

The Dynamics of Government.

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

Preliminary and incomplete.

Do not distribute without authors' permission.

Current draft: March 2002.

Abstract

What is the long-run size of the transfer system, and what does the path there look like? We model income redistribution as determined by voting among consumers of different types and income realizations. Taxation is distortionary because it discourages effort. Voters are fully rational: they realize that transfers do not come for free, and that they have implications also for future economic decisions and taxation outcomes. In our economy, our politically driven redistribution provides insurance, and we investigate to what extent the democratic process provides it appropriately. The general finding is that there tends to be too much redistribution in equilibrium. This is due to (i) a lack of political representation of future generations in current decisions and (ii) a lack of commitment in the democratic political system: current voters are not allowed to bind the hands of future voters.

^{*}John Hassler: IIES and CEPR; e-mail: John.Hassler@iies.su.se.

[†]Per Krusell: University of Rochester, IIES, and CEPR; e-mail: pekr@troi.cc.rochester.edu.

[‡]Kjetil Storesletten: IIES and CEPR; e-mail: Kjetil.Storesletten@iies.su.se.

[§]Fabrizio Zilibotti: IIES, University of Southampton, CERGE-EI, and CEPR; e-mail: Fabrizio.Zilibotti@iies.su.se.

1 Introduction

The theoretical literature on the dynamic evolution of the size of government is still very limited. Basic questions such as what factors (preferences, technology, political institutions, etc.) are important in determining the long-run size of the welfare state, its shorter-run dynamics and its welfare properties are hardly addressed at all on a theoretical level. These questions are, in our opinion, among the most important ones in political economy, or even public economics more generally. The variation in outcomes across different countries, and the changes that given countries experience over time, are substantial, and arguably of first order for understanding the welfare of citizens. Empirical analyses are helpful for beginning to understand outcomes, but it is hard to see how fundamental insights could be derived without a parallel development of theory. This is of course especially true for normative questions regarding what institutions are amenable to good outcomes.

In this paper, we construct a positive theory of redistributive government activity. In our model, government activity provides agents with insurance in a world where missing markets makes it impossible for agents to insure against income risk. The size of redistribution programs is determined as the political resolution of an ex-post conflict of interests between agents whose uncertainty is already realized: some are net benefactors from redistribution while others are net beneficiaries. More specifically, our economy is populated by overlapping generations of two-period-lived agents. In each generation, there are two types of agents: *entrepreneurs* and *workers*. Within each of these groups, all agents are ex-ante identical, but there is ex-post heterogeneity: some receive high income and some receive low income. Moreover, entrepreneurs differ from workers in two ways. First, entrepreneurs are risk-neutral whereas workers are risk-averse. Second, entrepreneurs are subject to moral hazard problems, as their income realization depends on some costly investment—effort—which is expended in the first period of their lives. Workers, on the other hand, make no private decision, and their success in the labor market is entirely determined by luck (or unobserved ability). These assumptions are stark, and they do reflect a need for tractability. However, we believe that they also provide a reasonable shortcut description of important real-world features: (i) in terms of its effect on productivity, the effort of some workers is more important than the effort of others; and (ii) it is likely that those agents whose effort matters are more well-endowed with wealth and therefore de facto less vulnerable to shocks.

Redistributive programs simply transfer income from richer to poorer agents, irrespective of types: we assume that the government cannot distinguish workers from entrepreneurs, so that benefits cannot be made type-dependent. In equilibrium, for workers as well as for entrepreneurs, rich agents oppose transfers, whereas poor agents support them. However, entrepreneurs would be against redistribution ex ante: behind the veil of ignorance—before their income is realized—all entrepreneurs dislike redistribution, since they are risk neutral and redistribution is distortionary. In contrast, workers would value insurance and support insurance even in ex-ante terms.

We focus on a probabilistic voting setup, as studied in Lindbeck and Weibull (1987). Compared to a majority-voting framework, this mechanism provides smoothness that is

analytically desirable: policy outcomes are based on maximizing “aggregate preferences” that are a smooth function—a weighted average in our case—of the utility levels of each consumer in the population. The effective weight associated to each group depends on both its size and its relative “preference intensity”—there is an ideology dimension, too, which plays no direct economic role but has implications for how easy a voter is to attract with economic policy.

In a society where all agents are risk-neutral (all agents are entrepreneurs) and where the distribution of ideology preferences is the same in all cohorts and independent of income realization, only group size matters. Throughout the paper we mainly focus on the case where the old agents have a stronger influence on outcomes than do young agents: they are easier to capture with competition in policy space than are younger agents. For this case, we show that the political equilibrium cannot sustain redistributive transfers in the long run: the long-run size of government is zero. However, in the short run—on the way toward the zero-government steady state—there is redistribution across cohorts: old consumers tax young consumers. We also demonstrate that this outcome is inefficient: no matter what weight a “social planner” would attach to the utility of different agents, any optimal tax sequence would be zero after the first period. The reason for this result is that the planner would (we assume) be able to choose future tax rates taking into account how they affect current effort decisions of entrepreneurs. In contrast, a democratic voting mechanism fundamentally respects future voters: current voters cannot bind future policy decisions. Thus, when a vote on taxes is taken, effort decisions in the past are bygones. Thus, it is not possible to implement the optimal allocation with our political mechanism: our political equilibrium is time-consistent—in the sense that nobody is surprised on the equilibrium path—but inefficient.

When we consider workers jointly with entrepreneurs, the results change qualitatively. The voting equilibrium now generally features positive redistribution in the long run. The reason why there is redistribution is that one group—workers who were unlucky and thus have higher marginal utility of income—has a higher value to candidates in the campaign: they are easier to sway with promises of transfers. This result would also hold even if workers were not subject to risk, but merely responded more to policy, such as if they were less ideological than entrepreneurs. The long-run size of transfers as well as the dynamics of redistribution hinge critically on how intensely the unemployed workers feel (i.e., on the marginal utility of unlucky workers or, expressed alternatively, on their ex-ante risk aversion). When preferences are less intense (i.e., workers are close to risk neutral), the dynamics of transfers are monotone, and long-run transfers small in relative terms. Conversely, when preferences are relatively intense, the steady state features higher transfers and the dynamic path toward the steady state is oscillatory.

The tendency for our economy to generate oscillatory outcomes—a welfare state that moves in waves toward a steady state—is noteworthy because it suggests interesting dynamics in a richer model: we do not consider aggregate shocks here, but it seems likely that the present model would predict cyclical responses to such shocks. The exact nature of the implied cycles—and how they would relate to politico-economic data—cannot be usefully ascertained without expanding on the setup, especially in terms of our generational struc-

ture and the nature of individual shocks. However, we believe that more realistic cohort descriptions would inherit the tendency for cyclical responses. The intuition behind why cycles can occur is simple. In a situation in which there are many unsuccessful agents (entrepreneurs), it is costly to provide a dollar of transfer—there are few benefactors and many beneficiaries, leading to a large effort distortion per dollar. This means that transfers in such a situation will be low, and that effort will be (relatively) high. Since effort has a permanent component (we assume it has no temporary component for simplicity), it will lead to a small group of unsuccessful entrepreneurs next period. With few unsuccessful entrepreneurs, transfers are cheap to provide, and equilibrium dictates high transfers, leading back to a large number of unsuccessful agents in the following period, and so on. The oscillation dies out over time.

Do the politically generated cycles reflect suboptimal allocation of resources? Yes and no. When we consider a planning outcome again—with commitment to future policy—redistribution is overall lower than with our democratic mechanism. However, it fluctuates more! That is, the cyclical nature of policy is natural and desirable in our framework, and the political mechanism actually limits it. For some parameter values—when the planner puts sufficiently different weights on the old and the young—the oscillations are even exploding, leading to long-run allocations which are cyclical. Benefits have a natural lower bound (equal to zero), however, so that means that benefits cycle between 0 and a positive number in the long run in this parameter range. The reason for the cycle here is the same as explained above, but when commitment to future tax rates is possible it is even less costly to cycle: a high current tax can be counterbalanced with a low future tax, thus not distorting effort so much, since effort has a permanent effect and thus depends on future taxes.

Another reason why the planning outcome gives rise to lower transfers than does the political system is that the latter does not take the utility of future agents into account. In particular, as the current electorate votes for a high current transfer, the number of unsuccessful will be high tomorrow and therefore raise the cost of redistribution for future agents. A benevolent planner—one who takes future consumers into account—would internalize this channel. We consider an extension where the current voters are benevolent vis-a-vis future agents, and we show that it leads to higher redistribution. However, it does not generate exploding cycles as does the allocation with commitment. Thus, not being able to overcome the commitment problem in the political mechanism has qualitative importance in our model: it (suboptimally) dampens policy cycles.

On a purely methodological level, the main reason why the theoretical literature on the dynamics of government is scant is a lack of convenient analytical tools. Economic dynamics are perhaps hard, but there is a large body of work on the subject: for a given policy environment, it is textbook material how to analyze an economy's behavior over time when the economic actors are fully rational. Similarly, pure political theory has worked on dynamic policy determination (although this literature is less well developed). 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 (usually competitive) equilibrium

theory to ensure that all economic agents—consumers, firms and government—maximize their respective objective functions under rational expectations, and resource constraints. Prior to Hassler et al. (2001), —henceforth HRSZ—the only nontrivial dynamic models (that is, that are not repeated static frameworks or purely “backward-looking” setups) relied essentially on numerical solution (see, e.g., Krusell and Ríos-Rull (1999) and the discussions therein). HRSZ provided a tractable framework where voters are influenced both by the state of the economy—in this case, the current income distribution—and foresee effects of the current policy outcomes on both future income distributions and future votes, which they care about, and therefore take into account when voting currently. The present paper builds further on the HRSZ paper by introducing risk aversion and hence, a natural role for government as the provider of insurance. The extension is quite nontrivial but, with specific restrictions on the nature of the insurance problem—who is risk-averse and the form of the distortions created by welfare payments—it is possible to maintain tractability.

