The Dynamics of Government.*

John Hassler[†], **Per Krusell** [‡],
Kjetil Storesletten[§] **and Fabrizio Zilibotti**[¶]

March, 2005

Abstract

How does the size of the transfer system evolve in the short and the long run? We model income redistribution as determined by voting among individuals of different types and income realizations. Taxation is distortionary because it discourages effort to accumulate human capital. Voters cannot commit to future tax rates, but are fully rational: they realize that transfers have implications also for future economic decisions and taxation outcomes. In our economy, redistribution provides insurance and we investigate to what extent it is appropriately provided by the democratic process.

A general finding is that redistribution tends to be too persistent relative to what would have been chosen by a planner with commitment. The difference is larger, the lower is the political influence of young agents, the lower is the altruistic concern for future generations, and the lower is risk-aversion. Furthermore, there tends to be too much redistribution in the political equilibrium. Finally, we find the political mechanism to be important: settings with smooth preference aggregation—we analyze probabilistic voting here—produce less persistence and do not admit multiple expectational equilibria, which occur under majority-voting aggregation.

*We thank Costas Azariadis, Alex Cukierman, Allan Drazen, Ed Green, Robert King, Etienne Lehmann, Dirk Niepelt, Torsten Persson, Harald Uhlig, Joseph Zeira and many seminar participants for valuable comments.

[†] John Hassler: IIES and CEPR; e-mail: John.Hassler@iies.su.se.

[‡] Per Krusell: Princeton University, IIES, and CEPR; e-mail: pkrusell@princeton.edu.

[§] Kjetil Storesletten: University of Oslo, IIES and CEPR; e-mail: Kjetil.Storesletten@econ.uio.no.

[¶] Fabrizio Zilibotti: IIES and CEPR; e-mail: fabrizio.zilibotti@iies.su.se.

1 Introduction

The evolution of the modern welfare state—transfers across consumers that are put in place via the democratic process—is arguably an important determinant both of inequality and aggregate macroeconomic performance. Yet there are few theoretical analyses of its determinants. To a large extent, this is likely due to the complexity of dynamic models with endogenous policy. Some analyses therefore restrict attention to models that are static in essence, thus not allowing forward-looking decisions, such as investment, to interact with policy choice. Other analyses incorporate realistic investment decisions but instead derive all results with numerical methods. Our ultimate aim is to provide a framework which is quantitatively reasonable, but which also allows the mechanisms to be understood as well as possible.

In this paper, we develop a theory of the dynamics of the welfare state that allows nontrivial forward-looking in investment but which can be analyzed analytically. Voters are selfish and rational: transfer policy is democratically determined by citizens whose only aim is to influence policy so as to maximize their own utility. The tractability we gain, and which comes from simplifying assumptions on preferences and technology, allows us to isolate and analyze the origins of the endogenous policy dynamics. When these assumptions are relaxed in future work, we will also learn what features of preferences and technology are important for policy dynamics. There are two roles of redistribution in our model. First, it is a pure wealth transfer motivated by selfish concerns: *ex post*, some agents are rich and others poor, and to the extent that the poor have political influence, they achieve net redistribution in their favor. Second, transfers also provide agents with *ex-ante* insurance, because we consider a world where insurance markets for individual risk are missing. Moreover, redistribution is costly: it reduces effort. Costly effort is expended as a function of its benefits—the difference between a high-wage and a low-wage outcome—and transfers from the rich to the poor thus reduce this benefit. We assume agents to be identical *ex ante*, but that individual preferences over redistribution then diverge as agents age, since some become successful in life, while others are less fortunate. The political system we consider does not have any direct commitment mechanisms—current voters cannot bind the hands of future voters—and we focus on the case where reputational mechanisms are absent.

Our setup has implications both for the long-run (average) level of redistribution and for the redistribution dynamics. The main primitives in our model are *(i)* the political mechanism (we assume a probabilistic-voting setup à la Lindbeck and Weibull, 1987), *(ii)* the associated weights on different agents (such as on the old versus on the young), *(iii)* the amount of risk aversion of agents, *(iv)* a measure of the distortionary impact of taxation, and *(v)* the income distribution, whose endogenous evolution drives the dynamics by which the size of government evolves. The long-run size of government is higher with higher risk aversion, lower when the distortionary impact of taxes is large, and lower when the interests of future generations are better represented in the political process.

The dynamics of taxation depend on the size of the group of poor through two channels. First, a large such group makes the tax cost per unit of benefits high. Thus, redistribution is more costly the larger is the group of poor agents. This speaks for lower taxes and redistribution. We call this the *tax-base effect*; the optimum under commitment discussed above precisely builds on trading off these tax-base effects over time. Second, with probabilistic voting an increase in the number of poor voters leads to larger political power of the group favoring redistribution. This *constituency effect* thus goes in the other direction: a larger number of poor agents speaks for higher taxes. Depending on parameter values, we can either have a tax rate which is increasing or decreasing in the size of the group of poor. The tax-base and constituency effects are key in driving the dynamics of our equilibria.

When risk-aversion is high, so that insurance is highly socially valuable, the dynamics feature oscillations: redistribution is high (low) in periods when the number of beneficiaries is low (high). This prediction is broadly consistent with some empirical evidence. For example, DiTella and MacCullough (2002) find that unemployment benefits (replacement rates) fall as the unemployment rate increases in a panel of OECD countries. Similarly, Razin et al. (2002) document that pension benefits fall as the dependency ratios increase.

One might suspect that the oscillatory nature of the dynamics is the result of political distortions, as argued in different contexts by the literature on "political business cycles" (see Alesina, Roubini and Cohen 1997 for a survey). Surprisingly, we reach the opposite conclusion: the political mechanism exerts a *stabilizing* influence on the redistribution dynamics and, in fact, renders it too smooth. More precisely, we find that relative to a "constrained optimal" allocation, the political system dampens, and for some parameter values even completely eliminates, the cycles that would be present in the optimal allocation. Our political equilibrium always settles down to a steady state, though for some parameter values in an oscillatory manner. In contrast, the constrained optimum may entail oscillations that do not die out.

The fact that taxes oscillate in the Ramsey allocation is interesting *per se*, as it seems to contrast with the usual tax-smoothing wisdom (see, e.g., Barro, 1979). A more general analysis of this issue is developed in Hassler et al. (2004). The intuition is the following. Effort decisions associated with human capital accumulation (like schooling, learning on the job, etc.) yield returns over many years. Thus, the taxes relevant to an agent's effort decision is the present value of taxes on the product of the effort. For example, the decision to obtain an MBA degree is distorted by all taxes on income accruing during the lifetime following the degree. Thus, an agent contemplating a human capital accumulation decision early on in life—consider, as we do here, two-period-lived agents making a one-time effort choice at the beginning of life—does not particularly care about whether taxes fluctuate over time or are constant. It follows that optimal allocations tend to lead to cycles, because if there is a reason to tax at a high rate at a point in time, e.g., to redistribute to the initial old or to finance a one-time expenditure, then the distortionary impact of this tax hike can be reduced by lowering taxes in the next period.[1] However, since taxes are lowered the following period, taxes must, in order to keep the present value constant, be increased again two periods forward. The pattern is repeated: a one-time splash produces ripples. The political system reduces these (constrained-optimal) oscillations because it lacks commitment. It cannot automatically adjust future—or past—taxes to reduce the distortionary impact of redistribution and social insurance. Thus, oscillations are partly or fully offset with politically determined taxation.

We emphasize the lack of commitment in the political mechanism by focusing entirely on equilibria which are limits of the corresponding finite-horizon economies. The absence of reputation mechanisms is operationalized by focusing on (in our case, linear) Markov-perfect equilibria. Of course, if the horizon is literally infinite and there is sufficiently low discounting, one could construct a large variety of equilibria (for this approach, see, e.g., Bernheim and Nataraj, 2002). We think, however, that it is useful to carefully examine the implications of a complete lack of commitment. Moreover, in models with state variables, there are channels allowing current voters to influence the future, thus not replicating commitment but imperfectly replacing it, as in the strategic-debt literature (see e.g. Persson and Svensson, 1989). Here, the state variable is the initial group of unlucky agents: a large such group tends to lead to high redistribution in the current period (assuming that equilibrium redistribution is driven by the constituency effect). As a consequence, next period's

[1] We abstract from the ability of governments to run deficits here, but introducing deficit financing does not eliminate tax-driven cycles; see Hassler et al. (2004) for details.

redistribution can be influenced today by using current taxes to influence current effort and, hence, affecting the set of unlucky agents at the beginning of the next period.

Once the attention has been limited to Markov equilibria, we still face the important question of whether our political Markov equilibria are unique: can we expect "stability" in the size of government in democracies? A similar setup, considered in Hassler et al. (2003a), assumes Downsian majority voting and finds that Markov equilibria are not unique: in one equilibrium the welfare state survives, while in another it collapses. In that case, multiplicity arises from a stark feature of majority voting models; the equilibrium tax rate increases discontinuously as the number of poor exceeds 50%. This opens the possibility of voting "strategically" over redistribution to induce future changes in the majority. The probabilistic-voting mechanism, in contrast, features a smooth mapping from group sizes to tax outcomes. In fact, we show here that multiple equilibria cannot occur with probabilistic voting in the finite- and infinite-horizon equilibria of our baseline setup, whereas they can with majority voting.

All results are analytical, due to the assumption that the investment decision – effort exerted to increase the probability of becoming productive in the future – is modelled with a quadratic cost and a linear benefit. This approach is similar to that addressing strategic concerns in the literature on time-consistent policies and differential games (see e.g. Cohen and Michel, 1988; and Hansen and Sargent, 2004). The combination of politics and economics is what poses a difficulty; one needs to model strategic voting interactions, where political agents consider the consequences of their choice on future political outcomes, as well as appeal to dynamic equilibrium theory to ensure that all economic agents—consumers, firms and government—maximize their respective objective functions under rational expectations and resource feasibility. Most nontrivial dynamic models (that is, that are not repeated static frameworks or purely "backward-looking" setups) rely on numerical solution (see, e.g., Krusell and Ríos-Rull (1999)). Hassler et al. (2003a) provided a tractable linear-quadratic framework where voters are influenced both by the state of the economy—the current income distribution—and foresee effects of the current policy outcomes on both future income distributions and future voting outcomes, about which they care. The present paper uses a model similar to that of Hassler et al. (2003a), but extends it in a technically non-trivial and economically important way by introducing a social insurance motive, and considering a richer voting model, i.e., probabilistic voting.[2]

The paper is organized as follows. Section 2 presents the economic structure of the model, and Section 3 describes the political decision making and analyzes politically determined redistribution. Section 3.4 discusses uniqueness of equilibrium under a finite horizon, and the connection between our Markov-perfect equilibrium and the limit of finite-horizon equilibria. Section 3.5 studies the case where voters are altruistic toward future generations (but cannot commit to future policy). Finally, Section 4 analyzes the constrained optimum: the allocation chosen by a planner who cares about future generations and has commitment. Section 5 concludes. Most proofs are provided in a technical appendix available upon request.[3]

[2]In Hassler et al. (2003a), redistribution is, by construction, socially wasteful, as it distorts incentives, while agents are risk neutral, so that insurance has no value. Therefore, the constrained-optimal allocation always entails zero redistribution, and the paper does not yield interesting normative implications. Apart from the different assumption about the political mechanism (majority vs. probabilistic voting), the model presented here encompasses Hassler et al. (2003a) as the particular case in which agents are risk neutral, as will be shown in the subsequent discussion.

