The Political Economy of Redistribution under Democracy*

Jess Benhabib and Adam Przeworski
Department of Economics   Department of Politics
New York University
(Very Preliminary-Comments Welcome)

September 12, 2003

Abstract

We ask what redistributions of income and assets are feasible in a democracy, given the initial assets and their distribution. The question is motivated by the possibility that if redistribution is insufficient for the poor or excessive for the rich, they may turn against democracy. In turn, if no redistribution simultaneously satisfies the poor and the wealthy, democracy cannot be sustained. Hence, the corollary question concerns the conditions under which democracy is sustainable. Since decisions to save are endogenous, we solve explicitly for the current growth rates given any time path of future tax rates. We find that the optimal path of redistribution chosen by the median voter under the constraint of rebellion by the poor or the wealthy consists of redistributing as much as possible as soon as possible. However, this path is time inconsistent unless voters punish governments that deviate from their promises. Democracies survive in wealthy societies, with a lower average capital stock when they are more equal.

*We thank Onur Ozgur for comments and for technical assistance.

1 Introduction

We ask what redistributions of income and assets are feasible in a democracy, given the initial assets and their distribution. The question is motivated by the possibility that if redistribution is insufficient for the poor or excessive for the rich, they may turn against democracy. In turn, if no redistribution simultaneously satisfies the poor and the wealthy, democracy cannot be sustained. Hence, the corollary question concerns the conditions under which democracy is sustainable.

In a simple model of production and accumulation, where agents are heterogeneous in their initial wealth, the median voter chooses a sequence of redistributive tax rates. Decisions to save are endogenous, which means that they depend on future tax rates and thus future growth rates. We solve explicitly for the current growth rates given any time path of future tax rates. We assume that decisions about redistribution are made in elections and show that no coalition of poor and wealthy leaves both better off than the decision of the median voter (see Theorem 4). Moreover, the identity of the median voter does not change over time (For a model where it may, see Besley and Coate (1998)). Hence, the same median voter is decisive at each time with regard to the entire time path of future redistribution. The median voter, however, must consider the eventuality of rebellion by the poor or the wealthy. Either of these groups can establish a dictatorship and impose by force its preferred path of redistribution. The cost of dictatorship, in turn, is the loss of freedom. We assume, in the spirit of Sen (1991), that the utility of any amount of consumption is lower when people are not free to live the lives of their choosing.

Since the median voter need not choose a constant tax rate, we encounter all the richness and all the difficulty of the optimal tax literature (Chamley (1985), (1986), Judd (1985), Benhabib and Rustichini (1997), Benhabib, Rustichini and Velasco (2001), Chari and Kehoe(1999)). A further complication arises from the redistribution of tax receipts back to the agents as transfers, which the agents must incorporate into their decisions to save. The equilibrium must therefore be solved for all feasible tax sequences, and the median voter optimizes over these sequences taking into account equilibrium consequences (see Theorems 1 and 2). In contrast to earlier results in the related literature which provide necessary conditions that all equilibria must satisfy (Chamley (1986) or Chari and Kehoe (2000)), we characterize the unique equilibrium, albeit in a framework which is different and in some dimensions more restrictive. (see Theorem 2, Remark 2, and Corollary 1.) Hence the length and complexity of our proofs.

Results are the following:

(1) Democracies survive in wealthy countries. Since the freedom of poor people is already circumscribed by poverty, people have more to lose from dictatorship when they are wealthier. As a result, they are less prone to turn against democracy in affluent countries. (see Theorems 5 and 6 , and also see the discussion preceding Theorem 6).

(2) Democracies survive at lower average capital stock in more equal societies.(See Theorem 6 and Remark 8.)

(3) If the capital stock is high enough at the time democracy is inaugurated, the median voter can equalize wealth in the first period, and implement zero taxes henceforth. (see Theorems 3 and 6).

(4) The optimal path of redistribution chosen by the median voter under the constraint of rebellion by the poor or the wealthy consists of redistributing as much as possible as soon as possible. While redistribution is a conflictive issue, since future taxes reduce current growth, everyone wants redistribution to stop at some time. (See Theorem 4) However, this path is time inconsistent unless voters punish governments that deviate from their promises (See Corollary 4 and Theorem 7).

(5) Since the median voter, as well as the poor and the wealthy dictators, want redistribution to end as soon as possible, political regimes affect growth rates only during the period when future taxes are expected to be positive.