The paper is organized as follows. Section 2 presents the model and the equilibrium concept. Section 3 focuses on risk neutrality and characterizes both political outcomes and Ramsey optimal outcomes. The risk neutrality simplification makes the analysis more tractable and serves as a useful contrast to the main case, involving risk aversion, studied in section 4. Section 5 concludes.

2 The model

The model economy has a continuum of two-period lived agents. Each generation consists of two types of ex-ante different agents: *entrepreneurs*, whose proportion in the population is μ , and *workers*, whose proportion is $1 - \mu$. Entrepreneurs are assumed to be risk-neutral, whereas workers are risk averse. All agents of a given type are born identical, but the subsequent earnings are stochastic. “Successful” agents earn a high wage, $w \in (0, 1]$, in both periods of their life, whereas “unsuccessful” agents earn a low wage, normalized to zero. At birth, entrepreneurs undertake a costly investment, increasing the probability of subsequent success. The cost of investment, which can be interpreted as the disutility of educational effort, is e^2 , where e is the probability of success. The probability of success of workers is entirely determined by luck. In particular, a worker becomes successful with probability s . The law of large number holds, so that the fractions of successful entrepreneurs and workers become e and s , respectively.

Workers’ ex-ante preferences are parameterized by the following utility function

$$V_t^{yw} = E_t[u(c_t) + \beta u(c_{t+1})],$$

where y stands for young, w stands for worker, $\beta \in [0, 1]$ is the discount factor and

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

Thus, the marginal utility of the workers drops discretely at the threshold consumption level x and is constant everywhere else. The kink in preferences allows us to maintain

analytical tractability while allowing risk aversion. The entrepreneur's utility function is linear in income and quadratic in effort: $V_t^{ye} = E_t[c_t + \beta c_{t+1}] - e^2$.

The government is assumed to redistribute income from successful to unsuccessful agents. In particular, in each period, a transfer $bw \in [0, \bar{b}w]$ to each low-income agent is determined. The transfer is financed by collecting a lump-sum tax τ , and the government budget is assumed to balance every period. The transfer, with its associated tax rate, is determined before the young entrepreneurs decide on their investment. By assumption, we rule out age- and type-dependent taxes and transfers.

For tractability, we impose a joint constraint on the threshold x and on the range of admissible benefits such that the consumption of the unsuccessful cannot exceed the threshold x , while the consumption of the successful cannot fall below x . More formally, we set an upper bound on benefits, \bar{b} and define x in terms of \bar{b} .¹ The assumptions about the range of admissible b 's and the kink of the utility function insure that for any $b \in [0, \bar{b}]$, the marginal utility of successful workers be one, whereas the marginal utility of unsuccessful workers be $a \geq 1$.

We further assume throughout that $(1+r)^{-1} = \beta$, where r is the net interest rate. Under this assumption, the savings decisions can be abstracted from. We also assume that there is no private insurance market available for the workers. Thus, social insurance has no competition by assumption. In terms of government policy variables and private decisions, the utilities of the agents alive at time t are given as follows:

$$\begin{aligned}\tilde{V}^{oes}(b_t, \tau_t) &= w - \tau_t \\ \tilde{V}^{oeu}(b_t, \tau_t) &= b_t w - \tau_t \\ \tilde{V}^{ye}(e_t, b_t, b_{t+1}, \tau_t, \tau_{t+1}) &= e_t(1+\beta)w + (1-e_t)(b_t + \beta b_{t+1})w - e_t^2 - \tau_t - \beta\tau_{t+1} \\ \tilde{V}^{ows}(b_t, \tau_t) &= w - \tau_t \\ \tilde{V}^{owu}(b_t, \tau_t) &= a(b_t w - \tau_t) - (a-1)x \\ \tilde{V}^{yw}(b_t, b_{t+1}, \tau_t, \tau_{t+1}) &= s(1+\beta)w + (1-s)a(b_t + \beta b_{t+1})w \\ &\quad - (s + (1-s)a)(\tau_t + \beta\tau_{t+1}) - (1-s)(a-1)(1+\beta)x\end{aligned}$$

where the superscripts s and u denote successful and unsuccessful, respectively. Note that the utility of young agents is computed prior to individual success or failure.

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

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

Since agents are *ex-ante* identical, entrepreneurs of the same cohort choose the same level of effort. This implies that the proportion of old unsuccessful entrepreneurs in period $t+1$ is given by $u_{t+1} = 1 - e_t(b_t, b_{t+1})$.

¹We set $\bar{b} \equiv \frac{1}{\mu w \beta} \left(\sqrt{((\mu+2)^2 + 4\mu w \beta)} - (\mu+2) \right) < 1$ and set $x = \bar{b}w - ((1-s)(1-\mu) + \frac{\mu}{2}(1+(1+\beta)\frac{w}{2} + \bar{b}\frac{w}{2}))\bar{b}w$. We show later that these constraints jointly suffice.

The government budget constraint yields $2\tau_t = (2(1-s)(1-\mu) + \mu u_t + \mu(1-e_t^*))wb_t$: total tax receipts equal expenditures on the two generations of unsuccessful workers, $2(1-s)(1-\mu)$, on the old unsuccessful entrepreneurs, μu_t , and on the young unsuccessful entrepreneurs, $\mu(1-e_t^*)$. Hence we have

$$\tau_t = \tau(b_t, b_{t+1}, u_t) = \left((1-s)(1-\mu) + \frac{\mu}{2} \left(u_t + 1 - (1+\beta) \frac{w}{2} + (b_t + \beta b_{t+1}) \frac{w}{2} \right) \right) b_t w. \quad (1)$$

The marginal tax cost of redistribution in period t , $\frac{\partial \tau_t}{\partial b_t}$, 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} .

We can now revisit the constraints on x and \bar{b} . The values of x and \bar{b} are set so that $x = \bar{b}w - \tau(\bar{b}, 0, 0) = w - \tau(\bar{b}, \bar{b}, 0)$. These restrictions are sufficient to ensure that x is an upper bound of the consumption of an unsuccessful agent and a lower bound of the consumption of a successful agent. They imply that the marginal utility of agents belonging to each group is constant under all feasible redistribution policies. We will later impose further conditions on risk aversion a , ensuring that the upper constraint \bar{b} will never be binding in a political equilibrium.

All old successful agents prefer zero benefits, since redistribution implies positive taxes without providing them with any benefits. The old unsuccessful agents, in contrast, are better off with some redistribution, even though their preferences for redistribution may be non-monotonic, as the marginal cost of redistribution is increasing. Concerning the preferences of the young, note that benefits may lead to some intergenerational redistribution between the old to the young depending on the proportion of old unsuccessful agents relative to that of young unsuccessful agents. Additionally, young workers value redistribution for its insurance property, which is instead worthless to young entrepreneurs. In particular, in steady-state, the young entrepreneurs do not value redistribution and are only hurt by its distortionary effect, whereas young workers trade off the insurance value of redistribution versus its cost in terms of productive efficiency.

For later use, it is convenient to define

$$\begin{aligned} & V^{ye}(b_t, b_{t+1}, b_{t+2}, u_t, u_{t+1}) & (2) \\ & = \tilde{V}^{ye}(e(b_t, b_{t+1}), b_t, b_{t+1}, \tau(b_t, b_{t+1}, u_t), \tau(b_{t+1}, b_{t+2}, u_{t+1})), \\ & V^{yw}(b_t, b_{t+1}, b_{t+2}, u_t, u_{t+1}) \\ & = \tilde{V}^{yw}(b_t, b_{t+1}, \tau(b_t, b_{t+1}, u_t), \tau(b_{t+1}, b_{t+2}, u_{t+1})), \\ V^{oj}(b_t, b_{t+1}, u_t) & = \tilde{V}^{oj}(b_t, \tau(b_t, b_{t+1}, u_t)), \quad \text{for } j \in \{es, eu, ws, wu\}. \end{aligned}$$

2.1 Political equilibrium

The benefit policy is chosen in each period through a voting mechanism. In the benchmark case, we assume that agents vote at the end of each period, after all uncertainty about individual earnings has been realized, over the redistribution policy next period. 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. We later extend the analysis to the case in which both the young and the old vote on current benefits. In this case, the young vote behind the veil of ignorance about their wage realization. We refer to the latter as the general case.

We assume a two-candidate political model of probabilistic voting a la Lindbeck-Weibull (1987) and restrict attention to Markov-perfect equilibria. Moreover, we require that these equilibria be smooth in the aggregate state variable: the proportion of current unsuccessful old agents (u_t). This restriction should be viewed as a vehicle for ruling out trigger strategies in voting behavior, thus instead focusing on the limit of the associated finite-horizon equilibria.

While a formal discussion of the political model is deferred to an appendix (to be included), we recall here the main features of this model. Agents cast their votes on one of two candidates, who maximize their probability of becoming elected. Individuals 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. We refer to this additional dimension as “ideology” Persson and Tabellini (2000). 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.²

In this model, both candidates set, in equilibrium, the same platform over redistribution, and each has a fifty percent probability of winning. More importantly, the impact of each group on the economic policy outcome in equilibrium increases the less dispersed are ideology preferences in this group or, equivalently, the less concerned the voters in this group are with the ideological dimension relative to the economic policy variables. Intuitively, if agents in a group have a lower concern for ideology, a candidate making a small change in redistribution in favor of this 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. In our model, furthermore, it is possible to show that in equilibrium, candidates simply maximize a weighted sum of individuals’ utilities, where the weights are determined by the relative intensity for ideology; if ideology intensity/the distribution of ideology intensity is the same in two groups, their relative weights are one. Thus, we can work with a “political aggregator”—a function which current policy choice must maximize—represented by the weighted utility of our groups of voters, with the weights being population shares. This result is important for tractability.

The political equilibrium is defined as follows.