[3]See http://www.iies.su.se/~zilibott/jmeappendix.pdf.

2 The model

2.1 Population, preferences, technology, and policy

The model economy has a continuum of two-period lived agents, who work in both periods. Upon birth, agents are subject to an ability shock. With probability μ, an agent is high-skilled, and with probability $1-\mu$, she is low-skilled. We label high-skilled agents "entrepreneurs" and low-skilled agents "workers". Entrepreneurs undertake a risky investment in human capital, yielding a stochastic return. With probability e, the investment is successful and the entrepreneur earns labor income $w+\underline{w}$ in each period, where $w \leq 1$. With probability $1-e$, the investment is unsuccessful, and the labor income is $\underline{w}$ in each period. The cost of investment is e^2, and we interpret it as the disutility of educational effort. Workers earn an income normalized to zero, which cannot be affected by human capital investments.

To make the problem interesting, we assume that the component $\underline{w}$ of the entrepreneurial income is not verifiable. Therefore, insurance agencies, whether private or public, cannot condition payments on agents' skills, but can only discriminate between successful entrepreneurs (with a verifiable income equal to w) and the rest of the population (with a verifiable income normalized to zero).

Agents' preferences are given by

$$V_t^y = E_t[v(c_t) + \beta v(c_{t+1}) - e_t^2],$$

where $\beta \in [0,1]$ is the discount factor and

$$v(c) = \begin{cases} ac - (a-1)x & \text{if } c < x \\ c & \text{if } c \geq x \end{cases},$$

with $a \geq 1$. Thus, felicity is concave and piecewise linear in consumption; marginal utility drops discretely at a threshold consumption level x and is constant everywhere else, as shown in Figure 1. The kink in preferences helps us maintaining analytical tractability while allowing *ex-ante* risk aversion. The parameter a regulates the concavity of the utility function; if $a=1$ agents are risk neutral, while if $a>1$ they are risk averse.

We assume that $\underline{w} > x$, implying that, after the realization of the ability shock, high-skill agents are effectively risk-neutral. In particular, the marginal utility of income for high-skill agents is equal to unity, independent of their income realization. Since agents cannot sign contracts before their skill level is realized, this implies that no private insurance market can exist. The government can, however, increase the *ex-ante* utility of agents through redistributive programs providing insurance "behind the veil of ignorance". Like private insurers, governments can only condition transfers on observable income. Unlike private insurers, however, they can force agents to be part of the insurance scheme by setting compulsory taxes. In particular, in each period, the government can levy a lump-sum tax τ on all agents and transfer the proceeds to individuals with low observable income (either workers or unsuccessful entrepreneurs).[4] We denote the transfer rate $b \in [0,1]$, implying that all agents but the successful entrepreneurs receive an amount bw. The government budget is assumed to balance in every period, and the government cannot issue age-dependent taxes and transfers (see Hassler et al. (2003b) for an extension where age-dependent programs are allowed).

[4] The assumption of lump-sum taxes is immaterial. It can be shown that the model is isomorphic to one where transfers are financed by taxation levied on the observable component of labor income. The proof is available upon request.

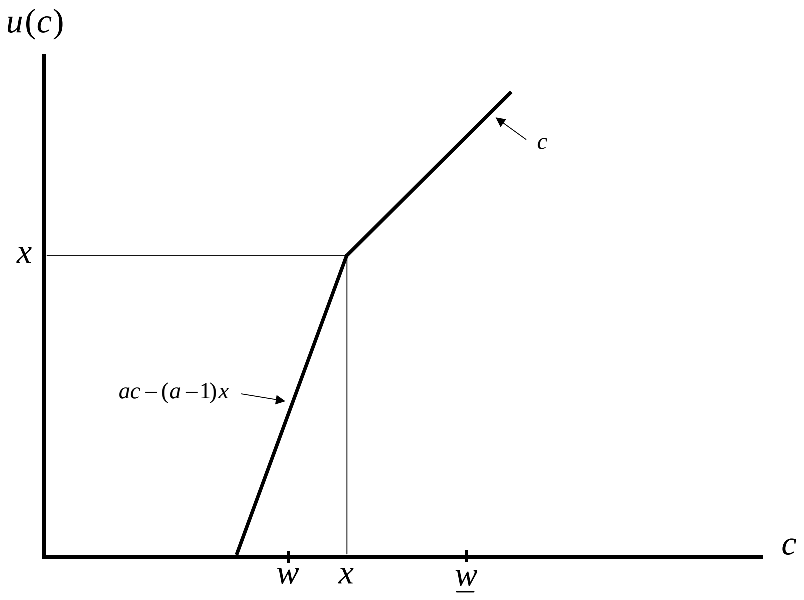

Figure 1. Worker felicity

Furthermore, we assume that $w < x$. This assumption simplifies the analysis, since it implies that the marginal utility of low-skill agents is $a > 1$, irrespective of the redistribution policy (recall that $b \leq 1$).[5] Thus, in summary, we assume that

$$w < x < \underline{w}. \tag{1}$$

Finally, we assume that the subjective discount rate, $(1 - \beta) / \beta$, equals the market interest rate. Under this assumption, the savings decisions can be abstracted from, since income is the same in both periods of life for all individuals.

2.2 Discussion of assumptions

We assume that only high skilled individuals have hidden income and make an effort choice and that individual ability is revealed already at the beginning of life. These assumptions are stark, but we believe that they provide a reasonable shortcut description of important real-world features: (i) in terms of their effect on productivity, the effort and human capital investments of some workers are more important than those of others; (ii) it is likely that those agents with high entrepreneurial ability are also well-endowed in other dimensions, therefore having higher income than workers also if they are less successful; and (iii) already before entering college, individuals have a good idea of their prospects in life.[6]

Our model abstracts from physical capital. The effects of redistribution on the accumulation of physical capital may of course be important, but the distortion to human capital accumulation considered here captures the same kind of dynamic trade-offs that are present in a standard consumption-savings decision.

[5] In a previous version of this paper, we assumed that even workers had stochastic income, but that this income process was exogenous and not influenced by investments. This analysis led to qualitatively similar results.

[6] For example, Keane and Wolpin (1997) argue that up to 90% of the variance of individuals' lifetime utility can be explained by information known when they are 16 years old.

Why are insurance markets missing in our model, and why is there a role for government-provided redistribution? In reality, parents are able to shelter their offspring against some types of verifiable ability shocks. In the model, we abstract from this possibility by assuming that parents are not altruistic. However, even though parental altruism should deliver some intergenerational insurance (for example, that successful parents make transfers to unskilled children), we believe that such insurance will never be perfect. Thus, democratic constitutions, allowing a possibility to vote over transfers, provide redistribution with an *ex-ante* insurance value that neither altruism nor private insurance markets can deliver.

The policy instruments available to the government are quite limited by design; in most political-economy setups—and this one is no exception—the policy instruments are restricted so as to yield a nontrivial and interesting choice situation for voters/the government. We do not allow budget deficits and surpluses and, more importantly, we restrict the ability of governments to target transfers to specific groups. In particular, we assume that unlucky entrepreneurs and workers are pooled in the same program, while it would be beneficial to separate them. Although this is an extreme characterization, it captures the realistic feature that welfare state programs are plagued by informational problems reducing their effectiveness and increasing their cost. The absence of government debt is instead due to tractability consideration, and we plan to extend our analysis in this direction in future research.

2.3 The determination of effort as a function of government policy

Ignoring irrelevant constants, the utilities of the agents alive at time t can be expressed as a function of government policy variables (benefits and taxes) and human capital investments:

$$\begin{aligned}
V_t^{oes} &= w - \tau_t \\
V_t^{oeu} &= b_t w - \tau_t \\
V_t^{ow} &= a(b_t w - \tau_t) \\
V_t^{ye} &= e_t\,(1+\beta)\,w + (1-e_t)\,(b_t + \beta b_{t+1})\,w - e_t^2 - (\tau_t + \beta\tau_{t+1}) \\
V_t^{yw} &= a\,(b_t w - \tau_t + \beta\,(b_{t+1} w - \tau_{t+1}))\,,
\end{aligned} \tag{2}$$

where superscripts oes, oeu, ow, ye and yw denote old successful entrepreneurs, old unsuccessful entrepreneurs, old workers, young entrepreneurs and young workers, respectively.

The optimal investment choice of the young entrepreneurs, given b_t and b_{t+1}, is

$$e_t^* = e\,(b_t, b_{t+1}) \equiv \frac{1+\beta-(b_t+\beta b_{t+1})}{2} w. \tag{3}$$

Since the realization of the investment is i.i.d. across entrepreneurs, and they all choose the same level of effort, $e_t\,(b_t, b_{t+1})$ is also the proportion of entrepreneurs who become successful. Moreover, since success is persistent, this is also the proportion of successful old entrepreneurs in period $t+1$. It is useful to denote the proportion of unsuccessful entrepreneurs by $u_{t+1} = 1 - e_t\,(b_t, b_{t+1})$.[7]

The government budget constraint is $2\tau_t = (2\,(1-\mu) + \mu u_t + \mu\,(1-e_t^*))\,w b_t$. Using (3), we have

$$\begin{aligned}
\tau_t &= \tau\,(b_t, b_{t+1}, u_t) \\
&= \left(1 + \frac{\mu}{2}\left(u_t - 1 - (1+\beta)\,\frac{w}{2} + (b_t + \beta b_{t+1})\,\frac{w}{2}\right)\right) b_t w.
\end{aligned} \tag{4}$$

[7]The restrictions $0 \le b \le 1$ and $w \le 1$ imply that $u_{t+1} \in \left[\frac{1-\beta}{2}, 1\right]$.

The marginal tax cost of redistribution in period t, $\partial\tau/\partial b_t$, increases in u_t (because more old entrepreneurs are benefit recipients) and in b_t and b_{t+1} (because more young entrepreneurs become unsuccessful). Since the old in period t cannot enjoy any benefits in period $t+1$, their equilibrium utility will therefore be decreasing in b_{t+1}.

Preliminary remarks about preferences for redistribution are as follows. The old successful entrepreneurs prefer zero benefits, since redistribution implies positive taxes without providing them with any benefits. Benefit recipients (workers and unsuccessful entrepreneurs), in contrast, are better off with some redistribution, even though their preferences for redistribution may be non-monotonic, as net benefits may be falling with b at high levels of taxation, due to a Laffer curve effect. Note also that the Laffer curve is dynamic, depending both on historical investment levels and expectations about future taxation.

