Democratic public good provision

Hassler, John ; Storesletten, Kjetil ; Zilibotti, Fabrizio

Abstract: This paper analyzes an overlapping generation model of redistribution and public good provision under repeated voting. Expenditures are financed through age-dependent taxation that distorts human capital investment. Taxes redistribute income both across skill groups and across generations. We focus on politico-economic Markov equilibria and contrast these with the Ramsey allocation under commitment. The model features indeterminate equilibria, with a key role of forward-looking strategic voting. Due to the lack of commitment to future policies, the tax burden may be on the wrong side of the dynamic Laffer curve. Moreover, restrictions on government policies can in some cases be welfare improving.

DOI: https://doi.org/10.1016/j.jet.2005.07.013

Originally published at:
Hassler, John; Storesletten, Kjetil; Zilibotti, Fabrizio (2007). Democratic public good provision. Journal of Economic Theory, 133(1):127-151.
DOI: https://doi.org/10.1016/j.jet.2005.07.013

Democratic Public Good Provision

John Hassler[*], **Kjetil Storesletten**[†], **and Fabrizio Zilibotti**[‡]

February 2003

Abstract

This paper analyzes an overlapping generation model of public good provision under repeated voting. The public good is financed through age-dependent taxation that distorts human capital investment. Taxes redistribute income both across different skill groups and across generations. We contrast the political equilibria with the Ramsey allocation, and analyze the sources of inefficiency. The political equilibria can feature both under- and over-provision of public good, as well an inefficient life-cycle profile of taxes.

JEL codes: D72, D78, E62, H21, H41, H53.

Key words: Markov equilibrium, multiple equilibria, public good, political economy, Ramsey allocation, redistribution, repeated voting, taxation.

[*] Institute for International Economic Studies and CEPR; e-mail: john.hassler@iies.su.se

[†] Institute for International Economic Studies and CEPR; e-mail: kjetil.storesletten@iies.su.se

[‡] University College London, IFS and CEPR; e-mail: : F.Zilibotti@ucl.ac.uk.

1 Introduction

This paper analyzes the dynamics of public good provision when this is determined in repeated elections. We consider an environment where – like in many real world applications – public goods benefit all agents equally, but the burden of financing is unevenly distributed. Hence, there are conflicts of interests about provision of public goods. In particular, we compare the political provision of the public good with the Ramsey second-best allocation, focusing on *(i)* the amount of public good provision and *(ii)* the distribution of the tax burden across age groups.

The distinctive feature of our theory is the presence of non-trivial dynamic interactions between economics and politics. In particular, *(i)* policies are decided through period-by-period voting and *(ii)* they affect investment decisions. These two elements make political decisions substantially richer than in static models, since the current political choice affects future income distribution and future voting outcomes, via private investment decisions. Thus, while some policies tend to regenerate their constituency over time, others carry the seed of their destruction. This, in turn, opens a scope for strategic voting, i.e., agents may use their current political choice to influence the future constituency of the policy.

The model economy is populated by overlapping generations of agents endowed with quasi-linear preferences over a private good, a pure public good and effort. Each period, agents vote on two separate income tax rates, for old and young agents, respectively, and the proceeds are used to finance the current provision of the public good. When young, agents make a human capital investment increasing their expected earnings. The initial investment and the subsequent realization of luck make agents more heterogenous as they age and hence, "lock-in" their preferences over taxation. Thus, taxation distorts human capital investment and has persistent effects on the distributions of earnings which, in turn, determines the political demand for taxation. It is the ability to use current policy to manipulate future distributions and future policies that introduces a motive for strategic voting.[1]

We start by characterizing constrained-efficient allocations. In particular, we solve for the Ramsey allocations whereby a utilitarian planner sets an infinite tax sequence, with full commitment, under the constraint of running a balanced budget. The planner's public good provision is affected both by its intrinsic value and by her desire to redistribute income between generations. The latter objective depends on the extent to which the planner

[1] Note that our use of the term "strategic" voting parallels that of the "strategic debt"-literature (see Persson and Tabellini, 2000) where governments use public debt in order to constrain the policies of future governments.

discounts future generations' utilities: if her discount rate is higher (lower) than that of private agents, she will let the young (old) carry the brunt of the tax burden.

Next, we move to the positive analysis, and characterize the set of Markov political equilibria under majority voting. We show that there exist multiple equilibria – one "sincere" equilibrium with high taxes and a range of "strategic" equilibria with low taxes. In the former, agents (optimally) do not use current policy to strategically manipulate future policies. In the latter, voters (again, optimally) restrain taxation on the young in order to induce a future majority that will not tax the old, thereby strengthening the incentives for the young to invest and thus enlarge the current tax base for the public good.

Generally, equilibria do not lie in the (second-best) Pareto frontier. A first reason for this is that in political equilibria, taxation is only driven by the desire to redistribute, so that the equilibrium allocations do not depend on the value of the public good, which is the case in the Ramsey allocation. Thus, when the value of the public good is sufficiently low, the political equilibria over-provide public goods, relative to any Ramsey-allocation. Conversely, if this value is sufficiently high, political equilibria under-provide public goods. A second, more subtle, reason is that the political system does not allow agents to commit to future taxation, which leads to an inefficient distribution of the burden of taxation between the young and the old.

In a related paper, Hassler, Rodríguez-Mora, Storesletten and Zilibotti (2003), henceforth "HRSZ", show that the process of "democracy over time" can lead to the perpetual survival of an *inefficient* transfer system, i.e., a system that no agent, at birth, would like to have in place, but that a majority imposes over the minority, once the uncertainty over lifetime earnings has been revealed. The current paper differs from HRSZ in several important respects.

First, the only government policy considered in HRSZ is transfers between groups. In the data, transfers constitute only a limited share of public redistribution. The lion share of government expenditures is allocated to activities with important public good components, such as education, defence, infrastructure, etc. Since agents are risk-neutral, in *ex-ante* terms, government policy is entirely wasteful in HRSZ, which precludes interesting normative analysis. The focal point of the current paper is precisely the normative aspects of political equilibria.

Second, in HRSZ, the fiscal tools for inter-generational redistribution are limited since polices are restricted to be age-independent. In reality, many government programs such as health care, schooling, etc., are targeted at specific age groups, and taxation also has important age-dependent elements. Allowing for age-dependent taxes has major implications for the dynamics. In particular, in HRSZ, the initial distribution of agents has permanent political effects; for instance, a welfare state can never arise if there is an initial majority of net contributors to the system. In contrast, in the current paper, the identity of the initial median voter only determines the initial taxation of the old, which has no effect on the future distribution of voters. As a result, the analytical characterization is simpler than in HRSZ, which is of independent methodological interest.

Third, in HRSZ, all agents are born identical, and the conflict of interests only arises *ex-post*, i.e., after the realization of stochastic returns to human capital investment. This implies, unrealistically, that all young agents expect to lose from redistribution. In this paper, however, we introduce *ex-ante* heterogeneity in ability, which allows us to address policy-relevant issues such as the fate of poor agents in different equilibria and under different political set-ups.

Finally, this paper generalizes the political mechanism in HRSZ by considering *(i)* probabilistic voting together with majority voting, and *(ii)* a different timing of the voting game, such that the young exert political influence.

We study a framework where the political demand for redistribution is driving the equilibrium provision of public goods. The argument that redistribution can be a motive for public good provision dates back to Wilson (1991) and Mirrlees (1994). They show that if agents are heterogenous and lump-sum taxes are available, then it can be optimal to over-provide public goods, relative to a Samuelson-Lindahl benchmark, in order to achieve redistribution.

There are examples of papers examining dynamic aspects of politically-driven public-good provision, such as the strategic use of debt-literature (Persson and Svensson, 1989, and Alesina and Tabellini, 1990). There, politicians have preferences over the type of public good provided or the amount of public good provision, which differ from the preferences of future policy makers, and they use debt and current provision of public goods strategically in order to manipulate the future provision of public goods.

The paper is organized as follows. Section 2 presents the model and Section 3 characterizes the Ramsey-optimal allocations. Section 4 presents the political equilibrium concept, and Section 5 analyzes political equilibria under majority voting. In Section 6, we extend the voting set-up to allow young voters to have larger influence on political outcomes and we also consider a qualitatively different voting mechanism, probabilistic voting. Section 7 concludes. Proofs are provided in the Appendix.

2 The model

The model economy consists of a continuum of two-period lived agents. Each generation has a unit mass. Agents are, at birth, of two types: high-ability and low-ability, in proportion μ and $1-\mu$, respectively. High-ability agents can affect their prospects in life by an initial (educational) effort choice. In particular, they either become rich or poor, and the costly investment increases their probability of becoming rich, denoted by e. The cost of investment is quadratic and equal to e^2. Rich agents earn a high wage, normalized to unity, in both periods, whereas poor agents earn a low wage, normalized to zero. Low-ability agents make no investment choice; irrespective of their private actions, they are deemed to poverty.

There is a government whose role it is to provide a public good, g, financed by income taxes levied on the rich. The tax rates are allowed to be age-dependent, τ^O for the old and τ^Y for the young, but restricted to lie between zero and 100%. Since only the rich pay taxes, public good provision is redistributive. The tax rates are determined before the young agents decide on their investment and the government budget balances in every period.

The expected utility of agents alive at time t is given as follows:

$$\begin{aligned}
\tilde{V}^{os}\left(g_t,\tau_t^O\right) &= 1-\tau_t^O+ag_t \\
\tilde{V}^{ou}\left(g_t,\tau_t^O\right) &= \tilde{V}^{ol}\left(g_t,\tau_t^O\right)=ag_t \\
\tilde{V}^{y}\left(e_t,g_t,g_{t+1},\tau_t^Y,\tau_{t+1}^O\right) &= e_t\left(1-\tau_t^Y+\beta\left(1-\tau_{t+1}^O\right)\right)+a\left(g_t+\beta g_{t+1}\right)-e_t^2 \\
\tilde{V}^{yl}\left(g_t,g_{t+1}\right) &= a\left(g_t+\beta g_{t+1}\right),
\end{aligned}$$

where $\tilde{V}^{os}$, $\tilde{V}^{ou}$and $\tilde{V}^{y}$ denote the utility of old rich, old poor and high-ability young, respectively, whereas $\tilde{V}^{ol}$ and $\tilde{V}^{yl}$ denote the utility of the old and young low-ability agents. Note that $\tilde{V}^{y}$ is computed prior to individual success or failure. $\beta\in[0,1]$ is the discount factor and a is the marginal utility of the public good. The assumption that this marginal utility is constant is made for tractability. It is straightforward to show that the solution to the optimal investment problem of the young, given τ_t^Y and τ_{t+1}^O, is

$$e^*\left(\tau_t^Y,\tau_{t+1}^O\right)=\left(1+\beta-\left(\tau_t^Y+\beta\tau_{t+1}^O\right)\right)/2.$$

Since high-ability agents are *ex-ante* identical, agents of the same cohort choose the same investment, implying that the proportion of old poor in period $t+1$ is given by

$$u_{t+1}=1-e^*\left(\tau_t^Y,\tau_{t+1}^O\right)=\left(1-\beta+\tau_t^Y+\beta\tau_{t+1}^O\right)/2. \tag{1}$$

Thus, the proportion of old poor at $t+1$ depends on the tax rate on the young at t, and the tax rate of the old at $t+1$. The price of the public good is normalized to one, implying that the balanced government budget can be expressed as

$$g_t = \mu\left(1-u_t\right)\tau_t^O + \mu e^*\left(\tau_t^Y, \tau_{t+1}^O\right)\tau_t^Y. \tag{2}$$

3 Efficient allocations

In this section, we characterize the Ramsey allocation. This is defined as the sequences of taxes, $\left\{\tau_t^Y, \tau_t^O\right\}_{t=0}^{\infty}$, chosen by a benevolent social planner who solves

$$\max_{\left\{\tau_t^Y, \tau_t^O\right\}_{t=0}^{\infty}} \left\{\beta\mu\left(1-u_0\right)\tilde{V}^{os}\left(g_0, \tau_0^O\right) + \beta\left(\mu u_0 + 1 - \mu\right)\tilde{V}^{ou}\left(g_0, \tau_0^O\right) + \right. \tag{3}$$
$$\left. \sum_{t=0}^{\infty}\lambda^t\left(\tilde{V}^y\left(e_t, g_t, g_{t+1}, \tau_t^Y, \tau_{t+1}^O\right) + \left(1-\mu\right)\tilde{V}^{yl}\left(g_t, g_{t+1}\right)\right)\right\},$$

subject to the implementability constraints, (1) and (2), and $\tau_t^O, \tau_t^Y \in [0,1] \ \ \forall t$. Note that the planner discounts the future generations' utilities with a "discount factor" $\lambda \in [0,1)$. Furthermore, the utility of the initial generation of old agents is weighted by the rate β. This implies that a planner who sets $\lambda = \beta$ attaches the same weight to the felicity of the young born at t and that of the old born at $t-1$. If $\lambda < \beta$ ($\lambda > \beta$), the planner attaches a smaller (larger) weight to the former than to the latter.