The first result is consistent with the well established fact that the probability that a democracy would survive rises steeply in per capita income (Przeworski and Limongi 1977, Przeworski et al. 2000). Between 1951 and 1999, the probability that a democracy would die during any particular year in countries where it emerged with per capita income under \$1000 was 0.0819, which implies that their expected life was about twelve years. For countries where democracy was inaugurated when they had incomes between \$1001 and \$3000, this probability was 0.0248, for an expected duration of about forty years. Between \$3001 and \$6055, the probability was 0.0099, which translates into about 101 years of expected life. And no democracy ever fell in a country with a per capita income higher than that of Argentina in 1975, \$6,055 (1985 PPP from PWT5.6). This is a startling fact, given that throughout history about seventy democracies collapsed in poorer countries. In contrast, thirty-seven democracies spent over 1000 years in more developed countries and not one died.

We do not have systematic knowledge about redistribution. All we know is that the distribution of market (pre-fisc) incomes seems to be extremely stable over time, implying that major redistributions of assets are in fact rare. It appears that there are no countries which over the long run equalized market incomes without some kind of cataclysm.

The strongest evidence comes from Li, Squire, and Zou (1997), who report that about 90 percent of total variance in the Gini coefficients is explained by the variation across countries, while few countries show any time trends. Massive redistributions occurred only as a result of (1) expropriation of large property as a consequence of foreign occupation (Japanese in Korea, Soviets in Eastern Europe), (2) revolution (Soviet Union, China), (3) destruction of large fortunes caused by war (France according to Piketty 2000), or (4) massive emigration of the poor (Norway, Sweden).

That political regimes, dichotomized as democracies and dictatorships, do not affect the rates of growth of total (as distinct from per capita) income is now generally accepted (Barro 1989, Helliwell 1994, Przeworski et al. 2000). Young democracies, however, tend to grow at a slower rate than mature ones (Przeworski et al. 2000), although we do not know whether this is because they are expected to tax for a longer period than either kind of dictatorship.

The result about front-loading redistribution is standard and surprisingly robust in the optimal taxation theory. Yet it appears to contradict the observed patterns: prima facie observation shows that democracies tend to tax year after a year and, if anything, the tax rate increases as they grow older.[1] In the light of the theoretical intuition, this fact is deeply puzzling. After all, even the wealthy, who lose from redistribution, would want it to occur as quickly as possible. A situation in which asset distribution and factor incomes remain more or less the same and some redistribution occurs year after a year is costly to everyone over any longer run. We have no answer to this puzzle. One reason may be time inconsistency. While the median voter pushes the wealthy to the rebellion constraint early into the life of democracy, growth generates some slack with regard to this constraint. Hence, the median voter is tempted to renege on her promises and to vote for a positive amount of redistribution. Yet if reneging causes economic agents to distrust promises and, therefore, to reduce their saving rates, such a proposal would not be electorally victorious or a government that implements it by surprise would suffer electoral defeat, which means in turn that it would not be made or implemented. The second reason, due to Meltzer and Richard (1981), is that because of suffrage extensions and of ageing, the median voter in fact became poorer in the course of history of democracies. But then as-

[1] We regressed in different ways tax revenues as a proportion of GDP (from World Development Indicators) and government expenditures (from Penn World Tables) on regime age, initial or current per capita income, electoral turnout, and proportion of the population over 65. The coefficient on regime age is never negative, regardless of the estimator.

set redistribution should come in spurts following extensions of suffrage, and there is no evidence that it did. Finally, the third possibility is that unequal distribution of non-alienable assets regenerates inequality of alienable assets even after their initial equalization (see Mukherjee and Ray 2002)[2]. The last alternative seems particularly plausible in the light of the experience of communist countries, which nationalized almost the entire stock of physical capital and yet continually faced inequality due to differences in human capital and continued to tax incomes.

The paper is organized as follows. We begin with the description of the economy and determine the optimal consumption path as a function of future growth rates, which in turn depend on the future tax rates. We then solve for equilibrium growth rates for arbitrary future tax sequences. With these instruments, we then determine the optimal tax sequences that will be chosen by the median voter, first without any political constraints. To check that the median voter is indeed decisive with regard to future tax rates, we show that no coalition between the poor and the rich could make both better off than the proposal of the median. Then we focus on the political constraints, assuming that either the poor or the rich can rebel against the proposal of the median voter when her program generates too little or too much redistribution. This analysis allows us to determine the conditions, in terms of the initial capital stock and its distribution, under which democracy is and is not sustainable. Finally, since the optimal program of the constrained median voter turns out to be time inconsistent, we consider the conditions under which the median voter would not want to depart from the optimal program.

2 The Economy

2.1 Production

We have a linear production economy where output y_t at time t is produced with capital k_t according to the production function

$$y_t = rk_t$$

[2]In principle, what matters is the equal ownership of assets in total, not asset by asset, unless the initial equality is undone by differential bequest and capital markets are imperfect. Since income streams associated with assets like talent or human capital can be computed an initial wealth endowment, it should, in theory, be possible to equalize wealth by giving less phsical assets to those under-endowed with non-alienable assets, provide there are enough physical assets to escape corner solutions. Thus what should be taxed is endowments, human and physical, rather than incomes. However in practice it would be an impossible task to compute, or to elicit truthful estimates of non-alienable assets.

with $r > 1$.