After the ability shock is realized, young workers like redistribution more than do young entrepreneurs. However, it should be noted that the government transfer programs entail some intergenerational redistribution, since the proportion of old and young successful entrepreneurs may differ. The preferences of the different groups of young agents will therefore depend on the balance between inter- and intra-generational effects.

3 Political equilibrium

3.1 The political game

In the political equilibrium, the benefit policy is chosen through voting each period. In the benchmark case, we assume that agents vote over next period's redistribution at the end of each period, after the uncertainty about individual entrepreneurial earnings has been realized. Since the old have no interest at stake, they are assumed to abstain. This is equivalent to assuming that agents vote over the current benefit policy before the effort choice of the entrepreneurs is made, and that only the old agents are entitled to vote (see Hassler et al., 2003a). We later extend the analysis to the case when both the young and the old vote on current benefits.

3.1.1 Probabilistic voting

We assume a two-candidate political model of probabilistic voting à la Lindbeck and Weibull (1987) and restrict attention to Markov-perfect equilibria. In this model, whose features are extensively discussed in Persson and Tabellini (2000) and which are therefore not detailed here, agents cast their votes on one of two candidates, who maximize their probability of becoming elected. Voters have heterogeneous preferences not only over redistribution, but also over some non-economic-policy dimension that is orthogonal to redistribution and over which the candidates cannot make binding commitments. As in Persson and Tabellini (2000), we refer to this additional dimension as "ideology". Voters differ in their evaluation of the candidates' ideology and their preferences over this dimension are subject to an aggregate shock whose realization is unknown to the candidates when platforms over redistribution are set.[8] In the equilibrium of this model, both candidates choose the same platform over redistribution and each of them has a fifty percent probability of winning. More importantly, the impact of each group on the equilibrium policy outcome increases with the relative weight in utility of the policy variable. Intuitively, if agents in a group have a lower concern for ideology, a candidate making a small change in redistribution in favor of this

[8] Since candidates have no intrinsic preferences over redistribution, they are assumed to implement their promised platform.

group will trigger a larger increase in her political support. In other terms, groups with many "swing-voters" are more attractive to power-seeking candidates, and exert a stronger influence on the equilibrium political outcome. Thus, we will assume that the relative concern for ideology versus redistribution is the same within cohorts, but may vary between cohorts. Under this assumption, it is straightforward to show that in equilibrium, the candidates' platforms simply maximize a weighted sum of individual utilities, where the weights are the same for all agents within a cohort but may differ between cohorts. Thus, the equilibrium policy maximizes a "political objective function" which is a weighted average utility of all voters. We will consider the cases when the political weight on the old is normalized to unity and the weight on the young is $\omega \in [0,1]$.

3.1.2 Definition of equilibrium

The "political" aggregation of the different preferences is summarized by the following function

$$\begin{aligned} & V\left(b_t, b_{t+1}, b_{t+2}, u_t, u_{t+1}\right) \\ \equiv \; & \mu\left((1-u_t)V_t^{oeu} + u_t V_t^{oes}\right) + (1-\mu)\, V_t^{ow} + \omega\left(\mu V_t^{ye} + (1-\mu)\, V_t^{yw}\right), \end{aligned} \tag{5}$$

where the equilibrium functions $e_t = e\left(b_t, b_{t+1}\right)$, as in (3), and $\tau_t = \tau\left(b_t, b_{t+1}, u_t\right)$ and $\tau_{t+1} = \tau\left(b_{t+1}, b_{t+2}, u_{t+1}\right)$, as in (4), have been substituted into each function V_t^j defined in equations (2).

The function V entails the assumption that all agents within a given generation exert the same political influence, irrespective of type. In the general case where two generations participate in each election, however, we allow for age-specific differences in the concern for the ideological dimension. This is parameterized by $\omega \in [0,1]$. In particular, $\omega < 1$ means that on average, the old care less about ideology and have more "swing-voters" than the young. Hence, their preferences carry more weight in the political objective function, V. The opposite would be true if $\omega > 1$, a case that we do not consider. When $\omega = 1$, all voters are equally represented.

We construct equilibria with linear policy functions (except for kinks implied by bounds on benefits) in the aggregate state variable: the proportion of current unsuccessful old entrepreneurs, u_t. The political equilibrium is defined as follows.

Definition 1 *A political equilibrium is defined as a pair of functions $\langle B, U\rangle$, where $B:[0,1]\to[0,1]$ is a public policy rule, $b_t = B\left(u_t\right)$, and $U:[0,1]\to[0,1]$ is a private decision rule, $u_{t+1} = U\left(b_t\right)$, such that, given the political weight $\omega \in [0,1]$ on the young, the following functional equations hold:*

1. $B\left(u_t\right) = \arg\max_{b_t\in[0,1]} V\left(b_t, b_{t+1}, b_{t+2}, u_t, u_{t+1}\right)$ subject to $u_{t+1} = U\left(b_t\right)$, $b_{t+1} = B\left(U\left(b_t\right)\right)$, and $b_{t+2} = B(U(B\left(U\left(b_t\right)\right)))$, and

2. $U\left(b_t\right) = 1 - e\left(b_t, b_{t+1}\right)$ with $b_{t+1} = B\left(U\left(b_t\right)\right)$.

The first equilibrium condition requires the political mechanism to choose b_t to maximize V, taking into account that future redistribution depends on the current policy choice via the equilibrium private decision rule and future equilibrium public policy rules. Furthermore, it requires $B(u_t)$ to be a fixed point in the first functional equation of Definition 1. In other words, suppose that agents believe future benefits to be set according to the function $b_{t+j} = B\left(u_{t+j}\right)$. Then, we require the same function $B(u_t)$ to define optimal benefits today. It should be noted that in the case $\omega = 0$, the political objective in (5) depends only on b_t, b_{t+1}, and u_t.

The second equilibrium condition states that young individuals choose their investment optimally, given b_t and b_{t+1}, and that agents have rational expectations about future benefits and distributions of types. In general, U could be a function of both u_t and b_t, but in our particular model u_t has no direct effect on the investment choice of the young. Thus, in our equilibria the equilibrium investment choice of the young is fully determined by the current benefit level.

3.2 An economy where agents are risk-neutral

In this section, we consider the particular case when $a = 1$, i.e., when agents are risk-neutral and therefore have the same preference intensity for economic policy *ex post*. Here, the welfare state entails no insurance value.

It is instructive to see how the equilibrium is constructed. Let us therefore sketch the method used to find the political equilibrium in the simplest case when $\omega = 0$, leaving the details for the cases discussed later in the paper to a technical appendix available upon request. Since in this case neither b_{t+2} nor u_{t+1} enter the aggregate political preferences, we can rewrite (with a slight abuse of notation) the political objective, (5), as

$$V\left(b_t, b_{t+1}, u_t\right) = \frac{\mu}{2}\left(u_t b_t w - \left(1 - e\left(b_t, b_{t+1}\right)\right) b_t w\right) + \mu(1 - u_t) w. \tag{6}$$

Note that the last term is exogenous from the voter's perspective: it is predetermined. Omitting this and the proportionality factor $\mu/2$, and noting that $1 - e\left(b_t, b_{t+1}\right) = u_{t+1}$, the political objective can therefore be written as $(u_t - u_{t+1}) b_t w$: positive benefits help the current old only if the number of unsuccessful old entrepreneurs exceeds the number of unsuccessful young entrepreneurs. Naturally, the latter is determined by policy. Thus, the workers do not enter this expression: since they are of equal number in each cohort, any transfers between them will net to zero.

Disregarding constants and using the expression for $e\left(b_t, b_{t+1}\right)$, the political objective in (6) can be written as

$$u_t b_t w - \left(1 - (1+\beta)\frac{w}{2} + \frac{w}{2}\left(b_t + \beta b_{t+1}\right)\right) b_t w. \tag{7}$$

We need to find two functions $B(u_t)$ and $U(b_t)$, satisfying the two equilibrium conditions in Definition 3. Guided by the linear-quadratic form of the objective function, we guess on the functional form for B: $B(u_t) = \alpha_0 + \alpha_1 u_t$, for some yet undetermined coefficients α_0 and α_1. Using this guess, the second equilibrium condition can be written as

$$U\left(b_t\right) = 1 - \frac{1 + \beta - \left(b_t + \beta\left(\alpha_0 + \alpha_1 U\left(b_t\right)\right)\right)}{2} w. \tag{8}$$

Solving for $U(b_t)$, we obtain

$$U\left(b_t\right) = \frac{2 - w\left(1 + \beta\left(1 - \alpha_0\right)\right) + b_t w}{2 - \beta \alpha_1 w}. \tag{9}$$

Substituting the expression for $U(b_t)$ and the guess of $B(u_t)$ into the first-order condition and solving for b_t gives

$$b_t = \frac{1}{2w}\left(-2 + w\left(1 + \beta\left(1 - \alpha_0\right)\right)\right) + \frac{2 - \beta\alpha_1 w}{2w} u_t,$$

which verifies the tentative guess as a fixed-point of equilibrium condition 1 if $\alpha_1 = \frac{2}{w(2+\beta)}$ and $\alpha_0 = -\alpha_1\left(1 - \frac{1}{2}(1+\beta) w\right)$, heuristically establishing the following proposition.[9]

[9] Given the quadratic objective, it is straightforward to check that the first-order condition will be sufficient for a maximum. What remains is to check that the constraint $b_t \in [0, 1]$ is satisfied along the proposed equilibrium. This check is carried out in the technical appendix.

Proposition 1 *Assume $a = 1$ and $\omega = 0$ (risk neutrality, "n", and only the old vote, "o"). The political equilibrium is characterized as follows:*

$$B^{no}(u_t) = \begin{cases} \frac{2}{w(2+\beta)}(u_t - u^{no}) & \text{if } u_t \geq u^{no} \\ 0 & \text{else} \end{cases}$$

$$U^{no}(b_t) = u^{no} + \frac{w}{2}\left(1 + \frac{\beta}{2}\right) b_t,$$

and the equilibrium law of motion is, for any $u_0 > u^{no}$,

$$u_{t+1} = u^{no} + \frac{1}{2}(u_t - u^{no}).$$

The economy converges monotonically to a unique steady state with $b = b^{no} = 0$ and $u = u^{no} = 1 - e(0,0)$. For $u_0 \leq u^{no}$, $u_t = u^{no}$ $\forall t > 0$.

Proof. In addition to what is stated in the text, the constraint $b_t \in [0,1]$ remains to be verified. The policy function $B(u_t) = [2/(w(2+\beta))] \cdot (u_t - u^{no})$ is positive for any $u_t \geq u^{no} = 1 - (1+\beta) w/2$. However, if $u_t < u^{no}$, the restriction $b_t \geq 0$ will bind. Thus, the guess in the text must be modified to $B(u_t) = [2/(w(2+\beta))] \cdot (u_t - u^{no})$ if $u_t \geq u^{no}$, and $B(u_t) = 0$ otherwise. This new guess will still maximize the political objective, (7). To see this, note that for any feasible policy (b_t, b_{t+1}), $u_{t+1} \geq u^{no}$. Thus, when agents use the equilibrium policy rule to forecast $b_{t+1} = B(u_{t+1})$, only the part $B(u) = [2/(w(2+\beta))] \cdot (u - u^{no})$ is relevant. For the same reason, $U(b)$ in (8) is unaffected. ∎

In the equilibrium of Proposition 1, redistribution occurs along the transition path, i.e., as long as $u_0 > u^{no}$. In the long run, however, there is no redistribution. Convergence is monotonic, and the dynamics are characterized by a positive root equal to $1/2$ (note that, since b is a linear function of u, then, in equilibrium $b_{t+1} = b_t/2$.). The speed of convergence is thus independent of w and β.