The objective function of the planner, (3), can be rewritten as follows;

$$\tilde{L}\left(u_0\right) = \max_{\left\{\tau_t^Y \in [0,1], \tau_t^O \in [0,1]\right\}_{t=0}^{\infty}} \left\{\beta\mu\left(1-u_0\right)\left(1-\tau_0^O\right) + \left(\beta+\lambda\right)a\mu\left(1-u_0\right)\tau_0^O \right. \tag{4}$$
$$+\sum_{t=0}^{\infty}\lambda^t\left(\lambda\mu\left(e^*\left(\tau_t^Y, \tau_{t+1}^O\right)\left(1-\tau_t^Y + \beta\left(1-\tau_{t+1}^O\right)\right) - \left(e^*\left(\tau_t^Y, \tau_{t+1}^O\right)\right)^2\right)\right.$$
$$\left.+\left(\beta+\lambda\right)a\mu e^*\left(\tau_t^Y, \tau_{t+1}^O\right)\tau_t^Y + \lambda\left(\beta+\lambda\right)a\mu e^*\left(\tau_t^Y, \tau_{t+1}^O\right)\tau_{t+1}^O\right),$$

where e^* is given by (1), and the terms $\mu e^*\left(\tau_t^Y, \tau_{t+1}^O\right)\tau_t^Y$ and $\mu e^*\left(\tau_t^Y, \tau_{t+1}^O\right)\tau_t^O$ denote the amount of public good financed by the young and by the old at t, respectively. Note that the objective function is proportional to the parameter μ, which implies that the Ramsey taxes will be invariant to this parameter. The reason is that public good provision g_t is proportional to μ, since only the rich are taxed (see equation (2)).

The formulation in (4) emphasizes that the Ramsey problem, while non-recursive in nature, admits a recursive representation whereby the planner commits to a tax sequence τ_t^Y and τ_{t+1}^O in each period t. In addition, the planner solves a different problem in the first period, when he can choose τ_0^O after the old have already made their effort choice, i.e., given u_0. It is convenient to rewrite the program as follows:

$$\tilde{L}(u_0) = \max_{\{\tau_0^O \in [0,1]\}} \left\{ \beta\mu (1-u_0)\left(1-\tau_0^O\right) + (\beta+\lambda) a\mu (1-u_0)\tau_0^O + \frac{L}{1-\lambda} \right\}, \qquad (5)$$

where

$$\begin{aligned} L = & \max_{\{\tau^Y \in [0,1], \tau^O \in [0,1]\}} \left\{ \lambda\mu \left(e^*\left(\tau^Y,\tau^O\right)\left(1-\tau^Y+\beta\left(1-\tau^O\right)\right) - \left(e^*\left(\tau^Y,\tau^O\right)\right)^2 \right) + \right. \\ & \left. (\beta+\lambda) a\mu e^*\left(\tau^Y,\tau^O\right)\tau^Y + \lambda(\beta+\lambda) a\mu e^*\left(\tau^Y,\tau^O\right)\tau^O \right\}. \qquad (6) \end{aligned}$$

In plain words, after the initial choice of τ_0^O, the problem reduces to a sequence of identical static optimizations over τ^Y and τ^O, without any state variable.[2]

The solution to the program (5) is straightforward, since τ_0^O has no distortionary effects and there is no interaction with future variables. Each dollar of taxes levied on a proportion $\mu(1-u_0)$ of old agents provides all living agents, both young and old, with $a\mu(1-u_0)$ utils from public good consumption. The planner sets 100% taxation if $a > \min\{\beta,\lambda\}/(\beta+\lambda)$, and zero taxation for a smaller than this amount.

Next, consider the continuation program (6). As mentioned above, λ parameterizes the inter-generational bias of the planner. If $\lambda \to 1$, she maximizes the average expected utility of young agents, i.e., $\mu V^y + (1-\mu)V^{yl}$. This is a counterpart to the "golden rule" in our model. If instead $\lambda \to 0$, the planner fully discounts future generations and maximizes the average *ex-post* utility of old agents. That is, she maximizes $e^*\left(\tau^Y,\tau^O\right)\tau^Y$; the amount of public good that the old can extract by taxing the young. The knife-edge case $\lambda=\beta$ corresponds to a perfectly utilitarian planner with the same regard for the *ex-post* utility of old agents as for the *ex-ante* utility of future generations.

The next Proposition characterizes the Ramsey allocation.

Proposition 1

If $a \leq \min\{\beta,\lambda\}/(\beta+\lambda)$, then the Ramsey allocation, defined by the recursive formulation (5)-(6), has $\tau_0^O=\tau^Y=\tau^O=0$ and no public good provision. Otherwise, the Ramsey allocation has $\tau_0^O=1$, a constant sequence of taxes, τ^Y and τ^O, and a present value of public good provision, PVg_R (net of the tax revenues from the initial old $(1-u_0)\tau_0^O$), given by the following:

[2] More specifically, for any $t \geq 0$, we have $\tau_t^Y = \tau^Y$ and $\tau_{t+1}^O = \tau^O$.

1. *If* $\beta > \lambda$, *then,*

$$\tau^Y = (1+\beta)\frac{a-\frac{\lambda}{\beta+\lambda}}{2a-\frac{\lambda}{\beta+\lambda}},$$
$$\tau^O = 0,$$
$$PVg_R = \frac{(1+\beta)^2}{2}\frac{\mu}{1-\lambda}\cdot f\left(a,\frac{\lambda}{\beta+\lambda}\right),$$

where $f(a,x) \equiv a(a-x)(2a-x)^{-2} \geq 0.$

2. *If* $\beta < \lambda$, *then*

   (a) *If* $a \in \left[\frac{\beta}{(\beta+\lambda)}, \frac{\beta}{(1-\beta)(\beta+\lambda)}\right]$, *then*

$$\tau^Y = 0,$$
$$\tau^O = \frac{1+\beta}{\beta}\frac{a-\frac{\beta}{\beta+\lambda}}{2a-\frac{\beta}{\beta+\lambda}},$$
$$PVg_R = \frac{(1+\beta)^2}{2}\frac{\lambda}{\beta}\frac{\mu}{1-\lambda}\cdot f\left(a,\frac{\beta}{\beta+\lambda}\right).$$

   (b) *If* $a \in \left[\frac{\beta}{(1-\beta)(\beta+\lambda)}, \frac{\lambda}{(1-\lambda)(\beta+\lambda)}\right]$, *then*

$$\tau^Y = 0,$$
$$\tau^O = 1,$$
$$PVg_R = \frac{\lambda}{2}\frac{\mu}{1-\lambda}.$$

   (c) *If* $a > \frac{\lambda}{(1-\lambda)(\beta+\lambda)}$, *then*

$$\tau^Y = \frac{a(1-\lambda)-\frac{\lambda}{\beta+\lambda}}{2a-\frac{\lambda}{\beta+\lambda}},$$
$$\tau^O = 1,$$
$$PVg_R = \frac{(1+\lambda)^2}{2}\frac{\mu}{1-\lambda}\cdot f\left(a,\frac{\lambda}{\beta+\lambda}\right).$$

3. *If* $\beta = \lambda$, *then*

$$\tau^Y + \beta\tau^O = \frac{(2a-1)(1+\beta)}{4a-1},$$
$$PVg_R = \frac{(1+\lambda)^2}{2}\frac{\mu}{1-\lambda}\cdot f\left(a,\frac{1}{2}\right).$$

< Figure 1 about here >

Figure 1 shows the tax rates in the Ramsey allocations as a function of λ, which illustrates all cases of Proposition 1. On the left-hand side of β, i.e. part *1* of the proposition, the tax burden falls on the young (in both panels). On the right-hand side of β (part *2*), most of the tax burden falls on the old. Part *2a* is represented by an intermediate range of λ in the left panel (small a), part *2b* is represented by the flat segments for T^O on the right-hand side of β (in both panels), and part *2c* is represented by an intermediate range of λ in the right panel (large a). Finally, for $\lambda = \beta$ (part *3*), the tax rates are indeterminate.

To get some intuition for Proposition 1, consider the knife-edge utilitarian case, $\beta = \lambda$. In this case, the planner chooses $x \equiv \tau^Y + \beta\tau^O$ (i.e., the present value of taxation faced by each young agent) so as to maximize

$$\bar{L}(x) = \left(e^*(x)(1 + \beta - x) - (e^*(x))^2\right) + 2ae^*(x)\,x,$$

where the first term captures the cost of taxation while the second term captures its benefits. Using the envelope theorem, the optimal solution simplifies to $e^*(x) = a(1 + \beta - 2x)$, equalizing the marginal cost of taxation with the marginal benefit of public good provision. In this knife-edge case, the planner is indifferent as to how to split the tax burden between the young and the old (see also Figure 1). In contrast, the present value of public good provision is determinate. When $\beta \neq \lambda$, however, the split ceases to be indeterminate, since it has inter-generational distributional effects, about which the planner cares. In particular, shifting the tax burden towards the young (old) increases (decreases) the *ex-post* utility of the old agents at the expense (in favor of) the future generations.

Consider, in particular, the two polar opposite cases. If $\lambda \to 0$, then the planner wishes to maximizes the public good financed by the young, i.e., she maximizes $e^*\left(\tau^Y, \tau^O\right)\tau^Y$ (see equation (6)). This is achieved by setting $\tau^O = 0$ and choosing τ^Y so as to attain, conditional on such a choice, the top of the Laffer curve.[3] Conversely, if $\lambda \to 1$, the planner maximizes the *ex-ante* utility of the young, and prefers to tax the old. The reason is that individual agents discount the effects of future taxation when deciding on their investment, thereby implying that taxes on the old are less distortionary.

< Figure 2 about here >

Figure 2 displays the present value of public good provision PVg_R as a function of a in the three different cases of Proposition 1. The present value of public good provision in the Ramsey allocation has the intuitive property that for any β and λ, it is monotone

[3] Another way of illustrating the result in this limit case is the following. When $\lambda = 0$, the planner only cares about the utility of the old at time zero. Thus, she sets $\tau_1^O = 0$, since future taxation is of no value for the current old, but distorts the incentives of the young.

increasing in the marginal utility, a (this is shown in the proof of Proposition 1). Moreover, for sufficiently low a, public good provision is zero. Note that even though the present value of public good provision is monotone increasing in a, the steady-state provision of public good is not necessarily monotone.

An interesting question is what allocation would be chosen by a planner caring only about the low-ability agents in society (a "Rawlsian" planner). It turns out that this coincides with the Ramsey allocation in Proposition 1 when the public good becomes arbitrarily important, i.e. the case when $a \to \infty$. In this case, the planner seeks to maximize the public good provision or, equivalently, maximize the tax revenues. The solution to this program is characterized in the following corollary and stems from letting $a \to \infty$ in part *2c* of Proposition 1.

Corollary 1

The Rawlsian planner allocation has $\tau_0^O = 1$, and the following public good provision and constant sequence of taxes, τ^Y and τ^O:

1. *If $\beta > \lambda$, then $\tau^Y = (1+\beta)/2$, $\tau^O = 0$, and $PVg = \frac{(1+\beta)^2}{8}\frac{\mu}{1-\lambda}$;*
2. *If $\beta < \lambda$, then $\tau^Y = (1-\lambda)/2$, $\tau^O = 1$, and $PVg = \frac{(1+\lambda)^2}{8}\frac{\mu}{(1-\lambda)}$;*
3. *If $\beta = \lambda$, then $\tau^Y + \beta\tau^O = (1+\beta)/2$ and $PVg = \frac{(1+\lambda)^2}{8}\frac{\mu}{(1-\lambda)}$.*

The proof follows immediately from Proposition 1 and is omitted.