2.2 Initial Wealth Distribution

There are n agents, indexed by i. In the initial period t_0, they each own a share of the capital stock, $v_{t_0}^i$ and $\sum_{i=1}^n v_{t_0}^i = 1$. We denote the shares of capital owned by agent i at time s as v_s^i. We define $k_t^i = v_t^i k_t$.

2.3 Redistribution Through Taxation

Taxes are redistributive. The tax rate on assets at time t is denoted as τ_t, and in each period tax collections are distributed to the agents in proportion n^{-1} of the total. We assume that $\tau_s \in [0, \tilde{\tau}]$, where $\tilde{\tau} \leq 1$ for all s.

$$y_t^i = (1 - \tau_t)\, r v_t^i k_t \,+\, n^{-1} \tau_t r k_t = (1 - \tau_t)\, r k_t^i + n^{-1} \tau_t r k_t.$$

2.4 Agents

Since redistribution has incentive effects, we have to first consider the general accumulation problem for the agents. Agents use their capital to produce income, pay proportional taxes on assets, and consume. Their preferences are assumed to be isoelastic, and their value function is:

$$V^i(k_t^i) = \max_{c_i} \frac{(c_t^i)^{1-\sigma} - 1}{(1-\sigma)} + \beta V^i(r(1-\tau_t)k_t^i \,+\, q_t^i - c_t^i)$$

where q_s^i is the redistributive transfer agent i receives at time s.

We start by deriving the consumption function for the agent. We allow for time varying taxes. The first-order condition of the agent for an interior solution is:

$$c_{t+1}^i = c_t^i (\beta r (1 - \tau_t))^{\frac{1}{\sigma}} \tag{1}$$

Note that consumption will grow if $(\beta r (1-\tau_s))^{\frac{1}{\sigma}} > 1$. Forward iteration of the budget constraint $k_{t+1}^i = r(1-\tau_t)k_t^i - (c_t^i - q_t^i)$ implies, provided $\tau_{t+s} < 1$ for $s = 1, 2 ...$(See Assumption 2 below), that:

$$c_0^i - q_0^i + \sum_{j=1}^{t} (c_j^i - q_j^i) \left[\prod_{s=1}^{j} (r(1-\tau_s))^{-1} \right] + \prod_{s=1}^{t} (r(1-\tau_s))^{-1} k_{t+1}^i$$
$$= (r(1-\tau_0))k_0^i \tag{2}$$

We also have, from the no Ponzi and transversality conditions,

$$\lim_{t \to \infty} (\prod_{s=1}^{t} (r(1-\tau_s))^{-1}) k_{t+1}^i = 0 \tag{3}$$

so that (2) becomes

$$c_0^i - q_0^i + \sum_{j=1}^{\infty} \left(c_j^i - q_j^i\right) \prod_{s=1}^{j} \left((r\,(1-\tau_s))^{-1}\right) = (r\,(1-\tau_0))k_0^i \tag{4}$$

Iterating the first order conditions for the agent we get

$$c_j^i = c_0^i \prod_{s=1}^{j} (\beta r\,(1-\tau_s))^{\frac{1}{\sigma}} \tag{5}$$

which, substituted into (4) gives

$$c_0^i - q_0^i + \sum_{j=1}^{\infty} \left(c_0^i \prod_{s=1}^{j} (\beta r\,(1-\tau_s))^{\frac{1}{\sigma}} - q_j^i \right) \prod_{s=1}^{j} \left((r\,(1-\tau_s))^{-1}\right) = r\,(1-\tau_0)\,k_0^i$$

Solving for c_0^i, we get

$$\begin{aligned} c_0^i = & \frac{(r\,(1-\tau_0))k_0^i}{\left(1 + \sum_{j=1}^{\infty} \prod_{s=1}^{j} \beta^{\frac{1}{\sigma}} \left(r\,(1-\tau_s)\right)^{\frac{1-\sigma}{\sigma}}\right)} \\ & + \frac{q_0^i + \sum_{j=1}^{\infty} q_j^i \prod_{s=1}^{j} \left((r\,(1-\tau_s))^{-1}\right)}{\left(1 + \sum_{j=1}^{\infty} \prod_{s=1}^{j} \beta^{\frac{1}{\sigma}} \left(r\,(1-\tau_s)\right)^{\frac{1-\sigma}{\sigma}}\right)} \end{aligned}$$