Figure 2 represents the equilibrium policy function and the law of motion of the state variable. The left-hand panel shows that when $u_t > u^{no}$, redistribution is positive in equilibrium. Moreover, the equilibrium level of b_t increases linearly with u_t. The right-hand panel illustrates how the equilibrium law of motion implies monotonic asymptotic convergence to the steady state, as long as $u_0 > u^{no}$.

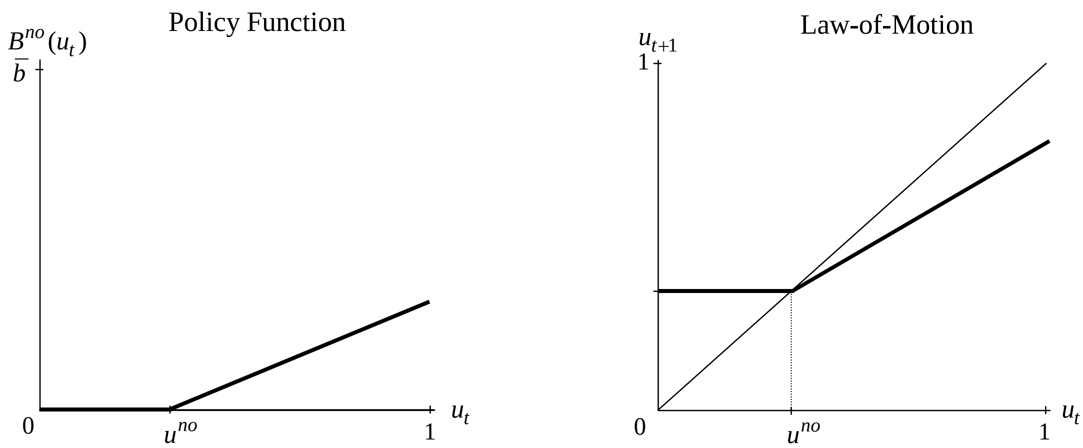

Figure 2. Risk neutrality and only the old vote.

Our results can be interpreted as follows: when only the old influence the political outcome, the equilibrium redistribution, $B^{no}(u_t)$, maximizes the average income of the old. This implies maximizing the intergenerational transfer from young to old individuals without any concern for intra-generational redistribution. Intergenerational transfers benefitting the current voters can, however, be achieved by setting $b_t > 0$ only if the proportion of old unsuccessful agents is higher than the proportion of young unsuccessful, i.e., if $u_t > u_{t+1}$. In particular, no redistribution can occur in steady state. The results of Proposition 1 generalize to the case of $\omega \in [0, 1]$.

Turning to the participation of young voters, we have

Proposition 2 *For $a = 1$ and any $\omega \in [0, 1]$, the political equilibrium is characterized as follows:*

$$
\begin{aligned}
B^n(u_t) &= \begin{cases} \frac{2Z}{w(1+\beta Z)}(u_t - u^n) & \text{if } u_t \geq u^n \\ 0 & \text{else} \end{cases} \\
U^n(b_t) &= u^n + \frac{w}{2}(1+\beta Z)\, b_t,
\end{aligned}
$$

and the equilibrium law of motion is, for any $u_0 > u^n$,

$$u_{t+1} = u^n + Z(u_t - u^n),$$

where $Z \in [0, 1/2]$ is a decreasing function of ω. The economy converges monotonically to a unique steady state with $b = b^n = 0$ and $u = u^n = 1 - e(0, 0)$. For $u_0 \leq u^n$, $u_t = u^n$ $\forall t > 0$.

In the case when both young and old agents vote on current benefits, the equilibrium has the same qualitative features as in the benchmark case ($\omega = 0$), provided that $\omega < 1$.[10] In particular, redistribution occurs along the transition path, but there is no welfare state in the long run. Since $\omega < 1$, the old are politically preponderant and the political equilibrium therefore favors redistribution from the young to the old. Such redistribution can be achieved via positive benefits if and only if $u_t > u_{t+1}$. Therefore, redistribution is positive only along the transition to the steady state. For any $\omega < 1$, dynamics are characterized by a positive root $Z \leq 1/2$. The higher is ω, the lower are the transfers and the flatter are the equilibrium policy function and the law of motion in Figure 1. This is due to the fact that the young exert political pressure against redistribution. With $\omega = 1$, benefits are zero regardless of u_t and the system immediately jumps to the steady state.

Hassler et al. (2003a) find that in a model of majority voting, the welfare state can survive in the long run, even though agents are risk-neutral. However, the results here show that under probabilistic voting, redistribution must die off in the long run. For the same economic environment, the long-run state of the transfer system can critically depend on the form of the democratic process. Moreover, the transitional dynamics are here characterized by monotonic rather than oscillatory convergence.

3.3 The case of risk-averse agents

In this section, we show that the political equilibrium features the long-run survival of the welfare state under probabilistic voting, provided that a positive proportion of agents in society are risk-averse. The convergence to the steady state may be oscillatory or monotonic, depending on the extent of risk aversion and the political influence of the young.

[10]When $\omega = 1$, then $Z = 0$, so that the economy converges immediately to a steady-state with zero redistribution. Moreover, when $\omega = 0$, then $Z = 1/2$, so that this Proposition nests the result of Proposition 1. An implicit expression for Z can be found in the technical appendix containing the complete proofs.

As in the previous subsection, we will initially assume that the young agents have no influence in the voting process, i.e., that $\omega = 0$. When $a \geq 1$, the political objective function, $V(b_t, b_{t+1}, u_t)$, can be expressed (up to scaling and excluding constants) as follows:

$$\begin{aligned} V(b_t, b_{t+1}, u_t) &= \frac{2R}{1+R}(1-u_t)\, b_t \\ &\quad + u_t b_t - \left(1 - (1+\beta)\frac{w}{2} + \frac{w}{2}(b_t + \beta b_{t+1})\right) b_t, \end{aligned} \tag{10}$$

where $R \equiv (1-\mu)(a-1) \geq 0$ is the population weighted marginal utility of income in excess of the marginal utility of entrepreneurs (unity), an aggregate measure of society's desire to redistribute. This political objective function is derived as in the case of risk neutrality, differing by the first term (the second line, as above, equals $(u_t - u_{t+1})b_t$), which reflects a positive effect from redistribution whenever aggregate risk aversion is positive: the higher marginal utility of workers makes any redistributed dollar pay off more, the higher is R. The part $(1-u_t)$ is the fraction of old successful entrepreneurs, representing the size of the inelastic tax base. This term inversely reflects the distortionary cost of redistribution. We can now characterize the equilibrium as follows.

Proposition 3 *Assume $\omega = 0$ and risk aversion ("a"). Then, $\exists R_{\max} > 1$ such that, if $R \in [0, R_{\max}]$, the political equilibrium is characterized as follows:*

$$\begin{aligned} B^{ao}(u_t) &= \begin{cases} b^{ao} + \frac{2Z}{w(1+\beta Z)}(u_t - u^{ao}) & \text{if } u_t \geq u^{ao} - \frac{w(1+\beta Z)}{2Z} b^{ao} \\ 0 & \text{else} \end{cases} \\ U^{ao}(b_t) &= u^{ao} + \frac{w}{2}(1+\beta Z)(b_t - b^{ao}), \end{aligned}$$

and the equilibrium law of motion is, for any $u_0 > u^{ao} - w(1+\beta Z)/(2Z) \cdot b^{ao}$,

$$u_{t+1} = u^{ao} + Z(u_t - u^{ao}),$$

where $Z = (1-R)/[2(1+R)] \in [-1/2, 1/2]$. Given u_0, the economy converges to a unique steady state, $b^{ao} = 4R(1+\beta)/[(1+3R)(2+\beta)]$ and $u^{ao} = 1 - e(b^{ao}, b^{ao})$. Convergence is oscillating if $R > 1$, monotone if $R < 1$, and immediate if $R = 1$.

Proof. (sketch) As in the proof of Proposition 1, we start by guessing that B^{ao} has a linear form, i.e., $B^{ao}(u_t) = \alpha_0 + \alpha_1 u_t$. Then, using equation (9), we obtain:

$$B^{ao}(U^{ao}(b_t)) = \alpha_0 + \alpha_1 \frac{2 - (1+\beta(1-\alpha_0))\, w}{2 - \beta\alpha_1 w} + b_t \frac{\alpha_1 w}{2 - \beta\alpha_1 w}.$$

Plugging in this expression in the first-order condition of the maximization of (5) with respect to b_t, and solving for b_t, yields

$$b_t = \frac{(2 - \alpha_1 w\beta)\, R}{w(1+R)} - \frac{1}{w} + \frac{1 + \beta(1-\alpha_0)}{2} + \frac{1}{2}\frac{(1-R)(2-\beta\alpha_1 w)}{(1+R)\, w} u_t,$$

which is linear, as conjectured. Equating the unknown coefficients α_0 and α_1 and substituting in the resulting values into the guess $B^{ao}(u_t) = \alpha_0 + \alpha_1 u_t$, yields the equilibrium policy function $B^{ao}(u_t)$ provided in the text. Moreover, plugging in the values of α_0 and α_1 into equation (9), and simplifying terms, yields the equilibrium expression for $U^{ao}(b_t)$.

To complete the proof, it must be shown that if $R < R_{\max}$, the constraint $b \leq 1$ is never binding in equilibrium. This is necessary for the equilibrium policy function and the private decision rule to be linear and, hence, for the analytical characterization of the political equilibrium to be valid. Formally, $R_{\max}$ is defined as the R such that $B^{ao}(0) = 1$; see the technical appendix for further details. ■

Proposition 3 establishes that the dynamics of redistribution involve convergence to a unique steady state, characterized by a positive benefit rate, provided that some agents are risk averse ($R > 0$). Note that, in equilibrium, benefits inherit the same dynamics as u_t: $b_{t+1} = b^{ao} + Z(b_t - b^{ao})$. Steady-state benefits, b^{ao}, increase in risk aversion and in the share of workers, while they decrease in the wage rate since the distortionary effect of benefits increases with the return to effort.

The equilibrium policy function and the dynamics of u_t are depicted in Figure 3. As long as $R < 1$, dynamics are characterized by a positive root, implying monotone convergence. If instead $R > 1$, the benefit rate is a decreasing function of u_t and the root Z is negative, implying convergence following an oscillatory pattern. In the particular case where $R = 1$, convergence to the steady state occurs in one period.

