DISCUSSION PAPER SERIES

No. 2905

THE SURVIVAL OF THE WELFARE STATE

John Hassler, José V Rodríguez Mora,
Kjetil Storesletten and Fabrizio Zilibotti

INTERNATIONAL MACROECONOMICS AND PUBLIC POLICY

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP2905.asp**

ISSN 0265-8003

THE SURVIVAL OF THE WELFARE STATE

John Hassler, IIES, Stockholm University and CEPR
José V Rodríguez Mora, Universitat Pompeu Fabra, Barcelona and CEPR
Kjetil Storesletten, IIES, Stockholm University and CEPR
Fabrizio Zilibotti, IIES, Stockholm University and CEPR

Discussion Paper No. 2905
August 2001

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 2905

August 2001

ABSTRACT

The Survival of the Welfare State*

This Paper provides an analytical characterization of Markov perfect equilibria in a model with repeated majority voting, where agents vote over income redistribution. The key feature of the theory is that the future constituency of redistributive policies depends positively on the current level of redistribution, since this affects both private investments and the future distribution of voters. Agents vote rationally, and fully anticipate the effects of their political choice on both private incentives and future voting outcomes. The equilibrium features multiple steady states, one with and one without a welfare state. The theory can explain why welfare state institutions, originally introduced in response to specific shocks (e.g. the Great Depression), have been so persistent.

JEL Classification: D72, E62, H11, H31 and P16
Keywords: education, Markov equilibrium, multiple equilibria, policy-behaviour complementarity, redistribution, repeated voting, wage inequality and welfare state

John Hassler
Institute for International Economic Studies (IIES),
Stockholm Universitet
S-10691 Stockholm
SWEDEN
Tel: (46 8) 162 070
Fax: (46 8) 161 443
Email: john.hassler@iies.su.se

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=126285

José V Rodríguez Mora
Department of Economics
Universitat Pompeu Fabra
Ramón Trías Fargas 25-27
08005 Barcelona
SPAIN
Tel: (34 93) 542 1755
Fax: (34 93) 542 1746
Email: sevimora@upf.es

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=139713

Kjetil Storesletten
Institute for International Economic Studies (IIES),
Stockholms Universitet
S-10691 Stockholm
SWEDEN
Tel: (46 8) 163 075
Fax: (46 8) 161 443
Email: kjetil@iies.su.se

Fabrizio Zilibotti
Institute for International Economic Studies (IIES),
Office A805
Stockholm Universitet
S-10691 Stockholm
SWEDEN
Tel: (46 8) 162 225
Fax: (46 8) 161 443
Email: fabrizio.zilibotti@iies.su.se

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=135389

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=118864

* Hassler, Storesletten and Zilibotti thank The Bank of Sweden Tercentenary Foundation for financial support (Project Name: ‘The Roles of Government and Economic Policy-Making in the Future Europe’).

Submitted 04 July 2001

NON-TECHNICAL SUMMARY

The rise of the welfare state in industrialized countries in the 20th century has led to an unprecedented change in the size and scope of governments. These developments were, to a large extent, triggered by specific historic episodes, such as the Great Depression or World War II, when large masses of people were impoverished, thereby creating a demand for public intervention. For instance, government spending as a fraction of GNP doubled in the 1930s in the US as well as in the major European countries, e.g. France, Italy and the UK. Government programmes, however, did not diminish after economies had recovered from the shocks. Instead, redistributive policies persisted and were further expanded in the late 1960s before levelling off in the 1980s. Interestingly, it seems as if neo-conservative policies reducing the scope of the welfare state, such as those implemented in the Anglo-Saxon countries during the 1980s, also feature some persistence. For instance, subsequent Labour governments have not yet reversed Mrs. Thatcher's reforms in the UK.

This Paper proposes a positive theory of the rise and persistence of redistributive policies. There are three aspects that can account for the existence of a welfare state. The first is that a government can deliver the insurance that missing markets fail to provide (e.g., unemployment insurance). The second is that welfare state institutions promote activities that would be underprovided in a *laissez-faire* society, due to the existence of externalities or capital market imperfections (e.g. education). The third is that of demands for redistribution driven by the conflict of interests between different groups in society. This Paper focuses on the last of these three, ignoring the other two.

In particular, we construct a model where welfare state institutions do not promote economic performance, but are, instead, distortionary. Still, some agents may have a strong interest in demanding redistribution. The analysis aims at enlightening the circumstances under which this political demand for redistribution will prevail over the concern about distortionary effects. The environment is one where individuals act on two qualitatively different arenas. Privately, they behave atomistically, and choose the most profitable action, taking the actions of other agents as given. Collectively, when voting on a common policy for everyone, they take into account the effect of their decisions on private and political behaviour.

Decisions on the two arenas, private and public, influence each other in a dynamic way. Solving explicitly for models which account for dynamic strategic interactions without resorting to *ad hoc* assumptions (e.g. partial myopia or once-and-for all voting) or numerical techniques has proven to be a formidable task in the existing literature. Our model is analytically tractable and enables us to overcome the technical difficulties encountered by the previous literature. This will facilitate empirical testing of the models in future

work. We therefore regard the methodological contribution as part of the value added of the Paper.

A first main finding is that the political mechanism can sustain the welfare state over time. In particular, if the economy starts with a pro-redistribution majority, high levels of redistribution will be sustained over time, whereas there will never be a welfare state if the economy starts with an anti-redistribution majority. Moreover, if a temporary shock creates a temporary political majority in favour of redistribution, the model predicts that the support for the welfare state will continue when the effect of the shock has vanished. If, for instance, a large proportion of the current electorate is impoverished by a temporary aggregate shock, a political majority in favour of the welfare state materializes. Our theory predicts that the institutions promoted by this majority can survive, as they reinforce their own constituency over time. This can be regarded as a parable of the Great Depression, when large-scale welfare programmes were first initiated in a number of Western countries. Similar arguments can be made for other events having triggered changes in redistribution policies, such as the universal suffrage, the European reconstruction after World War II or the sudden increase of the threat of communism in the early 1950s.

A second finding is that, under some parameter restrictions, there exists another equilibrium featuring the permanent dismantling of an existing welfare state. This equilibrium hinges on the rational behaviour of voters in repeated elections, and would not exist if agents myopically took future election outcomes as exogenous. In this equilibrium, an initial pro-welfare state majority strategically restrains its demand for redistribution, in order to induce a future anti-welfare state majority. The expectation that the welfare state will vanish strengthens the incentives of the young to invest, which broadens the tax base and reduces current taxes. In conclusion, our theory predicts that welfare state institutions create their support over time, but that this support is potentially fragile, and can be disrupted forever by even a single episode.

One interesting implication of our model is that an increase in the wage premium weakens the political support for the welfare state. The reason is that agents then have stronger private incentives to invest in education, thereby increasing the mass of the upper tail of the income distribution. This, in turn, implies a reduction in the constituency of redistributive policies. This prediction is consistent with the observations that, first, welfare state policies were sustained and expanded in the post-war period in times characterized by a narrowing wage structure. Second, conservative governments proposing drastic reductions in social policies, especially in Anglo-Saxon countries, were elected and re-elected in the 1980s, a decade associated with a significant increase in both wage inequality and educational investment. Finally, political changes were less dramatic in continental European countries, where the increase in wage inequality was smaller than in the UK and the US.

1 Introduction

The rise of the welfare state in industrialized countries in the 20th century has led to an unprecedented change in the size and scope of governments. These developments were, to a large extent, triggered by specific historical episodes, such as the Great Depression or World War II, when large masses of people were impoverished, thereby creating a demand for public intervention. For instance, between 1929 and 1934, government spending as a fraction of GNP doubled in the US as well as in the major European countries, e.g., France and Italy (see Tables 1 and 2). A change of the same magnitude occurred in the UK in the 1930s, although most of the expansion took place after 1934. Government programs, however, did not diminish after the economies had recovered from the shocks. Instead, redistributive policies persisted and were further expanded in the late 1960s before leveling off in the 1980s. Interestingly, it seems as if neoconservative policies reducing the scope of the welfare state, such as those implemented in the Anglo-Saxon countries during the 1980s, also feature some persistence. Mrs. Thatcher's reforms in the UK, for instance, have not yet been reversed by subsequent Labor governments.[1]

In this paper, we argue that the persistence of redistributive programs is the natural outcome of a dynamic political game. This persistence might imply the indefinite survival of an inefficient level of redistribution. To substantiate this claim, we construct a Downsian majority voting model where agents vote repeatedly over the level of redistribution. We consider economies populated by two-period lived agents who are *ex-ante* identical, but *ex-post* heterogeneous. The overlapping generation structure is intended to capture the idea that, as life goes by, uncertainty about lifetime income is resolved. While young individuals are born identical and have common *ex-ante* preferences, the old individuals have heterogeneous preferences for redistribution, since *ex-post*, the resolution of uncertainty has turned some of them into high-income ("successful") individuals and others into low-income ("unsuccessful") individuals.

A key assumption is that young individuals can affect their chances of becoming successful by making a private (human capital) investment when young. The optimal investment

[1] Table 1 reports a variety of measurements of US government activity. The first column reports the GNP share of government purchases of goods and services. The second column reports total government outlays, including transfers, net of defense, while the third and fourth columns focus on social expenditure that has a major redistributive effect, but excludes social security. In all cases, government programs exhibit large increases (about 100%) between 1927 and 1932, followed by a constant share until 1965, a further expansion between 1965-75, and thereafter a leveling off. Table 2 shows that similar trends are present in the major Western European countries. Note that Table 2 reports final government expenditure as a share of GNP, excluding transfers, as comparable international figures for the early 20th century are not available.

is negatively affected by the extent of the redistribution, which is set period-by-period in political elections. Voters are fully rational, and take into account the effects of policies on current investments and on the future distribution of voters.

The purpose of this paper is to analyze whether the *ex-post* conflict over redistribution can, by itself, lead to the perpetual survival of the welfare state. To this end, we assume that individuals are risk-neutral, abstracting from a standard alternative motivation for the welfare state, i.e., that a government can deliver the insurance that missing markets fail to provide. Our assumption of risk neutrality and the fact that redistribution is distortionary, implies that the welfare state would not survive if the future path of redistribution were set *ex-ante* by a utilitarian planner attaching any arbitrary sequence of positive weights on current and future generations. In this sense, the survival of a welfare state would constitute a "political failure" (as defined by Besley and Coate (1998)). Similarly, there would be no welfare state if young agents could commit to vote in a particular way in the future. However, as such commitments are not feasible in democratic systems, *ex-post* conflicts influence political outcomes.

A first main finding is that the political mechanism can sustain the welfare state. In particular, if the economy starts with a pro-redistribution majority, high levels of redistribution will be sustained over time, whereas there will never be a welfare state if the economy starts with an anti-redistribution majority. Moreover, if a temporary shock (e.g., the Great Depression) creates a temporary political majority in favor of redistribution, the model predicts that the support for the welfare state will continue when the effect of the shock has vanished. More formally, the model features multiple steady-states, due to a self-reinforcing mechanism linking private and collective choices, to which we refer as *policy-behavior complementarity*. The point is that a high provision of a particular policy today affects individual behavior in a way that increases the future constituency for that policy. In our case, high current redistribution reduces investments, implying that a larger share of future voters will benefit from redistributive policies. In our model, however, there is also a negative feed-back in that high redistribution reduces the future tax-base. This increases the future cost of financing redistribution, and diminishes, ceteris paribus, future demand for redistribution. When the policy-behavior complementarity effect prevails, multiple steady-states can exist. In related papers, Hassler et al. (1999, 2001), we explore other examples of this mechanism in settings where unemployment interacts with the provision of unemployment insurance.

A second finding is that, under some parameter restrictions, there exists another equilibrium featuring the permanent dismantling of an existing welfare state. This equilibrium hinges on the rational behavior of voters in repeated elections (as laid out in Krusell, Quadrini and Ríos-Rull (1997)), and would not exist if agents myopically took future election outcomes as exogenous. In this equilibrium, an initial pro-welfare state majority strategically restrains its demand for redistribution, in order to induce a future anti-welfare state majority. The expectation that the welfare state will vanish strengthens the incentives of the young to invest, which broadens the tax-base and reduces current taxes.

One interesting implication of our model is that an increase in the wage premium weakens the political support for the welfare state. The reason is that agents then have stronger private incentives to invest in education, thereby increasing the mass of the upper tail of the income distribution. This, in turn, implies a reduction in the constituency of redistributive policies. This prediction is consistent with the observations that, first, welfare state policies were sustained and expanded in the post-war period in times characterized by a narrowing wage structure (see Goldin and Margo (1992)). Second, conservative governments proposing drastic reductions in social policies, especially in Anglo-Saxon countries, were elected and re-elected in the 1980's, a decade associated with a significant increase in both wage inequality and educational investment (see Katz and Murphy (1992)). Finally, political changes were less dramatic in Continental European countries, where the increase in wage inequality was smaller than in the U.K. and the U.S.

Several papers have analyzed the political economy of redistribution. A classic example is Meltzer and Richard (1981), who construct a static model where the level of redistribution is determined by the median voter and taxes distort labor supply. Salient aspects of voting and the distortions created by redistribution are inherently dynamic, however. Earlier dynamic models of redistribution either assumed myopic voting behavior, as in Alesina and Rodrik (1994), or focused on the case with no strategic interaction between voters at different dates (e.g., Persson and Tabellini, 1994 and Benabou, 1996). An alternative route has been to incorporate repeated voting with strategic interactions, but relying on numerical analysis (e.g., Krusell and Ríos-Rull (1996, 1999), Krusell, Quadrini and Ríos-Rull (1996), and Bassetto (1999)). In contrast, we provide analytical characterizations of Markov perfect equilibria when voters use current political decisions to manipulate future political outcomes. To the best of our knowledge, the only paper working out an analytical solution to the Markov perfect equilibria of a dynamic political economy model is Grossman and Helpman (1998). They analyze the political determination of redistribution in a growth model with overlapping generations, lobbies and an AK technology. While they focus on intergenerational redistribution only, we focus on both intra and intergenerational redistribution. More important, agents make no private decisions in their model, and thus, there is no feedback between public policy and individual behavior, a mechanism central to our analysis. Finally, their model exhibits equilibrium indeterminacy since, in equilibrium, the government is indifferent to the level of redistribution, due to linear preferences and technology. This limits the scope for deriving positive and normative implications.