Both the Ramsey and the Rawlsian allocation assume that the planner can commit to future tax sequences. We conclude this section by characterizing the choice of a planner without a commitment technology. This *time-consistent planning allocation* is interesting, because the lack of commitment is an inherent feature of political equilibria, as discussed in section 4.

The following Proposition characterizes the time-consistent Ramsey allocation.

Proposition 2

If $a \le \min\{\beta, \lambda\}/(\beta+\lambda)$, then the time-consistent Ramsey allocation has $\tau_0^O = \tau^Y = \tau^O = 0$ and no public good provision. Otherwise, the time-consistent Ramsey allocation features $\tau_0^O = \tau^O = 1$ and

1. *If $a \leqq \frac{\lambda}{(1-\lambda)(\beta+\lambda)}$, then $\tau^Y = 0$ and $PVg_{TC} = \frac{\lambda}{2}\frac{\mu}{1-\lambda}$.*
2. *If $a > \frac{\lambda}{(1-\lambda)(\beta+\lambda)}$, then $\tau^Y = \frac{a(1-\lambda)-\frac{\lambda}{\beta+\lambda}}{2a-\frac{\lambda}{\beta+\lambda}} \in \left[0, \frac{1-\lambda}{2}\right]$ and $PVg_{TC} = \frac{(1+\lambda)^2}{2}\frac{\mu}{1-\lambda} \cdot f\left(a, \frac{\lambda}{\beta+\lambda}\right)$, where f is defined in Proposition 1.*

If $a \geq \beta / (\beta + \lambda)$ then, *ex-post*, the planner would like to set $\tau^O = 1$. Thus, the proof amounts to solving (6) for τ^Y, subject to $\tau^O = 1$. Time inconsistency arises whenever the Ramsey allocation prescribes less than full taxation of the old. Hence, this problem arises when she either has a pro-old bias (i.e., for small λ), or when the marginal utility of public good consumption is sufficiently low, i.e., $a < \beta / [(1 - \beta)(\beta + \lambda)]$. In the latter case, the planner would promise, *ex-ante*, a lower taxation of the old, but would have an incentive, after investment decisions are sunk, to increase the taxation to 100%.

Finally, the time-consistent Rawlsian planner allocation is given by the time-consistent planner allocation in Proposition 2 when $a \to \infty$, namely $\tau_0^O = \tau^O = 1$ and $\tau^Y = (1 - \lambda)/2$.

4 Political equilibrium

In this section, we characterize the equilibrium tax sequence when policies are set through repeated voting. Agents vote on taxation period-by-period. As a benchmark, we assume that elections are held at the end of each period, and that the politician who is elected sets the tax rates for the following period. The old have no interests at stake and are assumed to abstain, a timing which emphasizes the predominance of *ex-post* interests in determining political outcomes. If agents voted at the beginning of the period in our two-period framework, fifty percent of voters would not yet have decided on their human capital investment. In any generalization with more than two periods, however, it would be natural to assume that the majority of voters had already undertaken such investment. Therefore, our timing assumption is a natural way of avoiding this artifact of the two-period model, which is observationally equivalent to assuming that agents vote over current taxes at the beginning of each period, but that only the old vote. For expositional ease, we keep the latter interpretation in the presentation. In section 6, we explore the case when all agents vote over current taxation at the beginning of each period.

We focus on Markov perfect equilibria, where the state of the economy is summarized by the proportion of current poor old agents (u_t). A convenient feature of the model is that the public good provision, g_t, is additively separable in two components, i.e., $g_t = G^O(\tau_t^O, u_t) + G^Y(\tau_t^Y, \tau_{t+1}^O)$, where $G^O(\tau_t^O, u_t) = \mu(1 - u_t)\tau_t^O$ and $G^Y(\tau_t^Y, \tau_{t+1}^O) = \mu e^*(\tau_t^Y, \tau_{t+1}^O)\tau_t^Y$ represent the part of the public good financed by the old and the young, respectively. This implies that the preferences of all voters at t can be expressed as a term depending on τ_t^O and u_t (but no future variables) and a term depending on τ_t^Y and τ_{t+1}^O (but no variable which is predetermined at t). An implication of this separability is that the state of the economy, u_t, only affects the political choice of τ_t^O. Preferences over τ_t^Y and τ_{t+1}^O are, instead, entirely forward-looking and independent of the state variable.

More explicitly, the utility of the old rich and poor, respectively, can be rewritten as

$$\begin{aligned} V^{os}\left(u_t, \tau_t^O, \tau_t^Y, \tau_{t+1}^O\right) &= W^{os}\left(u_t, \tau_t^O\right) + Z^o\left(\tau_t^Y, \tau_{t+1}^O\right) \\ V^{ou}\left(u_t, \tau_t^O, \tau_t^Y, \tau_{t+1}^O\right) &= W^{ou}\left(u_t, \tau_t^O\right) + Z^o\left(\tau_t^Y, \tau_{t+1}^O\right), \end{aligned} \tag{7}$$

where

$$W^{ou}\left(u_t, \tau_t^O\right) = a\mu\left(1-u_t\right)\tau_t^O \tag{8}$$

$$W^{os}\left(u_t, \tau_t^O\right) = 1-\tau_t^O + W^{ou}\left(u_t, \tau_t^O\right) \tag{9}$$

$$Z^o\left(\tau_t^Y, \tau_{t+1}^O\right) = a\mu e^*\left(\tau_t^Y, \tau_{t+1}^O\right)\tau_t^Y. \tag{10}$$

For expositional simplicity, we mainly focus on the case $a < 1/\mu$ (although in Proposition 3 below, Markov equilibria are characterized for any value of a). This implies that whatever the level of u_t, the old rich would never tax themselves in order to finance the public good provision. Thus, there is an intra-generational political conflict among the old voters: the rich prefer zero taxes on the old, while the poor prefer 100% taxation. However, irrespective of a, the old agents' preferences over τ_t^Y are perfectly aligned: they wish to maximize $Z^o(\tau_t^Y, \tau_{t+1}^O)$, i.e., to attain the top of the Laffer curve.

This separability simplifies the definition of the equilibrium. We use the superscript *dec* to refer to the 'decisive' voter. The identity of such a voter depends on the political mechanism and the state of the economy and will be specified as we proceed.

Definition 1 *A (Markov perfect) political equilibrium is defined as a triplet of functions* $\langle T^O, T^Y, U\rangle$*, where* $T^O : [0,1] \to [0,1]$ *and* T^Y *are two public policy rules,* $\tau_t^O = T^O(u_t)$ *and* $\tau_t^Y = T^Y$*, and* $U : [0,1] \to [0,1]$ *is a private decision rule,* $u_{t+1} = U\left(\tau_t^Y\right)$*, such that the following functional equations hold:*

1. $T^O(u_t) = \arg\max_{\tau_t^O \in [0,1]} W^{dec}\left(\tau_t^O, u_t\right)$.
2. $T^Y = \arg\max_{\tau_t^Y \in [0,1]} Z^o\left(\tau_t^Y, \tau_{t+1}^O\right)$ *subject to* $\tau_{t+1}^O = T^O\left(U\left(\tau_t^Y\right)\right)$.
3. $U\left(\tau_t^Y\right) = 1 - e^*\left(\tau_t^Y, \tau_{t+1}^O\right)$, *with* $\tau_{t+1}^O = T^O\left(U\left(\tau_t^Y\right)\right)$.

The first equilibrium condition requires that τ_t^O maximizes the utility of the decisive old voter, W^{dec}. As discussed above, the optimal choice of τ_t^O is independent of τ_t^Y and any future variables. The second equilibrium condition requires that even τ_t^Y maximizes the objective function of the decisive voter. Given our voting set-up, we have that $Z^{dec} = Z^o$. Rational voters understand that their choice over current redistribution affects future redistribution via the private decision rule and public policy rules. In equilibrium, $\tau_t^Y = T^Y$

is constant, since W^{dec} does not depend on the state variable. The third equilibrium condition implies that all young individuals optimally choose their investment, given τ_t^Y and τ_{t+1}^O, and that agents hold rational expectations about future taxes and distributions of types.

Note that the three equilibrium conditions can be solved recursively. First, condition 1 defines a one-to-one mapping from the state variable, u_t, to the equilibrium choice of taxation of the old, i.e., $\tau_t^O = T^O(u_t)$. Second, condition 3 defines a functional equation whose solution is the optimal investment schedule, as a function of the current taxation of the young. More explicitly, suppose that agents believe that the future taxation of the old is set according to $T^O(.)$, and that future private investment decisions are taken according to $U\left(\tau_t^Y\right)$. Then, the optimal current investment decisions are also set according to the same function, $U\left(\tau_t^Y\right)$. There is no guarantee that this functional equation has a unique solution; in fact, we will show that for a range of parameter values, multiple fixed points exist. Finally, consider condition 2. Suppose agents believe the future taxation of the old to be set according to the function $T^O(u_t)$ and the mapping from current taxation of the young to future distribution to be given by $U\left(\tau_t^Y\right)$. Condition 2 then defines an optimal choice of taxation today, as a function of the state variable u_t. The strategic voting motive arises from agents recognizing that τ_t^Y affects τ_{t+1}^O via equilibrium policies $T^O(.)$ and $U(.)$.

5 Equilibrium under majority voting.

Assume the political decisions are made though majority voting where voters elect an office-seeking politician proposing income tax rates on the old and the young. In general, the presence of multiple issues can give rise to Condorcet cycles. This problem does not arise in our set-up, however, since there are only two groups of voters (rich and poor old agents). Therefore, under majority voting, the winning politician chooses τ_t^Y and τ_t^O, so as to maximize, respectively, $Z^o(\tau_t^Y, T^O\left(U\left(\tau_t^Y\right)\right))$ and

$$W^{dec}\left(\tau_t^O, u_t\right) = \begin{cases} W^{os}(\tau_t^O, u_t) & \text{if } \mu(1-u_t) \geq 0.5 \\ \\ W^{ou}(\tau_t^O, u_t) & \text{if } \mu(1-u_t) < 0.5. \end{cases} \tag{11}$$

Given (11), the equilibrium mapping $T^O(u)$ (see equilibrium condition 1) is as follows:

$$T^O(u) = \begin{cases} 0 & \text{if } u_t \leq 1 - \frac{1}{2\mu} \\ \\ 1 & \text{if } u_t > 1 - \frac{1}{2\mu}. \end{cases} \tag{12}$$

In words, a majority of rich high-ability agents sets $\tau^O_{t+1} = 0$, whereas a majority of poor (high-ability and low-ability) agents sets $\tau^O_{t+1} = 1$.

Next, we rewrite equilibrium condition 3 by substituting in the optimal effort $e^*\left(\tau^Y_t, \tau^O_{t+1}\right)$. This yields the following functional equation

$$U\left(\tau^Y\right) = \left(1 - \beta + \tau^Y + \beta T^O\left(U\left(\tau^Y\right)\right)\right)/2, \tag{13}$$

where $T^O(.) \in \{0,1\}$ is given by (12). Let U^* denote the set of solutions to this functional equation. That is, each element $U \in U^*$ is a function $U : [0,1] \to [0,1]$ solving (13). Since $T^O(.) \in \{0,1\}$, it immediately follows from (13) that any function $U \in U^*$ must satisfy

$$U\left(\tau^Y\right) \in \left\{\left(1 - \beta + \tau^Y\right)/2, \left(1 + \tau^Y\right)/2\right\} \quad \forall \tau^Y \in [0,1].$$

We can now state the following Lemma

Lemma 1 *Let $U \in U^*$. Then, for any $\tau^Y > \beta - (1-\mu)/\mu$, the function U satisfies $U\left(\tau^Y\right) = \left(1 + \tau^Y\right)/2 > 1/2$ and $T^O\left(U\left(\tau^Y\right)\right) = 1$.*

An immediate implication of Lemma 1 is that, if $\beta < (1-\mu)/\mu$, the set U^* is a singleton, $U^* = \{U^*_-\}$, where the function U^*_- is defined as

$$U^*_-\left(\tau^Y\right) = \left(1 + \tau^Y\right)/2, \ \forall \tau^Y \in [0,1]. \tag{14}$$

Next, consider the set of solutions to (13) with the property that θ is the largest value of τ^Y producing a majority of rich agents in the following period. More formally, given a $\theta \in [0,1]$,

$$U^*_\theta \equiv \left\{U \in U^* \mid U(\theta) = (1 - \beta + \theta)/2 \le 1 - 1/(2\mu) \text{ and } U\left(\tau^Y\right) = \left(1 + \tau^Y\right)/2 \ \forall \tau^Y \in (\theta, 1]\right\}. \tag{15}$$