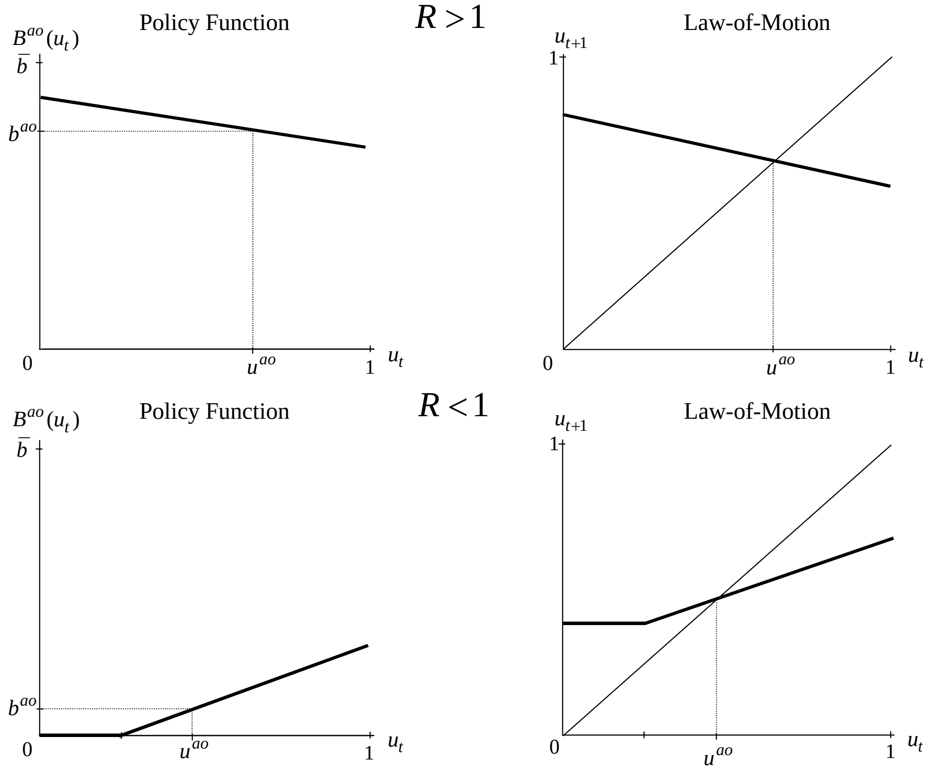

Figure 3. The equilibrium policy function $B(u_t)$ and u_t dynamics under risk aversion ($R \geq 0$).

The dynamics are characterized by two opposing forces. On the one hand, the larger is the current share of unsuccessful entrepreneurs, u_t, the higher is the tax cost (and, hence, the distortion) per unit of benefits. This is captured by the fact that the first term of (10) falls in u_t, which reflects the higher dependency ratio associated with a higher share of unsuccessful old entrepreneurs. We label this the *tax-base effect.* Through this effect, a higher u_t reduces the marginal (political) value of benefits, which tends to generate a negative relationship between b and u. On the other hand, the larger is u_t, the larger is the second term of (10), which reflects a stronger political pressure for redistribution since more individual entrepreneurs benefit from redistribution. We label this the *constituency effect* and note that a higher u_t increases the marginal political value of redistribution, which tends to generate a positive relationship between b and u. When aggregate risk aversion is low, the latter effect dominates, while the opposite is true when aggregate risk aversion is high. The reason for this is that when R is high, the political influence of the entrepreneurs diminishes as workers, on average, become more sensitive to the issue of redistribution, due to their higher individual risk aversion. Since intensity of preferences plays a key role in probabilistic voting models, this implies that the policy implemented in equilibrium more closely reflects the will of the average worker, namely attaining more redistribution. Consequently, the policy outcome becomes less sensitive to the share of *unsuccessful entrepreneurs* who want positive redistribution. Thus, the dynamics are dominated by the cost effect. In sum, higher aggregate risk aversion therefore *increases* steady state benefits and reinforces the *tax-base effect*.

Proposition 3 can be generalized to the case where the young participate in the political decision: $\omega \in [0,1]$. The equilibrium has the same form as that in Proposition 3. However, the expression for Z is complicated and we only state its main properties here.

Proposition 4 *Assume that $0 \leq R \leq R_{\max}$ and $\omega \in [0,1]$. The political equilibrium is then characterized as follows:*

$$\begin{aligned} B^a(u_t) &= \begin{cases} b^a + \frac{2Z}{w(1+\beta Z)}(u_t - u^a) & \text{if } u_t \geq u^a - \frac{w(1+\beta Z)}{2Z} b^a \\ 0 & \text{else} \end{cases} \\ U^a(b_t) &= u^a + \frac{w}{2}(1+\beta Z)(b_t - b^a), \end{aligned}$$

and the equilibrium law of motion is, for any $u_0 > u^a - w(1+\beta Z)/(2Z) \cdot b^a$,

$$u_{t+1} = u^a + Z(u_t - u^a),$$

where $Z \in (-4/7, 1/2]$ is decreasing in ω. Given u_0, the economy converges to a unique steady state, $b^a > 0$ and $u^a > 0$ (expressions in the technical appendix). Convergence is oscillating if $R > (1-\omega)/(1+\omega)$, monotone if $R < (1-\omega)/(1+\omega)$, and immediate if $R = (1-\omega)/(1+\omega)$.

We note that an increase in the political participation of the young decreases the slope of the policy function. In particular, the sign of the slope coefficient and whether the dynamics are oscillatory depend on whether $R \lesseqgtr (1-\omega)/(1+\omega)$. When the dynamics are monotone, the persistence of redistributive policies falls in R (it can be shown that $dZ/dR < 0$ when $Z > 0$). This condition nests the result of Proposition 3 that the policy function is upward- (downward-)sloping if and only if $R < 1$ ($R > 1$) when $\omega = 0$. If instead the young are as politically influential as the old ($\omega = 1$), the policy function becomes downward-sloping and the dynamics are oscillatory for any positive level of risk aversion.

As far as the young are concerned, both the cost effect and the intergenerational redistribution motive imply that benefits should be falling in u_t (recall that the larger is u_t, the larger is the transfer from the young to the old). Therefore, as the influence of the young increases, the intergenerational redistribution motive is mitigated. If $\omega < 1$, the old retain some political preponderance, and intergenerational transfers towards the old carry some weight in the political decision. If $\omega = 0$, however, this motive disappears and the dynamics of redistribution are determined by the cost effect alone. Since the tax-base effect implies a negative relation between benefits and the number of old unsuccessful entrepreneurs, a stronger influence of young voters reduces the slope of the policy function and tends to make dynamics oscillatory.

Unfortunately, due to the complicated expression for Z, we have not been able to sign the effect of an increase in the participation of the young on steady-state redistribution, although numerical analysis suggests that an increase in ω reduces redistribution in the long run.

3.4 Finite-horizon results

In this section we seek to answer two related questions.[11] First, we verify that the Markov-perfect equilibria derived above are indeed limits of finite-horizon equilibria. Second, and more substantially, we wish to find out whether there can be more than one finite-horizon equilibrium, i.e., whether there can be a role for "coordination", and perhaps "reputation", even in finite-horizon versions of this model. The uniqueness question is a substantial one not only formally, but in a very applied sense: it touches on the "stability" of government redistribution schemes, which was recently challenged in Hassler et al. (2003a). In that paper, a simple version of the present model with majority voting was shown to robustly produce multiple equilibria, independent of the time horizon. That is, a "belief in the welfare system" seemed necessary to support the system. For brevity, we will not cover all cases in this section; we concentrate on our baseline setup where only the old vote and $w = 1$. First, we will study the limit of the finite-horizon case and then discuss uniqueness.

We assume the economic environment to be identical to that of previous sections, except in a final period T where the newborn young make an effort investment but only live one period. In the finite-horizon economy, the equilibrium policy function will, in general, be time-dependent. For $t < T$, equilibrium condition 1 is thus modified to

Definition 2 $B^t(u_t) = \arg\max_{b_t \in [0,\bar{b}]} V(b_t, b_{t+1}, u_t)$ *subject to* $b_{t+1} = B^{t+1}(u_{t+1})$ *with* $u_{t+1} = 1 - e(b_t, b_{t+1})$.

Guessing preliminarily that $B^{t+1}(u)$ is linear, i.e., $B^{t+1}(u) = A_{t+1} + B_{t+1}u$, it is straightforward to show that an interior solution to the maximization problem at period t yields a linear policy function $B^t(u) = A_t + B_t u$. With $j \equiv T - t$ denoting the number of periods until the last date, the system

$$\begin{bmatrix} A_j \\ B_j \end{bmatrix} = \begin{bmatrix} -\frac{\beta}{2} & -\frac{R}{1+R}\beta \\ 0 & -\frac{1-R}{1+R}\frac{\beta}{2} \end{bmatrix} \begin{bmatrix} A_{j-1} \\ B_{j-1} \end{bmatrix} + \begin{bmatrix} \frac{2R}{1+R} - \frac{1-\beta}{2} \\ \frac{1-R}{1+R} \end{bmatrix}, \tag{11}$$

determines the coefficients A_j and B_j.[12]The solution to this linear system of difference equations is

$$\begin{aligned} A_j &= \left(-\frac{\beta}{2}\right)^j (A_0 - A) + \left(-\frac{\beta}{2}\right)^j \left(1 - \left(\frac{1-R}{1+R}\right)^j\right)(B_0 - B) + A, \\ B_j &= \left(-\frac{\beta}{2}\frac{1-R}{1+R}\right)^j (B_0 - B) + B, \end{aligned} \tag{12}$$

where A_0 and B_0 now define the policy rule in the last period ($t = T$). As this system is stable (the roots are $-\frac{\beta}{2}$and $-\frac{\beta}{2}\frac{1-R}{1+R}$), as $j \to \infty$, the coefficients of the policy rule converge to

$$\begin{aligned} A &\equiv \frac{8R}{(2+\beta)(2+\beta+R(2-\beta))} - \frac{1-\beta}{2+\beta}, \\ B &\equiv 2\frac{1-R}{2(1+R)+\beta(1-R)}, \end{aligned}$$

and the policy rule is identical in the limit to infinite horizon case detailed in Proposition 3 for the case $\omega = 0$ and $w = 1$.

[11]This section discusses methodological aspects, and the reader may decide to skip it without any loss of continuity.

[12]This equation follows from using $b_{t+1} = A_{j-1} + B_{j-1}u_{t+1}$ in the political objective function (10).

Next, we derive the final-period policy function, $B^T(u)$ and show that this is unique and linear, with $b_T = A_0 + B_0 u_T$, for all u_T in the reachable range $\left[\frac{1-\beta}{2}, 1\right]$. This provides the initial condition for the difference equation (12). Given the above result, it follows that the finite horizon equilibrium is unique.