A number of previous contributions have explored cases of persistent political failures. In Coate and Morris (1999), as in our paper, the introduction of a particular policy affects individual behavior in a way that sustains future demand for this policy. Their analysis differs substantially from ours, however. First, the authors focus on the possibility for special interest groups to bribe politicians, whereas the majority of electors are constantly opposed to the subsidy, which implies that there are no dynamic voting interactions. Second, in their paper, persistence hinges on the existence of relocation costs. Absent these, the steady state would be unique. In a growth setting with repeated voting, Krusell and Ríos-Rull (1996) construct a model where agents having made irreversible investments in learning a specific technology oppose the introduction of new technologies by advocating barriers. They can, however, obtain numerical solutions only.

In a series of papers, Acemoglu and Robinson provide a positive theory explaining the emergence of the modern welfare state. In Acemoglu and Robinson (2000*b*), they document that general suffrage was accompanied by a substantial increase in redistribution policies during the 19th century. The reason why dominant classes extended franchise despite knowing that this would imply increased redistribution, is that this provided a commitment device for continued redistribution in return for less social unrest. While providing a convincing explanation of earlier events, their paper does not address more recent episodes where the welfare state expanded when universal suffrage had already been implemented. Their argument can be regarded as complementary with ours.[2] Acemoglu and Robinson (2000*a*) study why governments use inefficient instruments to redistribute, arguing that these instruments are chosen so as to credibly ensure that redistribution will be continued in the future and that its constituency will not dry up.

The view that political demand for redistribution may arise as a result of previous redistribution is not new in the literature. For instance, Lindbeck (1997) conjectures that the sustained strong political pressure for redistribution and government spending in Sweden might be due to the fact that government transfers represent the main source of income for a large share of the Swedish electorate. Moreover, Lindbeck (1995) and Lindbeck, Nyberg and Weibull (1999) stress a related reason for policy persistence, namely changes in social norms due to a large mass of agents becoming dependent on safety nets. Our paper shows that persistence can be derived from a more standard political economy mechanism,

[2] There is another dimension in which their paper and ours are complementary; our paper reinforces the credibility of the commitment in Acemoglu and Robinson (2000*b*), since initial redistribution policies would be endogenously perpetuated by a sequence of pro-welfare majorities.

where individual preferences are not affected by policies, but political demand for redistribution changes due to the effect of the welfare state on the distribution of voters. More similar in this respect to our approach is Benabou (2000) who constructs a model where redistribution ameliorates capital market imperfections. In his model, political support for redistribution is high when the efficiency-enhancing effect dominates the purely redistributive one. This occurs when inequality is sufficiently small, and thus, on the one hand, low inequality induces high redistribution. On the other hand, high redistribution sustains low inequality, and hence, multiple steady-states are possible. Apart from the different set-up and objectives of the analysis (his paper is motivated by the observation that inequality and redistribution are negatively correlated across countries), a major difference is that, in his model, there are no strategic voting interactions between generations, which is a key element in our analysis. Another related contribution is Piketty (1995) where multiple steady-states arise as a result of the process of social learning about the trade-off between efficiency and incentives.

Finally, less directly related to our contribution are a series of papers providing positive theories of social security in repeated voting models (Cooley and Soares (1999), Galasso and Ruiz (1999), and Boldrin and Rustichini (2000)). In these papers, intergenerational redistribution is sustained by trigger strategies in infinite horizon games. In contrast, our results would survive in a finite horizon environment.

The plan of the paper is the following. Section 2 describes the model. Section 3 characterizes the political equilibria, and Section 4 explores two extensions of the basic set-up. Section 5 concludes. All proofs are in the appendix.

2 The model

The model economy consists of a continuum of risk-neutral, two-period lived agents. Each generation has a unit mass. All agents are born identical, but the subsequent earnings are stochastic. "Successful" agents earn a high wage, normalized to unity, in both periods of their life, whereas "unsuccessful" agents earn a low wage, normalized to zero. At birth, each agent undertakes a costly investment, thereby 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.[3]

The dynamics of redistribution from successful to unsuccessful agents is the focal point

[3]It is important for the analysis that agents earn income in both periods. The assumption that first and second period income are perfectly correlated is, however, not essential – the qualitative results will be preserved provided that earnings in the two periods are positively correlated.

of the paper. Each period, a transfer $b \in [0,1]$ to each low-income agent is determined, financed by collecting a lump-sum tax τ, and the government budget is assumed to balance every period. The transfer, and the associated tax rate, are determined before the young agents decide on their investment. By assumption, we rule out age-dependent taxes and transfers.

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

$$\begin{aligned}
\tilde{V}^{os}(b_t, b_{t+1}, \tau_t) &= 1 - \tau_t \\
\tilde{V}^{ou}(b_t, b_{t+1}, \tau_t) &= b_t - \tau_t \\
\tilde{V}^{y}(e_t, b_t, b_{t+1}, \tau_t, \tau_{t+1}) &= e_t(1+\beta) + (1-e_t)(b_t + \beta b_{t+1}) - e_t^2 - \tau_t - \beta\tau_{t+1},
\end{aligned} \tag{1}$$

where $\tilde{V}^{os}$, $\tilde{V}^{ou}$, and $\tilde{V}^{y}$ denote the objective of old successful, old unsuccessful, and young agents, respectively. $\tilde{V}^{y}$ is computed prior to individual success or failure and $\beta \in [0,1]$ is the discount factor. It is straightforward to show that the solution to the optimal investment problem of the young, given b_t and b_{t+1}, is $e_t^*(b_t, b_{t+1}) = (1 + \beta - (b_t + \beta b_{t+1}))/2$.

Since agents are *ex-ante* identical, agents of the same cohort choose the same investment, which implies that the proportion of old unsuccessful in period $t+1$ is given by

$$u_{t+1} = 1 - e_t^* = \frac{1 - \beta + b_t + \beta b_{t+1}}{2}. \tag{2}$$

Thus, the future proportion of old unsuccessful depends on benefits in period t and $t+1$. To balance the budget, tax revenues must amount to $2\tau_t = (u_t + u_{t+1})\, b_t$, yielding

$$\tau_t = \frac{1 - \beta + b_t + \beta b_{t+1} + 2u_t}{4} b_t. \tag{3}$$

By substituting for τ_t and e_t^* in equations (1), the indirect utility functions can be written as:

$$\begin{aligned}
V^{os}(b_t, b_{t+1}, u_t) &= 1 - \frac{(1-\beta) + (b_t + \beta b_{t+1}) + 2u_t}{4} b_t, \\
V^{ou}(b_t, b_{t+1}, u_t) &= b_t - \frac{(1-\beta) + (b_t + \beta b_{t+1}) + 2u_t}{4} b_t, \\
V^{y}(b_t, b_{t+1}, b_{t+2}, u_t) &= \frac{(1+\beta)^2}{4} + \frac{(1-\beta) - 2u_t}{4} b_t - \frac{b_{t+1} + \beta b_{t+2}}{4} \beta b_{t+1}.
\end{aligned} \tag{4}$$

Note that the marginal tax cost of redistribution in period t, given by $(1 - \beta + b_t + \beta b_{t+1} + 2u_t)/4$, increases in u_t (because more old agents are benefit recipients) and in b_t and b_{t+1} (because more young agents become unsuccessful). Since the old in period t cannot enjoy benefits in period $t+1$, their utility is decreasing in b_{t+1}.

The old successful agents obviously prefer zero benefits, since redistribution implies positive taxes without providing 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 positive benefits lead to positive (negative) intergenerational redistribution from the old to the young, if the number of old unsuccessful is sufficiently small (large). Holding future benefits constant, the young therefore prefer positive redistribution if and only if $u_t < (1-\beta)/2$.

Before proceeding to the main analysis, we note that any Pareto efficient allocation is characterized by zero redistribution in every period except, possibly, in the first period.[4] The reason for this is that redistribution distorts the effort choice of the young, but has no insurance value as agents are risk-neutral.

3 Political equilibrium

The purpose of this paper is to explore the impact of the *ex-post* conflict of interest between groups on the dynamics of redistribution. More specifically, can an "inefficient" welfare state survive over time? In answering this question, we restrict the attention to Markov perfect equilibria, where the state of the economy is summarized by the proportion of current unsuccessful old agents (u_t). The political equilibrium is defined as follows.

Definition 1 *A (Markov perfect) 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,* $b_t = B(u_t)$, *and* $U : [0,1] \to [0,1]$ *is a private decision rule,* $u_{t+1} = 1 - e_t = U(b_t)$, *such that the following functional equations hold:*

1. $B(u_t) = \arg\max_{b_t} V(b_t, b_{t+1}, b_{t+2}, u_t)$ *subject to* $b_{t+1} = B(U(b_t))$, $b_{t+2} = B(U(B(U(b_t))))$, *and* $b_t \in [0,1]$, *and* $V(b_t, b_{t+1}, b_{t+2}, u_t)$ *is defined as the indirect utility of the current decisive voter.*

2. $U(b_t) = (1 - \beta + b_t + \beta b_{t+1})/2$, *with* $b_{t+1} = B(U(b_t))$.

[4] More formally, we define the class of Pareto optimal sequences of benefits, $\{b_t\}_{t=1}^{\infty}$, as those which would be chosen by a social planner whose objective function is given by

$$\max_{\{b_t\}_{t=1}^{\infty}} \left\{ \lambda_{0s}(1-u_1) V^{os}(b_1, b_2, u_1) + \lambda_{0u} u_1 V^{ou}(b_1, b_2, u_1) + \sum_{t=1}^{\infty} \lambda_t V^y(b_t, b_{t+1}, b_{t+2}, u_t) \right\},$$

subject to $b_t \in [0,1]$ $\forall t$, where the planner weights λ_{0s}, λ_{0u}, and $\{\lambda_t\}_{t=0}^{\infty}$ are positive and satisfy $\sum_{t=0}^{\infty} \lambda_t + \lambda_{0s} + \lambda_{0u} = 1$. It is straightforward to show that the planner would choose zero benefits after the first period, for any arbitrary sequence of (positive) planner weights. Moreover, a utilitarian planner with equal weights on all initially living individuals would set $b_0 = 0$, whatever u_0. The proof is available upon request.

The first equilibrium condition requires that b_t maximizes the objective function of the decisive (median) voter 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 future benefits to be set according to the function $b_{t+j} = B(u_{t+j})$. Then, we require that the same function $B(u_t)$ defines optimal benefits today.

The second equilibrium condition implies 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. In our particular model, however, u_t has no direct effect on the investment choice of the young. Thus, we focus on equilibria where their equilibrium investment choice is fully determined by the current benefit level.

3.1 Dictatorship

It is expositionally convenient to start the analysis by describing the equilibrium under the assumption that the political power permanently rests in the hands of one of the two groups of old agents in the society. Then, we extend the analysis to the case of majority voting.

We define "plutocracy" (PL) and "dictatorship of the proletariat" (DP), as the regimes where the level of redistribution is chosen at the beginning of each period by the currently living successful and unsuccessful old agents, respectively. Formally, under DP, $V(b_t, b_{t+1}, b_{t+2}, u_t) \equiv V^{ou}(b_t, b_{t+1}, u_t)$, whereas, under PL, $V(b_t, b_{t+1}, b_{t+2}, u_t) \equiv V^{os}(b_t, b_{t+1}, u_t)$. The equilibrium under dictatorship is characterized in the following Proposition:[5]

Proposition 1 *The PL equilibrium, $\langle B^{pl}, U^{pl} \rangle$, is characterized as follows;*

$$\begin{aligned} B^{pl}(u_t) &= 0 \\ U^{pl}(b_t) &= \frac{1-\beta+b_t}{2}. \end{aligned} \tag{5}$$

[5] To give a hint of how we have obtained these and all subsequent policy functions, start by assuming B to be linear in u_t, ignoring the constraints that $b \in [0,1]$. Then, as the young are risk-neutral with quadratic effort costs, the function U, satisfying condition 2 in equilibrium definition 1, is linear in b_t. Moreover, the indirect utility is also linear-quadratic in b_t, once B and U have been substituted into (4). It turns out that, in the absence of constraints, the optimal choice of b_t is indeed linear in u_t. Imposing condition 1 in equilibrium definition 1, it is straightforward to solve for the coefficients in B. What remains is to impose the constraints on b_t, and check that no deviations from this constrained linear rule can be optimal. Such deviations may in some cases be optimal, as e.g. explained in footnote 6, in which case the policy rule B must be modified in a non-trivial way.