In words, for any function U being an element of U^*_θ, θ is the largest tax rate on the young for which it is rational to expect a future majority of rich agents. The expectations could also be such that a majority of poor agents materialize if τ^Y is larger *or equal* to some threshold tax-rate θ. We therefore define

$$\hat{U}^*_\theta \equiv \left\{U \in U^* \mid \lim_{\tau^Y \to \theta^-} U\left(\tau^Y\right) = (1 - \beta + \theta)/2 \le 1 - 1/(2\mu) \text{ and } U\left(\tau^Y\right) = \left(1 + \tau^Y\right)/2 \ \forall \tau^Y \in [\theta, 1]\right\}.$$

From the definition in (15), it follows that the sets U^*_θ and $\hat{U}^*_\theta$ are non-empty if and only if $U(\theta) \le 1 - 1/(2\mu)$, i.e., if and only if $0 \le \theta \le \beta - (1-\mu)/\mu$. Therefore, from (15), it follows that the union of the set of subsets U^*_θ, $\hat{U}^*_\theta$ over $\theta \in [0, \beta - (1-\mu)/\mu]$ and U^*_- partitions the set U^*, i.e.,

$$U^* = U^*_- \cup \bigcup_{\theta \in [0, \beta - (1-\mu)/\mu]} \left\{U^*_\theta \cup \hat{U}^*_\theta\right\}. \tag{16}$$

Next, consider the second equilibrium condition in Definition 1. The preferences of the decisive voter over τ^Y depend on the private investment behavior, $U\left(\tau^Y\right)$. Consider, first, the case $U\left(\tau^Y\right) = U_-^*$. Voters then realize that they cannot affect the identity of the future median voter, since a majority of old poor agents will emerge, irrespective of their choice of τ^Y. In this case,

$$T^Y \mid_{U_-^*} = \arg \max_{\tau_t^Y \in [0,1]} Z^o\left(\tau_t^Y, 1\right) = 1/2. \tag{17}$$

Thus, the equilibrium features $\tau^Y = 1/2$ and $\tau^O = 1$ (except in the first period, when $\tau_0^O = 1$ if the poor are in majority and zero otherwise). This equilibrium is "sincere" in the sense that agents do not use current policy strategically in order to reduce future redistribution, but instead indulge in high current taxation.

Consider now the case $U_\theta \in U_\theta^*$ for some $\theta \in [0, \beta - (1-\mu)/\mu]$. In contrast to the above case, agents now have the option to affect the identity of the median voter in the next period. Namely, they can choose τ_t^Y so as to trigger a future majority of rich who will set $\tau_{t+1}^O = 0$. This may be attractive, since Z^o is decreasing in τ_{t+1}^O. The next lemma characterizes the set of strategic equilibria.

Lemma 2 *Assume $a < 1/\mu$. For any $\theta \in \left[\tilde{\theta}(\beta), \beta - (1-\mu)/\mu\right]$, where $\tilde{\theta}(\beta) \equiv \left(1 + \beta - \sqrt{\beta(2+\beta)}\right)/2$, there exists a "strategic" equilibrium $\langle T^O, T^Y, U\rangle$ such that $U\left(\tau^Y\right) = U_\theta \in U_\theta^*$ (where $U_\theta \leq 1 - 1/(2\mu)$), $T^Y = \theta$, and $T^O(u_t)$ is given by (12).*[4]

Lemma 2 shows that, conditional on pursuing the strategy of inducing $\tau_{t+1}^O = 0$, the decisive voter opts for the maximum taxation of the young, consistent with a future majority of rich, i.e., she sets $\tau_t^Y = \theta$. To understand the requirement on θ, observe that, given $U \in U_\theta^*$, inducing $\tau_{t+1}^O = 0$ is optimal if and only if

$$Z^o(\theta, 0) \geq \max_{\tau_t^Y \in [0,1]} Z^o\left(\tau_t^Y, 1\right),$$

which is equivalent to the requirement

$$\theta \geq \left(1 + \beta - \sqrt{\beta(2+\beta)}\right)/2 = \tilde{\theta}(\beta).$$

Lemma 2 implies that necessary and sufficient conditions for a strategic equilibrium to exist are that the set $\left[\tilde{\theta}(\beta), \beta - (1-\mu)/\mu\right]$ be non-empty and $a < 1/\mu$. The former holds if and only if

$$\beta \geq \frac{(2-\mu)^2}{4\mu} \geq \frac{1}{4}. \tag{18}$$

[4] Note that a strategic equilibrium cannot feature $U \in \hat{U}_\theta^*$, since there is no solution to the maximization problem in part 2 of Definition 1.

Since a "sincere" equilibrium always exists, $a < 1/\mu$ and (18) are necessary and sufficient conditions for multiple equilibria. Hence, for smaller values of β, the unique equilibrium features $u > 1 - 1/(2\mu)$ and $\tau^O = 1$.

The next Proposition summarizes the results discussed so far (proof in the text).

Proposition 3 *Assume majority voting. Then,*

1. *if $\beta < (2-\mu)^2/(4\mu)$ and $a \leq 1/\mu$, the Markov equilibrium is unique and is characterized as follows: $T^O(u_t)$ is given as in (12), $T^Y = 1/2$, and either $U(\tau^Y) \in U^*_-$ or $U(\tau^Y) \in U^*_\theta \cup \hat{U}^*_\theta$ with $\theta < \tilde{\theta}(\beta)$. Along the equilibrium path, $\forall j \in [0,\infty)$, $\tau^Y_j = 1/2$, $\tau^O_{j+1} = 1$, $u_j = 3/4 > 1 - 1/(2\mu)$, and $g_j = 3\mu/8$.*

2. *If $\beta \geq (2-\mu)^2/(4\mu)$ and $a \leq 1/\mu$, there exist multiple Markov equilibria. These are of two types:*

   (a) The first equilibrium (sincere equilibrium) is characterized as in part 1 of the proposition.

   (b) The second set of equilibria (strategic equilibria) is characterized as follows: $T^O(u_t)$ is given as in (12), $T^Y = \theta$, and $U(\tau^Y) \in U^*_\theta$, where $\theta \in [\tilde{\theta}(\beta), \beta - (1-\mu)/\mu]$. Along the equilibrium path, $\forall j \in [0,\infty)$, $\tau^Y_{t+j} = \theta$, $\tau^O_{t+1+j} = 0$, $u_j = 1/2 - (\beta - \theta)/2 \leq 1 - 1/(2\mu)$, and $g_j = \mu(1+\beta-\theta)\theta/2$.

3. *If $a > 1/\mu$, the Markov equilibrium is unique and identical to the sincere equilibrium (in part 1 of the proposition) with $U = U^*_-$, except that $T^O \equiv 1$.*

The key element in Proposition 3 is the existence of multiple self-fulfilling private investment rules, $U(\tau^Y)$. Figures 3, 4 and 5 provide, respectively, an illustration of a sincere and a strategic equilibrium.

< Figures 3, 4 and 5 about here >

The left-hand panels of the figures correspond to a sincere equilibrium while the right-hand panels illustrate a strategic equilibrium for one particular θ. Figure 3 shows the tax rate on the old ($T^O(u_t)$) and the young (T^Y) as functions of u_t. Two important properties of the equilibrium tax rates should be noted. First, $T^O(u_t)$ is the same in both equilibria and, in accordance with equation (12), prescribes that zero taxation be chosen if there is a majority of old rich agents (i.e., $u_t \leq 1 - 1/(2\mu)$), and 100% taxation be chosen otherwise. Second, T^Y differs across equilibria; in particular, $T^Y = 1/2 \geq \tilde{\theta}(\beta)$ in the sincere equilibrium, while $T^Y = \theta$ in the strategic equilibrium (where θ may be larger or smaller than 1/2, depending on the parameters).

Figure 4 plots the decision rules, $U\left(\tau^Y\right)$. Note the central role of expectations about how young agents react to current taxation. In the sincere equilibrium (left-hand panel), agents expect that a majority of old rich can only materialize at a low current taxation of the young ($U\left(\tau^Y\right) \in U^*_\theta$ with $\theta < \tilde{\theta}(\beta)$). In strategic equilibria (right-hand panel), agents expect that a majority of old rich can materialize for a relatively high current taxation of the young ($U\left(\tau^Y\right) \in U^*_\theta$ with $\theta \geq \tilde{\theta}(\beta)$).

It is instructive to study how this role of expectations affects the utility of the old from taxing the young, Z^o. This is shown in Figure 5. First, in the sincere equilibrium illustrated here (left-hand panel), it is feasible to induce a majority of old rich agents in the next period, but in order to achieve this, current taxes on the young must be suppressed below $\tilde{\theta}(\beta)$. That would be inferior to opting for $\tau^Y = 1/2$, which attains the top of the Laffer curve, given the expectation that $\tau^O_{t+1} = 1$. However, when $\theta \geq \tilde{\theta}(\beta)$ (and $U\left(\tau^Y\right) \in U^*_\theta$), the right-hand panel of Figure 5, it is beneficial to induce a majority of rich agents in the next period. Note also that the old are indifferent between the two equilibria when $\theta = \tilde{\theta}(\beta)$.

Consider now the public good provision in political equilibria. First, political equilibria always deliver a positive provision in steady-state, bounded from below by $\mu/8$, the lowest provision in strategic equilibria. The public good provision in sincere equilibria always exceeds this minimum provision in strategic equilibria (since $\mu/8 < 3\mu/8$). The reason why political equilibria provide public goods for any $a > 0$, is that the political conflict over redistribution is the only force driving equilibrium taxation (except for the fact that strategic equilibria vanish for $a \geq 1/\mu$). Equivalently, the provision of public good is the only way for the old to obtain inter-generational redistribution from the young. Moreover, with majority voting, votes are not weighted by marginal utility, so in equilibrium, taxes and public good provision are actually independent of a.[5]

However, although sincere equilibria feature a higher taxation of the old, they do not necessarily deliver a larger public good provision than strategic equilibria.[6] The reason is that the maximum of the Laffer curve with respect to τ^Y moves to the right when agents expect zero future taxation of the old. Assume, for example, that $\mu = 1$ and $\beta = 1$. Then, the strategic equilibrium with the largest public good provision has $g = 1/2 > 3/8$, where

[5] In Section 6.2, we consider a different voting mechanism, probabilistic voting, where political choices are affected by the value of public goods and the provision turns out to be generally increasing in a.

[6] It is straightforward to show that the maximum present value, computed using a discount factor λ, of public good provision in strategic equilibria is larger than its sincere equilibrium counterpart whenever $\beta > \mu\lambda/2 + (2-\mu)^2/(4\mu)$.

the latter is the level of public good in the sincere equilibrium.

The normative implications for public good provision in democracies depend on the marginal utility of the public good a. In particular, since the political equilibria are, as discussed above, independent of a, while the Ramsey allocation is increasing in a and in fact commands zero public good whenever $a \leq \min\{\beta, \lambda\} / (\beta + \lambda)$, it immediately follows that the political allocations over-supply public goods for sufficiently low a (see Proposition 1).

Another important normative result can also be proved; for sufficiently large a, the political equilibrium does not deliver enough public goods compared to the Ramsey allocations.[7] One reason for this is that the maximum amount of tax revenues that can be sustained is larger in the Ramsey allocation than in the political equilibria, since one cannot commit to future taxation in the political equilibria. Indeed, in the time-consistent allocation, the public good provision converges to $\frac{\mu}{1-\lambda}(1+\lambda)^2/8$ as $a \to \infty$, which is less than the (Rawlsian) Ramsey allocation.

We now examine normative aspects of political equilibria in more detail. First, the sincere equilibrium bears some resemblance to the Ramsey allocations in the case of $\lambda > \beta$, where the main tax burden falls on the old (see Proposition 1). There are two differences, though. First, when a is low, the Ramsey planner does not tax the old at the maximum rate – the same discrepancy as that between the Ramsey and the time-consistent Ramsey allocation of Proposition 2. Second, in the Ramsey allocation with $\lambda > \beta$, the young are subject to a tax $\tau^Y \leq (1-\lambda)/2$, while they are taxed at a 50% rate in the sincere equilibrium.