We propose a parametrization of the final period which we regard as reasonable, although the argument does not hinge on the specific choice. In particular, we assume that the young born in the final period live one period only, and that they make an effort choice. To make the final period comparable to the previous ones, we compensate for the fact that the young in the last period obtain return on effort only in one period by scaling down their disutility of effort by $(1+\beta)^{-1}$. Thus, the effort cost is equal to $e^2/(1+\beta)$, implying that the optimal effort is equivalent to that which agents living two periods would have chosen had they faced the benefit level b_T in both periods. This optimal effort level is $e_T^* = (1-b_T)\frac{1+\beta}{2}$, implying that taxes in the final period are given by

$$\tau_T = \frac{1}{2}\left(2(1-\mu) + \mu u_T + \mu\left(\frac{1-\beta}{2} + b_T\frac{1+\beta}{2}\right)\right) b_T.$$

Substituting τ_T into the utility function yields the following political objective function:

$$V_T(b_T, u_T) = \mu((1-u_T) + u_T b_T) + ((1-\mu) b_T + R b_T) - (1+R)\tau_T.$$

Maximizing $V_T(b_T, u_T)$ with respect to b_T yields $b_T^* = A_0 + B_0 u_T$, with

$$\begin{aligned} A_0 &= -\frac{1}{2}\frac{1-\beta}{1+\beta} + \frac{2R}{(1+R)(1+\beta)}, \\ B_0 &= \frac{1}{1+\beta}\frac{1-R}{1+R}. \end{aligned}$$

For the sake of simplicity, the analysis has so far ignored the constraint that $b \in [0,1]$. Characterizing the sequence of policy functions when these constraints may bind is more complicated and the details of the results depend critically on the exact form of the effort function in the last period, about which we do not have strong prior information.[13]

3.5 Voting with benevolence toward future generations

In this section, we shall study the case where agents (the old) vote with some altruistic concern: they vote to maximize a welfare function which is a weighted average of their own felicity and the welfare of the next generation.[14] This case is interesting since it allows us to relax the assumption that agents have no concern for future generations, while retaining that the political mechanism

[13] That economies where the constraint that $b \in [0,1]$ never binds exist can easily be shown in some special cases, such as when $R = 1$. In this case, $B_j = B = 0$ for all $j \geq 0$, and, for all t,

$$B^t(u) = A\left(1 - \left(-\frac{\beta}{2}\right)^j \frac{\beta}{2(1+\beta)}\right) > 0,$$

since $A = (1 + \beta + 2s(1-\mu)/\mu)/(2+\beta)$. So the constraint is not binding when $R = 1$. By continuity, the same argument carries over for values of R sufficiently close to one. For more general values of the parameters, we also encountered no multiplicity, though some of our analysis here relies on numerical methods.

[14] Agents, however, do not display altruism in their private behavior. For instance, they continue to not insure their children against the ability shock. An alternative interpretation of the present setup is that of a time-consistent benevolent planner choosing policy with some weight on future generations.

lacks a commitment technology. In the next section, we will compare these results with the Ramsey allocation, where redistribution is set by a benevolent planner with access to a commitment technology. In particular, we will stress the different dynamics of redistribution in the two cases.

To derive a recursive formulation of the problem, we define the "weighted average felicity" across both young and old agents at time t as

$$\begin{aligned} F\left(u_t, b_t, b_{t+1}\right) \;&\equiv\; \beta\left(1-\mu\right) V_t^{ow} \\ &+\beta\mu\left(\left(1-u_t\right) V_t^{oes} + u_t V_t^{oeu}\right) + \lambda\left(F^y\left(e\left(b_t, b_{t+1}\right), b_t, \tau_t\right)\right), \end{aligned} \tag{13}$$

where

$$\begin{aligned} F^y\left(e\left(b_t, b_{t+1}\right), b_t, \tau_t\right) \;&\equiv\; \left(1-\mu\right) a\left(b_t - \tau_t\right) \\ &+\mu\left(e\left(b_t, b_{t+1}\right)\left(w - \tau_t\right) + \left(1 - e\left(b_t, b_{t+1}\right)\right)\left(b_t - \tau_t\right) - e\left(b_t, b_{t+1}\right)^2\right), \end{aligned}$$

and $\tau_t = \tau(b_t,\, b_{t+1}, u_t)$.

The parameter λ is a measure of intergenerational altruism. In particular, $\beta \in [0,1]$ and $\lambda \in [0,\beta]$ are the weights on the old and young currently alive. When $\beta = \lambda$, the old are "perfectly altruistic" and value equally their old-age felicity and that of the young. If, on the other hand, $\lambda = 0$, we obtain the case analyzed in Proposition 1. We restrict attention to economies where $\lambda \leq \beta$, so that altruism cannot "exceed 100%".

Since the problem is autonomous when policies are in the Markov class, a recursive formulation of the political objective can be written as

$$\begin{aligned} W\left(u_t\right) \;&\equiv\; \max_{b_t \in [0,1]} \left\{F\left(u_t, b_t, b_{t+1}\right) + \lambda W\left(u_{t+1}\right)\right\}, \\ &\quad \text{subject to} \\ b_{t+1} \;&=\; B\left(U\left(b_t\right)\right),\ u_{t+1} = U\left(b_t\right). \end{aligned} \tag{14}$$

In direct analogy with our equilibrium definition above, we provide

Definition 3 *A political equilibrium with altruistic voting is defined as a set of functions* $\langle B, U, W\rangle$, *where* $B : [0,1] \rightarrow [0,1]$ *is a public policy rule,* $b_t = B\left(u_t\right)$, $U : [0,1] \rightarrow [0,1]$ *is a private decision rule,* $u_{t+1} = U\left(b_t\right)$, *and* $W : [0,1] \rightarrow$ *is a value function.*

1. $B\left(u_t\right) = \arg\max_{b_t}\left\{F\left(u_t, b_t, b_{t+1}\right) + \lambda W\left(u_{t+1}\right)\right\}$ subject to $u_{t+1} = U\left(b_t\right),\ b_{t+1} = B\left(U\left(b_t\right)\right)$ and $b_t \in [0,1]$,
2. $U\left(b_t\right) = 1 - e\left(b_t, b_{t+1}\right)$ with $b_{t+1} = B\left(U\left(b_t\right)\right)$,
3. $W\left(.\right)$ satisfies the Bellman equation (14).

The following can then be established.

Proposition 5 *Assume that* $\lambda \leq \beta$ *(altruism, "al") and* $0 \leq R \leq R_{\max}$. *The political equilibrium with altruistic voting is characterized as follows:*

$$\begin{aligned} B^{alo}\left(u_t\right) \;&=\; \begin{cases} b^{alo} + \frac{2Z}{w(1+\beta Z)}\left(u_t - u^{alo}\right) & \text{if } u_t \geq u^{alo} - \frac{w(1+\beta Z)}{2Z} b^{alo} \\ 0 & \text{else} \end{cases} \\ U^{alo}\left(b_t\right) \;&=\; u^{alo} + \frac{w}{2}\left(1+\beta Z\right)\left(b_t - b^{alo}\right), \end{aligned}$$

and the equilibrium law of motion is, for any $u_0 > u^{alo} - w(1+\beta Z)/(2Z) \cdot b^{alo}$,

$$u_{t+1} = u^{alo} + Z\left(u_t - u^{alo}\right),$$

where $Z \in [-1/(2+\lambda), 1/2]$ is decreasing in R and λ. Given u_0, the economy converges to a unique steady state, $b^{alo} \leq b^{ao}$ and $u^{alo} \leq u^{ao}$ (expressions in the technical appendix). Convergence is oscillating if $R > (\beta - \lambda)/(\beta + \lambda)$, monotone if $R < (\beta - \lambda)/(\beta + \lambda)$, and immediate if $R = (\beta - \lambda)/(\beta + \lambda)$.

The proposition establishes that the slope coefficient of the policy function is decreasing in λ. Namely, the altruistic motive in the political equilibrium reduces the interest in intergenerational redistribution, and tends to make efficiency considerations more important (similarly to when we considered political participation of the young). This strengthens the tax-base effect, and tends to make dynamics oscillatory. In particular, if $\lambda = \beta$ and $R > 0$, the dynamics are oscillatory, indicating that the tax-base effect dominates the constituency effect. However, oscillations need not occur: for $\lambda < \beta$, and R sufficiently small, dynamics are monotone. In general, however, voters' altruism toward future generations is a force toward oscillatory dynamics.[15]

4 The Ramsey allocation with commitment

In this section, we show that optimal policy—to be precisely defined below—in the present model necessarily involves oscillations in taxation and redistribution. Specifically, we characterize the full commitment solution in the case where the planner's weights on generation t is β^t, i.e., the planner discounts the felicity of different generations at the same rate as that the private agents use for discounting their own felicity over time. In this particular case, the solution is relatively simple. We also briefly extend the analysis to the case where the planner discounts the future with a general factor $\lambda \leq \beta$. In the general case, however, the optimal tax sequence is more involved. Since the main focus of this paper is on political economy, we limit attention in this case to long-run properties.

4.1 Statement of the commitment problem

The choice set of the planner is the set of sequences of benefits, $\{b_t\}_{t=0}^{\infty}$, that are feasible for some sequence of taxes and associated private effort choices. We assume that the planner can commit to future benefits; we refer to this problem as the Ramsey problem and to its solution as the Ramsey allocation. The planner is assumed to be perfectly utilitarian when evaluating the utility of *ex-ante* identical agents. To simplify, we assume that she discounts future generations at a constant rate by attaching a weight λ^t to agents born at time t. In the analysis of the political allocation, we will assume balanced budgets in each period and, to compare the two allocations, we also impose this assumption on the Ramsey allocation.[16] The planner chooses the sequence of b_t $\forall t \geq 0$ in order to

[15] It is natural to expect that steady-state redistribution should be decreasing in λ, since more altruistic agents care more about the negative externality that current redistribution imposes on future generations via the future tax base. Although we have not been able to formally establish this relation, we have not found any numerical counterexample.