Definition 1 *A probabilistic-voting political equilibrium is defined as a pair of functions $\langle B, U \rangle$, where $B : [0, 1] \rightarrow [0, 1]$ is a public policy rule, $b_t = B(u_t)$, and $U : [0, 1] \rightarrow [0, 1]$ is*

²Since candidates have no intrinsic preferences over redistribution, they are assumed to implement their promised platform over redistribution.

a private decision rule, $u_{t+1} = 1 - e_t^* = U(b_t)$, such that the following functional equations hold:

A) *Benchmark case (only the old vote):*

1. $B(u_t) = \arg \max_{b_t} V(b_t, b_{t+1}, u_t)$ subject to $b_{t+1} = B(U(b_t))$, and $b_t \in [0, 1]$,
2. $U(b_t) = 1 - (1 + \beta - (b_t + \beta b_{t+1})) \cdot w/2$, with $b_{t+1} = B(U(b_t))$,

where

$$V(b_t, b_{t+1}, u_t) \equiv (1 - u_t) \cdot \mu V^{oeu}(b_t, b_{t+1}, u_t) + u_t \cdot \mu V^{oes}(b_t, b_{t+1}, u_t) + \quad (3)$$

$$(1 - s) \cdot (1 - \mu) V^{owu}(b_t, b_{t+1}, u_t) + s \cdot (1 - \mu) V^{ows}(b_t, b_{t+1}, u_t) \quad (4)$$

B) *General case (all agents vote, with a weight $\omega \in (0, 1]$ on each young):*

1. $B(u_t) = \arg \max_{b_t} V_g(b_t, b_{t+1}, b_{t+2}, u_t, u_{t+1})$ subject to $u_{t+1} = U(b_t)$, $b_{t+2} = B(U(B(U(b_t))))$, $b_{t+1} = B(U(b_t))$, and $b_t \in [0, 1]$,
2. $U(b_t) = 1 - (1 + \beta - (b_t + \beta b_{t+1})) \cdot w/2$, with $b_{t+1} = B(U(b_t))$,

where

$$V_g(b_t, b_{t+1}, b_{t+2}, u_t, u_{t+1}) \equiv V(b_t, b_{t+1}, u_t)$$

$$+ \omega (\mu V^{ye}(b_t, b_{t+1}, b_{t+2}, u_t, u_{t+1}))$$

$$+ (1 - \mu) V^{yw}(b_t, b_{t+1}, b_{t+2}, u_t, u_{t+1}).$$

The first equilibrium condition requires that the political mechanism chooses 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 functional equation (1). In other words, suppose that agents believe that future benefits are set according to the function $b_{t+j} = B(u_{t+j})$. Then we require that the same function $B(u_t)$ define optimal benefits today.

The second equilibrium condition states that all young individuals choose their investment optimally, given b_t and b_{t+1} and that agents hold rational expectations about future benefits and distributions of types. In general, U might 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, we focus on equilibria where the equilibrium investment choice of the young is fully determined by the current benefit level.

The function V entails the assumption that all agents within a given generation exert the same political influence, irrespectively of their type. In the general case where two generations participate into 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 the old care less, on average, 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. Clearly, the general case encompasses the benchmark case, and the two are identical when $\omega = 0$.

3 An economy where all agents are risk-neutral

In this section, we consider the particular case when $a = 1$, i.e., when all agents in the society are risk-neutral and have the same preference intensity for economic policy. Here, the welfare state clearly entails no insurance value. This section sets the stage for the risk-aversion case discussed in section 4.

3.1 Political equilibrium

We start by analyzing the benchmark case ($\omega = 0$): only the old agents vote. Substituting the discounted value functions (2) into the political objective function (3), we have

$$\begin{aligned}
& \mu u_t (b_t w - \tau_t) + \mu (1 - u_t) (w - \tau_t) + (1 - \mu) s (w - \tau_t) + (1 - \mu) (1 - s) (b_t w - \tau_t) \\
= & \mu u_t b_t w + (1 - \mu) (1 - s) b_t w + [\mu (1 - u_t) w + (1 - \mu) s w] - \tau_t \\
= & \mu u_t b_t w + (1 - \mu) (1 - s) b_t w + [\mu (1 - u_t) w + (1 - \mu) s w] \\
& - \left((1 - \mu) (1 - s) + \frac{\mu}{2} u_t + \frac{\mu}{2} (1 - e_t^*) \right) b_t w \\
= & \frac{\mu}{2} \{ u_t b_t w - (1 - e_t^*) b_t w \} + [\mu (1 - u_t) w + (1 - \mu) s w].
\end{aligned}$$

Notice that the expression in brackets to the right in the last expression is exogenous from the perspective of the voter: it is predetermined. Therefore, by omitting the exogenous term and rescaling, the objective function to be maximized by the two political candidates can be expressed simply as

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

Equivalently, this expression can 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. The latter, of course, 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.

The following proposition characterizes the equilibrium in the benchmark case.³

³Proofs are, with the exception of that of proposition 1, contained in the appendix. However, the proofs of propositions 2 and 4 are similar to the proof of proposition 1.

Proposition 1 Assume $\omega = 0$ (only the old vote) and $a = 1$. The PV political equilibrium, $\langle B^{pv}, U^{pv} \rangle$, is characterized as follows:

$$B^{pv}(u_t) = \begin{cases} \frac{2}{w(2+\beta)}(u_t - u^{pv}) & \text{if } u_t \geq u^{pv} \\ 0 & \text{if } u_t < u^{pv} \end{cases}$$

$$U^{pv}(b_t) = u^{pv} + b_t \frac{w}{2} \left(1 + \frac{\beta}{2}\right),$$

where $u^{pv} \equiv 1 - \frac{w}{2}(1 + \beta)$. Given any u_0 , the economy converges monotonically to a unique steady-state, such that $b = b^{pv} \equiv 0$ and $u = u^{pv}$ following the equilibrium law of motion;

$$u_{t+1} = \begin{cases} u^{pv} + \frac{1}{2}(u_t - u^{pv}) & \text{if } u_t \geq u^{pv} \\ u^{pv} & \text{if } u_t < u^{pv}. \end{cases}$$

Proof. As explained above, when $a = 1$ and $\omega = 0$, the political objective function can be written as

$$V(b_t, b_{t+1}, u_t) = u_t b_t w - \left(1 - (1 + \beta) \frac{w}{2} + \frac{w}{2} (b_t + \beta b_{t+1})\right) b_t w. \quad (5)$$

The first-order condition, when b_{t+1} is written in terms of its equilibrium dependence on b_t , is

$$0 = u_t w - \left(1 - (1 + \beta) \frac{w}{2} + \frac{w}{2} (b_t + \beta B(U(b_t)))\right) w - (1 + \beta B'(U(b_t)) U'(b_t)) b_t \frac{w^2}{2}.$$

This equation defines b_t as a function of u_t , given perceptions about how future benefits are set.⁴ We thus need to find two functions $B^{pv}(u_t)$ and $U^{pv}(b_t)$, satisfying the two equilibrium conditions in Definition 1, i.e., satisfying this first-order condition and the equation determining optimal effort. These two equations are, in general, a (differential, since the derivatives of the functions appear) functional equation system: the equilibrium functions need to be such as to make the equations hold for all u_t . Fortunately, the functions admit closed-form solution for our model. We first guess on the functional form for B : $B(u) = X + Yu$. This guess will not work globally, because b_t cannot be negative, so we later adjust the guess to $\max\{B(u), 0\}$. The second equilibrium condition can now be written as

$$U(b_t) = 1 - \frac{1 + \beta - (b_t + \beta(X + YU(b_t)))}{2} w. \quad (6)$$

Solving for $U(b_t)$ we obtain

$$U(b_t) = \frac{2 - w(1 + \beta(1 - X)) + b_t w}{2 - \beta Y w} \quad (7)$$

⁴Given the quadratic objective, it is straightforward to check that the first-order condition will be sufficient for a maximum.

and

$$U'(b_t) = \frac{w}{2 - \beta Y w}. \quad (8)$$

Substituting the linear guess and (6) into the FOC gives

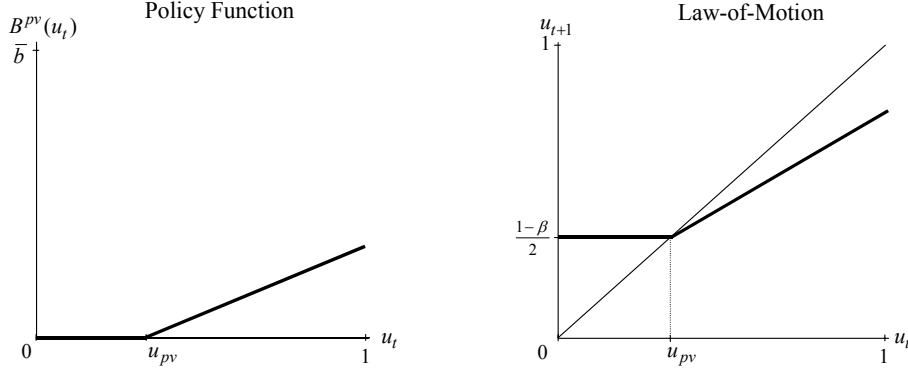
$$\begin{aligned} 0 &= u_t w - \left(1 - (1 + \beta) \frac{w}{2} + \frac{w}{2} \left(b_t + \beta \left(X + Y \left(\frac{2 - w(1 + \beta(1 - X)) + b_t w}{2 - \beta Y w} \right) \right) \right) \right) w \\ &\quad - \left(1 + \beta Y \left(\frac{w}{2 - \beta Y w} \right) \right) b_t \frac{w^2}{2} \\ b_t &= \frac{1}{2w} (-2 + w(1 + \beta(1 - X))) + \frac{2 - \beta Y w}{2w} u_t \end{aligned} \quad (9)$$

which verifies the tentative guess as a fixed-point of equilibrium condition 1 if $X = \frac{1+\beta}{2+\beta} - \frac{2}{w(2+\beta)}$ and $Y = \frac{2}{w(2+\beta)}$.