Thus the agent consumes a fraction λ_t of the sum of his net capital income $(r\,(1-\tau_t))k_t^i$ plus the value of transfers that he receives, discounted at $r\,(1-\tau_t)$:

$$\begin{aligned} c_t^i = & \frac{(r\,(1-\tau_t))k_t^i}{\left(1 + \sum_{j=t+1}^{\infty} \prod_{s=t+1}^{j} \beta^{\frac{1}{\sigma}} \left(r\,(1-\tau_s)\right)^{\frac{1-\sigma}{\sigma}}\right)} \\ & + \frac{q_t^i + \sum_{j=t+1}^{\infty} q_j^i \prod_{s=t+1}^{j} \left((r\,(1-\tau_s))^{-1}\right)}{\left(1 + \sum_{j=t+1}^{\infty} \prod_{s=t+1}^{j} \beta^{\frac{1}{\sigma}} \left(r\,(1-\tau_s)\right)^{\frac{1-\sigma}{\sigma}}\right)} \\ = & \lambda_t^i \left((r\,(1-\tau_t))k_t^i + \left(q_t^i + \sum_{j=t+1}^{\infty} q_j^i \prod_{s=t+1}^{j} \left((r\,(1-\tau_s))^{-1}\right) \right) \right) \end{aligned}$$

Note that $\lambda_t^i = \lambda_t$ so the propensity to consume is identical across agents.

The following assumptions assure that $0 < \lambda_t < 1$ for all $t \geq t_0$. Note that the assumption places no further restrictions on the tax rate in the initial period t_0.

Assumption 1 $\tau_s \leq \tilde{\tau}_s < 1$ *for all* $s = t_0 + 1, t_0 + 2, \ldots$ *where* t_0 *is the initial period.*

Assumption 2 $\beta^{\frac{1}{\sigma}} \left(r\left(1 - \tilde{\tau}_s\right)\right)^{\frac{1-\sigma}{\sigma}} < 1, \;\; \beta^{\frac{1}{\sigma}} r^{\frac{1-\sigma}{\sigma}} < 1.$

The excluded case where $\tau_{\bar{t}} = 1$, where $\bar{t}$ is the first time at which $\tau_t = 1$, can be dealt with separately. In that case, if $\sigma < 1$, the Euler equation (1) holds as an inequality and implies that all assets are consumed in period $\bar{t} - 1$. The growth rate reverts to zero forever, since transfers must become zero from then on as well. Furthermore note that if $\sigma \geq 1$, utility becomes unbounded below if consumption is forced to zero. Prior to $\bar{t}$ the Euler equation holds with equality and the consumption savings problem reduces to a standard finite horizon problem with all capital consumed at $\bar{t} - 1$ Since this case is of little interest we will rule it out by Assumption 2.

2.5 Endogenizing Transfers

Under the redistribution scheme

$$y_t^i = (1 - \tau_t)\, r k_t^i + n^{-1} \tau_t r k_t$$

so that

$$\sum_{i=1}^{n} y_t^i = \sum_{i=1}^{n} (1 - \tau_t)\, r v_t^i k_t + n^{-1} \tau_t r k_t = r k_t$$

Without loss of generality we define growth rates as $g_s = \frac{k_s}{k_{s-1}}$ so that

$$k_t = \left(\prod_{s=1}^{t} g_s \right) k_0$$

Let the transfers be defined as:

$$q_0^i = n^{-1} \tau_0 r k_0$$

$$q_t^i = n^{-1} \tau_t r k_t = n^{-1} \tau_t r \left(\prod_{s=1}^{t} g_s \right) k_0$$

where

$$\sum_{i=1}^{n} v_t^i = 1, \qquad \lambda_t^i = \lambda_t$$

Using the definition of the growth rates and transfers we can write

$$c_t^i = \lambda_t \left(r\left(1 - \tau_t\right) k_t^i + \left(q_t^i + \sum_{j=t+1}^{\infty} q_j^i \prod_{s=t+1}^{j} \left(\left(r\left(1 - \tau_s\right)\right)^{-1} \right) \right) \right)$$

(after some algebra) as

$$c_t^i = \lambda_t \left((1-\tau_t)\, r k_t^i + n^{-1} \left(\tau_t + \sum_{j=t+1}^{\infty} \tau_j \prod_{s=t+1}^{j} g_s (r\,(1-\tau_s))^{-1} \right) r k_t \right) \tag{6}$$