Next, consider the Rawlsian planner allocation in the limit case when $\lambda \to 1$, i.e. the allocation that maximizes the steady-state provision of public good. This requires $\tau^O = 1$ and $\tau^Y = 0$. In comparison, the sincere equilibrium imposes a larger tax burden and delivers less public goods (except for the initial period). The reason for this excess taxation of the young is that voters have a short life horizon. Taxing the young is a liability on future generations that is not internalized by the current generation of voters. As a result, each generation inherits a narrower tax-base than that required to maximize the provision of public good. Thus, in a dynamic sense, the sincere equilibrium is on the wrong side of the Laffer curve, due to the inability to commit future policy outcomes. Still, in the absence of transferable utility, a Pareto improvement is not possible since the initial old generation

[7] This can be seen by comparing the present value of public good provision, using a discount factor of λ. This figure is $NPVg_S = \frac{2\lambda+1}{8}\frac{\mu}{1-\lambda}$ for sincere equilibria and $NPVg_\theta \in \left[\frac{1}{8}\frac{\mu}{1-\lambda}, \frac{\beta}{2}\frac{\mu}{1-\lambda}\right]$ for strategic equilibria. It is straightforward to verify that, for sufficiently large a, these figures are smaller than their Ramsey counterparts of Proposition 1.

would have to be made worse off in order to improve future efficiency.

Consider now strategic equilibria. These resemble the Ramsey allocations with low λ, where the entire tax burden falls on the young. Note, though, that the Ramsey allocation of Proposition 1 (as well as the Rawlsian allocation of Corollary 1) commands $\tau^Y = (1+\beta)/2$, whereas, in a strategic equilibrium, $\tau^Y = \theta < (1+\beta)/2$. Therefore, if a is sufficiently large, strategic equilibria feature an inefficiently low provision of public good compared to the allocation maximizing the average utility of voters. The reasons is that old voters are subject to the political constraint that old rich agents remain in majority over time. However, if voters were to set $\tau_t^Y = (1+\beta)/2$, as in their preferred Ramsey allocation, then $e_t^* \leq 1 - (1+\beta)/4 \leq 1/2$ and a majority of rich agents would not arise in the next period. Thus, the political constraint that $\tau_t^Y \leq \theta$ limits the extent to which the current generation of voters can extract a surplus from future generations. Interestingly, the public good provision of the Rawlsian time-consistent planner is actually smaller than that of the sincere equilibrium when λ is sufficiently low, i.e. when the time-consistency problem is the largest (for example when $\beta = \mu = 1$ and $\lambda < 1$). The reason is that, while a majority of rich old agents do not want to tax themselves (for $a < 1/\mu$), the time-consistent planner starts taxing the old already for $a = \min\{\beta, \lambda\}/(\beta+\lambda)$, even though a commitment to not taxing the old would yield more public goods.

Finally, we examine which political equilibrium provides the different groups of living with the highest welfare. Provided that β is sufficiently high, old voters prefer a strategic equilibrium where $\theta = \beta - (1-\mu)/\mu$. Namely, if they could manipulate the expectations of the young, they would make them believe that the welfare state is very fragile. As far as the high-ability young agents are concerned, their expected utility is given by

$$\tilde{V}^y = \left(e^*\left(\tau^Y, \tau^O\right)\right)^2/2 + a(1+\beta)g.$$

As discussed above, the young prefer a tax structure qualitatively similar to that of the sincere equilibrium. However, that equilibrium features too high taxation compared with their preferred Ramsey allocation (i.e., when $\lambda \to 1$). Therefore, the result is *a priori* ambiguous. The expected utility in the two classes of equilibria is

$$\begin{aligned}\tilde{V}^y_{Sinc} &= 1/16 + 3\mu a(1+\beta)/8 \\ \tilde{V}^y_{Strat}(\theta) &= (1+\beta-\theta)^2/4 + \mu a(1+\beta)(1+\beta-\theta)\theta/2,\end{aligned}$$

and there exist parameters such that $\tilde{V}^y_{Sinc} > \tilde{V}^y_{Strat}(\theta)$, as well as the converse.[8]

[8] For instance, set $\mu = 1$ and let $a \to 1$, and consider strategic equilibria where $\theta = \beta$. In this case, $\tilde{V}^y_{Sinc} \to 1/16 + 3(1+\beta)/8$ whereas $\tilde{V}^y_{Strat} \to 1/4 + (1+\beta)\beta/2$, so $\tilde{V}^y_{Sinc} > \tilde{V}^y_{Strat}$ for $\beta < 0.5$, whereas $\tilde{V}^y_{Sinc} < \tilde{V}^y_{Strat}$ for $\beta > 0.5$.

The low-ability young, finally, only care about the provision of the public good. The stationary tax schedule maximizing g is given by $\tau^Y = 0$ and $\tau^O = 1$. Hence, the most preferred tax policy of low and high-ability young workers are similar, and are in fact identical when a is sufficiently large.

We can establish the following general result. If $\beta < 0.5$, and $\mu a \to 1$, then the sincere equilibrium is preferred by all agents in the economy, except the old rich. In particular, all agents prefer, *ex-ante*, to be in a sincere equilibrium.

Throughout the paper we focus on Markov equilibria. Other equilibria, however, exist. A more comprehensive analysis of the set of equilibria can be carried out using the methodology proposed by Abreu, Pearce and Stacchetti (1990). A key step in their methodology is to characterize the *worst* equilibrium. Once such equilibrium is found, one can easily retrieve the range of utilities that can be sustained by credible plans. In our model, characterizing the worst equilibrium is particularly simple, since the most severe punishment that the old voters in period t can be subject to is that the current young be taxed at 100% in period $t+1$. Therefore, if the old in period t deviate from some proposed equilibrium, they expect to be punished with $\tau^O_{t+1} = 1$, and, hence, the best deviation must involve setting $\tau^Y_t = 1/2$. The worst equilibrium is then identical to the sincere equilibrium. An implication of this observation is that, when $\beta < (2-\mu)^2 / (4\mu)$, the sincere equilibrium is unique even if one considers non-Markov equilibria (see Proposition 3).

6 Alternative timing: all agents vote on current tax rates.

In this section, we characterize equilibria in the case when both the old and the young vote on current tax rates, before investment decisions are made. In this case, the presence of multiple issues and more than two groups of voters may lead to Condorcet voting-cycles. To avoid this issue, we consider two settings where this problem does not arise. In the first, voters cast ballots separately on the two issues. In the second, we assume probabilistic voting, following Lindbeck and Weibull (1987) (see Persson and Tabellini, 2000, for a textbook description).

6.1 Majority voting.

In order to rule out Condorcet voting-cycles, this section considers an institutional set-up inspired by the notion of structure-induced equilibrium proposed by Shepsle (1979) (see also Conde-Ruiz and Galasso, 2002, for an application). To this end, we assume that agents vote separately and simultaneously for a politician in charge of setting taxation on the old and another in charge of setting taxation on the young.[9] We show the existence of a unique voting equilibrium in this environment by showing that two conditions hold:

1. For all $t \geq 0$, the vote over τ_t^O yields an unambiguous Condorcet winner, i.e., $\tau_t^O = 1$, irrespective of any past, present or future circumstance.

2. For all $t \geq 0$ the vote over τ_t^Y yields an unambiguous Condorcet winner, given the equilibrium sequence for $\left\{\tau_j^O\right\}_{j=0}^{\infty}$.

When both young and old vote, there will always be a strict majority supporting full taxation of the old rich, i.e., $\tau_t^O = 1$, since this taxation is non-distortionary and benefits all agents but the old rich. Since this is a dominant strategy for the majority in any period t, condition 1 is satisfied.

In order to show condition 2, we need to check that all agents' preferences are single-peaked with respect to τ_t^Y, conditional on $\left\{\tau_j^O = 1\right\}_{j=0}^{\infty}$. To this end, observe that the indirect utility of old agents is given by (7), while the indirect utility of young low-ability and high-ability agents can be written as

$$V^{yl} = W^{ou}\left(u_t, \tau_t^O\right) + Z^y\left(\tau_t^Y, \tau_{t+1}^O\right) + \beta Z^o\left(\tau_{t+1}^Y, \tau_{t+2}^O\right), \tag{19}$$

and

$$V^y = e^*\left(\tau_t^Y, \tau_{t+1}^O\right)\left(1 - \tau_t^Y + \beta\left(1 - \tau_{t+1}^O\right)\right) - e^*\left(\tau_t^Y, \tau_{t+1}^O\right)^2 + V^{yl}, \tag{20}$$

where $Z^y\left(\tau_t^Y, \tau_{t+1}^O\right) = a\mu e^*\left(\tau_t^Y, \tau_{t+1}^O\right)\left(\tau_t^Y + \beta\tau_{t+1}^O\right)$ and W^{ou} and Z^o are as in (8) and (10), respectively.

Consider now how τ_t^Y affects the utility for different groups. We know, by condition 1, that $\tau_{t+1}^O = 1$, independently of τ_t^Y. Therefore, there is no scope for strategic voting in this game. The following Lemma can then be established.

Lemma 3 *Conditional on $\tau_{t+1}^O = 1$, all agents have single-peaked preferences with respect to τ_t^Y. The most preferred alternative for old agents, both rich and poor, is $\tau_t^Y = 1/2$. The most preferred alternative for young low-ability agents is $\tau_t^Y = (1-\beta)/2$. The most preferred alternative for young high-ability agents is $\tau_t^Y = 0$.*

Next, we need to identify the median voter for the choice of τ_t^Y. In our model, 50% of the voters are old and, by Lemma 3, vote for $\tau_t^Y = 1/2$. The remaining 50%, however, would prefer a lower τ_t^Y. Thus, we must specify a tie-break rule. If the tie-break rule were

[9] As will be clear below, the order of the elections is of no importance in our model. Therefore, the assumption of simultaneous voting entails no loss of generality.

to favor the young (for example, if there is population growth), the young low-ability would be pivotal, and the unique equilibrium outcome would be $\tau^Y = (1-\beta)/2$. If the tie-break rule were, instead, to favor the old, the unique equilibrium would prescribe $\tau^Y = 1/2$.

In conclusion, this extension shows that a different timing of voting granting political influence to the young eliminates strategic equilibria, and results in a unique political equilibrium similar to the sincere voting equilibrium of Proposition 3. The reason for this result is that when the young participate in the determination of current period taxation, the old rich can never escape high taxation. Note that this would remain true even under parameters such that the efficient provision of public good would be zero, i.e., for small a's.

The result that empowering the young can increase the size of the welfare state contrasts with the results of HRSZ, where young voters would undermine the sustainability of a transfer system. The reason is that, in that model, voters only have access to one distortionary instrument – an age-independent transfer policy. Here, instead, the poor and the young have, *ex-post*, an incentive to tax the rich old, since their investment decision is sunk and, therefore, taxation entails no distortion *ex-post*. Note that if a is low, it would be in the interest of the young to limit future taxation of the old since, *ex-ante*, this tax is distortionary. However, this plan is time-inconsistent, and will not be implemented since the political mechanism does not provide agents with any commitment device.

A final remark concerns the robustness of our result to the specific political mechanism introduced in this section. In the case that $\mu < 1/2$, and the tie-break rule were to favor the young, the equilibrium outcome $\tau_t^Y = (1-\beta)/2$ and $\tau_t^O = 1$ would be a Condorcet winner even in a standard Downsian model where agents elect a single politician setting both tax rates jointly. To prove this, consider an alternative platform such that $\tau_t^O = 1-\varepsilon$ and $\tau_t^Y = (1-\beta)/2+\nu$, where ε and ν are non-negative deviations. Such a platform may, at best, attract the vote of all old agents, but young agents would continue to support the equilibrium platform. Under the specified tie-break rule, the young are pivotal and the equilibrium platform would still prevail. Consider, instead, the alternative $\tau_t^O = 1-\varepsilon$ and $\tau_t^Y = (1-\beta)/2-\nu$. At best, such a platform would attract the old rich (i.e., a subset of the old high-ability) and the young high ability agents. But, by assumption, low-ability agents are in majority ($\mu < 1/2$). Therefore, this alternative policy will not attract the support of a majority.