Finally, the constraint $b_t \in [0, 1]$ remains to be checked. Equation (9) yields an interior solution if $u_t \geq 1 - \frac{1}{2}(1 + \beta)w$. However, if $u_t < 1 - \frac{1}{2}(1 + \beta)w$, the restriction $b_t \geq 0$ will bind. Thus, we modify the guess to $B^{pv}(u_t) = \frac{\beta(1-X)-1}{2} + (1 - \frac{1}{2}\beta Y)u_t$ if $u_t \geq \frac{1-\beta}{2}$ and zero otherwise. The new guess will still maximize V^{pv} . To see this, note that (5) remains unchanged under the new guess of $B^{pv}(u_t)$, since the guess is modified only for values of u_{t+1} that are infeasible (this follows directly from the fact that, for any feasible pair (b_t, b_{t+1}) , $u_{t+1} = 1 - \frac{1+\beta-(b_t+\beta b_{t+1})}{2}w \geq 1 - \frac{1}{2}(1 + \beta)w$). For the same reason, $U^{pv}(b_t)$ in 6 is unaffected. Hence, $B^{pv}(u_t)$ maximizes V^{pv} under the constraint $b_t \in [0, 1]$, and is a fixed-point of the functional equation in the first part of the equilibrium definition, and $\langle B^{pv}, U^{pv} \rangle$ is a political equilibrium. ■

In the equilibrium of Proposition 1, redistribution occurs along the transition path, i.e., as long as $u_0 > u^{pv} \equiv 1 - \frac{1}{2}(1 + \beta)w$. In the long-run, however, there is no welfare state. Figure 1 represents the equilibrium policy function and law of motion, with the steady-states levels, b^{pv} and u^{pv} , respectively. The left-hand panel shows that, when $u_t > u^{pv}$, redistribution is positive in equilibrium. Moreover, the equilibrium level of b_t increases linearly with u_t . The right-hand panel shows that the equilibrium law of motion implies a monotonic asymptotic convergence to the steady-state as long as $u_0 > u^{pv}$.

These results can be interpreted as follows: when only the old influence the political outcome, the equilibrium redistribution, b^{pv} , maximizes the average income of the old. This implies maximizing the intergenerational transfer from young to old individuals without concern for intragenerational 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 qualitative results of Proposition 1 generalize to the case of $\omega \in [0, 1]$.

Proposition 2 *Assume $\omega \in [0, 1]$ and $a = 1$. The PV political equilibrium is then characterized as follows:*

$$B_g^{pv}(u_t) = \begin{cases} \frac{\rho}{w}(u_t - u^{pv}) & \text{if } u_t \geq u^{pv} \\ 0 & \text{else} \end{cases}$$

$$U_g^{pv}(b_t) = u^{pv} + b_t \left(\frac{w}{2 - \beta\rho} \right),$$

where $\rho \approx \frac{2(1-\omega)}{2+\beta}$ and $\rho = \frac{2(1-\omega)}{2+\beta}$ for $\omega = 1$ and $\omega = 0$.⁵ The equilibrium law of motion is

$$u_{t+1} = \begin{cases} u^{pv} + \frac{\rho}{2-\rho\beta}(u_t - u^{pv}) & \text{if } u_t \geq u^{pv} \\ u^{pv} & \text{else.} \end{cases}$$

Given any u_0 , the economy converges monotonically to a unique steady-state with $b = b^{pv} = 0$ and $u = u^{pv}$.

⁵The exact expression for ρ is

$$\rho \equiv \frac{1 - \omega}{1 + \beta\bar{Z} + \omega(1 + \beta\bar{Z})\beta\bar{Z}^2}$$

where

$$\bar{Z} \equiv 6^{-2/3} \frac{\left(\left(\left(9(1-\omega) + \left(3 \left(27(1-\omega(2-\omega)) + \frac{16}{\omega\beta} \right) \right)^{1/2} \right) \beta^2 \omega^2 \right)^{1/3} \right)^2 - 6^{1/3} 2\beta\omega}{\beta\omega \left(\left(9(1-\omega) + \left(3 \left(27(1-\omega(2-\omega)) + \frac{16}{\omega\beta} \right) \right)^{1/2} \right) \beta^2 \omega^2 \right)^{1/3}}.$$

Since ρ is continuous and β and ω are bounded between 0 and 1, numeric methods can be used to establish that $0.061 > \rho - \frac{2(1-\omega)}{2+\beta} \geq 0$.

In the general case, as long as $\omega < 1$, the equilibrium has the same qualitative features as in the benchmark case ($\omega = 0$). In particular, redistribution occurs only along the transition path, and there is no welfare state the long-run. Whenever $\omega < 1$, the political system favors redistribution from young to old. Such redistribution can be achieved via positive benefits if $u_t > u_{t+1}$. The higher is ω , the lower are transfers and the flatter are the equilibrium policy function and law of motion in figure 1. This is due to the fact that the young exert political pressure against redistribution. Using the approximation of ρ , we find that $u_{t+1} \approx u^{pv} + \frac{1-\omega}{2+\omega\beta} (u_t - u)$. With $\omega = 1$, the system jumps to the steady state immediately.

In section 4.1, we also characterize the set of voting equilibria when voters are benevolent vis-a-vis future consumers. We defer discussion of this case until then.

3.2 Efficient allocations under commitment

In this section, we characterize the set of constrained Pareto optimum allocations, i.e., the set of sequences of benefits, $\{b_t\}_{t=0}^{\infty}$, which would be chosen by a social planner. We assume that the planner can commit to future benefits; we sometimes refer to this problem as the Ramsey allocation. The planner is assumed to be perfectly utilitarian when evaluating the utility of ex-ante identical agents. Moreover, she discounts future generations at a constant rate by attaching a weight λ^t to agents born at time t .⁶ If she were to discount future generations at the same rate as agents discount the future, then $\lambda = \beta$. We focus on the utilitarian case of maximizing a weighted sum of utilities, where all agents are weighted alike. The program can be expressed as follows (subscripts on value functions denote dates of birth)

$$\begin{aligned}
W(u_0) = & \max_{\{b_t\}_{t=0}^{\infty}} \left\{ \beta(1-\mu) (sV_{-1}^{ows}(b_0, \tau_0) + (1-s)V_{-1}^{owu}(b_0, \tau_0)) + \right. \\
& + \beta\mu \left((1-u_0)V_{-1}^{oes}(b_0, \tau_0) + u_0V_{-1}^{oeu}(b_0, \tau_0) \right) \\
& \left. + \sum_{t=0}^{\infty} \lambda^{t+1} \left((1-\mu)V_t^{yw}(b_t, b_{t+1}, \tau_t, \tau_{t+1}) + \mu V_t^{ye}(e_t, b_t, b_{t+1}, \tau_t, \tau_{t+1}) \right) \right\}, \tag{10}
\end{aligned}$$

subject to $b_t \in [0, \bar{b}]$,

$$\begin{aligned}
\tau_t = & \begin{cases} \left((1-s)(1-\mu) + \frac{\mu}{2} \left(1 - (1+\beta) \frac{w}{2} + (b_0 + \beta b_1) \frac{w}{2} + u_0 \right) \right) w b_0, & \text{for } t = 0, \\ \left((1-s)(1-\mu) + \frac{\mu}{2} \left(1 - (1+\beta) \frac{w}{2} + (b_t + \beta b_{t+1}) \frac{w}{2} + 1 - e_{t-1} \right) \right) w b_t, & \text{for } t \geq 1, \end{cases} \\
e_t = & (1 + \beta - b_t - \beta b_{t+1}) \frac{w}{2}.
\end{aligned}$$

The following proposition characterizes the optimal benefits in the case of risk neutrality.

Proposition 3 *If agents are risk neutral ($a = 1$) and $\lambda > 0$, the solution to program (10) is given by $b_t = 0$ for all $t \geq 1$ and first-period benefits*

⁶Geometric discounting of future generations is assumed here for notational convenience only. The result in the next proposition holds for any sequence of weights.

$$b_0 = \begin{cases} 0 & \text{if } (u_0 - u^*)(\beta - \lambda) \leq 0 \\ \bar{b} & \frac{1}{w}(u_0 - u^*)\left(1 - \frac{\lambda}{\beta}\right) > \bar{b} \\ \frac{1}{w\beta}(u_0 - u^*)(\beta - \lambda) & \text{otherwise.} \end{cases}$$

Thus, the planner will always choose zero benefits after the first period, whatever the sequence of (positive) planner weights. Moreover, the solution to this program features positive redistribution in the first period ($b_0 > 0$) if the planner has more concern for the initial old than for the initial young ($\lambda < \beta$), and the initial old entrepreneurs are sufficiently unsuccessful ($u_0 > u^*$). In this case there is scope for intergenerational redistribution between the first two generations. For instance, if the number of initially unsuccessful (u_0) is large, the planner can obtain redistribution from the initial young to the initial old by choosing $b_0 > 0$. One important implication of propositions 3 and 1 is that probabilistic voting induces too much redistribution during the transition (expect maybe in the first period), but that the steady-state is efficient.

The efficient (Ramsey) allocations are, in general, time-inconsistent. This should be evident from the problem formulation: when the benefit in period $t > 0$ is chosen, the planner takes into account how it influences effort in period $t - 1$, but this is not true for time $t = 0$. Time inconsistency also appears starkly in the result: proposition 3 states that it is optimal to commit to zero after $t = 0$, although it may be optimal to have positive benefits in the *first* period ($b_0 > 0$). Hence, once the next period appears, the planner has a temptation to re-optimize and again choose positive benefits and plan for zero benefits from next period onwards. This is because $u_1 = u^* + \frac{w}{2}b_0 > u^*$, so that if the planner cares more about the old than about the young ($\beta > \lambda$), she has an incentive to deviate from the plan of the previous period by setting positive benefits in the second period and committing to zero benefits from period three onwards.

In section 4.1, we characterize the set of time-consistent constrained Pareto optimum allocations, defined as the benefits chosen by a planner who cares about future agents but who cannot commit to a particular future behavior. This case can also be viewed as one with voting when voters are benevolent vis-a-vis future consumers; it is an intermediate case between the benchmark voting equilibrium and the Ramsey equilibrium.

4 The case of risk-averse workers

HRSZ find that, in a model of majority voting, the welfare state can survive in the long-run even though agents are risk-neutral. The previous section shows, however, that under probabilistic voting redistribution must die off in the long run. That is, for the same economic environment, the long-run state of the transfer system can depend critically on the form of the democratic process. Moreover, the transitional dynamics here are characterized by monotonic rather than oscillatory convergence. 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.