Each agent's budget constraint implies

$$\begin{aligned} k_{t+1}^i &= (1-\tau_t)\, r k_t^i + n^{-1} \tau_t r k_t - c_t^i = (1-\tau_t)\, r k_t^i + n^{-1} \tau_t r k_t \\ &\quad - \lambda_t \left((1-\tau_t)\, r k_t^i + n^{-1} \left(\tau_t + \sum_{j=t+1}^{\infty} \tau_j \prod_{s=t+1}^{j} g_s (r\,(1-\tau_s))^{-1} \right) r k_t \right) \\ &= (1-\tau_t)\,(1-\lambda_t)\, r k_t^i \\ &\quad + \left(n^{-1} r \tau_t\, (1-\lambda_t) - n^{-1} r \lambda_t \left(\sum_{j=t+1}^{\infty} \tau_j \prod_{s=t+1}^{j} g_s (r\,(1-\tau_s))^{-1} \right) \right) k_t \end{aligned}$$

Summing over agents,

$$\begin{aligned} k_{t+1} = \sum_{i=1}^{n} k_{t+1}^i &= r \Bigg((1-\tau_t)\,(1-\lambda_t) + \tau_t\,(1-\lambda_t) \\ &\quad - \lambda_t \left(\sum_{j=t+1}^{\infty} \tau_j \prod_{s=t+1}^{j} g_s \left(r\,(1-\tau_s) \right)^{-1} \right) \Bigg) k_t \\ &= r \left(1 - \lambda_t - \lambda_t \left(\sum_{j=t+1}^{\infty} \tau_j \prod_{s=t+1}^{j} g_s (r\,(1-\tau_s))^{-1} \right) \right) k_t \end{aligned}$$

We have the equilibrium relation describing growth rates for our redistributive economy:

$$g_{t+1} = r \left(1 - \lambda_t - \lambda_t \left(\sum_{j=t+1}^{\infty} \tau_j \prod_{s=t+1}^{j} g_s (r\,(1-\tau_s))^{-1} \right) \right) \tag{7}$$

The solution of this equation, describing current growth rate as a function of future growth rates, would express the growth rates in terms of an arbitrary sequence of current and future taxes and will allow us to express the equilibrium of the economy for arbitrary tax sequences. Then a political system or mechanism would choose the equilibrium by selecting the tax sequence.

Note right away that if we confine ourselves to a tax sequence that remains constant after the first period, $\tau_s = \tau$ for $s > t_0$, the solution of the above equation is simple:

$$g_s = r\,(1-\lambda_s)\,(1-\tau) \tag{8}$$

Note that g_s is constant in this case since under constant taxes λ_s is also constant. We will not of course restrict our analysis to constant tax schemes.

2.6 Dynamics of Shares $\frac{k_t^i}{k_t}$

To characterize the equilibrium dynamics of the economy, we first describe the evolution of asset shares from an initial distribution, given the tax and redistribution schemes. From the agent's budget constraint we have:

$$\begin{aligned} k_{t+1}^i &= (1-\lambda_t)(1-\tau_t) r k_t^i + n^{-1} r k_t \\ &\times \left((1-\lambda_t)\tau_t - \lambda_t \left(\sum_{j=t+1}^{\infty} \tau_j \prod_{s=t+1}^{j} g_s \left((r(1-\tau_s))^{-1} \right) \right) \right) \end{aligned}$$

We can express this in terms of shares:

$$\begin{aligned} \frac{k_{t+1}^i}{k_{t+1}} \frac{k_{t+1}}{k_t} &= (1-\tau_t)(1-\lambda_t) r \frac{k_t^i}{k_t} + n^{-1} r \tau_t (1-\lambda_t) \\ &- n^{-1} r \lambda_t \left(\sum_{j=t+1}^{\infty} \tau_j \prod_{s=t+1}^{j} g_s (r(1-\tau_s))^{-1} \right) \end{aligned}$$

which can be written as

$$\begin{aligned} v_{t+1}^i g_{t+1} &= (1-\tau_t)(1-\lambda_t) r v_t^i + n^{-1} r \tau_t (1-\lambda_t) \\ &- n^{-1} r \lambda_t \left(\sum_{j=t+1}^{\infty} \tau_j \prod_{s=t+1}^{j} g_s (r(1-\tau_s))^{-1} \right) \end{aligned}$$

and can be solved for v_{t+1}^i as

$$\begin{aligned} v_{t+1}^i &= \left((g_{t+1})^{-1} r (1-\lambda_t)(1-\tau_t) \right) v_t^i \\ &+ (g_{t+1})^{-1} r n^{-1} \left(\tau_t (1-\lambda_t) - \lambda_t \sum_{j=t+1}^{\infty} \tau_j \prod_{s=t+1}^{j} g_s (r(1-\tau_s))^{-1} \right) \quad (9) \end{aligned}$$

where the law of motion for g_{t+1} is given in (7). Thus if we could solve for growth rates as functions of future taxes, we could also solve for the dynamics of the shares.

Note that the second and third terms on the right in the last expression for v_{t+1}^i are independent of i and accrue to everyone, while the first term changes proportionally to initial endowment $v_{t_0}^i$. Therefore even if shares change, their ordering is unaffected, and the median voter will be the same agent in each period.