6.2 Probabilistic voting.

We now consider the alternative institutional set-up of probabilistic voting. Under probabilistic voting, agents are assumed to evaluate candidates both on their proposed tax rates and on some exogenous ideological traits. Agents are assumed to have heterogeneous preferences in the ideological dimension, where individuals are distributed according to a differentiable distribution function. Compared to the standard Downsian framework, probabilistic voting yields a smoother aggregation of preferences and the outcome can be expressed as if policy outcomes were based on maximizing the weighted average of individual utilities, disregarding the ideological dimension. Weights depend on the relative strength of preferences in the ideological dimension, i.e., how easily voters can be persuaded to switch between candidates. Here, we will assume symmetry: within a cohort, the preferences over ideological traits are identically and uniformly distributed for all groups. However, we do allow for differences across age groups. In particular, we assume old agents' preferences to be at least as focused on taxation issues as those of young agents. This implies that the young have relatively less (or possibly equal) influence in the determination of taxation. We regard this as a plausible case. For instance, Mulligan and Sala-i-Martin (1999) argue that the elderly have a preponderant influence on the determination of redistribution policies, due to the old having a low opportunity cost of time. Moreover, voting turnout is, empirically, lower for younger agents (see Wolfinger and Rosenstone (1980)).

More formally, under probabilistic voting, the winning politician maximizes

$$W^{dec} = V^{pv} = \mu\left(1-u_t\right)V^{os} + \left(1-\mu\left(1-u_t\right)\right)V^{ou} + \omega\mu V^{y} + \omega\left(1-\mu\right)V^{yl}, \tag{21}$$

where $\omega \in [0,1]$ denotes the political weight on the young. If $\omega = 1$, all agents have the same influence on the political outcome. The polar opposite case, $\omega = 0$, represents a case where the young have no influence, and can be regarded as the probabilistic voting counterpart of the benchmark model of section 5. Substituting the expressions for V^{os}, V^{ou}, V^{y}, and V^{yl} into (21) yields

$$\begin{aligned} V^{pv} &= \mu\left(1-u_t\right)\left(1-\tau_t^O\right) + \left(1+\omega\right)W^{ou}\left(u_t,\tau_t^O\right) + Z^o\left(\tau_t^Y,\tau_{t+1}^O\right) \\ &+\omega\left(Z^y\left(\tau_t^Y,\tau_{t+1}^O\right) + \beta Z^o\left(\tau_{t+1}^Y,\tau_{t+2}^O\right)\right) \\ &+\omega\mu\left(e^*\left(\tau_t^Y,\tau_{t+1}^O\right)\left(1-\tau_t^Y+\beta\left(1-\tau_{t+1}^O\right)\right) - e^*\left(\tau_t^Y,\tau_{t+1}^O\right)^2\right). \end{aligned}$$

Note that, as under majority voting, τ_t^O is the only policy variable interacting with the state variable, u_t. Furthermore, preferences are additively separable in τ_t^O versus other taxes. A useful preliminary observation is that the partial derivative $\partial V^{pv}/\partial\tau_t^O = -\mu\left(1-u_t\right) + \left(1+\omega\right)a\mu\left(1-u_t\right)$ is positive (negative) if $\left(1+\omega\right)a > (<)\, 1$, implying that

$$\tau^O = T^O = \begin{cases} 1 & \text{if } a > \frac{1}{1+\omega} \\ 0 & \text{if } a \le \frac{1}{1+\omega}. \end{cases} \tag{22}$$

Note that, under probabilistic voting, T^O is invariant with respect to u_t. The reason is that the objective function V^{pv} does not embed any concern for intra-generational redistribution. Moreover, an increase in ω reduces the threshold above which the old are taxed at 100%. The intuition is straightforward: the young gain from taxing the old and the larger their political influence, the more likely is the outcome $\tau^O = 1$.

Setting τ^O as in (22), and maximizing V^{pv} with respect to τ_t^Y, yields the following result (proved in the appendix):

$$\tau^Y = T^Y = \begin{cases} \frac{1}{2} - \omega \frac{a\beta + 1/2}{2a + \omega(2a-1)} & a \geq \max\left\{\frac{1}{1+\omega}, \frac{\omega}{1+\omega(1-\beta)}\right\} \\ \frac{(1+\beta)(a+\omega(a-1))}{2a(1+\omega)-\omega} & a \in \left(\frac{\omega}{1+\omega}, \frac{1}{1+\omega}\right] \\ 0 & \text{else.} \end{cases} \tag{23}$$

< Figure 6 about here >

Figure 6 displays the equilibrium tax rates as a function of the marginal utility of the public good, a. For a lower range, $a \leq \omega/(1+\omega)$, the young are subject to zero tax. For an intermediate range, $a \in [\omega/(1+\omega), 1/(1+\omega)]$, the tax rate τ^Y is increasing in a. For $a > 1/(1+\omega)$, however, agents forecast that $\tau_{t+1}^O = 1$, and the optimal tax rate on young τ_t^Y falls discontinuously and then increases. The lower panel shows the case when the constraint $\tau^Y \geq 0$ is binding (for $a > 1/(1+\omega)$), and the upper panel shows the unconstrained case.

This equilibrium is qualitatively similar to the time-consistent planner (see Proposition 2). In particular, both the time-consistent allocation and the probabilistic voting equilibrium feature a falling tax on the young τ^Y as the weight on future generations (in the planner case) and the weight on the young (in the probabilistic voting equilibrium) increase. Moreover, both the time-consistent allocation and the probabilistic voting equilibrium feature $\tau^O = 0$ for low a and $\tau^O = 1$ for large a.

It should also be noted that, while the equilibrium public good provision depends on a, it is not necessarily monotone increasing. In particular, the public good provision is non-decreasing in a within each of the segments of a; $a \in [0, 1/(1+\omega)]$ and $a \in \left[\frac{1}{1+\omega}, \infty\right)$, but it may experience a discontinuous fall as a increases past $1/(1+\omega)$.[10]

Finally, in contrast to the case in subsection 6.1 above, τ^Y can actually be lower than τ^O. This happens in the range $a \in \left(\frac{\omega}{1+\omega}, \frac{1}{1+\omega}\right]$ (see Figure 6). The reason is that in this intermediate range of a, where $\tau^O = 0$, taxes on the young transfer resources to the old,

[10] For example, as $\omega \to 0$, the first-period provision of public good, net of the taxation of the initial old, converges to $(1+\beta)^2 \mu/8$ for $a \leq 1/(1+\omega)$ and to $\mu/8$ for $a > 1/(1+\omega)$. Thus, for sufficiently small discount factor λ, the public good provision falls as a increases past $1/(1+\omega)$.

who are more influential in the political process when $\omega < 1$. The range for a where τ^O can be smaller than τ^Y vanishes as ω approaches unity, as this removes the motive for inter-generational transfers from the young to the old.

7 Conclusion

In spite of the increasing literature integrating politico-economic elements into macroeconomic models, most of the existing contributions are silent on the role of forward-looking politico-economic dynamics. Others incorporate them into the model, but rely on numerical techniques (see e.g. Krusell and Ríos-Rull, 1996). In this paper we have analytically characterized the set of Markov equilibria of a model with repeated voting over public good provision, financed by distortionary taxation. The politico-economic equilibrium allocations have been contrasted with a set of Ramsey-optimal allocations.

We have shown that democracy can deliver too much or too little public goods, relative to the Ramsey benchmark. The main reasons are that, first, in political equilibria with majority voting, taxes and public good provision are exclusively driven by redistributive concerns. Second, voters cannot commit to future taxation, which limits the amount of tax revenue that can be raised. For example, even when poor agents who do not pay taxes are in majority, the provision of public goods can be less than that chosen by a Rawlsian planner.

In related work, Hassler, Krusell, Storesletten and Zilibotti (2003) analyze the dynamics of redistribution in an economy where risk averse agents cannot insure each other because of market incompleteness. In future research, we plan to study the effect of productivity shocks on the dynamics of income taxation and the quantitative properties of this model at business cycle frequencies, as well as the role of government debt.

References

Abreu, D., Pearce, D. and Stacchetti, E. (1990), 'Toward a theory of discounted repeated games with imperfect monitoring', *Econometrica* **58**(5), 1041–63.

Alesina, A. and Tabellini, G. (1990), 'A positive theory of fiscal deficits and government debt', *Review of Economic Studies* **57**(3), 403–14.

Hassler, J., Krusell, P., Storesletten, K. and Zilibotti, F. (2003), The dynamics of government: a positive analysis, Mimeo, Stockholm University.

Hassler, J., Rodríguez Mora, J. V., Storesletten, K. and Zilibotti, F. (2003), 'The survival of the welfare state', *American Economic Review*. Forthcoming.

Krusell, P. and Ríos-Rull, J. V. (1996), 'Vested interests in a positive theory of stagnation and growth', *Review of Economic Studies* **63**(2), 301–29.

Lindbeck, A. and Weibull, J. W. (1987), 'Balanced-budget redistribution as political equilibrium', *Public Choice* **52**(3), 273–297.

Mirrlees, J. A. (1994), *Optimal Taxation and Government Finance*, pp. 213–31. in J. M. Quigley and E. Smolensky Ashenfelter (eds.) *Modern Public Finance* Harvard University Press: Cambridge, MA.

Persson, T. and Svensson, L. E. O. (1989), 'Why a stubborn conservative would run a deficit: Policy with time-inconsistent preferences', *Quarterly Journal of Economics* **104**(2), 325–45.

Persson, T. and Tabellini, G. (2000), *Political Economics - Explaining Economic Policy*, MIT Press, Cambridge.

Ruiz, J. I. C. and Galasso, V. (2002), 'Early retirement', *Reveiw of Economic Dynamics*. forthcoming.

Shepsle, K. A. (1979), 'Institutional arrangements and equilibrium in multidimensional voting models', *American Journal of Political Science* **23**(1), 27–59.

Wilson, J. D. (1991), 'Optimal public good provision with limited lump-sum taxation', *American Economic Review* **81**(1), 153–66.

8 Appendix

Proof of Proposition 1. The proof of the claim that $\tau_0^O = 1$ is immediate, given the parametric condition. The second part descends from the solution of the following maximization problem

$$\begin{aligned}\max_{\{\tau^Y\in[0,1],\tau^O\in[0,1]\}} \bar{L}\left(\tau^Y,\tau^O\right) &= \Big\{\lambda\mu\left(e^*\left(\tau^Y,\tau^O\right)\left(1+\beta-\tau^Y-\beta\tau^O\right)-\left(e^*\left(\tau^Y,\tau^O\right)\right)^2\right)+ \\ &\quad \left(\beta+\lambda\right)a\mu\left(1+\beta-\tau^Y-\beta\tau^O\right)\left(\tau^Y+\lambda\tau^O\right)/2\Big\}.\end{aligned}$$

The solution must satisfy the following FOCs:

$$\begin{aligned}\frac{\partial\bar{L}}{\partial\tau^Y}-\xi^Y+\theta^Y &= 0 \\ \frac{\partial\bar{L}}{\partial\tau^O}-\xi^O+\theta^O &= 0,\end{aligned}$$

where θ^Y,θ^O are the Kuhn-Tucker multiplier associated with the constraints $\tau^Y \geq 0$ and $\tau^O \geq 0$, respectively, whereas ξ^Y and ξ^O are the Kuhn-Tucker multiplier associated with the constraints $\tau^Y \leq 1$ and $\tau^O \leq 1$, respectively.

Assume, first, that $\tau^O = \xi^Y = \xi^O = \theta^Y = 0$. Then, one obtains

$$\begin{aligned}\tau^Y &= \frac{\left(1+\beta\right)\left(a\left(\beta+\lambda\right)-\lambda\right)}{2a\left(\beta+\lambda\right)-\lambda}, \\ \theta^O &= \frac{1}{2}\mu a^2\frac{\left(1+\beta\right)\left(\beta-\lambda\right)\left(\beta+\lambda\right)^2}{2a\left(\beta+\lambda\right)-\lambda},\end{aligned}$$

where $\tau^Y < 1$ and $\theta^O > 0$ as long as $\beta > \lambda$. This establishes part *1.*

Assume, next, that $\tau^Y = \xi^Y = \xi^O = \theta^O = 0$. Then, one obtains

$$\begin{aligned}\tau^O &= \frac{\left(1+\beta\right)\left(a\left(\beta+\lambda\right)-\beta\right)}{\beta\left(2a\left(\beta+\lambda\right)-\beta\right)}, \\ \theta^Y &= \frac{1}{2}\mu a^2\frac{\left(1+\beta\right)\left(\lambda-\beta\right)\left(\beta+\lambda\right)^2}{\beta\left(2a\left(\beta+\lambda\right)-\beta\right)},\end{aligned}$$

where $\theta^Y > 0$ as long as $\beta < \lambda$. Furthermore, it is immediate to verify that $\tau^O < 1$ as long as $a < \frac{\beta}{(1-\beta)(\beta+\lambda)}$ (part *2a*). If $a \geq \frac{\beta}{(1-\beta)(\beta+\lambda)}$, then $\tau^O = 1$ and $\xi^O > 0$. It remains to be checked, however, whether $\tau^Y = 0$ continues to be a solution. Given $\tau^O = 1$, we obtain

$$\frac{\partial\bar{L}}{\partial\tau^Y}-\xi^Y+\theta^Y = \frac{1}{2}\left(-\lambda\mu\left(1-\tau^Y\right)+\left(\beta+\lambda\right)a\mu\left(1-\lambda-2\tau^Y\right)\right)+\theta^Y = 0.$$

The solution features $\tau^Y = 0$ and $\theta^Y > 0$ if and only if $a \leq \frac{\lambda}{(1-\lambda)(\beta+\lambda)}$, which establishes part *2b*. For larger values of a, the solution instead features $\theta^Y = 0$ and

$$\tau^Y = \frac{a\left(1-\lambda\right)-\frac{\lambda}{\beta+\lambda}}{2a-\frac{\lambda}{\beta+\lambda}},$$

as in part *2c*. This establishes part *2*.