The results we derive also apply qualitatively for a slightly different version of our model. Suppose all workers (lucky as well as unlucky) simply have a higher utility intensity—consumption multiplied by $a > 1$ —while having the same ideology preferences as those of the entrepreneurs (alternatively, they are less ideological and thus more attractive as swing voters). Such an assumption will translate into a more fundamental reason for redistribution—from the candidates’ perspective, the reason is political competition, leading to taxation decisions that target workers—and, consequently, the transfer system will play a more important role both in the long and in the short run.

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 exogenous terms) as follows:

$$V(b_t, b_{t+1}, u_t) = \frac{m-1}{m} \frac{1-l-\mu u_t}{\mu} 2b_t w \quad (11)$$

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

where $l \equiv (1-\mu)(1-s)$ is the share of unsuccessful workers and $m \equiv 1+l(a-1)$ is the average marginal utility of income. This equation is derived as in the case of risk neutrality, and it differs by including the first term (the second term, as before, equals $(u_t - u_{t+1})b_t w$). This first term reflects a positive effect from redistribution whenever $m > 1$ —the higher marginal utility of unlucky workers makes any redistributed dollar pay off more the higher is m . The part $(1-l-\mu u_t)/\mu$ is the number of successful old, representing the size of the inelastic taxbase, relative to the number of young workers. This terms reflects the cost of redistribution. We can now characterize the equilibrium as follows.

Proposition 4 *Assume $1 \leq a \leq a_{\max}$ and $\omega = 0$ (only the old vote). Then, the PV political equilibrium, $\langle B^{pv}, U^{pv} \rangle$, is characterized as follows:*

$$B^{pv}(u_t) = \begin{cases} b_{pv} + \frac{\rho}{w} (u_t - u_{pv}) & \text{if } u_t \geq u_{pv} - \frac{w}{\rho} b_{pv} \\ 0 & \text{otherwise} \end{cases}$$

$$U^{pv}(b_t) = u_{pv} + \frac{w}{2-\beta\rho} (b_t - b_{pv}),$$

where

$$\rho = \frac{2(2-m)}{2+\beta+(2-\beta)(m-1)},$$

$$b_{pv} = \frac{4(2(1-\mu)s + \mu w(1+\beta))}{\mu w(2+\beta)(1+3(m-1))} (m-1),$$

$$u_{pv} = 1 - \frac{w}{2} (1+\beta)(1-b_{pv}).$$

If $m > 1$, then $b^{pv} > 0$ and $B^{pv}(u_t) > 0 \forall t \geq 1$ (redistribution is positive after at most one period).

The upper bound risk aversion, $a_{\max} > 1$, is defined by the inverse function $a_{\max} = \phi^{-1}(\bar{b})$, where $\phi(a) = \max\{B^{pv}(0), B^{pv}(1)\}$, $\phi(1) < \bar{b}$, and $\phi'(a) > 0$.

The equilibrium law of motion is as follows;

$$u_{t+1} = u_{pv} + \frac{1}{2} \frac{2-m}{m} (u_t - u_{pv}),$$

Given u_0 , the economy converges to a unique steady-state, such that $b = b_{pv}$ and $u = u_{pv}$. The model features monotonic asymptotic convergence if $m \in [1, 2)$, oscillatory asymptotic convergence if $m \in (2, m_{\max}]$, where $m_{\max} \equiv 1 + l(a_{\max} - 1)$, and immediate convergence to the steady-state if $m = 2$.

Imposing an upper bound on risk aversion ($a \leq a_{\max}$) ensures that the constraint $b \leq \bar{b}$ is not binding in equilibrium. This is necessary in order for the equilibrium policy function and private decision rule to be linear and, hence, for the analytical characterization of the political equilibrium to be viable.

Proposition 4 establishes that the dynamics of redistribution converge to a unique steady-state characterized by a positive benefit rate, provided some agents are risk averse ($m > 1$). Note, in particular, that $b_{pv} = 0$ if $s = 1$. In this case, there is no insurance in equilibrium, since everyone has the same marginal utility of consumption. The comparative statics (see the appendix) shows that b_{pv} is increasing in the degree of risk aversion (in a) and in the share of workers in the population. Moreover, b_{pv} is decreasing in the wage rate because the distortionary effect of benefits increase as the return to effort increases. The share of unsuccessful workers, however, has an ambiguous effect on redistribution.

The equilibrium policy function and the dynamics of u_t are depicted in figure 2. As long as $m < 2$, the benefit rate behaves monotonically along transition, i.e., $B(u)$ is positively sloped. If $m > 2$ instead, the benefit rate is a decreasing function of u_t , and convergence follows an oscillatory pattern. In the particular case where $m = 2$, convergence to the steady-state occurs in one period.

These dynamics result from two opposite forces. On the one hand, the larger the current share of unlucky entrepreneurs, u_t , the higher the tax cost per unit of benefits (see the term with a negative sign in front of u_t in the political objective function (11)). Thus, efficiency considerations tend to generate a negatively sloped relationship between b and u . On the other hand, the larger is u_t , the larger is the political pressure for redistribution (see the terms in the political objective with positive signs in front of u_t). This political power effect tends to generate a positively sloped relationship between b and u . When risk aversion is low, the latter effect dominates. When risk aversion is high, however, the former effect dominates. The reason is that, when a is high, the political influence of the entrepreneurs diminish, as workers become, on average, more sensitive to the issue of redistribution, due to their high marginal utility of income. In a probabilistic voting model, this implies that the policy implemented in equilibrium reflects the will of the average worker more closely, i.e., there is more redistribution. Moreover, and this is what

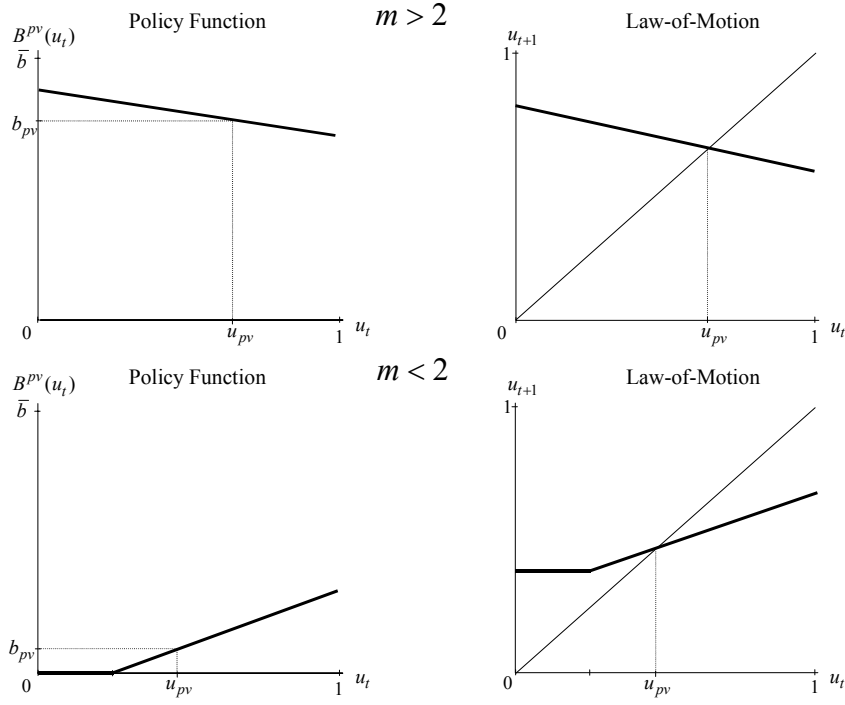


Figure 1: Figure 2. Equilibrium policy function $B(u_t)$ and u_t dynamics under risk aversion ($a > 1$).

effects transitional dynamics, the policy outcome becomes less sensitive to the share of entrepreneurs demanding redistribution. Thus, the dynamics are dominated by the cost effect.

So far we have ignored the constraint that $b_t \leq \bar{b}$. If this constraint were violated, the proposed equilibrium would have to be modified. As the slope Z is less than $1/2$ for all parameter choices, it is sufficient to explore for which parameter values $(1 - Z)b_{pv} + Z \cdot 0 < \bar{b}$. This condition holds whenever $a < a_{\max}$. Note also that since $b_{pv} \geq 0$, the constraint $b \geq 0$ is never binding in equilibrium for $t \geq 1$. Thus, neither of the constraints are binding in equilibrium as long as $a < a_{\max}$ is satisfied.

We now proceed to generalize Proposition 4 to the case in which the young participate in the political decision, $\omega \in [0, 1]$. Unfortunately, we cannot obtain closed-form solutions in this case, but it is possible to establish that the policy function and private decision rules are linear and to characterize some important properties of these functions. We maintain the assumption that $a < a_{\max}$ throughout.

Proposition 5 *Assume $1 \leq a \leq a_{\max}$ and $\omega \in [0, 1]$. The PV political equilibrium is then*

characterized as follows:

$$B_g^{pv}(u_t) = \begin{cases} b_{pv} + \frac{\rho}{w}(u_t - u_{pv}) & \text{if } u_t \geq u_{pv} - \frac{w}{\rho}b_{pv} \\ 0 & \text{otherwise} \end{cases}$$

$$U_g^{pv}(b_t) = u_{pv} + \frac{w}{2 - \beta\rho}(b_t - b_{pv}),$$

where $\rho = \frac{2Z}{1+Z\beta}$ and Z is the real root of

$$Z = \frac{1}{2} \frac{2 - m(1 + \omega)}{m(1 + \omega\beta Z^2) + \omega(1 + \beta Z)(m - 1)},$$

and where

$$\equiv \frac{b_{pv} (2(m-1)(2s(1-\mu) + \mu w(1+\beta))(1 + \omega(1 + \beta Z)))/(1 - Z)}{\mu w(2(m-1)(\beta + \omega(1 + \beta(1 + Z(2 + \beta(1 + Z)))))) + m(1 + \beta Z)(2 + \beta\omega(1 + 2Z(1 + Z)))}, \quad (12)$$

$$u_{pv} \equiv 1 - \frac{w}{2}(1 + \beta)(1 - b_{pv}). \quad (13)$$

The equilibrium law of motion is

$$u_{t+1} = \begin{cases} u_{pv} + \frac{\rho}{2 - \beta\rho}(u_t - u_{pv}) & \text{if } u_t \geq u_{pv} \\ u_{pv} & \text{else.} \end{cases}$$

Given any u_0 , the economy converges to a unique steady-state such that $b = b_{pv}$ and $u = u_{pv}$.

