, 1890-2005", *NBER Working Paper*, No. 12984.

[17] Holtz-Eakin, D., M. Lovely and M. Tosun (2004), "Generational Conflict, Fiscal Policy, and Economic Growth", *Journal of Macroeconomics*, 26, 1-23.

[18] Judd, K. (1985), "Redistributive Taxation in a Simple Perfect Foresight Model", *Journal of Public Economics*, 28, 59-83.

[19] Kaplow, L. (2008), *The Theory of Taxation and Public Economics*, Princeton University Press.

[20] Kocherlakota, N. (2010), *The New Dynamic Public Finance*, Princeton University Press.

[21] Kotlikoff, L. (2002), "Generational Policy", *Handbook of Public Economics*, 4, 1873-1932.

[22] Krause, A. (2006), "Redistributive Taxation and Public Education", *Journal of Public Economic Theory*, 8, 807-819.

[23] Krause, A. (2007), "Generational Incidence of Savings Taxation versus Capital-Income Taxation", *International Journal of Economic Theory*, 3, 113-129.

[24] Mirrlees, J. (1971), "An Exploration in the Theory of Optimum Income Taxation", *Review of Economic Studies*, 38, 175-208.

[25] Myles, G. (1995), *Public Economics*, Cambridge University Press.

[26] Ordover, J. and E. Phelps (1979), "The Concept of Optimal Taxation in the Overlapping-Generations Model of Capital and Wealth", *Journal of Public Economics*, 12, 1-26.

[27] Pirttila, J. and M. Tuomala (2001), "On Optimal Non-Linear Taxation and Public Good Provision in an Overlapping Generations Economy", *Journal of Public Economics*, 79, 485-501.

[28] Razin, A., E. Sadka and P. Swagel (2002), "The Aging Population and the Size of the Welfare State", *Journal of Political Economy*, 110, 900-918.

[29] Roell, A. (2012), "Voting over Nonlinear Income Tax Schedules", *mimeo*.

[30] Samuelson, P. (1975), "Optimum Social Security in a Life-Cycle Growth Model", *International Economic Review*, 16, 539-544.

[31] Simula, L. (2010), "Optimal Nonlinear Income Tax and Nonlinear Pricing: Optimality Conditions and Comparative Static Properties", *Social Choice and Welfare*, 35, 199-220.

[32] Stiglitz, J. (1982), "Self-Selection and Pareto Efficient Taxation", *Journal of Public Economics*, 17, 213-240.

[33] Weinzierl, M. (2011), "The Surprising Power of Age-Dependent Taxes", *Review of Economic Studies*, 78, 1490-1518.

TABLE 1

Numerical Analysis Without Public Spending

	Parameter values		
	$\eta = 0.50$	$a_1 = 1.00$	$n = 0.22$
	$\delta = 0.67$	$a_2 = 1.60$	

	$\alpha = 0.50$	$\alpha = 0.51$	$\alpha = 0.52$
τ_1	0.00000	−0.04860	−0.09973
τ_2	0.00000	−0.03518	−0.07156
T_1	0.12261	0.12600	0.12948
T_2	0.00000	0.00000	0.00000
U_1	−1.31405	−1.31737	−1.32162
U_2	−1.00612	−1.00359	−1.00172
c_1	0.62959	0.61933	0.60897
l_1	1.06916	1.07590	1.08284
x_1	0.51699	0.53328	0.54993
c_2	0.86238	0.85557	0.84864
l_2	1.23363	1.23853	1.24358
x_2	0.70814	0.72727	0.74673

TABLE 2

Numerical Analysis With Public Spending

	Parameter values		
	$\eta = 0.50$	$a_1 = 1 + e_1^{\kappa}$	$n = 0.22$
	$\delta = 0.67$	$a_2 = 1.6(1 + e_2^{\kappa})$	$\kappa = 0.25$
	$\alpha = 0.50$	$\alpha = 0.51$	$\alpha = 0.52$
e_1	0.01200	0.01195	0.01190
e_2	0.02615	0.02589	0.02564
b	0.22436	0.22826	0.23215
τ_1	0.00000	–0.04913	–0.10085
τ_2	0.00000	–0.03491	–0.07100
T_1	0.13513	0.13877	0.14249
T_2	0.00000	0.00000	0.00000
U_1	–2.42937	–2.43282	–2.43746
U_2	–2.12395	–2.12157	–2.12012
c_1	0.58220	0.56851	0.55490
l_1	1.05014	1.05690	1.06388
x_1	0.47808	0.48977	0.50161
c_2	0.81213	0.80014	0.78814
l_2	1.24124	1.24604	1.25098
x_2	0.66689	0.67997	0.69314