Given any u_0*, for all* $t \geq 1$, $u_t = u^{pl} \equiv \frac{1-\beta}{2}$.
The DP equilibrium, $\langle B^{dp}, U^{dp} \rangle$*, is characterized as follows;*

$$B^{dp}(u_t) = \begin{cases} \frac{3}{2} - u_t & \text{if } u_t > \bar{u}(\beta) \\ \frac{3(2+\beta)-\beta^2}{4-\beta^2} - \frac{2}{2-\beta} u_t & \text{if } u_t \in \left[\frac{3}{2} - \frac{2}{2+\beta}, \bar{u}(\beta)\right] \\ 1 & \text{if } u_t \in \left[0, \frac{3}{2} - \frac{2}{2+\beta}\right] \end{cases} \tag{6}$$

$$U^{dp}(b_t) = \begin{cases} \frac{\beta(1+\beta)+2}{2(2+\beta)} + \frac{2-\beta}{4} b_t & \text{if } b_t \in \left[\frac{2\beta}{2+\beta}, 1\right] \\ \frac{1+b_t}{2} & \text{if } b_t \in \left[0, \frac{2\beta}{2+\beta}\right) \end{cases} \tag{7}$$

where $\bar{u}(\beta) = \min\left\{\frac{\beta+6-\beta\sqrt{4-2\beta}}{2(2+\beta)}, 1\right\}$*. The equilibrium law of motion is as follows;*

$$u_{t+1} = \begin{cases} \frac{5}{4} - \frac{u_t}{2} & \text{if } u_t > \bar{u}(\beta) \\ \frac{1}{4}\left(5 + \frac{\beta^2}{2+\beta}\right) - \frac{u_t}{2} & \text{if } u_t \in \left[\frac{3}{2} - \frac{2}{(2+\beta)}, \bar{u}(\beta)\right] \\ \frac{\beta}{4} + \frac{2}{2+\beta} & \text{if } u_t \in \left[0, \frac{3}{2} - \frac{2}{(2+\beta)}\right] \end{cases} \tag{8}$$

Given any u_0*, the economy converges (with an oscillatory pattern) to a unique steady-state, such that:*

$$u = u^{dp} = \left(5 + \beta^2/(2+\beta)\right)/6, \tag{9}$$

$$b = b^{dp} = \frac{4}{3}\frac{1+\beta}{2+\beta}. \tag{10}$$

Figure 1 represents the equilibrium policy function and law of motion for the PL and DP equilibrium. In the PL case (upper panels), both functions are constant at $B(u_t) = 0$ and $u_{t+1} = (1-\beta)/2$, respectively. In the DP case (lower panels), the equilibrium redistribution is always positive, and is 100% as long as $u_t \leq \frac{3}{2} - \frac{2}{(2+\beta)}$. Note that in steady state, benefits are smaller than 100%, irrespective of β.[6]

3.2 Majority voting

We now assume that political decisions are taken through majority voting. Agents vote on the single issue of redistribution. It is straightforward to show that if all living agents

[6] Figure 1 represents a parametric case where $\bar{u}(\beta) = 1$. This occurs if $\beta \leq \left(\sqrt{17}-1\right)/4 \approx 0.78$. If, in contrast, $\bar{u}(\beta) < 1$, there is a range $u_t \in [\bar{u}, 1]$ such that the unsuccessful old in period t induce a u_{t+1} where the constraint $b_{t+1} \leq 1$ is binding. This creates a downward discontinuity of $B(u_t)$ at $\bar{u}$. The range $[\bar{u}, 1]$ is ephemeral, in the sense that in equilibrium, $u_t < \bar{u}$ $\forall t > 0$. Therefore, if the economy starts at $u_0 < \bar{u}$, the policy rule and law of motion are qualitatively identical to those in Figure 1. Otherwise, they differ for just one period. To understand the origin of the downward discontinuity, note that b_{t+1} is negatively related to b_t when the constraint $b_{t+1} \leq 1$ is not binding. Thus, the marginal distortion of current benefits shifts up at $\bar{u}$ as the constraint on b_{t+1} becomes binding. Therefore, $B(u_t)$ falls discontinuously at $u_t = \bar{u}$.

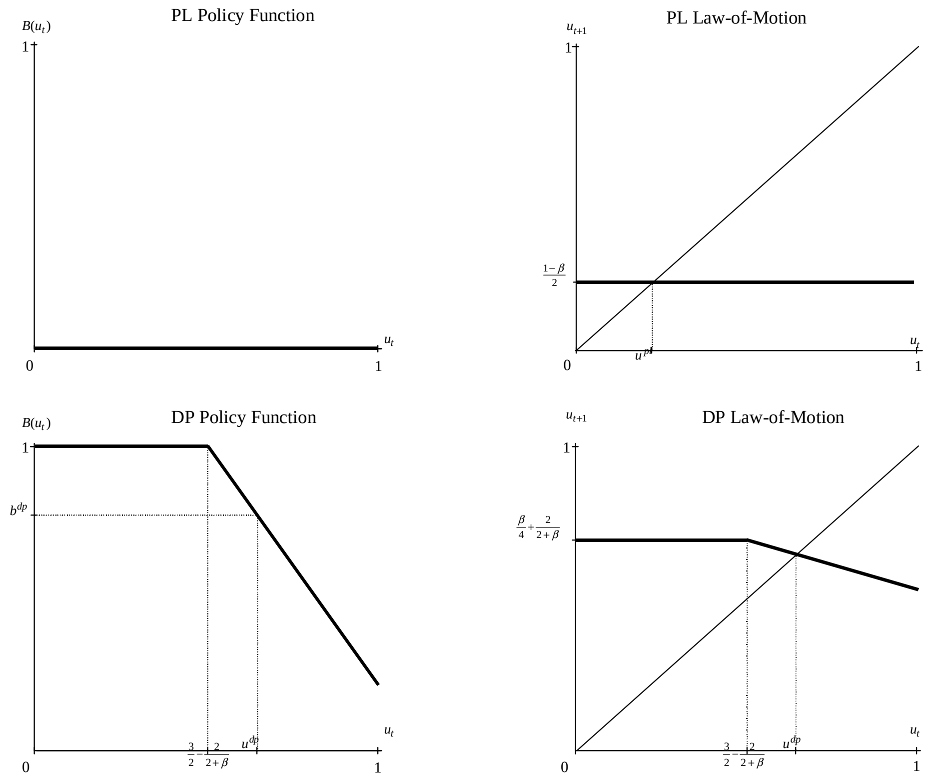

Figure 1: Equilibria under dictatorship; PL and DP

were to vote over current benefits, then, for all $t > 0$, benefits would be zero in equilibrium. Intuitively, as the young are still behind the veil of ignorance, they oppose distortionary redistribution.[7] Moreover, they are always pivotal, since they represent 50% of the population and the old successful oppose any redistribution.

However, we regard the assumption that voters behind the veil of ignorance are pivotal as unrealistic.[8] Thus, this paper explores the consequences of letting the political decisions be determined by the *ex-post* conflict of interests between individuals who know their type, i.e. the old. Two alternative assumptions can deliver a political preponderance of the old in our model. The first is to assume that young individuals have a lower voting turnout than the old, maintaining the benchmark setting that current benefits are set at the beginning of each period. This assumption can be defended empirically as the voting turnout is increasing with age. For example, Wolfinger and Rosenstone (1980) document that turnout in U.S. elections is sharply increasing in age, rising from 45% for the 20-years old to 75% for the 65 years old.[9] The alternative assumption is that elections are held at the end of each period, and agents vote over next period benefits. In particular, the young vote after the uncertainty about individual success has unraveled. In this case, the old, who will not be alive in the next period, have no interest at stake and are assumed to abstain from voting. Clearly, this is equivalent to assuming that the choice of current benefits is taken at the beginning of each period but only the old vote. We will treat this case as our benchmark.[10] As an extension, we explore the case when the young also vote, albeit with a lower turnout (and elections are held at the beginning of each period).

In the benchmark case, benefits maximize the indirect utility of the old successful (unsuccessful) if $u_t \leq 1/2$ ($u_t > 1/2$). As we shall see, majority voting can generate persistence in the equilibrium choice of redistribution. If the economy starts with a pro-welfare state majority ($u_t > 1/2$), there exists an equilibrium where the welfare state survives over time.

[7] Formally, the equilibrium features $B(u_t) = 0\ \forall u_t \geq \frac{1-\beta}{2}$ and $B(u_t) = 1$ otherwise. The economy attains the steady state, $u = \frac{1-\beta}{2}$, after two periods at most. Given this policy rule, the young want positive benefits whenever $u_t < \frac{1-\beta}{2}$, since, in this case, benefits redistribute from the old to the young. The set of states $u_t \in [0, \frac{1-\beta}{2})$ is ephemeral, however, and can never be attained after the initial period.

[8] In a model where agents live for more than two periods and make their investment decision only in the first period, for instance, the "young" would constitute a small proportion of the electorate and would most likely not be decisive. We have, however, not obtained analytical results in a model where agents live for more than two periods.

[9] Mulligan and Sala-i-Martin (1999) argue, with the aid of an interest group model, that the elderly have a preponderant weight on redistribution policies. In their paper, this arises due to the old having a low opportunity cost of time, which increase their willingness to lobby.

[10] For expositional ease, we maintain in the presentation the interpretation that agents vote over current benefits and only the old vote.

Conversely, if $u_t \leq 1/2$, the welfare state will never arise. The positive feedback mechanism giving rise to this policy persistence – policy-behavior complementarity – is that high (low) benefits today affect private incentives so as to induce a large (small) proportion of unsuccessful agents tomorrow, and therefore a broad (narrow) future constituency for redistribution.

There is, however, an important asymmetry between these two steady-states. The welfare state will never arise when there is an initial majority of successful agents, irrespective of the discount factor. In contrast, an initial majority of unsuccessful does not guarantee the eternal survival of the welfare state. In particular, the survival of a welfare state is the unique equilibrium only if the discount factor is sufficiently low. For higher discount factors, and given an initial majority of unsuccessful, there exist both an equilibrium where the welfare state survives forever, and an equilibrium where any existing welfare state is dismantled within two periods. The survival of the welfare state is in this case a matter of (self-fulfilling) expectations. If, on the one hand, the young believe that the welfare state will not survive, the ruling unsuccessful old agents will find it optimal to vote for moderate benefits and induce a future political majority that will vote for zero redistribution. The initial expectations of the young are thus fulfilled in equilibrium. On the other hand, if the young believe that the welfare state will survive, it is optimal for the old to vote for a higher level of redistribution. This, in turn, sustains a constituency for the welfare state and fulfills the expectations.

In Section 3.2.1 we characterize an equilibrium where the welfare state survives in the long-run, provided that $u_0 > 1/2$. In Section 3.2.2 we characterize equilibria where the welfare state disappears after one or two periods even if $u_0 > 1/2$.

3.2.1 Pro-welfare equilibrium

The following Proposition establishes the existence of an equilibrium with history dependent multiple steady-states. If the initial share of unsuccessful is higher than 1/2, the welfare state survives forever, otherwise it never arises. Essentially, the equilibrium functions B^{mv} and U^{mv} are found by splicing together the equivalent functions from the equilibrium under dictatorship in Proposition 1 (i.e., B^{pl} and B^{dp}, and U^{pl} and U^{dp}).

Proposition 2 *For any $\beta \in [0,1]$, there exists an equilibrium featuring multiple steady-states, $\langle B^{mv}, U^{mv} \rangle$, with the following characteristics:*

1.

$$B^{mv}(u_t) = \begin{cases} B^{dp}(u_t) & \text{if } u_t \in \left(\frac{1}{2}, 1\right] \\ B^{pl}(u_t) & \text{if } u_t \in \left[0, \frac{1}{2}\right] \end{cases} \tag{11}$$

$$U^{mv}(b_t) = \begin{cases} U^{dp}(b_t) & \text{if } b_t \in (0, 1] \\ U^{pl}(b_t) & \text{if } b_t = 0 \end{cases} \tag{12}$$

where $B^{dp}(u_t), B^{pl}(u_t), U^{dp}(b_t)$ and $U^{pl}(b_t)$ are defined in Proposition 1. This implies the following equilibrium law of motion;

$$u_{t+1} = \begin{cases} \frac{5}{4} - \frac{u_t}{2} & \text{if } u_t > \bar{u}(\beta) \\ \frac{1}{4}\left(5 + \frac{\beta^2}{2+\beta}\right) - \frac{u_t}{2} & \text{if } u_t \in \left[\frac{3}{2} - \frac{2}{(2+\beta)}, \bar{u}(\beta)\right] \\ \frac{\beta}{4} + \frac{2}{2+\beta} & \text{if } u_t \in \left(\frac{1}{2}, \frac{3}{2} - \frac{2}{(2+\beta)}\right] \\ \frac{1-\beta}{2} & \text{if } u_t \in \left[0, \frac{1}{2}\right]. \end{cases} \tag{13}$$

2. There are two locally stable steady-states. In particular,

(a) if $u_0 \leq 0.5$, the economy converges in one period to a steady-state equilibrium with $b = b^{pl} = 0$ and $u = u^{pl} = \frac{1-\beta}{2}$.

(b) if $0.5 < u_0 \leq 1$, the economy converges asymptotically to an equilibrium characterized by $b = b^{dp} \equiv \frac{4}{3}\frac{1+\beta}{2+\beta}$ and $u = u^{dp} \equiv \left(5 + \beta^2/(2+\beta)\right)/6$.

Figure 2 depicts the equilibrium policy function and law of motion for the case when $\bar{u}(\beta) = 1$ (i.e., $\beta \leq 0.78$). The left-hand panel shows that, when $u_t \leq 1/2$, then $b_t = 0$ in equilibrium. At $u_t = 1/2$, the policy function increases discontinuously, as the unsuccessful become the decisive group. In fact, for $u_t \in (1/2, 3/2 - 2/(2+\beta)]$, the equilibrium policy function prescribes 100% redistribution.

For $u_t \geq 3/2 - 2/(2+\beta)$, the equilibrium law of motion is downward sloping, reflecting the fact that the marginal cost of redistribution increases, as the current proportion of old successful agents falls. The right-hand panel shows that the law of motion has two fixed points for u, given by u^{pl} and u^{dp}, the steady-states under plutocracy and dictatorship of the proletariat, respectively. For all $u_t \leq 1/2$, the economy converges in one period to u^{pl}. For higher initial u_t, the relationship between u_{t+1} and u_t is downward sloping, implying that the economy converges to the steady-state u^{dp}, with a welfare state following an oscillatory pattern.

Proposition 2 delivers an explanation for the historical persistence of the welfare state. In particular, comparative statics can rationalize some important historical episodes. Suppose that a large proportion of the current electorate is impoverished by a temporary

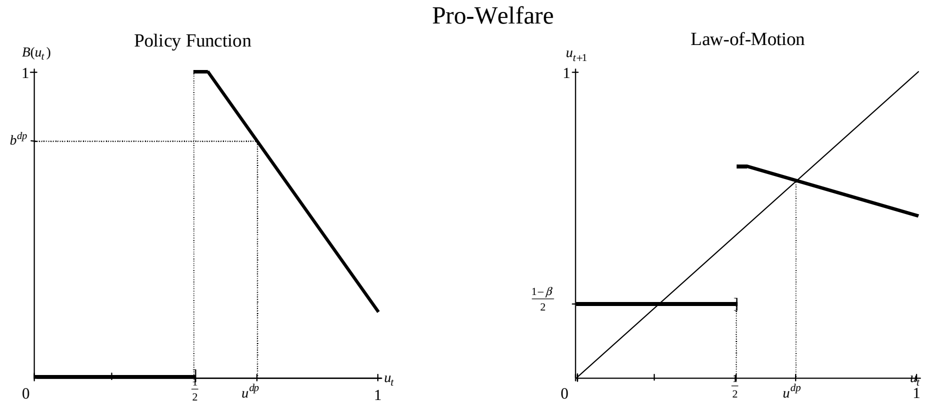

Figure 2: Equilibrium with majority voting; multiple steady-states.

unexpected aggregate shock, so that a pro-welfare political majority is induced. Our model predicts that the institutions promoted by this majority can survive, as they reinforce their own constituency over time. This can be regarded as a parable of the Great Depression, when large scale welfare programs were first initiated in a number of Western countries. Similar arguments can be made for other events having triggered changes in redistribution policies, such as the universal suffrage, the European reconstruction after World War II or the sudden increase of the threat of communism in the early 1950's.