Finally, consider the knife-edge case where $\beta = \lambda$. Then, the objective function can be written as

$$\bar{L}(x) = \beta\mu\left(e^*(x)(1+\beta-x) - (e^*(x))^2\right) + \beta a\mu(1+\beta-x)x,$$

where $x \equiv \tau^Y + \beta\tau^O$. The FOCs yields the result in part *3*.

As to the public good provision, it is straightforward to compute the steady-state public good provision in each of the cases by substituting in for tax rates from Proposition 1. PVg_R is then computed by discounting the infinite sum of these taxes (note that τ^Y start accumulating in the first period while τ^O start accumulating in the second period). Given PVg_R, it is straightforward to show that $df/da > 0$ for $a \geq x$ since

$$f'(a) = (2a-x)^{-3}x^2 > 0.$$

This establishes that PVg_R is increasing for $\beta \geq \lambda$ and, for $\beta < \lambda$, within each of the three segments $a \in \left[\frac{\beta}{(\beta+\lambda)}, \frac{\beta}{(1-\beta)(\beta+\lambda)}\right]$, $a \in \left[\frac{\beta}{(1-\beta)(\beta+\lambda)}, \frac{\lambda}{(1-\lambda)(\beta+\lambda)}\right]$, and $a \in \left[\frac{\beta}{(1-\beta)(\beta+\lambda)}, \infty\right)$. It remains to be shown that PVg_R is increasing when switching segments. This holds since

$$\begin{aligned} &\lim_{a\to\frac{\beta}{(1-\beta)(\beta+\lambda)}}\left\{\frac{(1+\beta)^2}{1-\lambda}\frac{\lambda}{\beta}\frac{\mu}{2}\cdot f\left(a,\frac{\beta}{\beta+\lambda}\right)\right\} \\ = &\lim_{a\to\frac{\lambda}{(1-\lambda)(\beta+\lambda)}}\left\{\frac{\mu}{2}\frac{(1+\lambda)^2}{1-\lambda}\cdot f\left(a,\frac{\beta}{\beta+\lambda}\right)\right\} = \frac{\lambda}{1-\lambda}\frac{\mu}{2}. \end{aligned}$$

Finally, $PVg_R \geq 0$ since $\lim_{a\to x} f(a,x) = 0$.

Proof of Lemma 1. We prove the Lemma by establishing a contradiction. Suppose that, in the range $\tau^Y > \beta - (1-\mu)/\mu$, there exists a $\bar{\tau}^Y$ such that $T^O\left(U\left(\bar{\tau}^Y\right)\right) = 0$. This would imply that $U\left(\bar{\tau}^Y\right) = \left(1-\beta+\bar{\tau}^Y\right)/2 > 1 - 1/(2\mu)$. From equation (12), it then follows that $T^O\left(U\left(\bar{\tau}^Y\right)\right) = 1$, in contradiction with the presumption. We have therefore established that, for all $\tau^Y > \beta - (1-\mu)/\mu$, it must be that $T^O\left(U\left(\tau^Y\right)\right) = 1$. Hence, (13) implies that $U\left(\tau^Y\right) = \left(1+\tau^Y\right)/2$, completing the proof of the Lemma.

Proof of Lemma 2. We start by showing that given $U_\theta \in U^*_\theta$, the decisive voter strictly prefers $\tau^Y = \theta$ to any $\tau^Y < \theta$. To prove this, observe that $Z^o\left(\tau^Y, 0\right) = a\mu\tau^Y_t\left(1+\beta-\tau^Y\right)/2$ is a hump-shaped function of τ^Y, with a maximum at $(1+\beta)/2 \geq \beta$. Since $\theta \leq \beta - (1-\mu)/\mu < \beta$, it follows that $Z^o\left(\tau^Y, 0\right)$ is necessarily increasing in τ^Y in the range $\tau^Y \in [0,\theta]$, and voters never choose $\tau^Y < \theta$. Next, suppose $\tilde{\theta}(\beta) \leq \beta - (1-\mu)/\mu$ and consider a function $U_\theta \in U_\theta^*$ for some $\theta \leq \beta - (1-\mu)/\mu$. Thus, it is optimal to set $\tau_t^Y = \theta$ if and only if

$$Z^o(\theta, 0) = a\mu\theta(1+\beta-\theta)/2 \geq a\mu/8 = Z^o(1/2, 1) = \max_{\tau_t^Y \in [0,1]} Z^o\left(\tau_t^Y, 1\right).$$

This inequality holds iff $\theta \geq \left(1+\beta-\sqrt{\beta(2+\beta)}\right)/2 = \tilde{\theta}(\beta)$. From the definition of U_θ^*, it follows that $U_\theta(\theta) = (1-\beta+\theta)/2 \leq 1 - 1/(2\mu)$. The function T^O follows from equation (12). **QED**

Proof of Proposition 3. Given $\tau_{t+1}^O = 1$, equation (17) implies that it is optimal to set $\tau_t^Y = 1/2$. Moreover, if $U \in U_\theta^*$ with $\theta < \tilde{\theta}(\beta)$, it is not optimal to opt for $\tau_t^Y = \theta$ and induce $\tau_{t+1}^O = 1$, since $Z^o(\theta, 0) < \max_{\tau_t^Y \in [0,1]} Z^o\left(\tau_t^Y, 1\right)$. It follows that $u_{t+1} = 3/4$ for all $t \geq 0$, which confirms that this is an equilibrium. The uniqueness of the sincere equilibrium when $\beta < (1-\mu)/\mu + \mu/4$ follows from the discussion preceding equation (18). This completes parts *1* and *2a* of the proposition. Part *2b* follows from Lemma 2.

Proof of the result in Section 6.2. The political objective function is

$$\begin{aligned} W &= (1+\omega)a\left(\mu(1-u_t)\tau_t^O + \frac{\mu}{2}\left(1+\beta-\tau_t^Y-\beta\tau_{t+1}^O\right)\tau_t^Y\right) \\ &\quad +\mu(1-u_t)\left(1-\tau_t^O\right) + \omega\beta a\left(\frac{\mu}{2}\left(1+\beta-\tau_t^Y-\beta\tau_{t+1}^O\right)\tau_{t+1}^O + \frac{\mu}{2}\left(1+\beta-\tau_{t+1}^Y-\beta\tau_{t+2}^O\right)\tau_{t+1}^Y\right) \\ &\quad +\frac{\omega\mu}{2}\left(\frac{1}{2}\left(1+\beta-\tau_t^Y\right)^2 - \beta\tau_{t+1}^O\left(1+\beta-\tau_t^Y-\frac{1}{2}\beta\tau_{t+1}^O\right)\right). \end{aligned}$$

The choice of taxes on the old is derived in the main text. For the choice of taxes on the young, we note that

$$\frac{dW}{d\tau_t^Y} = \frac{1}{2}\mu\left((1+\beta)(a+\omega(a-1)) - (2a(1+\omega)-\omega)\tau_t^Y - \beta(a(1+\omega)+\omega(a-1))\tau_{t+1}^O\right).$$

Consider first the case when $a > \frac{1}{1+\omega}$. Then, $\tau_{t+1}^O = 1$ and

$$\begin{aligned} \frac{dW}{d\tau_t^Y} &= \frac{1}{2}\mu\left((1+\beta)(a+\omega(a-1)) - (2a(1+\omega)-\omega)\tau_t^Y - \beta(a(1+\omega)+\omega(a-1))\right) = 0 \\ &\rightarrow \tau_t^Y = \frac{1}{2} - \omega\frac{a\beta+1/2}{2a+\omega(2a-1)} \in \left[0, \frac{5}{12}\right]. \end{aligned}$$

For this to be non-negative, we require $a \geq \frac{\omega}{1+\omega(1-\beta)}$ (in addition to $a > \frac{1}{1+\omega}$), and we note that $\frac{\omega}{1+\omega(1-\beta)}$ is implied by $\frac{1}{1+\omega}$ iff $\beta < \frac{1-\omega^2}{\omega}$, which always holds if $\omega < -\frac{1}{2} + \frac{1}{2}\sqrt{5} \approx 0.618$.

Now, when $a < \frac{1}{1+\omega}$, we have $\tau_{t+1}^{O} = 0$ and

$$\begin{aligned} \frac{dW}{d\tau_t^Y} &= \frac{1}{2}\mu\left((1+\beta)(a+\omega(a-1)) - (2a(1+\omega)-\omega)\tau_t^Y\right) = 0 \\ &\rightarrow \tau_t^Y = \frac{(1+\beta)(a+\omega(a-1))}{(2a(1+\omega)-\omega)} \in \left(0, \frac{1}{2}\right]. \end{aligned}$$

This increases in a and is non-negative when $a \geq \frac{\omega}{1+\omega}$.

Proof of Lemma 3. First, recall that given $\tau_{t+1}^{O} = 1$,

$$\begin{aligned} V^{yl}\left(\tau_t^Y, \tau_{t+1}^O = 1, .\right) &= K^{yl} + Z^y\left(\tau_t^Y, 1\right) \\ V^{y}\left(\tau_t^Y, \tau_{t+1}^O = 1, .\right) &= e^*\left(\tau_t^Y, 1\right)\left(1-\tau_t^Y\right) - e^*\left(\tau_t^Y, 1\right)^2 + K^{yl} + Z^y\left(\tau_t^Y, 1\right) \\ V^{ol}\left(\tau_t^Y, \tau_{t+1}^O = 1, .\right) &= K^{o} + Z^y\left(\tau_t^Y, 1\right), \end{aligned}$$

where K^{yl} and K^o are independent of τ_t^Y. Then, observe that from the definition of Z^y and Z^o, it follows that

$$\begin{aligned} Z^y\left(\tau_t^Y, 1\right) &= \frac{a\mu}{2}\left(1-\tau_t^Y\right)\left(\tau_t^Y + \beta\right) \\ Z^o\left(\tau_t^Y, 1\right) &= \frac{a\mu}{2}\left(1-\tau_t^Y\right)\tau_t^Y. \end{aligned}$$

Standard methods establish that both $Z^y\left(\tau_t^Y, 1\right)$ and $Z^o\left(\tau_t^Y, 1\right)$ are concave and, hence, single-peaked with respect to τ_t^Y, and that they have unique maxima at $\tau_t^Y = (1-\beta)/2$ and $\tau_t^Y = 1/2$, respectively. The properties of Z^y and Z^y carry over to V^{yl} and V^o. Finally, concerning the preferences of the high-ability young, observe that

$$\frac{\partial V^y\left(\tau_t^Y, \tau_{t+1}^O = 1, .\right)}{\partial \tau_t^Y} = \frac{1}{2}\left(a\mu(1-\beta) - 1\right) + \left(\frac{1}{2} - a\mu\right)\tau_t^Y < 0, \tag{24}$$

for all $\tau_t^Y \in [0,1]$. To see why this derivative is negative, recall that, by assumption, $a\mu < 1$. Hence, $a\mu(1-\beta) < 1$. Thus, if $a\mu > 1/2$, the derivative $\partial V^y\left(\tau_t^Y, \tau_{t+1}^O = 1, .\right)/\partial \tau_t^Y$ is necessarily negative. If $a\mu < 1/2$, it is sufficient to show that the derivative is negative at $\tau_t^Y = 1$. The resulting expression is, in this case, $\partial V^y/\partial \tau_t^Y = a\mu((1-\beta)/2 - 1) < 0$. Hence, $\tau_t^y = 0$ is the maximal alternative for the young high-ability agents. QED.