We can define Λ_t^i and Ψ_t^i as follows:

$$c_t^i = \lambda_t \left(v_t^i + \left(n^{-1} - v_t^i\right)\tau_t + n^{-1}\sum_{j=t+1}^{\infty}\tau_j \prod_{s=t+1}^{j} g_s(r\left(1-\tau_s\right))^{-1}\right) rk_t = \Lambda_t^i k_t$$

$$k_{t+1}^i = \Bigg(\left(1-\lambda_t\right) v_t^i + \left(n^{-1} - v_t^i\right)\tau_t\left(1-\lambda_t\right)$$

$$- n^{-1}\lambda_t \sum_{j=t+1}^{\infty}\tau_j \prod_{s=t+1}^{j} g_s(r\left(1-\tau_s\right))^{-1}\Bigg) rk_t = \Psi_t^i k_t$$

Note that if $\tau_s = \tau$, then $\lambda_s = \lambda$; using (8) and (9) we can solve for:

$$v_{t+1}^i = g^{-1} r\left(1-\tau\right)\left(1-\lambda\right) v_t^i = v_t^i$$

Therefore, shares v_{t+1}^i remain constant. Ψ_t^i, Λ_t^i are also constant and $g_s = r\left(1-\tau\right)\left(1-\lambda\right)$. That is, in this linear economy with redistribution, initial ownership ratios remain constant if taxes are constant. Furthermore Ψ_t^i, Λ_t^i ,and g_s are also constant, irrespective of redistributions n^{-1}. This is because agents always consume all income other than capital income, just as in Bertola's (1993) results in a slightly different context (He allows for differentially productive labor by introducing increasing returns to scale.) To see this, note that if τ_s is constant, $g_s = r\left(1-\tau\right)\left(1-\lambda\right)$. Then, non-capital income minus the fraction of this income consumed (not including the fraction of capital income consumed), is given by:

$$n^{-1}r\tau\left(1-\lambda\right) - n^{-1}r\lambda\tau\left(\sum_{j=t+1}^{\infty}\prod_{s=t+1}^{j} g_s(r\left(1-\tau_s\right))^{-1}\right)$$
$$= n^{-1}r\tau\left(1-\lambda\right) - n^{-1}r\lambda\tau\left(\lambda^{-1}-1\right) = 0$$

Dynamics after a one step redistribution: Suppose the tax rate is τ_{t_0} at t_0, and 0 afterwards. Then the shares

$$v_{t_0+1}^i = g_{t_0+1}^{-1} r\left[\left(1-\lambda_{t_0}\right)\left(1-\tau_{t_0}\right)v_{t_0}^i + \tau_{t_0}\left(1-\lambda_{t_0}\right)n^{-1}\right]$$

In this case $g_{t_0+1} = r\left(1-\lambda_{t_0}\right) = \left(\beta r\right)^{\frac{1}{\sigma}}$. Therefore, if the initial share is $v_{t_0}^i$,

$$v_{t_0+1}^i = \left(1-\tau_{t_0}\right)v_{t_0}^i + \tau_{t_0} n^{-1}.$$
$$v_{t_0+s+1}^i = v_{t_0+s}^i$$

$\tau_t = 1$ produces complete non-distortionary equality in one step. That is $v^i_{t_0+s} = n^{-1}$ for all $s = 1, 2, \ldots$ Shares will remain constant however if $\tau_{t+s} = \tau$ for $s > 0$ even if $\tau < 1$.

In order to continue with our analysis and characterize the equilibrium, we need to solve the equation

$$g_{t+1} = r\left(1 - \lambda_t - \lambda_t\left(\sum_{j=t+1}^{\infty}\tau_j\prod_{s=t+1}^{j}g_s(r(1-\tau_s))^{-1}\right)\right)$$

which we turn to in the next section.

2.7 The Equilibrium Growth Rates $\mathbf{g}_t$ for Arbitrary Tax Sequences

Let x_{t+1} be the discounted value of tax revenues from $t+1$ on. We have

$$\begin{aligned} x_{t+1} &= \sum_{j=t+1}^{\infty}\tau_j\prod_{s=t+1}^{j}g_s\left(r\left(1-\tau_s\right)\right)^{-1} \\ g_{t+1} &= r\left(1-\lambda_t-\lambda_t x_{t+1}\right) \end{aligned}$$