Note that, conditional on a welfare state being in place ($u_t > 0.5$), our model predicts a negative relationship between redistribution and the size of the pro-welfare constituency (see figure 2, in the case of $u_t > 1/2$). This is due to the fact that the marginal cost of redistribution increases with u_t. This prediction is consistent with empirical findings of Razin, Sadka and Swagel (2000) that redistribution, as a fraction of GDP per worker, is decreasing as a function of the dependency ratio in Western European welfare states.

3.2.2 Anti-welfare equilibrium

So far, we have analyzed equilibria where an existing welfare state survives. As discussed above, there also exist rational expectations equilibria where the welfare state eventually breaks down. To understand this, it should be noted that old agents, including the unsuccessful, are strictly better off if there is no welfare state after their death. The reason is that future redistribution, while having no direct effect on the income of the old, distorts incentives for the young and increases current taxes (formally, b_{t+1} enters negatively in (4)). The old unsuccessful are, therefore, tempted to vote strategically so as to change the identity of the future median voter, thereby ensuring that the welfare state disappears. This has a cost, however; a future anti-welfare majority ($u_{t+1} \leq 1/2$) can be triggered only if current benefits are set sufficiently low. The pro-welfare equilibrium characterized in Section 3.2.1 had the property that, in order to induce a switch of majority, current benefits had to be set as low as zero (see (12)). Since $b_t = 0$ is never optimal for the old unsuccessful, the welfare state could not break down. In this section, we show that there exists another equilibrium, sustained by alternative self-fulfilling beliefs, where it is sufficient to set $b_t \leq \beta$ in order to induce the switch of majority and the end of the welfare state.

The source of multiple self-fulfilling expectations is that the equilibrium condition 2 in definition 1, $u_{t+1} = (1 - \beta + b_t + \beta B(u_{t+1}))/2$, may feature multiple solutions for u_{t+1}, provided $B(u_{t+1})$ is (locally) increasing. Since any candidate equilibrium policy function shifts upwards at $u_t = 1/2$, the policy function B is, in fact, locally increasing.[11] Intuitively, if agents believe that benefits next period will be high (low), then u_{t+1} will be large (small), which will fulfill the belief that b_{t+1} is high (low). In the equilibrium described in Proposition 2, agents' beliefs are coordinated on the *largest* of the solutions. We label these beliefs as pro-welfare expectations. Alternatively, agents' beliefs might be coordinated on the *lowest* of the (rational expectation) solutions to equilibrium condition 2. We label these beliefs as anti-welfare expectations. This section focuses on equilibria where agents hold anti-welfare expectations.

In particular, we prove two results. First, if β is sufficiently large ($\beta \geq \beta_M \simeq .555$), there exists an equilibrium featuring the termination of the welfare state in, at most, two periods (Proposition 3). Second, if β is sufficiently low ($\beta < 1/4$), then *no* equilibrium featuring the breakdown of the welfare state can exist (Proposition 4). This result is more general, as it holds under any expectations or class of policy functions. In this range of parameters, the old unsuccessful prefer $b_t > \beta$, an initial majority of unsuccessful regenerates itself and the welfare state is sustained perpetually.

The following Proposition establishes that an equilibrium with the breakdown of the welfare state exists, provided β is sufficiently large.

Proposition 3 *Let $\beta_M \approx 0.555$ be the root in $[0,1]$ to the equation $0 = 2 - \beta_M - 2\beta_M^2 - 2\left(\sqrt{\beta_M}\right)^3$. For all $\beta \geq \beta_M$, there exists an "anti-welfare equilibrium" (AWE) converging to zero redistribution in, at most, two periods.*

The characterization of this equilibrium is conceptually simple. The analytical description varies with β and is somewhat cumbersome. It is therefore deferred to appendix A (Proposition 6). Here, we provide a graphical description of two cases (see figure 3). The

[11] Proof that all equilibrium policy functions are upward discontinuous at $u_t = 1/2$ is available upon request.

first case, described in the two upper panels, corresponds to a range of high discount factors ($\beta > 0.618$). When $u_t \leq 1/2$, the equilibrium prescribes that $b_t = 0$, as before, while $b_t = \beta$, whenever $u_t > 1/2$. This implies that $u_{t+1} = 1/2$ (see the left-hand panel), namely, the old unsuccessful in majority at t moderate their demand for redistribution (i.e., set $b_t = \beta$), in order to induce a majority of successful next period, who, in turn, choose zero benefits. In other terms, the sequence of equilibrium redistribution is $b_0 = \beta$, and $b_{1+j} = 0 \ \ \forall j \geq 0$. Both moderate redistribution at $t = 0$ and the expectation of no redistribution at $t = 1$ induce the young to exert high investment at $t = 0$, thus implying low taxes at $t = 0$.

The second case, described in the two lower panels, corresponds to the range $\beta \in [0.570, 0.618]$. The main change, relative to case 1, is that now there is a range of intermediate levels of initial old unsuccessful, $u \in (1/2, u^a)$, where the proletarian majority chooses 100% redistribution. The sequence of equilibrium redistribution is, in this case, $b_0 = 1$, $b_1 = \beta$, and $b_{2+j} = 0 \ \forall j \geq 0$. The intuition why the welfare state is dismantled in two periods rather than one for lower β's is the following. A lower discount factor weakens the temptation for the old unsuccessful to vote for a benefit level sufficiently low to induce a plutocratic majority the next period (and, consequently, $b_{t+1} = 0$). This is because the distortionary effect of future benefits is falling in β (see equation (3)), and because a smaller β implies that benefits must be lowered for the switch to be feasible. Moreover, a large current tax base ($u_t < u^a$) further reduces the cost of current redistribution. Hence, for $u_t < u^a$, the old unsuccessful prefer maximal benefits, which, in turn, induce $u_{t+1} \geq u^a$ and the breakdown of the welfare state in period $t + 2$.[12]

It is worth noting that the existence of the anti-welfare equilibrium is intrinsically due to the rational forward-looking political behavior of the agents. If, alternatively, agents' voting behavior were myopic, taking the future path of redistribution as given (as in e.g. Alesina and Rodrik (1994)), the breakdown of the welfare state could not be an equilibrium. If, for instance, the ruling unsuccessful expected zero redistribution from next period onwards, irrespective of the current political outcome, they would prefer very high benefits in the current period. But then no anti-welfare majority would materialize, and the initial expectations would not be fulfilled. Hence, the destruction of the welfare state cannot be an equilibrium under myopic voting behavior.

Propositions 2 and 3 imply, jointly, that multiple self-fulfilling equilibria exist when the economy starts with a pro-welfare majority, provided that β is not too low. In one of these

[12] A third case, covered in Proposition 6 in appendix A but omitted from figure 3, arises when $\beta \in [0.555, 0.570]$. This case is qualitatively similar to that described in the lower panels of figure 3, with the only difference that there exists a range of u_t, where the unsuccessful majority prefers $b \in (\beta, 1)$. Nevertheless, it sets benefits high enough to induce a shift of majority in two periods.

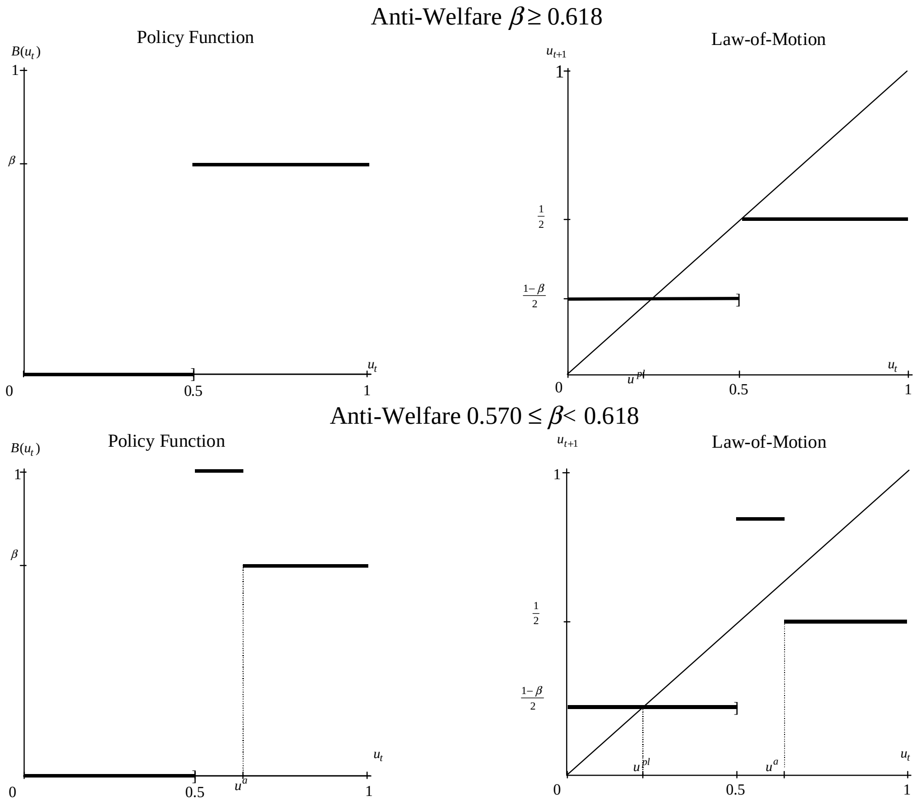

Figure 3: Equilibria under anti-welfare expectations and low rates of discounting

equilibria, the welfare state survives, while it is terminated in the other. As noted above, the source of this multiplicity is that on the one hand, the investment choice of the young depends on the expected future redistribution while, on the other hand, the future political choice of redistribution depends on the current investment of the young. If the young believe that the welfare state will (not) survive, they will choose low (high) investment, implying a large (small) future constituency for the welfare state.

Finally, we can prove a stronger result; when agents discount the future sufficiently, there is no equilibrium featuring the dismantling of an existing welfare state, even if agents hold anti-welfare expectations.

Proposition 4 *Assume that $u_0 > 1/2$ and $\beta < 1/4$. Then, any equilibrium path will feature $u_t > 1/2$ and $b_t > 1/4$, $\forall t \geq 0$.*

This result holds for any set of rational expectations and policy rule (including non-Markovian strategies). The reason is that when $\beta < 1/4$, then $V^{ou}(\beta, 0, u_t) < \max_{b_t \in (\beta,1)} V^{ou}(b_t, 1, u_t)$. The left-hand-side is an upper bound for the value of inducing a breakdown and the right-hand-side is a lower bound for the value of sustaining a pro-welfare majority. We conclude that a strategic termination of the welfare state cannot occur if agents put sufficiently low weight on the future. Hence, for β sufficiently low, *any* equilibrium will feature perpetual survival of the welfare state, given an initial majority of unsuccessful.

4 Extensions.

4.1 Wage inequality and political support for the Welfare State.

In this section, we analyze the effects of changes in productivity inequality on the political equilibrium. To this aim, we extend the model and allow the wage perceived by the successful agents to differ from $w = 1$. Note that the wage w parameterizes the degree of technological inequality (a high w implying high inequality). The average wage level has no effect on the equilibrium of our model, which justifies maintaining the convenient normalization that the wage of the unsuccessful agents is equal to zero. Also, now b_t denotes the *benefit rate*, which implies that unsuccessful agents earn a before-tax income of $b_t w$, and that the constraint $b_t \in [0, 1]$ is maintained.

The optimal effort choice in the extended model implies

$$u_{t+1} = 1 - e_t^* = \max\left\{\frac{2 - (1+\beta)\, w + b_t w + \beta b_{t+1} w}{2}, 0\right\},$$

where the constraint $u_{t+1} \geq 0$ is never binding if $w \leq 1$. The tax rate consistent with the balanced government budget constraint is

$$\tau_t = \frac{\max\{2 - (1+\beta)\,w + b_t w + \beta b_{t+1} w, 0\} + 2u_t}{4} w b_t.$$

It can immediately be established that the political equilibrium necessarily features a welfare state as long as $w < 1/(1+\beta)$. In this case, given any initial u_t, $u_{t+1} > 1/2$ irrespective of the redistribution policy chosen by the first generation, and the unsuccessful are always in majority from the second period onwards. The "dictatorship of proletariat" is, in this case, the unique Markov equilibrium, irrespective of initial conditions or expectations.

For a range of intermediate values of w, the equilibrium features multiple steady-states, and the welfare state never arises if $u_0 \leq 1/2$, while it survives perpetually if $u_0 > 1/2$. An extension of the argument of Proposition 4 ensures that no anti-welfare equilibrium arises for this intermediate range of wage inequality, i.e. a majority of unsuccessful agents do not strategically hand over power to the next generation of successful. Next, for another range of intermediate, larger values of w, there exist both an equilibrium featuring the break-down of the welfare state and one featuring its perpetual survival, as analyzed in Section 3.2.2 (recall that if $w = 1$, multiple equilibria arise provided that β is sufficiently large). Finally, for an open set of sufficiently large values of w, there exists no equilibrium featuring a Welfare State for more than one period.[13]

As this discussion suggests, the model has predictions about the effect of technology-driven changes in wage inequality. Assume, for instance, that there is an unexpected increase in the premium to education a result, agents increase their investment in education, and a larger proportion will, *ex-post*, be against redistributional policies. Thus, the initial impact of technological inequality is magnified by reduced support for the welfare state. This prediction is in line with important events characterizing the last quarter of the 20th century. The skill-biased technical change that, as documented by a number of authors, started in the 1970's, was followed by the success of conservative governments, whose political platform included a reduction of the redistributive role of governments, especially in

[13] To see why the Welfare State cannot survive when w is sufficiently large, consider the case of $\beta = 0$. This is the easiest case for the survival of the Welfare State, since the old unsuccessful in majority have no strategic motive to induce the break-down of the Welfare State. In this case, if $u_0 > 1/2$, the equilibrium benefit rate is given by

$$\arg\max_b \left[b - (\max\{2 - w(1-b), 0\} + 2u)\, b/4\right] \geq \frac{1}{2} + \frac{1-u}{w},$$

that is decreasing in w, so the equilibrium b_0 is bounded from below by $1/2$. But, then, $u_1 = \max\{1 - w(1-b_0)/2, 0\} < 1/2$ for sufficiently large w. Hence, irrespective of the initial value of u, there will be a anti-welfare majority in period one when w is sufficiently large.