Figure 1

Ramsey-optimal Tax Rates

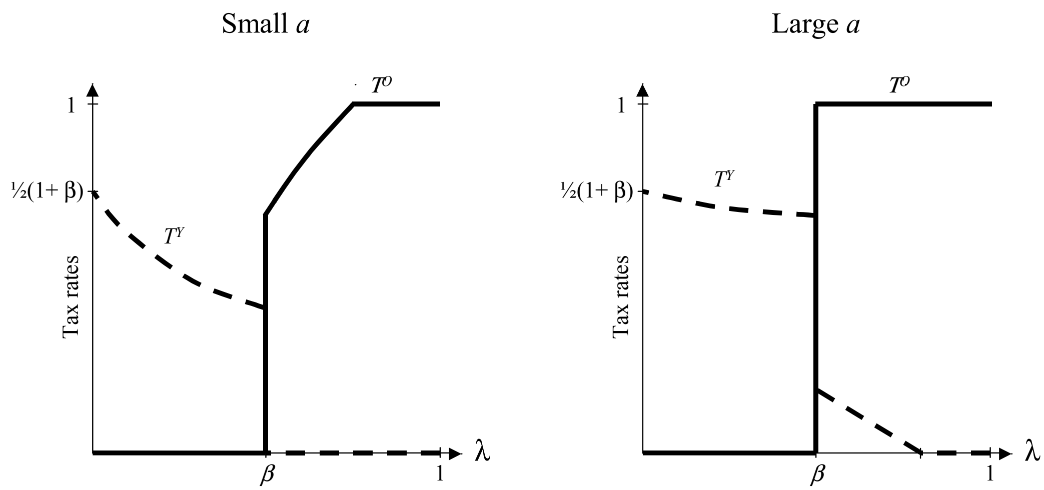

Figure 1: The figures represents tax rates, T^Y and T^O, as a function of the planner's discount factor λ in the Ramsey allocations of Proposition 1. The left-hand (right-hand) panel shows the optimal tax rates when the marginal value of public goods, a, is relatively small (large). In the figures, the private discount factor and the fraction of high-ability agents are held constant at $\beta = 1/2$, and $\mu = 1$.

Figure 2

Present Value of Public Goods

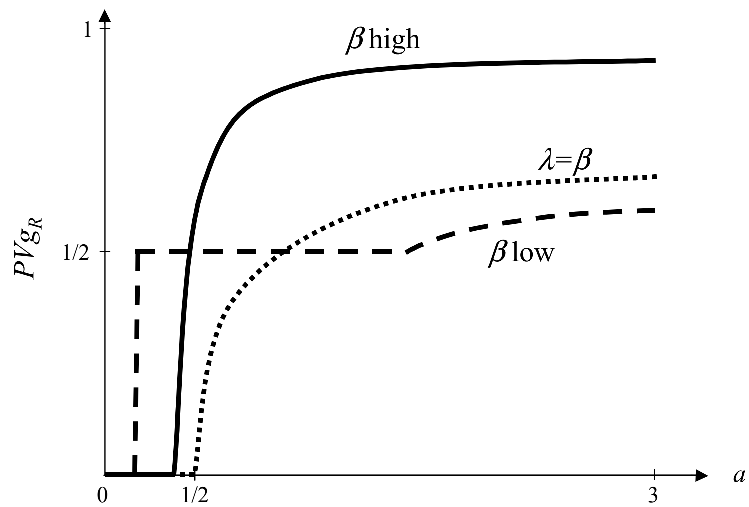

Figure 2: The figure represents PVg_R, the present value of public good provision as a function of a in the Ramsey allocations of Proposition 1. The planner's discount factor is held constant, and the three graphs represent three different values of β, displaying the different cases in Proposition 1. The parameter values behind the figure are $\lambda = 1/2$ and $\beta \in \{0.1, 0.5, 0.9\}$.

Figure 3
Equilibrium Tax Rates under Majority Voting

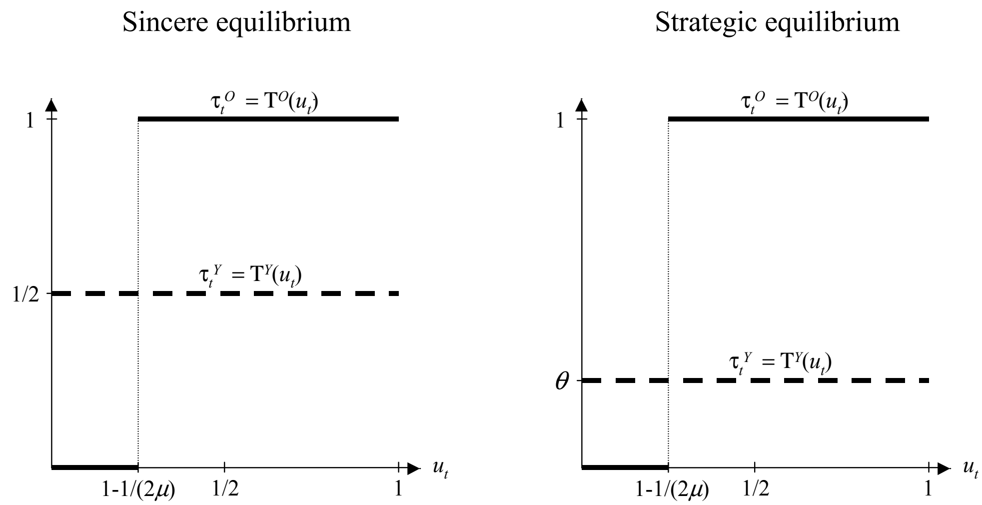

Figure 3: The figures represent the equilibrium tax functions, T^Y and $T^O(u_t)$, under majority voting. The left-hand (right-hand) panel corresponds to a sincere (strategic) equilibrium of proposition 3. The parameter values behind the figures are $\mu = 2/3$ and $\beta = 3/4$, and in the right-hand panel, θ is set to $\theta = \beta - (1 - \mu)/\mu = 1/4$.

Figure 4

Equilibrium Decision Rules under Majority Voting

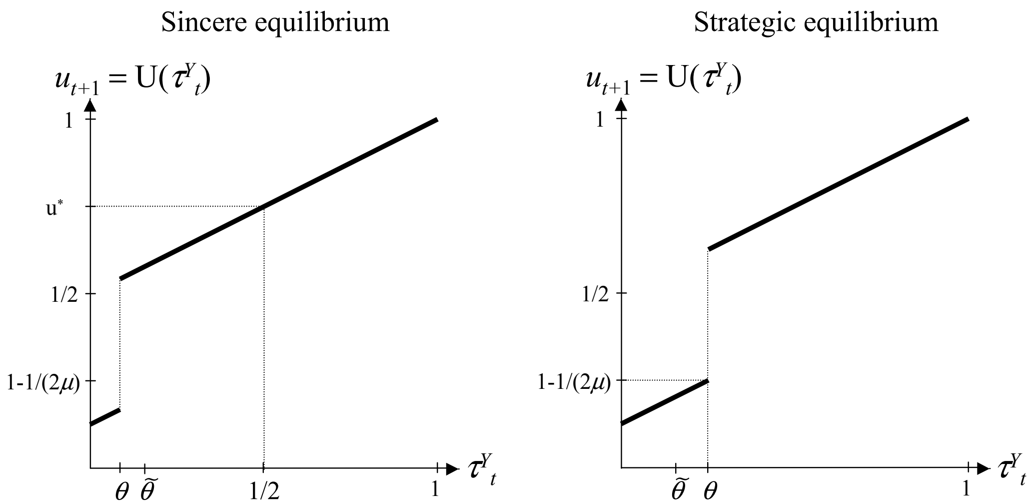

Figure 4: The figures represent the equilibrium decision rule $u_{t+1} = U\left(\tau_t^Y\right)$ under majority voting. The left-hand (right-hand) panel corresponds to a sincere (strategic) equilibrium of proposition 3. The parameter values behind the figures are $\mu = 2/3$ and $\beta = 3/4$, and in the right-hand panel, θ is set to $\theta = \beta - (1 - \mu)/\mu = 1/4$.

Figure 5

$Z^o\left(\tau_t^Y, \tau_{t+1}^O\right)$ ***under Majority Voting***

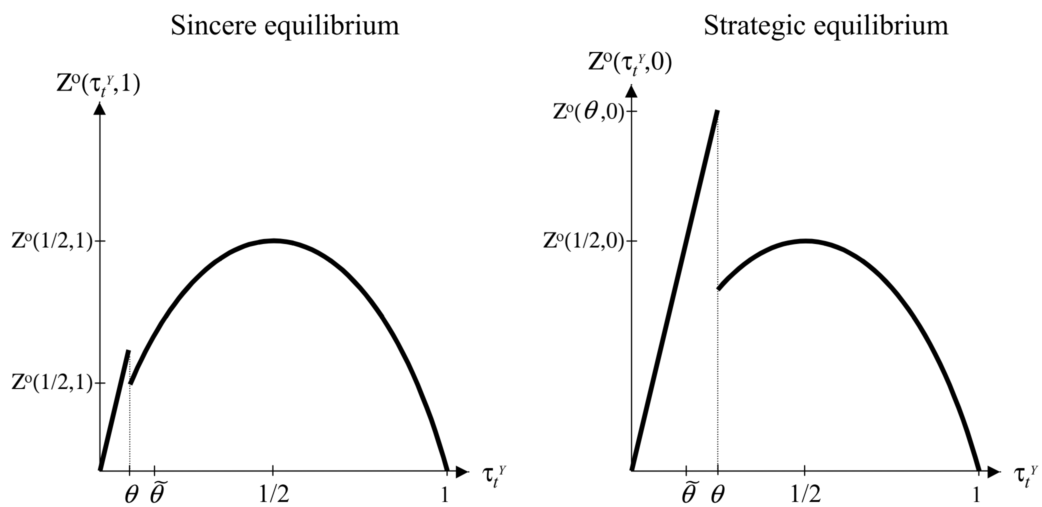

Figure 5: The figures represent the utility of old agents from taxing the young, Z^o, under majority voting. The left-hand (right-hand) panel corresponds to a sincere (strategic) equilibrium of proposition 3. The parameter values behind the figures are $\mu = 2/3$ and $\beta = 3/4$, and in the right-hand panel, θ is set to $\theta = \beta - (1 - \mu)/\mu = 1/4$.

Figure 6

Equilibrium Tax Rules under Probabilistic Voting

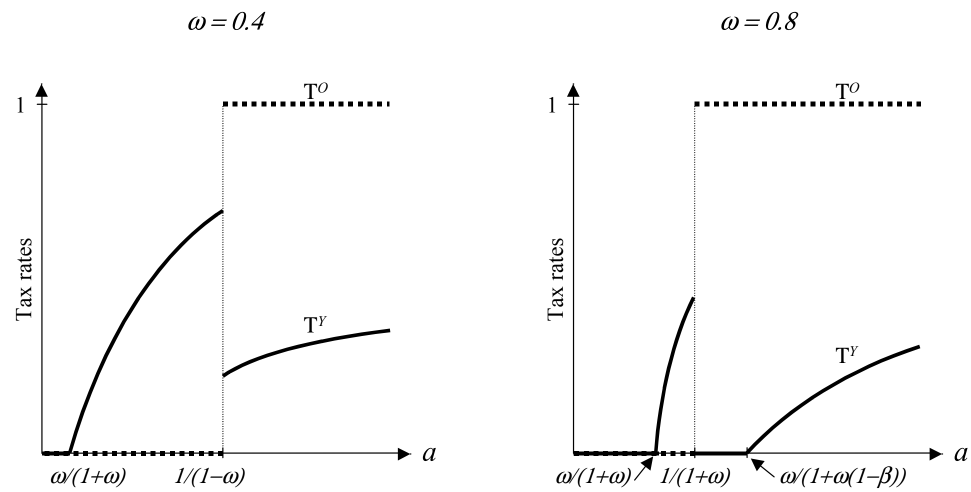

Figure 6: The graphs represent values for T^Y and T^O for various values of the marginal value of the public good, a. The left-hand panel is a case when $1/(1+\omega) > \omega/(1+\omega(1-\beta))$, and the right-hand panel represents the opposite case. The parameter values are $\beta = 1/2$ and $\omega = 0.4$ and $\omega = 0.8$ for left-hand and right-hand panel, respectively.