Then from the definition of x_{t+1} :

$$\begin{aligned} x_{t+1} &= x_{t+2}g_{t+1}\left(r\left(1-\tau_{t+1}\right)\right)^{-1} + \tau_{t+1}g_{t+1}\left(r\left(1-\tau_{t+1}\right)\right)^{-1} \\ &= g_{t+1}\left(x_{t+2}+\tau_{t+1}\right)\left(r\left(1-\tau_{t+1}\right)\right)^{-1} \end{aligned}$$

so that

$$g_{t+1} = \frac{x_{t+1}\left(r\left(1-\tau_{t+1}\right)\right)}{\left(x_{t+2}+\tau_{t+1}\right)} = r\left(1-\lambda_t-\lambda_t x_{t+1}\right)$$

Solving,

$$x_{t+1}r\left(1-\tau_{t+1}\right) = r\left(1-\lambda_t-\lambda_t x_{t+1}\right)\left(x_{t+2}+\tau_{t+1}\right)$$

We can express x_{t+1} as

$$x_{t+1} = \frac{r\left(1-\lambda_t\right)\left(x_{t+2}+\tau_{t+1}\right)}{r\left(1-\tau_{t+1}\right)+r\lambda_t\left(x_{t+2}+\tau_{t+1}\right)} = \frac{\left(1-\lambda_t\right)}{\left(1-\tau_{t+1}\right)\left(x_{t+2}+\tau_{t+1}\right)^{-1}+\lambda_t}$$

$$x_{t+1} = \frac{\left(1-\lambda_t\right)}{\left(1-\tau_{t+1}\right)\left(\frac{r(1-\lambda_{t+1})}{r(1-\tau_{t+2})(x_{t+3}+\tau_{t+2})^{-1}+r\lambda_{t+1}}+\tau_{t+1}\right)^{-1}+\lambda_t}$$

This in turn can be expressed as a continued fraction:

$$x_{t+1} = \frac{\left(1-\lambda_t\right)}{\left(1-\tau_{t+1}\right)\left(\frac{(1-\lambda_{t+1})}{(1-\tau_{t+2})\left(\frac{(1-\lambda_{t+2})}{(1-\tau_{t+3})(x_{t+4}+\tau_{t+3})^{-1}+\lambda_{t+2}}+\tau_{t+2}\right)^{-1}+\lambda_{t+1}}+\tau_{t+1}\right)^{-1}+\lambda_t}$$

that is

$$x_{t+1}=\frac{(1-\lambda_t)}{\lambda_t+\frac{(1-\tau_{t+1})}{\tau_{t+1}+\frac{(1-\lambda_{t+1})}{\lambda_{t+1}+\frac{(1-\tau_{t+2})}{\tau_{t+2}+\frac{(1-\lambda_{t+2})}{\lambda_{t+2}+\cdots}}}}}=\frac{a_1}{b_1+\frac{a_2}{b_2+\frac{a_3}{b_3+\frac{a_4}{b_4+\frac{a_5}{b5+\cdots}}}}}\geq 0$$

The non-negativity of x_{t+1} follows from the non-negativity of all the elements of the continued fraction. Expressed in terms of the coefficients and variables of our model, we can define $\{a_i, b_i\}$ for $i\ =, 1, 2, ...$ as follows.

Definition 1 *Let*

$$\begin{aligned} a_i &= 1-\lambda_{t+\frac{i-1}{2}}, \quad b_i = \lambda_{t+\frac{i-1}{2}}, \text{ for } i \text{ odd} \\ a_i &= 1-\tau_{t+\frac{i}{2}}, \quad b_i = \tau_{t+\frac{i}{2}}, \quad \text{for } i \text{ even} \end{aligned}$$

Definition 2 *(The Recurrence Relation of Continued Fractions)*

$$\begin{aligned} &A(n,t) = b_n A(n-1,t) + a_n A(n-2,t) \\ &B(n,t) = b_n B(n-1,t) + a(n) B(n-2,t) \\ &A(1,t) = a_1 = 1-\lambda_t,\ A(0,t) = 0,\ A(-1,t) = 1, \\ &B(1,t) = b_1 = \lambda_{t,}\ B(0,t) = 1,\ B(-1,t) = 0 \end{aligned}$$

The second argument t indicates that the recurrence relation depends on the initial values $A(1,t) = a_1 = 1-\lambda_t$, and $B(1,t) = b_1 = \lambda_{t,}$ and are used to define x_t. We will suppress the first argument t from here on.

Theorem 1 $x_t = \lim_{n\to\infty} \frac{A(n)}{B(n)}$; *the limit exists and is finite if* $\tau_s \in [0,1]$. *Furthermore, equilibrium growth rates are given by*

$$0 \leq g_{t+1} = r(1-\lambda_t-\lambda_t x_{t+1}) = \frac{x_{t+1}\left(r\left(1-\tau_{t+1}\right)\right)}{\left(x_{t+2}+\tau_{t+1}\right)} \leq r$$

Proof: See Appendix.