Using the previous proposition, some results of comparative statics can be established.⁷ if $m < 2/(1 + \omega)$ then $Z \in (0, 1/2]$ and $\rho \in (0, 2/(2 + \beta)]$, if $m \geq 2/(1 + \omega)$, then $Z \in (-1, 0]$ and $\rho \in (-1/(2 + \beta), 0]$. In other words, an increase in the political participation of the young and/or an increase in risk aversion decreases the slope of the policy function. In particular, the sign of the slope coefficient depends on whether $m \leq 2/(1 + \omega)$. This condition nests the result of Proposition 4 that the policy function is upward-(downward-)sloping if and only if $m < 2$ ($m > 2$). If the young are as politically influential as the old ($\omega = 1$), then the policy function becomes downward-sloping for any positive level of risk aversion (recall, from Proposition 2, that the policy function is exactly flat when agents are risk-neutral, corresponding to the particular case of $m = 1$ in Proposition 5). In summary, for $m < 2$, there is a threshold participation level of the young such that, for higher participation, the policy function is downward-sloping, and for lower participation it is upward-sloping. When $m > 2$, instead, the policy function is always downward-sloping, and the participation of the young makes the function steeper. As before, finally, an increase

⁷It is easily established that the denominator $m(1 + \omega\beta Z^2) + \omega(1 + \beta Z)(m - 1)$ is always positive. Hence, the sign of Z depends on the sign of $m - 2/(1 + \omega)$.

in the average risk aversion, m , reduces the slope coefficient and increases steady-state redistribution.

Unfortunately, we have so far been unable to sign the effect of an increase of the participation of the young on steady-state redistribution, although numerical analysis suggests that an increase in ω reduces redistribution in the long-run.

4.1 Voting with benevolence vis-a-vis future generations

In this section, we characterize the allocation that is chosen by a benevolent social planner who has no access to a commitment technology. This planner can be viewed as a set of current old voters with positive weights on the utility of future voters. As before, the commitment problem arises from the fact that redistribution does not distort, ex-post, the choice of old agents who have already chosen their effort. Thus, the planner has systematically the temptation to promise a lower level of redistribution and to renege and choose more redistribution. This leads to lower equilibrium effort. Note that the probabilistic-voting allocation in section 4 is a special case of the present one if one sets the weight λ below to 0, i.e., if the planner does not care about the future agents (including the current young) at all. Thus, as in Klein and Ríos-Rull (2000), the planner plays a game against future versions of herself.

Formally, the time-consistent allocation is the solution of the following program:

$$\begin{aligned} W(u_0) = & \max_{\{b_0\}} \left\{ \beta(1-\mu) (sV_{-1}^{ows}(b_0, \tau_0) + (1-s)V_{-1}^{owu}(b_0, \tau_0)) + \right. \\ & + \beta\mu ((1-u_0)V_{-1}^{oes}(b_0, \tau_0) + u_0V_{-1}^{oeu}(b_0, \tau_0)) \\ & \left. + \sum_{t=0}^{\infty} \lambda^{t+1} ((1-\mu)V_t^{yw}(b_t, b_{t+1}, \tau_t, \tau_{t+1}) + \mu V_t^{ye}(e_t, b_t, b_{t+1}, \tau_t, \tau_{t+1})) \right\}, \end{aligned}$$

$$\text{subject to } b_0 \in [0, \bar{b}],$$

$$\tau_t = \begin{cases} ((1-s)(1-\mu) + \frac{\mu}{2} (1 - (1+\beta)\frac{w}{2} + (b_0 + \beta b_1)\frac{w}{2} + u_0)) w b_0, & \text{for } t = 0, \\ ((1-s)(1-\mu) + \frac{\mu}{2} (1 - (1+\beta)\frac{w}{2} + (b_t + \beta b_{t+1})\frac{w}{2} + 1 - e_{t-1})) b_t w, & \text{for } t \geq 1, \end{cases}$$

$$e_t = \frac{1 + \beta - b_t - \beta b_{t+1}}{2} w$$

$$b_t = B(u_t), \text{ for } t \geq 1$$

It is convenient to note the fact that $u_0 = 1 - (1 + \beta - b_{-1} - \beta B(b_{-1})) w / 2$, and to reformulate the problem in terms of the state variable b_{-1} instead of u_0 . Then, define $U(b_{-1}, b_0, b_1)$ as the “weighted average felicity” across both young and old agents at time zero. We let $\beta \in [0, 1]$ and $\lambda \in [0, 1]$, respectively, be the weight of the old and young agents alive at time zero in this felicity. Thus, $\beta = \lambda$ means that the planner is perfectly utilitarian and treats equally old and young agents who are alive in a certain period. Or, in the alternative interpretation, it means that old agents are perfectly altruistic and weight equally their old-age felicity and that of their offspring. If, on the other hand, $\lambda = 0$, we obtain the case of probabilistic voting analyzed in Proposition 1. We restrict attention to economies where $\lambda \leq \beta$, namely, where altruism cannot exceed 100%.

Now,

$$\begin{aligned}
U(b_{-1}, b_0, b_1) &= \beta(1-\mu)(sV_{-1}^{ows}(b_0, \tau_0) + (1-s)V_{-1}^{owu}(b_0, \tau_0)) + \\
&\quad + \beta\mu((1-u_0)V_{-1}^{oes}(b_0, \tau_0) + u_0V_{-1}^{oeu}(b_0, \tau_0)) \\
&\quad + \lambda((1-\mu)U_0^{yw}(b_0, b_t, \tau_0) + \mu U_0^{ye}(e_0, b_0, b_t, \tau_0)) \\
&= -(\beta + \lambda)m\tau_0 + \beta(1-\mu)(sw + (1-s)ab_0w) \\
&\quad + \beta\mu((1-u_0)w + u_0b_0w) \\
&\quad + \lambda((1-\mu)(sw + (1-s)ab_0w) + \mu(e_0w + (1-e_0)b_0w - e_0^2))
\end{aligned}$$

where U^{yw} and U^{ye} refer, respectively, to the expected felicity (not the PDV utilities) of the young workers and entrepreneurs, before their success is realized. Substituting in the expressions of τ_0 yields

$$\begin{aligned}
U(b_{-1}, b_0, b_1) &= -(\beta + \lambda)m\left(l + \frac{\mu}{2}\left(1 - (1 + \beta)\frac{w}{2} + (b_0 + \beta b_1)\frac{w}{2} + u_0\right)\right)wb_0 \\
&\quad + \beta(1-\mu)(sw + (1-s)ab_0w) + \beta\mu((1-u_0)w + u_0b_0w) \\
&\quad + \lambda(1-\mu)(sw + (1-s)ab_0w) + \lambda\mu(1 + \beta - b_0 - \beta b_1)\frac{w^2}{2} \\
&\quad + \lambda\mu\left(1 - (1 + \beta - b_0 - \beta b_1)\frac{w}{2}\right)wb_0 - \lambda\mu\left((1 + \beta - b_0 - \beta b_1)\frac{w}{2}\right)^2
\end{aligned} \tag{14}$$

In the time-consistent solution, $b_1 = B(b_0)$, and e_0 and u_0 are given, respectively, by

$$e_0 = ((1 + \beta - b_0 - \beta B(b_0))w/2) \tag{15}$$

$$u_0 = (1 - (1 + \beta - b_{-1} - \beta B(b_{-1}))w/2). \tag{16}$$

Note, in particular, that equation (16) implies that e_{-1} , and, hence u_0 , was set by entrepreneurs conditionally on the knowledge of b_{-1} , and on the expectation that b_0 would be set according to the equilibrium law of motion.

The problem admits a recursive formulation of the following type:

$$\begin{aligned}
W(b_{t-1}; B) &= \max_{b \in [0, \bar{b}]} \{U(b_{t-1}, b_t, b_{t+1}) + \lambda W(b_t; B)\}, \\
b_{t+1} &= B(b_t).
\end{aligned} \tag{17}$$

We guess that the time-consistent social planners choose benefits, from period one onwards, according to the following linear rule

$$b_{t+1} = B(b_t) = X + Zb_t.$$

Given the guess, we can substitute $b_1 = X + Zb_0$ and $u_0 = 1 - (1 + \beta - \beta X - (1 + \beta Z)b_{-1})w/2$ into the weighted-average felicity and rewrite the latter as

$$U(b_{-1}, b_0, B(b_0)) = Q + Ab_0^2 + Cb_0b_{-1} + Db_0 + Fb_{-1}$$

where

$$\begin{aligned}
A &= \mu(1 + \beta Z) \frac{w^2}{2} \left(-\frac{1}{2}(\beta + \lambda)m + \lambda \left(1 - \frac{1}{2}(1 + \beta Z) \right) \right) \\
C &= (1 + \beta Z) \frac{1}{2} \mu w^2 \left(\beta - \frac{1}{2}(\beta + \lambda)m \right) \\
D &= \lambda(1 + \beta Z) w^2 \mu \frac{1}{2} (1 + \beta - \beta X) + \left(w \mu \frac{1}{2} (1 + \beta - \beta X) + (1 - \mu)s \right) (\beta + \lambda) w (m - 1) \\
&\quad - \lambda \mu \frac{1}{2} (1 + \beta Z) w^2 \\
F &= \beta \mu (1 + \beta Z) w \\
Q &= \beta \mu (-\beta + \beta X) w + \beta (1 - \mu) s w + \lambda \mu \left(\frac{1}{2} (1 + \beta - \beta X) w^2 - \frac{1}{4} (1 + \beta - \beta X)^2 w^2 \right) \\
&\quad + \lambda (1 - \mu) s w
\end{aligned}$$

and

$$Z = -\frac{C}{2A + \lambda CZ}.$$

The first-order condition for the maximization of 17 delivers

$$\frac{dU(b_{-1}, b_0, B(b_0))}{db_0} + \lambda \frac{dW(b_0)}{db_0} = 0,$$

which, using the envelope condition, $\frac{dW(b_t)}{db_t} = \frac{\partial U(b_t, b_{t+1}, b_{t+2})}{\partial b_t}$, can be rewritten as

$$\frac{dU(b_{-1}, b_0, B(b_0))}{db_0} + \lambda \frac{\partial U(b_0, b_1, b_2)}{\partial b_0} = 0.$$