Anglo-Saxon countries. It is interesting to observe that not all industrialized countries went through similar political changes, though. This observation is consistent with the argument of our paper for two reasons. First, we predict that multiple self-fulfilling expectations exist. As a matter of fact, the investment in education increased more in the United States than in continental Europe, which is consistent with the expectation that more, respectively, less, redistributive policies will be sustained in future. Second, if other institutions (e.g., unions) compress the wage structure and prevent the productivity differences from giving rise to large wage inequalities, the investment incentives do not change significantly, and the constituency for the welfare state does not dry up in countries where these institutions are established. This can explain why countries experiencing a lower increase in pre-tax inequality also reformed their welfare state institutions less radically.

Finally, in our stylized model, all agents are ex-ante identical and there is no constraint preventing agents from making an educational investment. Thus, in *ex-ante* terms, there are no losers. In reality, agents differ in both ability and the extent to which capital market imperfections restrain their educational choice. If we extend the model in this dimension, it is clear that low-skill or poor agents that cannot match the new situation by increasing their educational effort are destined to suffer from both the increasing relative demand for skills and the induced loss of constituency for the welfare state.

4.2 Participation of young voters

We have shown in the Section 3 that a welfare state can persist when the old vote over current benefits (or, equivalently, when the young vote over next period benefits, after knowing their type). However, if young agents with an ex-ante perspective are pivotal, there can be no persistent welfare state. In this section, we analyze an intermediate case, allowing the young to participate in the political process before knowing their type, albeit with a lower turnout than the old. We denote by $\varepsilon \in [0, 1]$ the share of the young individuals participating in the voting process. The key insight is that our benchmark case, when only the old vote ($\varepsilon = 0$), is not knife-edge. Indeed, an equilibrium featuring the survival of the welfare state can also exist for $\varepsilon > 0$, provided that ε is not too large. Since we want to investigate the robustness of the equilibrium featuring a welfare state, we focus on the case where agents hold pro-welfare expectations and $u_0 > 1/2$. In other words, we explore whether a limited participation of the young makes the equilibrium in Proposition 2 break down.

If a share ε of the young vote, the old unsuccessful are in majority as long as $u_t > (1 - u_t) + \varepsilon$, i.e., if $u_t > (1 + \varepsilon)/2$. Similarly, if $u_t < (1 - \varepsilon)/2$, there is a majority of old successful. It is important to note that the partial turnout of the young affects the political equilibrium, even when the old unsuccessful are in absolute majority. In particular, when the young vote, it becomes more attractive for the old unsuccessful to vote strategically for low benefits so as to induce a switch of majority and the breakdown of the welfare state. The reason is that the young will always vote for zero benefits, which means that it is no longer necessary to set $b_t = 0$ in order to induce a majority voting for zero benefits in the next period, even if agents hold pro-welfare expectations. More formally, in order for the expectation that $b_{t+1} > 0$ (survival of the Welfare State) to be rational, the choice of b_t that maximizes utility for the old unsuccessful at t must now induce $u_{t+1} > (1+\varepsilon)/2$. Hence, for any $u_t > 1/2$, it is necessary that benefits exceed a break-down threshold, i.e., $b_t = B(u_t) > \varepsilon + \beta\left(1 - B^{dp}((1+\varepsilon)/2)\right) \equiv b^{bd}$. When only the old vote, then $b^{bd} = 0$ and any positive choice b_t is consistent with the survival of the Welfare State. If $\varepsilon > 0$, however, the minimum share of old successful required for an anti-redistribution majority decreases. Then, there exists a positive range of b_t triggering a switch of majority, and the pro-welfare equilibrium becomes more fragile. Nevertheless, for a range of sufficiently low ε's, this equilibrium is sustained, and the economy converges to the same steady-state with positive benefits as in the DP-equilibrium. These arguments are formalized in the following Proposition;

Proposition 5 *Suppose that a share ε of the young are allowed to vote under majority voting. Then, if $u_0 \in \left(\frac{1+\varepsilon}{2}, \frac{\beta}{4} + \frac{2}{2+\beta}\right]$ and $\varepsilon \leq \bar{\varepsilon}(\beta)$, there exists an equilibrium such that the economy converges asymptotically to a steady state with $b = b^{dp}$ and $u = u^{dp}$ following the law-of-motion*

$$u_{t+1} = \begin{cases} \frac{1}{4}\left(5 + \frac{\beta^2}{2+\beta}\right) - \frac{u_t}{2} & \text{if } u_t > \frac{3}{2} - \frac{2}{(2+\beta)} \\ \frac{\beta}{4} + \frac{2}{2+\beta} & \text{if } u_t \leq \frac{3}{2} - \frac{2}{(2+\beta)} \end{cases},$$

where

$$\bar{\varepsilon}(\beta) = \begin{cases} \frac{1}{4}\frac{\beta^2+8\beta+4-\sqrt{(\beta^4+34\beta^3+60\beta^2+24\beta)}}{\beta+2} & \text{if } \beta > \beta^y \\ \frac{-\frac{1}{8}\beta^3+\frac{1}{4}\beta^2+\frac{3}{2}\beta+1-\frac{1}{8}\sqrt{(\beta^6+30\beta^5-72\beta^4-80\beta^3+144\beta^2+96\beta)}}{2+\beta} & \text{if if } \beta \leq \beta^y, \end{cases}$$

and $\beta^y \equiv \frac{4}{51}\sqrt[3]{\left(586 + 102\sqrt{33}\right)} + \frac{16}{51\sqrt[3]{\left(586+102\sqrt{33}\right)}} - \frac{26}{51} \simeq .347.$

In the interest of space, Proposition 5 does not cover initial conditions $u_0 \leq (1+\varepsilon)/2$ and $u_0 > \beta/4 + 2/(2+\beta)$. These cases are, however, straightforward to analyze.[14]

[14] If $u_0 < (1-\beta)/2$, the young prefer positive benefits since they generate intergenerational transfers from the old to the young. If β were sufficiently low, this motive would be so strong that the young would vote for $b_t > b^{bd}$. If, in addition, the young were politically pivotal, they would induce a majority of unsuccessful

Note that the image of $\bar{\varepsilon}(\beta)$ is $[0.174, 0.5]$, where $\bar{\varepsilon}$ is decreasing in β. Hence, the welfare state can survive only if the voting turnout of the young is limited (and at most 50% in our model). This result might seem destructive at a first glance. However, a simple and plausible extension of our model would suffice to guarantee the survival of the welfare state for higher degrees of influence of the young, including $\varepsilon = 1$.

In our benchmark model, the cards are stacked against the welfare state, because the young are assumed to be ex-ante homogenous and, therefore, all prefer zero redistribution. Suppose, alternatively, that the young were heterogeneous in ability or effort cost. To be concrete, assume that low-skill workers are not subject to incentive problems, whereas high-skill workers are similar to the agents described in our model. Low-skill workers and unsuccessful high-skill workers receive the low wage (here, normalized to zero), whereas successful high-skill workers are paid the high wage. Then, a political alliance between low-skill workers (both young and old) and unsuccessful old high-skill workers arise, and they can form a majority in favor of the welfare state. In this paper we have, for simplicity, focused on a basic case where workers are ex-ante homogenous but, as this natural and realistic extension shows, our mechanism does not critically hinge on exogenous and unrealistic assumptions about the participation of different groups in the political process.

5 Conclusion

In this paper, we have analyzed the dynamics of redistribution under repeated majority voting, assuming agents to be fully rational and forward-looking. Following previous research, we have restricted the attention to Markov perfect equilibria. In contrast to most previous papers, however, we have obtained an analytical characterization of the equilibria. The key assumption delivering analytical solutions is that agents have linear utility in income and quadratic disutility of effort. Under this assumption, piecewise linear equilibrium policy functions can be found by a standard guess and verify technique.

In our model, redistribution from rich to poor agents has a distortionary effect and agents are risk neutral, attaching no *ex-ante* value to redistribution. Nevertheless, some agents want redistribution *ex-post*, and as we have seen, this may sustain welfare state institutions. Our theory is therefore a complement to the standard explanation of the

---

old individuals in period 1. However, the set of states $u_t \in [0, (1-\beta)/2)$, are ephemeral as benefits cannot be negative. A similar caveat applies for $u_0 > \beta/4 + 2/(2+\beta)$, because $u_t < \beta/4 + 2/(2+\beta)$ along any equilibrium path. In this case, the majority of old unsuccessful will vote for $b_0 = b^{bd}$, inducing a switch of majority and the breakdown of the welfare state (provided that ε and β are large enough). This set of states is also ephemeral under the assumed policy rule.

existence of a welfare state – that a government can deliver the insurance that missing markets fail to provide.

Our main result is that multiple steady-states exist. Then, initial conditions determine the long-run level of redistribution since the existence (non-existence) of a welfare state, created by an initial majority of unsuccessful (successful), distorts private investment decisions in a way that regenerates political support for redistribution. This happens, despite the fact that the cost of financing current redistribution increases with future redistribution and that the welfare state has no intrinsic role. Temporary shocks changing the political conditions can therefore lead to the make or break of a welfare state, and have permanent consequences.

Second, due to forward-looking voting and repeated elections, agents may choose benefits so as to strategically manipulate the identity of the future median voter. In particular, a pro-redistribution majority may induce the future breakdown of an existing welfare state, in order to sharpen the incentives of the currently young and to enjoy lower taxes. Thus, multiple self-fulfilling equilibria, some with the eternal survival and some with the termination of the welfare state, may co-exist.

This paper has focused on simple majority voting as the political mechanism delivering persistent political failure. In this setting, all political decision power rests in the hands of the majority. Using a similar model, Hassler, Storesletten and Zilibotti (2001) explore alternative political mechanisms, such as probabilistic voting with lobbies, featuring more political influence for minorities. They find that even in these alternative settings, the model can feature the perpetual survival of an inefficient welfare state. In future work, we plan to extend our analysis to incorporate an insurance motive for redistribution that interacts with the redistribution motive.

References

Acemoglu, D. and Robinson, J. A. (2000*a*), Inefficient redistribution, Mimeo, Massachusetts Institute of Technology.

Acemoglu, D. and Robinson, J. A. (2000*b*), 'Why did the west extend the franchise? democracy, inequality and growth in historical perspective', *Quarterly Journal of Economics* **115**, 1167–99.

Alesina, A. and Rodrik, D. (1994), 'Distributive politics and economic growth', *Quarterly Journal of Economics* **109**(2), 465–490.

B. R. Mitchell (1975), *European Historical Statistics: 1750-1970*, Columbia University Press, New York, NY.

Bassetto, M. (1999), 'Political economy of taxation in an overlapping-generations economy', *Federal Reserve Bank of Minneapolis Discussion Paper*.

Benabou, R. (1996), 'Inequality and growth', *NBER Macroeconomic Annuals* **1996**, 11–92.

Benabou, R. (2000), 'Unequal societies: Income distribution and the social contract', *American Economic Review* **90**, 96–129.

Besley, T. and Coate, S. (1998), 'Sources of inefficiency in a representative democracy: A dynamic analysis', *American Economic Review* **88**(1), 139–56.

Boldrin, M. and Rustichini, A. (2000), 'Political equilibria with social security', *Review of Economic Dynamics* **3**(1), 41–78.

Coate, S. and Morris, S. (1999), 'Policy persistence', *American Economic Review* **89**(5), 1327–36.

Cooley, T. F. and Soares, J. (1999), 'A positive theory of social security based on reputation', *Journal of Political Economy* **107**(1), 135–60.

Galasso, V. and Ruiz, J. I. C. (1999), Positive arithmetic of the welfare state, CEPR Discussion Paper No. 2202.

Goldin, C. and Margo, R. A. (1992), 'The great compression: The wage structure in the united states at mid- century', *Quarterly Journal of Economics* **107**(1), 1–34.

Grossman, G. M. and Helpman, E. (1998), 'Intergenerational redistribution with short-lived governments', *Economic Journal* **108**, 1299–1329.

Hassler, J., Rodríguez Mora, J. V., Storesletten, K. and Zilibotti, F. (1999), Equilibrium unemployment insurance, CEPR discussion paper 2126.

Hassler, J., Rodríguez Mora, J. V., Storesletten, K. and Zilibotti, F. (2001), *Unemployment, Specialization, and Collective Preferences for Social Insurance.* in D. Cohen, T. Piketty, and G. Saint-Paul (eds.) *The NewEconomics of Rising Inequalities*, Oxford University Press.

Hassler, J., Storesletten, K. and Zilibotti, F. (2001), Repeated voting on redistribution, mimeo Stockholm University.

Katz, L. F. and Murphy, K. M. (1992), 'Changes in relative wages, 1963-1987: Supply and demand factors', *Quarterly Journal of Economics* **107**(1), 35–78.

Krusell, P., Quadrini, V. and Ríos-Rull, J. V. (1996), 'Are consumption taxes really better than income taxes?', *Journal of Monetary Economics* **37**(3), 475–503.

Krusell, P., Quadrini, V. and Ríos-Rull, J. V. (1997), 'Politico-economic equilibrium and economic growth', *Journal of Economic Dynamics and Control* **21**(1), 243–272.

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

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. (1997), 'The Swedish experiment', *Journal of Economic Literature* **35**(3), 1273–1319.

Lindbeck, A., Nyberg, S. and Weibull, J. W. (1999), 'Social norms and economic incentives in the welfare state', *Quarterly Journal of Economics* **114**(1), 1–35.