[16] Interestingly, it can be shown that when $\beta = \lambda$, the balanced budget restriction does not bind for the Ramsey planner, who would choose the same allocation also if she was allowed to accumulate debt or savings. See Hassler et al. (2004).

maximize

$$W(u_0) \equiv \beta(1-\mu) V_0^{ow} + \beta\mu(1-u_0) V_0^{oes} + \beta\mu u_0 V_0^{oeu} + \sum_{t=0}^{\infty} \lambda^{t+1} \left(\mu V_t^{ye} + (1-\mu) V_t^{yw}\right), \tag{15}$$

subject to

$$\begin{aligned} b_t &\in [0,1], \\ \tau_t &= \begin{cases} \tau(b_t, b_{t+1}, u_t), \text{for } t=0, \\ \tau(b_t, b_{t+1}, 1-e(b_{t-1}, b_t)), \text{for } t \geq 1, \end{cases} \\ e_t &= e(b_t, b_{t+1}). \end{aligned} \tag{16}$$

4.2 Characterizing the solution: a recursive formulation

The planner's problem, (15), does not admit a standard recursive formulation since its solution is time-inconsistent. Intuitively, the choice of b_{t+1} takes into account how the effort choice at t is influenced, but this effort choice is bygone when the time comes to implement b_{t+1}. It is well known that Ramsey problems admit a two-stage formulation whereby future decisions, in stage two, can be described as coming from a recursive problem with an additional state variable whereas the time-zero decisions, in stage one, can be derived from a "static" problem whose payoffs are given by the value function associated with the solution to the recursive problem.[17] In this framework, we will show that the second-stage recursive problem is particularly simple in that it involves one state variable only: *next period's* level of transfers. This result follows from the fact that since individuals live for two periods only, a benevolent planner who can commit for one period only chooses the same level of redistribution as a planner who could commit for all future periods. Specifically, if the planner in period t chooses b_{t+1}, she would have chosen the same b_{t+1} if she had had the ability to commit at any period $s < t$. Furthermore, although the flow of felicity in period t is affected by both the predetermined variables u_t and b_t, the optimal choice of b_{t+1} is only affected by b_t. Therefore, the recursive program only has b_t as a state variable, with b_{t+1} being the choice variable. As for the initial choice, the planner is not subject to earlier pre-commitments and thus, chooses b_0 and b_1 simultaneously. We thus prove that

Lemma 1 *The utilitarian planner program (15) is equivalent to the following recursive program:*

$$W(u_0) = \max_{b_0 \in [0,1]} \{Y_0(u_0, b_0) + V(b_0)\} \tag{17}$$

$$V(b_t) = \max_{b_{t+1} \in [0,1]} \{Y(b_t, b_{t+1}) + \lambda V(b_{t+1})\} \quad \textit{for } t \geq 0, \tag{18}$$

where $Y_0(u_0, b_0)$ is a linear-quadratic function (see proof) and

$$\begin{aligned} Y(b_t, b_{t+1}) = \left(\frac{\mu w^2}{4}\right) \cdot \big\{ & 2\left((1+\beta)(\beta+\lambda) R - \beta^2\right) b_t \\ & - \left((1+\beta)(\beta+\lambda)(R+1) - \lambda - 2\beta^2\right) b_t^2 \\ & - \left((\beta+\lambda)^2 (R+1) - 2\lambda\beta\right) b_t b_{t+1} \\ & + 2\lambda\beta^2 b_{t+1} - \beta^2 \lambda b_{t+1}^2 \big] + Q, \end{aligned}$$

[17] See, e.g., Marcet and Marimon (1999), where the additional state variable is marginal utility or the Lagrange multiplier associated with the incentive constraint.

where Q is a constant defined in the proof. Moreover, the mapping $\Gamma(v) = \max_{b' \in [0,1]} \{Y(b, b') + \lambda v(b')\}$ is a contraction mapping with V as the unique fixed point.

Proof. Consider the Ramsey-problem as formulated in (15). Now, define the planner's period t felicity, i.e., the "weighted average felicity" across young and old agents at time t for $t \geq 1$, as

$$\begin{aligned} F(b_{t-1}, b_t, b_{t+1}) &\equiv \beta(1-\mu) V_t^{ow} + \beta\mu \left(e_{t-1}^* V_t^{oes} + \left(1 - e_{t-1}^*\right) V_t^{oeu}\right) \\ &\quad + \lambda\mu \left(we_t^* + (1-e_t^*) b_t w - (e_t^*)^2 - \tau_t\right) + \lambda(1-\mu) a (b_t w - \tau_t), \end{aligned} \tag{19}$$

subject to $e_j^* = e(b_j, b_{j+1})$ and $\tau_t = \tau(b_t, b_{t+1}, 1 - e(b_{t-1}, b_t))$. Note that the function $F(b_{t-1}, b_t, b_{t+1})$ is additively separable in (b_{t-1}, b_t) and (b_t, b_{t+1}). More formally, there exist (linear-quadratic) functions G and H such that $F(b_{t-1}, b_t, b_{t+1}) = G(b_{t-1}, b_t) + H(b_t, b_{t+1})$, where

$$G(b_{t-1}, b_t) = -\frac{\mu w^2}{4} \left(\left((\beta+\lambda)(R+1) - 2\beta\right) b_t b_{t-1} + 2\beta b_{t-1}\right),$$

$$\begin{aligned} H(b_t, b_{t+1}) &= \left(\frac{w^2 \mu}{4}\right) \cdot \left[2\left((1+\beta)(\beta+\lambda)(R+1) - (1+2\beta)\beta - \lambda\right) b_t\right. \\ &\quad - \left((1+\beta)(\beta+\lambda)(R+1) - \lambda - 2\beta^2\right) b_t^2 \\ &\quad \left. + 2\lambda\beta^2 b_{t+1} - \lambda\beta^2 b_{t+1}^2 - \beta(\beta+\lambda)(R+1) b_t b_{t+1}\right] + Q, \end{aligned}$$

where $Q \equiv \mu(1+\beta)(2\beta + \lambda(1-\beta)) w^2/4$. Define now $Y(b_t, b_{t+1}) \equiv \lambda G(b_t, b_{t+1}) + H(b_t, b_{t+1})$. Using this and the definition of the function $F_0(u_0, b_0, b_1)$ from equation (13), the planner problem under commitment (15) can be expressed as

$$\begin{aligned} W(u_0) &= \max_{\{b_t\}_{t=0}^{\infty}} \left\{F_0(u_0, b_0, b_1) + \sum_{t=1}^{\infty} \lambda^t F(b_{t-1}, b_t, b_{t+1})\right\} \\ &= \max_{\{b_t\}_{t=0}^{\infty}} \left\{F_0(b_0, b_1, u_0) - H(b_0, b_1) + \sum_{t=0}^{\infty} \lambda^t \left(\lambda G(b_t, b_{t+1}) + H(b_t, b_{t+1})\right)\right\} \\ &= \max_{\{b_t\}_{t=0}^{\infty}} \left\{Y_0(u_0, b_0) + \sum_{t=0}^{\infty} \lambda^t Y(b_t, b_{t+1})\right\}, \end{aligned} \tag{20}$$

where $Y_0 \equiv F_0(b_0, b_1, u_0) - H(b_0, b_1)$ is given by

$$\begin{aligned} Y_0(b_0, u_0) &= \frac{\mu w}{2} \left(\left((\beta+\lambda)(R+1) - 2\beta\right)\left(1 - u_0 - \frac{w}{2}(1+\beta)\right) + w\beta^2\right) b_0 \\ &\quad + \frac{\mu w^2}{4} \beta \left((\beta+\lambda)(R+1) - 2\beta\right) b_0^2 + \mu\beta w \left(1 - u_0 - \frac{w}{2}(1+\beta)\right). \end{aligned}$$

Clearly, we can rewrite (20) as

$$\begin{aligned} W(u_0) &= \max_{b_0} \left\{Y_0(u_0, b_0) + \max_{\{b_t\}_{t=1}^{\infty}} \sum_{t=0}^{\infty} \lambda^t Y(b_t, b_{t+1})\right\} \\ &= \max_{b_0} \left\{Y_0(u_0, b_0) + \max_{b_1} \left\{Y(b_0, b_1) + \max_{\{b_t\}_{t=2}^{\infty}} \sum_{t=1}^{\infty} \lambda^t Y(b_t, b_{t+1})\right\}\right\} \end{aligned} \tag{21}$$

Defining the value function $V(b_t) \equiv \max_{\{b_{t+s}\}_{s=1}^{\infty}} \sum_{s=0}^{\infty} \lambda^s Y(b_{t+s}, b_{t+s+1})$, standard recursion on (21) yields the functional Bellman equation (18) for $t \geq 1$. Since Y is bounded by the fact that $b \in [0,1]$ and since $0 \leq \lambda < 1$, the Bellman equation (18) is a contraction mapping with a unique solution, which must also be the solution to the sequential continuation problem (Theorem 4.3 in Stokey and Lucas, 1989). Given V, it follows from the sequential formulation (20) that benefits in the initial period can be determined from the static problem (17). ■

The recursive formulation in Lemma 1 shows that the optimal policy can be represented in terms of two policy rules. The first rule, which sets the initial choice of redistribution, maps the initial (predetermined) proportion of unsuccessful entrepreneurs into initial choices of redistribution. The second rule applies from period one onwards and maps previous period's benefits into current benefits: $b_t = f(b_{t-1})$.

The policy rule $b_t = f(b_{t-1})$ is not globally linear, except in the case $\lambda = \beta$. This case is a natural benchmark, as it implies that the planner discounts future felicities at the same rate as that by which agents discount future within their life horizon. In this case, it is possible to attain a simple closed-form solution, summarized in the following Proposition (proof in the technical appendix).[18]

Proposition 6 *The optimal solution to the planner program (15) in the case $\lambda = \beta$ is*

$$b_t = b^p - (b_{t-1} - b^p), \forall t \geq 1,$$

and

$$b_0 = \left(1 + \frac{1-u_0}{(1-\beta)\frac{w}{2}}\right) b^p,$$

where

$$b^p \equiv \frac{R}{1+2R}.$$

As is clear from the proposition, the period-0 choice of the planner depends negatively on the initial condition u_0, which represents the size of the inelastic tax base at zero (investments in period minus one are sunk when the planner sets the period-0 benefits). As long as $u_0 < 1$, the planner chooses initial benefits larger than or equal to b^p, and then oscillate forever between this level and another level at the other side of b^p. This is a particular case of the more general analysis of capital taxation developed in Hassler et al. (2004), to which we refer the reader for a more detailed analysis.

When $\lambda < \beta$, the dynamics continue to be oscillatory, but they are explosive. Consider the definition of the function $Y(b_t, b_{t+1})$ given in Lemma 1. If the constraint $b_{t+1} \in (0,1)$ is not binding, the optimal allocation must satisfy the following first-order condition: $Y_2(b_t, b_{t+1}) + \lambda Y_1(b_{t+1}, b_{t+2}) = 0$. This follows from a standard envelope argument and the first-order condition on the Bellman equation (18). Calculating the derivatives and simplifying terms yields the following dynamic system:

$$\eta_0 + \eta_1 b_t + \eta_2 b_{t+1} + \lambda \eta_1 b_{t+2} \leq 0, \tag{22}$$

$$\begin{aligned}
\eta_0 &\equiv 2\lambda(1+\beta)(\beta+\lambda) R \geq 0 \\
\eta_1 &\equiv -\left((\beta+\lambda)^2 (R+1) - 2\lambda\beta\right) < 0 \\
\eta_2 &\equiv -2\lambda\left((1+\beta)(\beta+\lambda)(R+1) - \lambda - \beta^2\right) < 0
\end{aligned}$$

[18] The proof strategy is as follows. We start by guessing that the value function V is linear-quadratic in b, $V(b) = A_0 + A_1 b + A_2 b^2$, where A_0, A_1, and A_2 are unknown coefficients. We then compute the envelope condition and the first-order condition and use them to verify the initial guess for particular values of the coefficients. The resulting value function satisfies the functional equation (18) and hence, by the contraction mapping theorem, is the unique solution. The policy function implied by the first-order condition is as reported in the proposition. Solving for the initial policy b_0 is then straightforward.