Hence,

$$2Ab_0 + (Cb_{-1} + D) + \lambda(F + C(X + Zb_0)) = 0$$

with solution

$$b_0 = -\frac{C}{2A + \lambda CZ} b_{-1} - \frac{\lambda CX + D + \lambda F}{2A + \lambda CZ}.$$

Equating coefficients, we have that the guess is verified as long as

$$\begin{aligned}
Z &= -\frac{C}{2A + \lambda CZ} \\
X &= -\frac{\lambda CX + D + \lambda \mu \beta (1 + \beta Z) w}{2A + \lambda CZ}.
\end{aligned}$$

Substituting in the expressions for the constants, (A, C, D, F) , we obtain

$$Z = \frac{m(\beta + \lambda) - 2\beta}{2m(\beta + \lambda) - 2\lambda + \lambda m(\beta + \lambda)Z} \equiv f(Z)$$

$$X = \frac{2(-\mu(1 + \beta Z)\beta(2 + w)\lambda - (w\mu(1 + \beta) + 2s(1 - \mu))(\beta + \lambda)(m - 1))}{w\mu((2\lambda - (2 + \lambda)(\lambda + \beta)m - (\beta + \lambda)\lambda mZ)(1 + \beta Z) - 2(\beta + \lambda)(m - 1)\beta)}$$

where the solution for Z must be such that $|Z| < 1$.

We can now establish the main proposition of this section.

Proposition 6 *Assume $\lambda \leq \beta$. The time-consistent (TC) planning solution is characterized as follows:*

$$B^{tc}(u_t) = \begin{cases} b_{tc} + \frac{\rho}{w}(u_t - u_{tc}) & \text{if } u_t \geq u_{tc} - \frac{w}{\rho}b_{tc} \\ 0 & \text{otherwise} \end{cases}$$

$$U^{tc}(b_t) = u_{tc} + \frac{w}{2 - \beta\rho}(b_t - b_{tc}),$$

where b_{tc} and u_{tc} are decreasing in λ (if $\lambda = 0$, then $b_{tc} = b_{pv}$ and $u_{tc} = u_{pv}$ as in Proposition 4). Furthermore, let $Z = \rho/(2 - \beta\rho)$, where Z is an increasing function of ρ . Then

$$Z = \frac{1}{\lambda m(\beta + \lambda)} \left(-(m(\beta + \lambda) - \lambda) + \sqrt{((1 - \lambda)m^2(\beta + \lambda)^2 - (1 - \beta)2\lambda m(\beta + \lambda) + \lambda^2)} \right),$$

where Z is decreasing in λ , and the following properties can be established:

1. if $m < 2\beta/(\beta + \lambda)$, then $Z \in (0, 1/2]$ and $\rho \in (0, 2/(2 + \beta)]$;
2. if $m > 2\beta/(\beta + \lambda)$, then $Z \in [-1, 0)$ and $\rho \in [-2/(1 - \beta), 0)$; and
3. if $m = 2\beta/(\beta + \lambda)$, then $Z = \rho = 0$.

The equilibrium law of motion is;

$$u_{t+1} = \begin{cases} u_{tc} + \frac{\rho}{2 - \beta\rho}(u_t - u_{tc}) & \text{if } u_t \geq u_{tc} \\ u_{tc} & \text{else} \end{cases}$$

Given any u_0 , the economy converges to a unique steady-state, such that $b = b_{tc} \leq b_{pv}$ and $u = u_{tc} \leq u_{pv}$.

The proposition establishes that the slope coefficient of the policy function is decreasing in λ . Namely, the time-consistent solution (or altruistic motive in the political equilibrium) reduces the intergenerational redistribution effect and tends to make the efficiency effect prevail (as when we considered the participation of the young). Thus, the transfer is now negatively related u_t for a larger set of parameters: the larger is the group of unsuccessful entrepreneurs, the more costly it is to generate a dollar of transfer.

4.2 Efficient allocations in the economy with risk-averse workers

In order to evaluate normatively the allocations implied by the political equilibrium, we now revisit the full-commitment, or Ramsey, allocation problem of section 3.2 in the case of positive risk aversion. The sequential planner problem, (10), does not admit a standard recursive formulation since, as discussed in section 4.1, the first best solution is time inconsistent.

An important equivalence result, however, can be established between the Ramsey problem and a recursive formulation where the planner can, in every period, commit to next period redistribution. Clearly, such equivalence does not hold for the choice of benefits in the initial period.

Lemma 7 *The solution to the sequential Ramsey problem, (10), is equivalent to the solution to the following recursive problems. For $t \geq 1$,*

$$W(b_{t-1}, b_t) = \max_{b_{t+1} \in [0, \bar{b}]} \{U(b_{t-1}, b_t, b_{t+1}) + \lambda W(b_t, b_{t+1})\}, \quad (18)$$

where $U(b_{t-1}, b_t, b_{t+1})$ is given by (14) with u_t being equal, respectively, to

$$u_t = 1 - (1 + \beta - b_{t-1} - \beta b_t) w/2,$$

for all $t \geq 1$.

Moreover, the choice b_0 and b_1 is the solution to the following problem

$$V_0(u_0) = \max_{b_0, b_1 \in [0, \bar{b}]} \left\{ \hat{U}(u_0, b_0, b_1) + \lambda W(b_0, b_1) \right\}, \quad (19)$$

where $\hat{U}(u_0, b_0, b_1)$ is also given by (14), but with u_0 predetermined.

Consider, first, the choice of b_2, b_3, \dots . The equivalence between (10) and (18) stems from the fact that, since agents only live for two periods, b_t has a distortionary effect on investment effort only in period $t - 1$ and t (i.e., on e_{t-1} and e_t) and, therefore affects taxation in periods $t - 1$, t and $t + 1$. In particular, in the sequential objective function in (10), b_t only interacts with b_{t-1} and b_{t+1} , and there is no interaction between b_{t-1} and b_{t+1} . Therefore, giving the planner ability to commit to redistribution one period in advance is sufficient to avoid the time inconsistency. As to the initial choice, the planner is not subject to earlier pre-commitments, and chooses b_0 and b_1 simultaneously.

We will now show that the problem from $t = 1$ and on can, in fact, be reformulated into a standard recursive form with one state variable. This characterization greatly simplifies the analysis. To this aim, note that the function $U(b_{t-1}, b_t, b_{t+1})$ is additively separable in (b_{t-1}, b_t) and (b_t, b_{t+1}) . More formally, there exist (linear-quadratic) functions F and H such that⁸

$$U(b_{t-1}, b_t, b_{t+1}) = F(b_{t-1}, b_t) + H(b_t, b_{t+1}).$$

⁸ $F(b_{t-1}, b_t) \equiv -b_{t-1} \frac{w^2}{2} (\beta \mu (1 - b_t) + (\beta + \lambda) m \frac{\mu}{2} b_t) \leq 0$ and $H(b_t, b_{t+1})$ is given in the appendix.

Since $F(b_{t-1}, b_t)$ does not affect the optimal choice of b_{t+1} , equation (18) can be rewritten as

$$W(b_{t-1}, b_t) - F(b_{t-1}, b_t) = \max_{b_{t+1} \in [0,1]} \{H(b_t, b_{t+1}) + \lambda W(b_t, b_{t+1})\}. \quad (20)$$

Since the RHS is a function of only b_t , we can now define the value function $V(b_t) \equiv W(b_{t-1}, b_t) - F(b_{t-1}, b_t)$ and rewrite (20) as

$$V(b_t) = \max_{b_{t+1} \in [0,1]} \{Y(b_t, b_{t+1}) + \lambda V(b_{t+1})\}$$

where $Y(b_t, b_{t+1}) \equiv H(b_t, b_{t+1}) + \lambda F(b_t, b_{t+1})$ is a linear-quadratic function of b_t and b_{t+1} given by

$$\begin{aligned} Y(b_t, b_{t+1}) &= w \left((\lambda + \beta)(1 - \mu)s + (\lambda + \beta(2 - \lambda))\mu \frac{w}{4}(1 + \beta) \right) \\ &\quad + \left((\beta + \lambda) \left(\frac{w}{2}(1 + \beta) - 1 \right) \mu(m - 1) - \frac{w}{2}\mu\beta^2 + (\beta + \lambda)l(a - m) \right) wb_t \\ &\quad + \left((\beta + \lambda)(m - 1) \left(s(1 - \mu) + \mu(1 + \beta) \frac{w}{2} \right) - \mu \frac{w}{2}\beta^2 \right) wb_t \\ &\quad - ((m - 1)(\beta + \lambda)(1 + \beta) + \beta(1 + \lambda - \beta))\mu \left(\frac{w}{2} \right)^2 b_t^2 \\ &\quad - \left(\frac{w}{2} \right)^2 \mu ((\beta^2 + \lambda^2)m + 2\lambda\beta(m - 1)) b_t b_{t+1} + \lambda\mu \left(\frac{w\beta}{2} \right)^2 b_{t+1} (2 - b_{t+1}). \end{aligned} \quad (21)$$

A similar argument can be applied to the choice of redistribution in the initial period. First, it can be shown that $\hat{U}(u_0, b_0, b_1) - H(b_0, b_1)$ is independent of b_1 . Thus, define

$$\begin{aligned} \hat{Y}(u_0, b_0) &\equiv \hat{U}(u_0, b_0, b_1) - H(b_0, b_1) \\ &= Q_2 + \beta \frac{w^2}{4} \frac{\mu}{(1 - \mu)} ((\beta + \lambda)m - 2\beta) b_0^2 - \frac{1}{2} \frac{\mu}{(1 - \mu)} ((\beta + \lambda)m - 2\beta) \cdot \\ &\quad \cdot \left(u_0 - \left(1 - (1 + \beta) \frac{w}{2} \right) \right) wb_0 + \beta^2 \frac{\mu}{(1 - \mu)} \frac{w^2}{2} b_0. \end{aligned} \quad (22)$$

Then, we can rewrite (19) as follows (see the proof for the intermediate steps):

$$V_0(u_0) = \max_{b_0 \in [0, \bar{b}]} \left\{ \hat{Y}(u_0, b_0) + V(b_0) \right\}.$$

Lemma 8 summarizes the findings so far. In particular, it states that there exists a simple transformation that allows a recursive representation of the full-commitment Ramsey problem.