Meltzer, A. H. and Richard, S. F. (1981), 'A rational theory of the size of government', *Journal of Political Economy* **89**, 914–927.

Mulligan, C. B. and Sala-i-Martin, X. (1999), Gerontocracy, retirement, and social security, NBER Working Paper 7117.

Persson, T. and Tabellini, G. (1994), 'Electoral competition with informed and uninformed voters', *American Economic Review* **84**, 600–21.

Piketty, T. (1995), 'Social mobility and redistributive politics', *Quarterly Journal of Economics* **110**(3), 551–584.

Razin, A., Sadka, E. and Swagel, P. (2000), The ageing population and the size of the welfare state, Mimeo, Tel-Aviv University.

U.S. Bureau of the Census (1975), *Historical Statistics of the United States, Colonial Times to 1970*, GPO, Washington, DC.

Wolfinger, R. E. and Rosenstone, S. J. (1980), *Who votes?*, Yale University Press.

Table 1

Size of US Government as % of GNP

	Government Purchases of Goods & Services	Total Expenditures − Defense	Education + Public Welfare + Income Insurance	Education + Public Welfare
1902	.	6.9	1.4	1.4
1913	.	7.5	1.6	1.6
1922	.	11.4	2.7	2.5
1927	.	11.2	2.8	2.5
1929	8.2	.	.	.
1932	14.0	20.2	5.2	4.8
1936	14.5	19.2	4.5	4.1
1940	14.0	18.9	5.7	4.2
1950	13.3	18.2	5.9	4.4
1955	18.6	16.9	5.3	4.0
1960	19.8	20.3	6.2	4.7
1965	20.0	21.9	6.6	5.3
1970	22.5	25.5	9.1	7.5
1970	22.8	25.1	9.0	7.4
1975	21.7	30.6	11.1	8.0
1980	19.9	30.7	10.3	7.1
1985	20.6	31.1	10.0	6.9
1990	20.2	32.6	10.4	7.7
1995	18.7	34.0	11.5	8.3

Sources: 1902–70: Tables F47–70 and Y533–566 in U.S. Bureau of the Census (1975). 1975–95: the 2000 Economic Report of the President. The post-1970 data are chained so as to match the Census data for 1970. Missing observations are denoted with a ".".

Table 2

Government Consumption Expenditures as % of GDP

	UK	France	Italy	Germany	Sweden
1899	6.2	4.4	8.5	5.8	4.5
1904	7.1	4.2	8.0	5.4	5.4
1909	6.9	4.1	8.2	6.9	5.2
1914	10.5	.	9.0	6.3	4.7
1924	9.0	6.0	12.2	7.0	7.0
1929	8.8	4.1	9.2	9.7	5.7
1934	9.3	9.2	18.7	12.0	7.6
1939	18.5	14.6	17.0	.	8.5
1949	17.8	15.5	14.6	11.2	13.1
1954	18.5	13.3	12.1	14.1	15.2
1959	16.6	14.0	12.1	13.7	16.3
1964	16.5	13.5	14.8	14.8	17.1
1969	17.6	13.6	14.6	15.6	20.9
1974	20.5	15.7	14.0	19.3	23.0
1979	20.1	17.9	14.8	19.7	28.7
1984	21.7	19.9	16.8	20.0	27.8
1989	19.5	22.3	19.8	18.8	26.2
1994	20.1	24.2	19.6	19.7	27.4
1998	18.2	23.5	18.5	19.0	26.7

Sources: 1899–1949: Table H4 in B. R. Mitchell (1975), *Central Government Expenditures.* 1954–98: the International Financial Statistics, *Government Consumption Expenditures* from the National Accounts. The pre-1954 data are chained so as to match the IMF data in 1954.

6 Appendix A. Characterization of the equilibrium with Anti-Welfare expectations.

Proposition 6 *An AWE, $\langle B^{aw}, U^{aw}\rangle$, has the following characteristics:*

1. *For* $\beta \geq \frac{\sqrt{5}-1}{2} \simeq 0.618$*;*

$$B^{aw}(u_t) = \begin{cases} \beta & \text{if } u_t > 1/2 \\ 0 & \text{if } u_t \in \left[0, \frac{1}{2}\right] \end{cases}$$
$$U^{aw}(b_t) = \begin{cases} \frac{1-\beta+\beta^2+b_t}{2} & \text{if } b_t > \beta \\ U^{pl}(b_t) & \text{else} \end{cases}$$

where $U^{pl}(b_t)$ *is defined in Proposition 1. This implies,*

$$u_{t+1} = \begin{cases} \frac{1}{2} & \text{if } u_t \in (0.5, 1] \\ \frac{1-\beta}{2} & \text{if } u_t \in [0, 0.5] \end{cases}$$

2. *For* $\beta \in \left[\beta_H, \frac{\sqrt{5}-1}{2}\right)$*;*

$$B^{aw}(u_t) = \begin{cases} \beta & \text{if } u_t \geq u^a(\beta) \\ 1 & \text{if } u_t \in (0.5, u^a(\beta)) \\ B^{pl}(u_t) & \text{if } u_t \in \left[0, \frac{1}{2}\right] \end{cases}$$
$$U^{aw}(b_t) = \begin{cases} \frac{1-\beta+b_t+\beta^2}{2} & \text{if } b_t > \beta \\ U^{pl}(b_t) & \text{else} \end{cases}$$
$$u_{t+1} = \begin{cases} \frac{1}{2} & \text{if } u_t \in [u^a(\beta), 1] \\ 1 - \frac{\beta(1-\beta)}{2} & \text{if } u_t \in (0.5, u^a(\beta)) \\ \frac{1-\beta}{2} & \text{if } u_t \in [0, 0.5] \end{cases}$$

where $u^a(\beta) \equiv 1 - \frac{\beta^2}{2(1-\beta)}$, *and* $\beta_H \simeq .570$ *is the solution in* $[0,1]$ *to*

$$\beta^3 - \beta^2 + 2\beta - 1 = 0.$$

3. For $\beta \in [\beta_M, \beta_H]$*;*

$$B^{aw}(u_t) = \begin{cases} \beta & if \quad u_t \in \left[u^d(\beta), 1\right] \\ \frac{3}{2} + \frac{\beta(1-\beta)}{2} - u_t & if \quad u_t \in \left[u^c(\beta), u^d(\beta)\right) \\ 1 & if \quad u_t \in (0.5, u^c(\beta)) \\ 0 & if \quad u_t \in [0, 0.5] \end{cases}$$

$$U^{aw}(b_t) = \begin{cases} \frac{1-\beta+b_t+\beta^2}{2} & if \quad b_t > \beta \\ U^{pl}(b_t) & else \end{cases}$$

$$u_{t+1} = \begin{cases} \frac{1}{2} & if \quad u_t \in \left[u^d(\beta), 1\right] \\ \frac{5-\beta(1-\beta)}{4} - \frac{1}{2}u_t & if \quad u_t \in \left[u^c(\beta), u^d(\beta)\right) \\ 1 - \frac{\beta(1-\beta)}{2} & if \quad u_t \in (0.5, u^c(\beta)) \\ \frac{1-\beta}{2} & if \quad u_t \in [0, 0.5] \end{cases}$$

where $u^c(\beta) \equiv \frac{1}{2} + \frac{\beta(1-\beta)}{2}$, $u^d(\beta) \equiv \frac{3}{2} - \frac{1}{2}\beta(1+\beta) - \left(\sqrt{\beta}\right)^3$, $\beta_M \approx 0.555$ *is the root in* $[0, 1]$ *to the equation* $0 = 2 - \beta_M - 2\beta_M^2 - 2\left(\sqrt{\beta_M}\right)^3$, *and* β_H *defined as above.*[15]

7 Appendix B. Proofs.

For notational convenience, the value functions are, in this section, rewritten as follows;

$$\begin{aligned} \hat{V}^j(b_t, u_t) &\equiv V^j(b_t, B(U(b_t)), u_t), \text{ for } j \in \{os, ou\}, \\ \hat{V}^y(b_t, u_t) &\equiv V^y(b_t, B(U(b_t)), B(U(B(U(b_t)))), u_t). \end{aligned}$$

7.1 Proof of Proposition 1

Proof. We must show that the pair $\langle B^i, U^i \rangle$, with $i \in \{pl, dp\}$, satisfy equilibrium conditions 1 and 2 in Definition 1. It is straightforward that, under PL, $\hat{V}^{os}$ is maximized by setting $b_t = 0$ in every period and that, consequently, $u_t = \frac{1-\beta}{2}$ for all $t \geq 1$.

Consider, next, the DP equilibrium. The strategy of the remainder of the proof is to show that, for all t and u_t, $\langle B^{dp}, U^{dp} \rangle$ satisfy

A) $B^{dp}(u_t) = \arg\max_{b_t} \left\{\hat{V}^{ou}(b_t, u_t)\right\}$, subject to $u_{t+1} = U^{dp}(b_t)$, $b_t \in [0, 1]$ and $b_{t+1} = B^{dp}(u_{t+1})$; and

B) $U^{dp}(b_t) = \left(1 - \beta + b_t + \beta B^{dp}\left(U^{dp}(b_t)\right)\right)/2$.

Given $u_{t+1} = U^{dp}(b_t)$, $b_t \in [0, 1]$ and $b_{t+1} = B^{dp}(u_{t+1}) = B^{dp}\left(U^{dp}(b_t)\right) =$

[15] The equation $0 = 2 - \beta_M - 2\beta_M^2 - 2\left(\sqrt{\beta_M}\right)^3$ originates from setting $\beta_M = \bar{b}(\beta_M)$, where $\bar{b}(\beta)$ is defined as the infimum of benefits b_t which will generate $u_{t+1} \geq u^d(\beta)$ and hence $b_{t+1} = \beta$.

$\left(\frac{3(2+\beta)-\beta^2}{4-\beta^2} - \frac{2}{2-\beta}\left(\frac{\beta(1+\beta)+2}{2(2+\beta)} + \frac{2-\beta}{4}b_t\right)\right)$, then $\hat{V}_t^{ou}$ can be expressed as;

$$\begin{aligned}\hat{V}_t^{ou}(b_t,u_t) &= b_t - \frac{(1-\beta)+\left(b_t+\beta B^{dp}\left(U^{dp}(b_t)\right)\right)+2u_t}{4}b_t \\ &= \begin{cases} b_t - \frac{1}{4}\left(1-\beta+b_t+\beta\left(\frac{3(2+\beta)-\beta^2}{4-\beta^2} - \frac{2}{2-\beta}\left(\frac{\beta(1+\beta)+2}{2(2+\beta)} + \frac{2-\beta}{4}b_t\right)\right)+2u_t\right)b_t & \text{if} \quad b_t \geq \frac{2\beta}{2+\beta} \\ b_t - \frac{1}{4}(1+b_t+2u_t)\,b_t & \text{else}\end{cases}\end{aligned}$$

(note that $b_{t+1}=1$ if (and only if) $b_t \leq 2\beta/(2+\beta)$). Maximizing $\hat{V}_t^{ou}$ over b_t yields;

$$b_t = \begin{cases} \frac{3}{2} - u_t & \text{if } u_t > \bar{u}(\beta) \\ \frac{3(2+\beta)-\beta^2}{4-\beta^2} - \frac{2}{2-\beta}u_t & \text{if } u_t \in \left[\frac{3}{2} - \frac{2}{(2+\beta)}, \bar{u}(\beta)\right] \\ 1 & \text{if } u_t \in \left[0, \frac{3}{2} - \frac{2}{(2+\beta)}\right]\end{cases} = B^{dp}(u_t)\,,$$

where $\bar{u}(\beta) = \frac{\beta+6-\beta\sqrt{4-2\beta}}{2(2+\beta)}$. This proves part (A) of the Proposition. To see the steps of this maximization in more details, define $\hat{V}^a(u_t)$ and $\hat{V}^b(u_t)$ as follows;

$$\begin{aligned}\hat{V}^a(u_t) &\equiv \max\left\{\hat{V}_t^{ou}(b_t,u_t)\,|_{b_t\in\left[0,\frac{2\beta}{2+\beta}\right]}\right\} \qquad (14)\\ &= \begin{cases}\frac{9}{16} - \frac{3}{4}u_t + \frac{1}{4}u_t^2 \equiv \hat{V}^{a,int}(u_t) & \text{if} \quad u_t > \frac{6-\beta}{2(2+\beta)} \\ \beta\frac{6+\beta-2u_t(2+\beta)}{2(2+\beta)^2} \equiv \hat{V}^{a,cor}(u_t) & \text{else}\end{cases} \\ \hat{V}^b(u_t) &\equiv \max\left\{\hat{V}_t^{ou}(b_t,u_t)\,|_{b_t\in\left[\frac{2\beta}{2+\beta},1\right]}\right\} \qquad (15)\\ &= \begin{cases}\frac{1}{8}\frac{\left(\beta^2-3\beta+2u_t(2+\beta)-6\right)^2}{(2-\beta)(2+\beta)^2} \equiv \hat{V}^{b,int}(u_t) & \text{if} \quad u_t \geq \frac{3}{2} - \frac{2}{(2+\beta)} \\ \frac{1}{8}\frac{8+\beta(6-\beta)}{2+\beta} - \frac{u_t}{2} \equiv \hat{V}^{b,cor}(u_t) & \text{else}\end{cases}\end{aligned}$$

where $\hat{V}^{a,cor}(u_t)$ and $\hat{V}^{b,cor}(u_t)$ result from corner solutions in the respective ranges (the corners being, respectively, equal to $b_t = \frac{2\beta}{2+\beta}$ and $b_t = 1$). First, standard algebra establishes that $\hat{V}^{b,int}(u_t) - \hat{V}^{a,cor}(u_t) = \frac{1}{8}\frac{\left(\beta^2-2\beta u_t-\beta+6-4u_t\right)^2}{(2-\beta)(2+\beta)^2} > 0$ and that, in the range where $u_t \leq \frac{3}{2} - \frac{2}{(2+\beta)}$, $\hat{V}^{b,cor}(u_t) - \hat{V}^{a,cor}(u_t) > \frac{1}{8}(2-\beta)\frac{4(1-\beta)+\beta^2}{(2+\beta)^2} > 0$. Thus, whenever $\hat{V}^a(u_t) = \hat{V}^{a,cor}(u_t)$, then $\hat{V}^b(u_t) > \hat{V}^a(u_t)$. Second, if $\beta < \frac{2}{3}$, then $\frac{6-\beta}{2(2+\beta)} > 1$ and $\hat{V}^a(u_t) = \hat{V}^{a,cor}(u_t)$ for all u_t. Thus, $\hat{V}^b(u_t) > \hat{V}^a(u_t)$ if $\beta < 2/3$. Third, note that, if $\beta \geq 2/3$, then there exists a range of u_t, where $\hat{V}^a(u_t) = \hat{V}^{a,int}(u_t)$. In this range, standard algebra establishes that $\hat{V}^{b,int}(u_t) > \hat{V}^{a,int}(u_t)$ for all u_t provided that $\beta < \left(\sqrt{17}-1\right)/4$. Thus, $\beta < \left(\sqrt{17}-1\right)/4$ implies that $\hat{V}^b(u_t) > \hat{V}^a(u_t)$ for all $u_t \in [0,1]$. Consider now the range of parameters such that $\beta \geq \left(\sqrt{17}-1\right)/4$. In this case, for all $u_t > \bar{u}(\beta) = \frac{\beta+6-\beta\sqrt{4-2\beta}}{2(2+\beta)}$, $\hat{V}^b(u_t) < \hat{V}^a(u_t)$. Thus, the choice of b_t maximizing $\hat{V}_t^{ou}$ is in the range $b_t \in \left[0, \frac{2\beta}{2+\beta}\right]$ and, in particular, it will be $argmax_{b_t}\left\{b_t - \frac{1}{4}(1+b_t+2u_t)\,b_t\right\} = 3/2 - u_t$.