This dynamic system is exactly linear, and it has the important property that the coefficients on b_t, b_{t+1}, and b_{t+2} are all negative.[19] Moreover, in the case where $\lambda < \beta$, both roots of the characteristic equation associated with (22) are smaller than minus one. Thus, the dynamics cannot converge to a steady-state. Based on numerical analysis, we conclude that benefits converge to a two-period cycle where the constraint $b \geq 0$ binds every second period.[20] Thus, the optimal plan never converges to a steady state, and in the long run, redistribution moves between zero and a positive level of redistribution.

It might seem surprising that the planner chooses an oscillating sequence, but it turns out that this minimizes the distortion associated with redistribution. As noted in the introduction, the reason is that the investments made by young agents in period t have an effect on the tax cost of redistribution both in period t and period $t+1$. More precisely, if benefits in period $t-1$ are large (small), young entrepreneurs will make a small (large) investment effort in that period. Thus, in period t, the old entrepreneurs will be relatively unsuccessful (successful), and there will be many (few) benefit recipients in that period. Therefore, the tax rate required to finance a certain benefit level will be relatively large (small). Since distortions to effort are convex in nature, redistribution is relatively costly (cheap) in period t, and the planner will set relatively small (large) benefits. Applying a similar logic for period $t+1$, it is clear why the optimal sequence of benefits might be oscillatory. Intuitively, the planner reduces the distortion of benefits in period t by choosing lower benefits in the next period, as the investment decision of the young in period t depends on redistribution both in period t and period $t+1$. When the planner sets benefits at a particular value, b_t, she takes into account the effects of this on felicity in $t-1$, t, and $t+1$. Therefore, the decision is both backward- and forward-looking. Without commitment, the backward-looking aspect disappears, having, as we will see below, qualitative consequences for redistribution dynamics.

In summary, this section has established that the Ramsey dynamics are oscillatory, and do not converge to a steady-state. This is in sharp contrast with the political equilibrium described in the previous sections. Even in the case where agents vote with altruism (section 3.5), the equilibrium always features convergence to a steady-state. Thus, an important conclusion of this paper is that the political mechanism dampens efficient fluctuations and, in some cases, even generates policy persistence (i.e., monotonic convergence).

5 Conclusion

Many political-economy questions call for analysis in dynamic settings. However, the literature lacks analytical frameworks where voting and economic decision making are both rational and forward-looking, and where the dynamic mechanisms are fully operative. In most of the existing

[19] The fact that all coefficients are negative follows from Y being strictly concave in each of its arguments separately, and from b_t and b_{t+1} displaying "substitutability". Intuitively, concavity follows from the convex cost function for effort and the fact that taxation is more costly on the margin, the higher is its level. The substitutability reflects the fact that effort depends on both b_t and b_{t+1}, so if one of these variables is high, the cost of increasing marginally the other is high.

[20] More precisely, in the long run the economy oscillates between zero and

$$\tilde{b} = \min\left\{\frac{R-1}{(1-\mu)R + \mu - \left(\frac{\lambda+\beta^2}{(1+\beta)(\beta+\lambda)}\right)}, 1\right\}.$$

We have also analyzed the case where $\lambda > \beta$, and have found that the dynamics then converge (as long as λ is not too large) in an oscillatory fashion to a steady-state. Details are available upon request.

papers with rational and optimizing agents, the dynamics are either muted by preferences that mimic myopia (under some conditions, logarithmic utility has this feature) or by a lack of dynamic decision variables (such as investment) that call for forward-looking expectations. Alternatively, the dynamics must be analyzed using numerical methods. In this paper, we have constructed a positive model where redistribution and social insurance take place in a dynamic setting: the taxation underlying these expenditures distorts human capital accumulation. In our economy, current taxation thus sets off nontrivial political and economic dynamics and the agents take these dynamics into account when making decisions. We are thus able to conduct both exercises of "comparative statics"—analyzing the effects of primitives on long-run outcomes—and of "comparative dynamics"—analyzing the effects of primitives on short-run outcomes. We focus on the case where the political system—a setting where policy decisions are made through probabilistic voting—cannot, either formally or through reputation effects, commit to future policy decisions.

The model is analytically tractable, making the mechanisms determining the dynamics transparent. Relative to "constrained-optimal" allocations, i.e., allocations which would result if a planner could set all taxes and transfers at time 0 to maximize some weighted utility of all agents, we find the political system to have a stabilizing role. In particular, the lack of commitment makes (optimal) oscillatory responses to disturbances become weaker or disappear. This effect can be qualitatively important: for a large range of parameter values, the constrained optimum prescribes limit cycles, whereas the political equilibrium never does.

We have identified two opposing mechanisms underlying the determination of redistribution and how it evolves over time. One of these mechanisms is the *constituency effect*. According to this effect, an increase in redistribution induces a change in individual actions, which in turn increase future demand for redistribution. Thus, the constituency effect tends to induce positive feedback, and therefore persistence, in the size of government.

We believe that such an effect may also operate in other areas of government activity. For example, an expansion of government employment may induce educational choices suited for government jobs and therefore make future reductions in investments politically costly; an effect stressed by, among others, Lindbeck (1995). In this paper, we have used probabilistic voting as the political aggregator of preferences. This voting model provides a smooth mapping from the distribution of preferences to political outcomes, which means that the constituency effect is smooth, operating over a large range of the domain of the state-variable. In particular, in this paper, the constituency effect generates persistence in the level of redistribution, but eventually redistribution always returns to a unique steady state.

In contrast to the smooth operation of probabilistic voting, Downsian majority voting may lead to abrupt changes in policy when the preferences of the median voter change. Therefore, the constituency effect can be stronger under majority voting than under probabilistic voting, not only leading to persistence but to complete hysteresis as in Hassler et al. (2003a), where a temporary shock to the demand for redistribution may lead to indefinitely high levels of redistribution. Our model therefore predicts that, *ceteris paribus*, countries with a political system closer in line with the smooth (discontinuous) preference aggregation of probabilistic voting (majority voting) should have weaker (stronger) policy persistence. Another difference between these institutions that follows from this logic and that we analyze in this paper is that majority voting can lead to expectational equilibria—beliefs that the government will (continue to) be large in the future can be self-fulfilling—whereas probabilistic voting cannot.

The second mechanism behind the dynamics of government we identify is the *tax-base effect*. According to this effect, positive redistribution today leads to higher future costs of redistribution, since higher levels of redistribution reduce investments and thereby shrink the future size of the tax base. Since higher costs of redistribution reduce the attractiveness and therefore the political viability of redistribution, the tax-base effect produces a negative feedback inducing oscillating dynamics. The tax-base effect is the only active channel underlying the constrained-optimal allocation.

We have showed that there are several factors that can strengthen the relative importance of the constituency and tax-base effects in our political equilibrium, thereby determining the extent to which equilibrium dynamics are persistent or oscillatory.

First, an increased political influence of individuals behind the veil of ignorance, or by young agents in our terminology, tends to increase the relative importance of the costs of redistribution and, thus, of the tax-base effect. This is a natural consequence of the fact that conflicts of redistribution strengthen with age when individuals are exposed to idiosyncratic shocks. Individuals with *ex-ante* coinciding interests on social insurance may later in life be divided into "losers" and "winners" from redistribution. As the political influence of the latter is diminished, the relative sizes of the winning and losing groups become less important as, instead, the common *ex-ante* interest wins the political upper hand. Furthermore, the distortionary costs of redistribution are partly borne in the future, since current redistribution reduces the future size of the tax base. This is a greater concern for young individuals with a longer remaining lifetime. However, whenever the *ex-post* interest is politically preponderant, a case which we deem to be the most likely, dynamics are monotone and redistribution persistent.

Second, more concern about the welfare of future generations also strengthens the tax-base effect. When the old voters are altruistic *vis-à-vis* the young, they appreciate the *ex-ante* interest and not only their own *ex-post* interest. Therefore, altruism *vis-à-vis* future generations tends to generate less persistence; in fact, when the old place the same weight on their own utility as on that of their offspring, dynamics are always oscillatory.

Third, an ability to commit future levels of redistribution strengthens the tax-base effect. Since future benefits distort current investment choices, thus increasing the current cost of redistribution, agents have an interest in curtailing future redistribution. This interest is particularly strong if current redistribution is chosen to be high. Therefore, commitment tends to induce a negative feedback and oscillating dynamics. In particular, when the discount factor on future cohorts equals the private intertemporal discount factor, dynamics are characterized by a unitary negative root, producing everlasting oscillations of constant amplitude.

Finally, we have found that higher risk aversion also strengthens the tax-base effect. In our model, we have separated the insurance value of redistribution from its distortive costs by assuming that individuals in need of redistribution do not make choices distorted by redistribution. In a more general setting, an increase in risk-aversion might also affect the marginal utility of unsuccessful agents whose investment decisions are sensitive to the amount of redistribution. In such a case, the constituency effect should also be strengthened by higher risk aversion, making the total effect on dynamics ambiguous.

In conclusion, we have identified several parameters affecting the dynamics of redistribution. We leave the task of confronting the model's predictions to data to future research.

References

Alesina, A., Roubini, N. and Cohen, G. (1997), *Political Cycles and the Macroeconomy*, MIT Press, Cambridge.

Barro, R. J. (1979), 'On the determination of the public debt', *Journal of Political Economy* **87**(5), 940–71.

Bernheim, D. and Nataraj, S. (2002), A solution concept for majority rule in dynamic settings, Mimeo, Stanford University.

Cohen, D. and Michel, P. (1988), 'How should control theory be used to calculate a time-consistent government policy?', *The Review of Economic Studies* **55**(2), 263–74.

Di Tella, R. and MacCulloch, R. (2002), 'The determination of unemployment benefits', *Journal of Labor Economics* **20**(2), 404–434.

Hassler, J., Krusell, P., Storesletten, K. and Zilibotti, F. (2004), On the optimal timing of capital taxes, Mimeo, Stockholm University.

Hassler, J., Rodríguez Mora, J. V., Storesletten, K. and Zilibotti, F. (2003a), 'The survival of the welfare state', *American Economic Review* **93**(1), 87–112.

Hassler, J., Storesletten, K. and Zilibotti, F. (2003b), Democratic public good provision, CEPR Discussion Paper, no. 4296.

Keane, M. P. and Wolpin, K. I. (1997), 'The career decisions of young men', *Journal of Political Economy* **105**, 473–522.

Krusell, P. and Ríos-Rull, J. V. (1999), 'On the size of the U.S. government: Political economy in the neoclassical growth model', *American Economic Review* **89**(5), 1156–81.

Lindbeck, A. (1995), 'Hazardous welfare-state dynamics', *American Economic Review* **85**(2), 9–15.

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

Marcet, A. and Marimon, R. (1999), Recursive contracts, Mimeo, Universitat Pompeu Fabra.

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.

Razin, A., Sadka, E. and Swagel, P. (2002), 'The ageing population and the size of the welfare state', *Journal of Political Economy* **110**(4), 900–918.

Sargent, T. and Hansen, L. P. (2004), Recursive models of dynamic linear economies, Mimeo, New York University.

Stokey, N. L. and Lucas, R. E. (1989), *Recursive Methods in Economic Dynamics*, Harvard University Press.