Lemma 8 *The utilitarian planner program (10) is equivalent to the following recursive program:*

$$\begin{aligned} V_0(u_0) &= \max_{b_0 \in [0, \bar{b}]} \left\{ \hat{Y}(u_0, b_0) + V(b_0) \right\} \\ V(b_t) &= \max_{b_{t+1} \in [0, \bar{b}]} \left\{ Y(b_t, b_{t+1}) + \lambda V(b_{t+1}) \right\} \quad \text{for } t \geq 0, \end{aligned}$$

where $Y(b_t, b_{t+1})$ and $\hat{Y}(u_0, b_0)$ are defined in equations (21) and (22). Moreover, the mapping $\max_{b' \in [0,1]} \{Y(b, b') + \lambda V(b')\}$ is a contraction mapping with V as the unique fixed point.

The recursive formulation of Lemma 8 suggests that the optimal policy can be represented in terms of two policy rules. The first rule applies from period one onwards and maps previous period benefits into current period benefits, $b_t = f(b_{t-1})$. The second rule, which refers to the initial choice of redistribution, maps initial (predetermined) proportion of unsuccessful entrepreneurs into initial choices of redistribution ($b_0 = f_0(u_0)$). It is possible, for all parameters, to obtain analytical characterizations of such policy rules. In some cases, these are relatively simple (in fact, they are linear). In other cases, however, they are more involved and do not admit closed-form solution. In all cases, we can characterize the long-run properties of the optimal benefit sequences.

It turns out that the optimal benefit sequence generically follows an oscillatory pattern from period one onwards. In some cases, the transition is saddle-path stable (in this case, the analytical characterization of the transition is simple). In other cases, the dynamics are explosive for some periods and eventually converge in finite time to two period-cycles with zero redistribution every other period. The case where $\lambda = \beta$ is of particular interest, since it has the interpretation that the planner gives equal weight to the felicity of all agents living in the same period. This case turns out to have the knife-edge property that redistribution perpetually oscillates between two levels without neither exploding nor imploding. Finally, the choice of redistribution in the first period always depends negatively on u_0 .

The next proposition characterizes the Ramsey plan.

Proposition 9 *The optimal solution to the planner program (10) can be represented as a first-order difference equation in b ,*

$$\begin{aligned} b_0 &= f_0(u_0) \\ b_{t+1} &= f(b_t), \quad \forall t \geq 0, \end{aligned}$$

where $f(\cdot)$ is monotone decreasing.

There exists a unique steady-state benefit level, given by $b^{plan} \equiv \min\{b^*, \bar{b}\} > 0$, where

$$b^* = \frac{4\lambda}{w\mu(1+\lambda)(\beta+\lambda)(M(1-\mu)+\mu)+2\lambda(1+\beta)(1-\mu)(M-1)} \left((s(1-\mu) + \mu\frac{w}{2}(1+\beta))(1-s)(1-\mu)(a-1) \right).$$

Moreover, this steady-state implies less redistribution than under probabilistic voting, i.e.

$$b^{plan} \leq b^{pv},$$

where the inequality is strict provided that $b^* < \bar{b}$, $a < 1$, and $l > 0$.

Define

$$\begin{aligned} \tilde{\lambda} &= -\frac{1}{2}(1-\beta) \left(\frac{2}{((1-\mu)M+\mu)} - 1 \right) \\ &\quad + \frac{1}{2} \sqrt{(1-\beta)^2 \left(\frac{2}{((1-\mu)M+\mu)} - 1 \right)^2 + 4\beta}, \end{aligned}$$

where $\tilde{\lambda} \in [\beta, 1]$. If $\lambda \in (\beta, \tilde{\lambda})$ then the optimal benefit sequence converges asymptotically to b^* . If $\lambda \in [0, \beta] \cup [\tilde{\lambda}, 1]$, then the optimal benefit sequence initially diverges from b^* , and eventually converges in finite time to a limit cycle of periodicity 2, with b_t oscillating between 0 and $b = \min \{\tilde{b}, \bar{b}\} \geq b^{plan}$, with \tilde{b} being given by

$$\tilde{b} = \frac{2}{w} (\beta + \lambda) (a - 1) (1 - s) \frac{(1 - \mu)}{\mu} \frac{(s(1 - \mu) + \mu \frac{w}{2} (1 + \beta))}{(((1 - \mu)M + \mu)(\beta + \lambda)(1 + \beta) - (\lambda + \beta^2))}.$$

A key result is that the equilibrium law of motion for benefits, f , is monotone decreasing. Thus, the dynamics are oscillatory, irrespectively whether the dynamics are saddle-path stable. It might seem surprising that the planner would not opt for benefit smoothing. It turns out that oscillating benefits precisely minimizes the distortion associated with redistribution. The reason is intimately linked to the fact that the investments young agents make in period t has an effect on the tax cost of redistribution both in period t and in period $t + 1$. More precisely, if benefits in period $t - 1$ were large (small), the young entrepreneurs will make a small (large) investment effort in that period. Thus, in period t , the old entrepreneurs will be relatively unsuccessful (successful), so there will be many (few) benefit recipients that period, and the cost of redistribution per dollar of transfer in period t will therefore be relative large (small). Thus, the planner will prefer relatively small (large) benefits in period t . 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 next period, as the investment decision of the young in period t depends on redistribution both in period t and in period $t + 1$. We suspect that this mechanism will be present in any overlapping-generations model where the human capital investment of the young has permanent income effects, so that it is distorted by future progressive redistribution. More complex generational overlap may mitigate the stark back-and-forth dynamics, however.

The proposition implies that for an intermediate range of planner weights, $\lambda \in (\beta, \tilde{\lambda})$, the optimal benefits converge to the steady-state value of transfers, b^{plan} , while for very high and very low planner discounting, the optimal benefits eventually end up cycling between $b = 0$ and $b = \tilde{b}$. It is worth noting that stable dynamics requires positive risk aversion (in order to obtain $\tilde{\lambda} > \beta$).

Proposition 4.2 has two important results for long-run benefits. First, the steady state associated with the planner program (10) with commitment involves positive redistribution. This result requires positive risk aversion and that the workers face some risk. Moreover, the planner must put some weight on future agents ($\lambda > 0$). Second, an important feature of the long-run benefits chosen by a utilitarian planner is that the benefits b^{plan} lie below the long-run benefits under probabilistic voting (b^{pv} in proposition 4). Note that this statement holds regardless of the rate the planner discounts future generations! The intuition for this involves two channels. First, the proportion of unsuccessful entrepreneurs inherited from the past is a burden for the current and future generations in their efforts to redistribute. The voting mechanism disregards the effects of current redistribution on future generations

other than those represented in the political process. Thus, current voters (who cannot commit future redistribution) choose to redistribute more than the utilitarian planner: the planner internalizes this effect here whereas the voting mechanism does not, leading to lower distortions under commitment. This effect was internalized also in the time-consistent planning solution without commitment of the previous section, since this solution does not require commitment. Second, high benefits at t distorts choices at $t - 1$, and the access to commitment allows the planner to internalize this effect as well, leading to lower equilibrium transfers. This did not occur in the time-consistent planning solution without commitment.

The comparative statics for b^* mirror those under probabilistic voting in section 4. In particular, b^* is increasing in risk aversion ($\partial b^*/\partial a > 0$), due to a larger insurance motive. Moreover, b^* is increasing in the fraction of workers in the economy, $\partial b^*/\partial \mu < 0$, for two reasons: the number of workers in need of insurance is increasing, and the distortion of entrepreneurial effort is of less importance as there are fewer of them around. Finally, b^* is decreasing in the wage rate w and in the planner's discount factor λ . The former is true because the distortion of entrepreneurial effort becomes more severe, and the latter holds because it is optimal to reduce the distortion when the planner increases the weight on future agents.

Further insights can be gained by looking at the transitional dynamics. Consider, first, the range $\lambda \in (\beta, \tilde{\lambda})$. For this range, we can provide an explicit solution for the entire benefit sequence (see proof of Proposition 9 for details). In particular, we can establish that

$$\begin{aligned} b_0 &= f_0(u_0) = A_0 + A_1 u_0 \\ b_{t+1} &= f(b_t) = b^* + Z(b_t - b^*), \quad \forall t \geq 0, \end{aligned}$$

where Z is given by

$$Z = \frac{1}{2\lambda\eta_1} \left(-\eta_2 + \sqrt{(\eta_2^2 - 4\lambda\eta_1^2)} \right) \in (-1, 0),$$

and

$$\begin{aligned} \eta_1 &= \frac{w^2}{4} \frac{\mu}{(1-\mu)} (2\lambda\beta - m(\beta + \lambda)^2) \\ \eta_2 &= \left(\frac{w^2}{2} \frac{\mu}{(1-\mu)} \lambda (\lambda + \beta^2 - m(\beta + \lambda)(1 + \beta)) \right). \end{aligned}$$

For the remaining range, there is no closed-form solution, but we have characterized the optimal rules f_0 and f numerically.

5 Conclusion

To be written.

References

- Hassler, J., Rodríguez Mora, J. V., Storesletten, K. and Zilibotti, F. (2001), The survival of the welfare state, CEPR working paper no. 2905.
- Klein, P. and Ríos Rull, J. V. (2000), Time-consistent optimal fiscal policy, University of Pennsylvania working paper.
- 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. and Weibull, J. W. (1987), ‘Balanced-budget redistribution as political equilibrium’, *Public Choice* **52**(3), 273–297.
- Persson, T. and Tabellini, G. (2000), *Political Economics - Explaining Economic Policy*, MIT Press, Cambridge.