To prove part (B), i.e., that $U^{dp}(b_t) = \left(1 - \beta + b_t + \beta B^{dp}\left(U^{dp}(b_t)\right)\right)/2$, observe that, from (7);

$$\left(1 - \beta + b_t + \beta B^{dp}\left(U^{dp}(b_t)\right)\right)/2 = \begin{cases} \frac{\beta(1+\beta)+2}{2(2+\beta)} + \frac{2-\beta}{4} b_t & \text{if } b_t \in \left[\frac{2\beta}{2+\beta}, 1\right] \\ \frac{1+b_t}{2} & \text{else} \end{cases} \tag{16}$$

Take, first, the range $b_t \in \left[\frac{2\beta}{2+\beta}, 1\right]$. Then;

$$\left(1 - \beta + b_t + \beta B^{dp}\left(\frac{\beta(1+\beta)+2}{2(2+\beta)} + \frac{2-\beta}{4} b_t\right)\right)/2 = \frac{\beta(1+\beta)+2}{2(2+\beta)} + \frac{2-\beta}{4} b_t$$

Next, note that $b_t \in \left[\frac{2\beta}{2+\beta}, 1\right] \Rightarrow \left(\frac{\beta(1+\beta)+2}{2(2+\beta)} + \frac{2-\beta}{4} b_t\right) \in \left[\frac{3}{2} - \frac{2}{(2+\beta)}, \bar{u}(\beta)\right]$. Hence, using (7);

$$\left(1 - \beta + b_t + \beta\left(\frac{3(2+\beta) - \beta^2}{4 - \beta^2} - \frac{2}{2-\beta}\left(\frac{\beta(1+\beta)+2}{2(2+\beta)} + \frac{2-\beta}{4} b_t\right)\right)\right)/2$$
$$= \frac{\beta(1+\beta)+2}{2(2+\beta)} + \frac{2-\beta}{4} b_t,$$

which was to be shown.

Take, next, the range $b_t \in \left[0, \frac{2\beta}{2+\beta}\right)$. In this case, $u_{t+1} = (1 + b_t)/2$, and, necessarily, $u_{t+1} \in \left[0, \frac{3}{2} - \frac{2}{(2+\beta)}\right]$. Thus, $B^{dp}(u_{t+1}) = 1$, which can be substituted into the left-hand-side of (16) to yield $(1 + b_t)/2$ which verifies the second part of (16) This concludes part (B) of the proof.

Finally, the characterization of the equilibrium law of motion of u_t, (8), and of the steady-state, (9)-(10), is straightforward. ■

7.2 Proof of Proposition 2

Proof. As in proof of Proposition 1, we must show that, for all t and u_t, $\langle B^{mv}, U^{mv}\rangle$ satisfy

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

B) $U^{mv}(b_t) = (1 - \beta + b_t + \beta B^{mv}(U^{mv}(b_t)))/2$.

As to part (A), consider first the range where $u_t > 1/2$:

$$\hat{V}_t^{mv}(b_t, u_t) = \hat{V}_t^{ou}(b_t, u_t)$$
$$= \begin{cases} b_t - \frac{1}{4}\left(1 - \beta + b_t + \beta\left(\frac{3(2+\beta)-\beta^2}{4-\beta^2} - \frac{2}{2-\beta}\left(\frac{\beta(1+\beta)+2}{2(2+\beta)} + \frac{2-\beta}{4} b_t\right)\right) + 2u_t\right) b_t & \text{if } \; b_t \geq \frac{2\beta}{2+\beta} \\ b_t - \frac{1}{4}(1 + b_t + 2u_t) b_t & \text{else} \end{cases}.$$

Note that $\hat{V}_t^{mv}(0, u_t) = 0$, which shows that setting $b_t = 0$ can never be optimal for the unsuccessful (recall that both (14) and (15) are strictly positive)

Next, if $u_t \leq 1/2$, then $V_t^{mv}(b_t, u_t) = V_t^{os}(b_t, u_t)$, and this is maximized by setting $b_t = 0$.

To prove part (B), i.e., that $U^{mv}(b_t) = (1 - \beta + b_t + \beta B^{mv}(U^{mv}(b_t)))/2$, observe that, from (11) and (12);

$$b_t = B^{mv}(u_t) = \begin{cases} B^{dp}(u_t) & \text{if } u_t \in \left(\frac{1}{2}, 1\right] \\ B^{pl}(u_t) & \text{if } u_t \in \left[0, \frac{1}{2}\right] \end{cases}$$

$$U^{mv}(b_t) = \begin{cases} U^{dp}(b_t) & \text{if } b_t \in (0, 1] \\ U^{pl}(b_t) & \text{if } b_t = 0 \end{cases}$$

while;

$$\frac{(1 - \beta + b_t + \beta B^{mv}(U^{mv}(b_t)))}{2}$$
$$= \begin{cases} \frac{\left(1-\beta+b_t+\beta B^{mv}\left(U^{dp}(b_t)\right)\right)}{2} & \text{if } b_t \in (0, 1] \\ \frac{\left(1-\beta+b_t+\beta B^{mv}\left(U^{pl}(b_t)\right)\right)}{2} & \text{if } b_t = 0 \end{cases}$$
$$= \begin{cases} \frac{\left(1-\beta+b_t+\beta B^{dp}\left(U^{dp}(b_t)\right)\right)}{2} & \text{if } b_t \in (0, 1] \\ \frac{1-\beta}{2} & \text{if } b_t = 0 \end{cases}$$

The first equality follows directly from ((12). The latter follows from (5)-(7)-(12) and, in particular, from the following two observations. First, for all $b_t \in (0, 1]$, $U^{dp}(b_t) > 1/2$. Hence, in this range $B^{mv}\left(U^{dp}(b_t)\right) = B^{dp}\left(U^{dp}(b_t)\right)$. Second, if $b_t = 0$, then $U^{pl}(b_t) = U^{pl}(b_t) = \frac{1-\beta}{2} < 1/2$. Hence, in this range, $B^{mv}\left(U^{pl}(b_t)\right) = B^{pl}\left(U^{pl}(b_t)\right) = 0$, and, consequently, $\left(1 - \beta + b_t + \beta B^{pl}\left(U^{pl}(b_t)\right)\right)/2 = (1 - \beta)/2$. Thus, in order to prove (B), we need to show that:

$$U(b_t) = \begin{cases} U^{dp}(b_t) = \frac{\left(1-\beta+b_t+\beta B^{dp}\left(U^{dp}(b_t)\right)\right)}{2} & \text{if } b_t \in (0, 1] \\ U^{pl}(b_t) = \frac{1-\beta}{2} & \text{if } b_t = 0 \end{cases}$$

But this follows immediately from Proposition 1, so part (B) is also established.

Finally, the characterization of the equilibrium law of motion of u_t, (13), and of the steady-state is straightforward. ■

7.3 Proof of Proposition 3 and 6

Proof. As in proof of Proposition 1, we must show that, for all t and u_t, $\langle B^{aw}, U^{aw} \rangle$ satisfy

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

B) $U^{aw}(b_t) = (1 - \beta + b_t + \beta B^{aw}(U^{aw}(b_t)))/2$.

1. Consider first the case when $\beta \geq \frac{\sqrt{5}-1}{2}$. As to part (A), consider first the range where $u_t > 1/2$.

$$\hat{V}_t^{aw}(b_t, u_t) = \hat{V}_t^{ou}(b_t, u_t) =$$

$$= \begin{cases} b_t - \frac{1}{4}\left(1-\beta+b_t+\beta\left(B^{aw}\left(U^{aw}(b_t)\right)\right)+2u_t\right)b_t & \text{if} \quad b_t > \beta \\ b_t - \frac{1}{4}\left(1-\beta+b_t+\beta\left(B^{pl}\left(U^{pl}(b_t)\right)\right)+2u_t\right)b_t & \text{else} \end{cases}.$$

$$= \begin{cases} b_t - \frac{1}{4}\left(1-\beta+b_t+\beta^2+2u_t\right)b_t & \text{if} \quad b_t > \beta \\ b_t - \frac{1}{4}\left(1-\beta+b_t+2u_t\right)b_t & \text{else} \end{cases}.$$

Simple algebra shows that $\hat{V}_t^{aw}(b_t, u_t)$ is increasing in b_t for $b_t \leq \beta$. Furthermore, $\hat{V}_t^{aw}(\beta, u_t) \geq \hat{V}_t^{aw}(b_t, u_t)$ for all $b_t > \beta, u_t > 1/2$ and $\beta \geq \frac{\sqrt{5}-1}{2}$. So $B^{aw}(u_t) = \beta$ for $u_t > 1/2$.

If $u_t \leq 1/2$, $\hat{V}_t^{aw}(b_t, u_t) = \hat{V}_t^{os}(b_t, u_t)$ which is decreasing in b_t so $B^{aw}(u_t) = 0$, for $u_t \leq 1/2$.

To prove part (B), i.e., that $U^{aw}(b_t) = \left(1-\beta+b_t+\beta B^{aw}\left(U^{aw}(b_t)\right)\right)/2$, observe that

$$\frac{\left(1-\beta+b_t+\beta B^{aw}\left(U^{aw}(b_t)\right)\right)}{2}$$

$$= \begin{cases} \frac{\left(1-\beta+b_t+\beta B^{aw}\left(U^{dp}(b_t)\right)\right)}{2} & \text{if} \quad b_t > \beta \\ \frac{\left(1-\beta+b_t+\beta B^{aw}\left(U^{pl}(b_t)\right)\right)}{2} & \text{else} \end{cases}$$

$$= \begin{cases} \frac{\left(1-\beta+b_t+\beta B^{dp}\left(\frac{\beta(1+\beta)+2}{2(2+\beta)}+\frac{2-\beta}{4}b_t\right)\right)}{2} & \text{if} \quad b_t > \beta \\ \frac{\left(1-\beta+b_t+\beta B^{pl}\left(\frac{1-\beta+b_t}{2}\right)\right)}{2} & \text{else} \end{cases}$$

$$= \begin{cases} \frac{\left(1-\beta+b_t+\beta^2\right)}{2} & \text{if} \quad b_t > \beta \\ \frac{\left(1-\beta+b_t\right)}{2} & \text{else} \end{cases} = U^{aw}(b_t)$$

where the second equality follows from the facts that $\left[\beta(1+\beta)+2\right]/\left[2(2+\beta)\right] + (2-\beta)b_t/4 > \frac{1}{2}$ for all $b_t \in [\beta, 1]$ and $(1-\beta+b_t)/2 < 1/2$ for all $b_t < \beta$.

2. Consider now the case when $\beta \in \left[\beta_H, \frac{\sqrt{5}-1}{2}\right)$. As to part (A), consider first the range where $u_t > 1/2$.

$$\hat{V}^{aw}(b_t, u_t) = \hat{V}^{ou}(b_t, u_t)$$

$$= b_t - \frac{1}{4}\left(1-\beta+b_t+\beta\left(B^{aw}\left(U^{aw}(b_t)\right)\right)+2u_t\right)b_t$$

$$= \begin{cases} b_t - \frac{1}{4}\left(1-\beta+b_t+\beta^2+2u_t\right)b_t & \text{if} \quad b_t > \beta \\ b_t - \frac{1}{4}\left(1-\beta+b_t+2u_t\right)b_t & \text{else} \end{cases}.$$

Simple algebra shows that $\hat{V}^{aw}(b_t, u_t)$ is increasing in b_t in both the region $b_t \leq \beta$ and $b_t > \beta$, and that at $b_t = \beta$ the value function has a discontinuous fall. Furthermore, $\hat{V}^{aw}(1, u_t) > \hat{V}^{aw}(\beta, u_t)$ when $u_t \in (0.5, u^a(\beta))$, $\hat{V}^{aw}(1, u_t) < \hat{V}^{aw}(\beta, u_t)$ when $u_t \in (u^a(\beta), 1]$, and $\hat{V}^{aw}(1, u^a(\beta)) = \hat{V}^{aw}(\beta, u^a(\beta))$, where $u^a(\beta)$ is defined in the text. Thus, $B^{aw}(u_t) = \beta$ for $u_t > u^a(\beta)$ and $B^{aw}(u_t) = 1$ for $u_t \in (0.5, u^a(\beta))$. If $u_t \leq 1/2$, $\hat{V}_t^{aw}(b_t, u_t) = \hat{V}_t^{os}(b_t, u_t)$ which is decreasing in b_t so $B^{aw}(u_t) = 0$, for $u_t \leq 1/2$.

To prove part (B), i.e., that $U^{aw}(b_t) = (1 - \beta + b_t + \beta B^{aw}(U^{aw}(b_t)))/2$, observe that

$$\begin{aligned} &\frac{(1 - \beta + b_t + \beta B^{aw}(U^{aw}(b_t)))}{2} \\ &= \begin{cases} \frac{1-\beta+b_t+\beta B^{aw}\left(\left(1-\beta+b_t+\beta^2\right)/2\right)}{2} & \text{if} \quad b_t > \beta \\ \frac{1-\beta+b_t+\beta B^{aw}\left(\frac{1-\beta+b_t}{2}\right)}{2} & \text{else} \end{cases} \\ &= \begin{cases} \frac{\left(1-\beta+b_t+\beta^2\right)}{2} & \text{if} \quad b_t > \beta \\ \frac{(1-\beta+b_t)}{2} & \text{else} \end{cases} = (U^{aw}(b_t)) \end{aligned}$$

where the second equality follows from the facts that $\left(1 - \beta + b_t + \beta^2\right)/2 \geq u^a(\beta)$ for all $b_t \in (\beta, 1]$ (since $\beta \geq \beta_H$), and that $(1 - \beta + b_t)/2 < 1/2$ for all $b_t < \beta$. QED

3. Consider now the case when $\beta \in [\beta_M, \beta_H)$. As to part (A), consider first the range where $u_t > 1/2$. Applying the equilibrium objects B^{aw} and U^{aw}, the indirect utility function can be written as

$$\begin{aligned} &\hat{V}^{aw}(b_t, u_t) = \hat{V}^{ou}(b_t, u_t) \\ &= b_t - \frac{1}{4}(1 - \beta + b_t + \beta(B^{aw}(U^{aw}(b_t))) + 2u_t) b_t \\ &= \begin{cases} b_t - \frac{1}{4}\left(1 - \beta + b_t + \beta\left(B^{aw}\left(\left(1 - \beta + b_t + \beta^2\right)/2\right)\right) + 2u_t\right) b_t & \text{if} \quad b_t > \beta \\ b_t - \frac{1}{4}\left(1 - \beta + b_t + \beta\left(B^{aw}\left(U^{pl}(b_t)\right)\right) + 2u_t\right) b_t & \text{else} \end{cases} \\ &= \begin{cases} b_t - \frac{1}{4}\left(1 - \beta + b_t + \beta^2 + 2u_t\right) b_t & \text{if} \quad b_t > \beta \\ b_t - \frac{1}{4}(1 - \beta + b_t + 2u_t) b_t & \text{else} \end{cases} \end{aligned}$$

Simple algebra shows that $\hat{V}^{aw}(b_t, u_t)$ is increasing in b_t in the region $b_t \leq \beta$. Moreover, conditional on $b_t \in (\beta, 1]$ and $\beta \in [\beta_M, \beta_H)$, the (constrained) optimal benefit level $\widetilde{b}$ would be

$$\widetilde{b}(u_t) = \begin{cases} 1 & \text{if} \quad u_t \in (0.5, u^c(\beta)] \\ \frac{3}{2} + \frac{\beta(1-\beta)}{2} - u_t & \text{if} \quad u_t \in (u^c(\beta), 1] \end{cases}.$$

where $u^c(\beta)$ is defined in the text. Hence, $\hat{V}^{aw}(\beta, u_t) \leq \hat{V}^{aw}\left(\widetilde{b}(u_t), u_t\right)$ if and only if $u_t \in \left(\frac{1}{2}, u^d(\beta)\right]$, with equality for $u_t = u^d(\beta)$, where, again, $u^d(\beta)$ is defined in the text. Finally, if $u_t \leq 1/2$, $\hat{V}_t^{aw}(b_t, u_t) = \hat{V}_t^{os}(b_t, u_t)$ which is decreasing in b_t so $B^{aw}(u_t) = 0$, for $u_t \leq 1/2$.

To prove part (B), i.e., that $U^{aw}(b_t) = (1 - \beta + b_t + \beta B^{aw}(U^{aw}(b_t)))/2$, for all $b_t \in [0,1]$, observe that

$$
\begin{aligned}
u_{t+1} &= \frac{(1 - \beta + b_t + \beta B^{aw}(U^{aw}(b_t)))}{2} \\
&= \begin{cases} \frac{1-\beta+b_t+\beta B^{aw}\left(\left(1-\beta+b_t+\beta^2\right)/2\right)}{2} & \text{if} \quad b_t > \beta \\ \frac{1-\beta+b_t+\beta B^{aw}\left(\frac{1-\beta+b_t}{2}\right)}{2} & \text{else} \end{cases} \\
&= \begin{cases} \frac{\left(1-\beta+b_t+\beta^2\right)}{2} & \text{if} \quad b_t > \beta \\ \frac{(1-\beta+b_t)}{2} & \text{else} \end{cases} \\
&= U^{aw}(b_t) \ \ \forall \, b_t \in [0,1],
\end{aligned}
$$

where the second equality follows from the fact that if $\beta \geq \beta_M$, then $\left(1 - \beta + b_t + \beta^2\right)/2 \geq u^d(\beta)$, which implies that $B^{aw}\left(\left(1 - \beta + b_t + \beta^2\right)/2\right) = \beta$ if $b_t > \beta$. This concludes the proof for the case when $\beta \in [\beta_M, \beta_H)$.

■

7.4 Proof of Proposition 4

Proof. First, note from equilibrium definition 2 that $b_t \leq \beta$ is a necessary condition for $u_{t+1} \leq 1/2$. Thus, since $u_{t+1} \leq$ implies $b_{t+1} = 0$, the value of inducing a welfare state breakdown is bounded from above by

$$V^{ou}(\beta, 0, u_t) = \beta - \frac{1 + 2u_t}{4}\beta,$$

since $V^{ou}(b_t, 0, u_t)$ is increasing in b_t.

On the other hand, sustaining the welfare state implies $b_{t+1} > 0$, which has a negative impact on the indirect utility of the current old. However, since $b_{t+1} \leq 1$, the indirect utility of sustaining the welfare state is bounded from below by

$$\max_{b_t} V^{ou}(b_t, 1, u_t) \text{ s.t. } b_t > \beta,$$

where the constraint ensures that $u_{t+1} > 1/2$. Simple algebra shows that for $\beta < 1/2$ this expression is solved by $b_t = 3/2 - u_t$ yielding

$$V^{ou}(3/2 - u_t, 1, u_t) = \frac{9}{16} - \frac{3}{4}u_t + \frac{1}{4}u_t^2.$$

Finally, we note that

$$\frac{9}{16} - \frac{3}{4}u_t + \frac{1}{4}u_t^2 - \left(\beta - \frac{1+2u_t}{4}\beta\right) > 0$$

for all u_t when $\beta < 1/4$. Thus, inducing a welfare state break-down necessarily implies lower indirect utility for a majority of unsuccessful than any alternative. ■

7.5 Proof of Proposition 5

Proof. Assume, tentatively, that

$$b_t = B^{yv}(u_t) \equiv \begin{cases} B^{dp}(b_t) & \text{if } u_t \in \left(\frac{1+\varepsilon}{2}, U^{dp}(1)\right] \\ 0 & \text{if } u_t = \left[0, \frac{1+\varepsilon}{2}\right] \end{cases}.$$

As noted in the text, if $b_t \leq b^{bd}$, $\frac{1-\beta}{2} < u_{t+1} \leq \frac{1+\varepsilon}{2}$ and $b_{t+1} = 0$. If instead $b_t > b^{bd}$, $U^{dp}(1) \geq u_{t+1} > \frac{1+\varepsilon}{2}$. Thus, under the tentative assumptions,

$$u_{t+1} = U^{yv}(b_t) \equiv \begin{cases} U^{dp}(b_t) & \text{if } b_t \in \left(b^{bd}, 1\right] \\ U^{pl}(b_t) & \text{if } b_t = \left[0, b^{bd}\right] \end{cases}$$

satisfies equilibrium condition 2.

Now, we need to show that the tentative assumptions satisfy equilibrium condition 1, i.e., that

A) $\forall\ u_t \in \left(\frac{1+\varepsilon}{2}, U^{dp}(1)\right]$, $B^{dp}(u_t) = \arg\max_{b_t}\left\{\hat{V}^{ou}(b_t, u_t)\right\}$, subject to $u_{t+1} = U^{yv}(b_t)$, $b_t \in [0,1]$ and $b_{t+1} = B^{yv}(u_{t+1})$,

B) $\forall\ u_t \leq \frac{1+\varepsilon}{2}$, $0 = \arg\max_{b_t}\left\{\hat{V}^{y}(b_t, u_t)\right\}$, subject to $u_{t+1} = U^{yv}(b_t)$, $b_t \in [0,1]$ and $b_{t+1} = B^{yv}(u_{t+1})$.

To check condition A, first note that $B^{dp}(u_t) \geq \frac{1}{2}\frac{3\beta+2}{2+\beta}$ for $u_t \in \left(\frac{1+\varepsilon}{2}, U^{dp}(1)\right]$, and

$$b^{bd}. = \begin{cases} 2\frac{\varepsilon(2+\beta)-\beta^2}{(2-\beta)(2+\beta)} & \text{if } \varepsilon \geq 2\frac{\beta}{2+\beta} \\ \varepsilon & \text{if } \varepsilon < 2\frac{\beta}{2+\beta} \end{cases}, \tag{17}$$

implying that $B^{dp}(u_t) > b^{bd}$ if $u_t \in \left(\frac{1+\varepsilon}{2}, U^{dp}(1)\right]$ and $\varepsilon < \bar{\varepsilon}(\beta) \leq 1/2$. Thus, by Proposition (1), $b_t = B^{dp}(u_t)$ is optimal in the range $b_t \in (b^{bd}, 1]$. Furthermore, for $b_t \leq b^{bd}$ the indirect utility of the old unsuccessful, given by $b_t - \frac{(1-\beta)+b_t+2u_t}{4}b_t$, is maximized at $b_t = b^{bd}$ for $u_t \in \left(\frac{1+\varepsilon}{2}, U^{dp}(1)\right]$ and $\varepsilon < \bar{\varepsilon}(\beta)$.

For A) to be satisfied, we must ensure that $b_t = B^{dp}(u_t)$ is preferred to $b_t = b^{bd}$ by the old unsuccessful. The former yields

$$B^{dp}(u_t) - \frac{(1-\beta) + \left(B^{dp}(u_t) + \beta B^{dp}\left(U^{dp}\left(B^{dp}(u_t)\right)\right)\right) + 2u_t}{4} B^{dp}(u_t) \equiv \hat{V}^{con}(u_t),$$

and the latter,

$$b^{bd} - \frac{(1-\beta) + b^{bd} + 2u_t}{4} b^{bd} \equiv \hat{V}^{bd}(\varepsilon, u_t). \tag{18}$$

Using the envelope theorem, we have that $\frac{d\left(\hat{V}^{con}(u_t) - \hat{V}^{bd}(\varepsilon,u_t)\right)}{du_t} = -\frac{B^{dp}(u_t) - b^{bd}}{2} < 0$, i.e., $\hat{V}^{con}(u_t) - \hat{V}^{bd}(\varepsilon, u_t)$ decreases in u_t. Thus, if $\hat{V}^{con}(u_t) \geq \hat{V}^{bd}(\varepsilon, u_t)$ at $u_t = U^{dp}(1)$, it is satisfied for all $u_t \in \left(\frac{1+\varepsilon}{2}, U^{dp}(1)\right]$. Now, using the definition of $\hat{V}^{con}$, we get

$$\hat{V}^{con}(U^{dp}(1)) = \frac{1}{32}(3\beta + 2)\frac{4(1+\beta) - 3\beta^2}{(2+\beta)^2}.$$

Furthermore, using (17) and (18) we find that

$$\hat{V}^{bd}(\varepsilon, U^{dp}(1)) = \begin{cases} \frac{1}{4}\left(\varepsilon(2+\beta) - \beta^2\right)\frac{8(1-\varepsilon)+\beta(12-\beta(2+\beta)-4\varepsilon)}{(2-\beta)^2(2+\beta)^2} & \text{if } \varepsilon \geq 2\frac{\beta}{2+\beta} \\ \frac{1}{8}\varepsilon\frac{4(1-\varepsilon)+\beta(8+\beta)-2\varepsilon\beta}{2+\beta} & \text{if } \varepsilon < 2\frac{\beta}{2+\beta} \end{cases}.$$

Solving $\hat{V}^{con}\left(U^{dp}(1)\right) = \hat{V}^{bd}(\varepsilon, U^{dp}(1))$, yields $\varepsilon = \bar{\varepsilon}(\beta)$, which is smaller (larger) than $2\frac{\beta}{2+\beta}$ if β larger (smaller) than β^y. Since $\hat{V}^{bd}(\varepsilon, U^{dp}(1))$ increases in ε for all $\varepsilon \leq \bar{\varepsilon}(\beta)$, $\hat{V}^{con}\left(U^{dp}(1)\right) > \hat{V}^{bd}(\varepsilon, U^{dp}(1))$ for all $\varepsilon < \bar{\varepsilon}(\beta)$ and condition A) is satisfied.

Now consider condition B). First, note the old successful always vote $b_t = 0$. Then, consider the young having an indirect utility of

$$\hat{V}^{y}(b_t, b_{t+1}, b_{t+2}, u_t) = \frac{(1+\beta)^2}{4} + \frac{1-\beta-2u_t}{4}b_t - \frac{b_{t+1}+\beta b_{t+2}}{4}\beta b_{t+1}.$$

Clearly, the young prefer zero benefits whenever $u_t \geq \frac{1-\beta}{2}$. Furthermore, since $u_{t+1} = \frac{1-\beta+b_t+\beta b_{t+1}}{2} \geq \frac{1-\beta}{2}$, the young prefer zero benefits at all $t > 0$ and by assumption, $u_0 \geq \frac{1+\varepsilon}{2} > \frac{1-\beta}{2}$. Thus, b_t is optimally chosen to zero whenever $u_t \leq \frac{1+\varepsilon}{2}$, which proves condition B) and concludes the proof. ■