It is easy to show in fact that when $\tau_{t+s} = \tau$, the solution for g_{t+1} given in Theorem 1 reduces to $g = r(1-\lambda)(1-\tau)$. To see this notice that

$$\lim_{k\to\infty} Q^k = \lim_{k\to\infty}\begin{bmatrix} 1-\lambda(1-\tau) & \lambda(1-\tau) \\ \tau & 1-\tau \end{bmatrix}^k = \begin{bmatrix} \frac{\tau}{\lambda(1-\tau)+\tau} & \frac{\lambda(1-\tau)}{\lambda(1-\tau)+\tau} \\ \frac{\tau}{\lambda(1-\tau)+\tau} & \frac{\lambda(1-\tau)}{\lambda(1-\tau)+\tau} \end{bmatrix} \quad (10)$$

This is easy to compute because the matrix Q is stochastic, with row sums equal to unity, so it converges to the matrix ev^T where e is a column vector of ones, and v is the characteristic vector belonging to the unit root of the transpose of Q, normalized so that $v^Te = 1$. Given $C(1) = A(0) = 0,\ A(1) = a_1 = (1-\lambda),\quad D(1) = B(0) = 1,\ B(1) = b_1 = \lambda$,

$$\begin{bmatrix} A(\infty) \\ C(\infty) \end{bmatrix} = \begin{bmatrix} \frac{\tau}{\lambda(1-\tau)+\tau} & \frac{\lambda(1-\tau)}{\lambda(1-\tau)+\tau} \\ \frac{\tau}{\lambda(1-\tau)+\tau} & \frac{\lambda(1-\tau)}{\lambda(1-\tau)+\tau} \end{bmatrix} \begin{pmatrix} A(1) \\ C(1) \end{pmatrix},$$
$$\begin{bmatrix} B(\infty) \\ D(\infty) \end{bmatrix} = \begin{bmatrix} \frac{\tau}{\lambda(1-\tau)+\tau} & \frac{\lambda(1-\tau)}{\lambda(1-\tau)+\tau} \\ \frac{\tau}{\lambda(1-\tau)+\tau} & \frac{\lambda(1-\tau)}{\lambda(1-\tau)+\tau} \end{bmatrix} \begin{pmatrix} B(1) \\ D(1) \end{pmatrix}$$

$$x = \frac{A(\infty)}{B(\infty)} = \frac{\tau(1-\lambda)}{\lambda\tau + \lambda(1-\tau)} = \frac{\tau(1-\lambda)}{\lambda} = \tau\left(\lambda^{-1} - 1\right)$$

but

$$g = r(1-\lambda-\lambda x) = r\left(1-\lambda-\lambda\tau\left(\lambda^{-1}-1\right)\right) = r(1-\lambda)(1-\tau)$$

as expected.

For another example, the growth rates for a tax sequence that has a tax of τ_{t_0} in the initial period t_0 and zero tax rates afterwards is $g = r(1-\lambda_{t_0}) = r\left(\beta^{\frac{1}{\sigma}} r^{\frac{1-\sigma}{\sigma}}\right) = (\beta r)^{\frac{1}{\sigma}}$, because the first period taxes are non-distortionary. However, if the tax rate at t_0+1 is τ_{t_0+1} and zero afterwards, we can compute

$$\begin{bmatrix} A(\infty) \\ C(\infty) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1-\lambda_{t_0}(1-\tau_{t_0+1}) & \lambda_{t_0}(1-\tau_{t_0+1}) \\ \tau_{t_0+1} & 1-\tau_{t_0+1} \end{bmatrix} \times \begin{pmatrix} 1-\lambda_{t_0} \\ 0 \end{pmatrix}$$

$$\begin{bmatrix} B(\infty) \\ D(\infty) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1-\lambda_{t_0}(1-\tau_{t_0+1}) & \lambda_{t_0}(1-\tau_{t_0+1}) \\ \tau_{t_0+1} & 1-\tau_{t_0+1} \end{bmatrix} \times \begin{pmatrix} \lambda_{t_0} \\ 1 \end{pmatrix} \tag{11}$$

where $\lambda_{t_0} = \left(1 + \sum_{j=t_0+1}^{\infty} \prod_{s=t_0+1}^{j} (\beta^{\frac{1}{\sigma}}(r(1-\tau_s))^{\frac{1-\sigma}{\sigma}}\right)^{-1}$ with tax sequence $\{\tau_{t_0}, \tau_{t_0+1}, 0, 0, \ldots\}$.

We will make use of the following Lemma in section 3.2.

Lemma 1 *If $\sigma \geq 1$, then $g_s \geq \delta > 0$.*

Proof. By the definition for the solution of x_{s+1}, we have $g_{s+1} = \frac{x_{s+1}(r(1-\tau_{s+1}))}{(x_{s+2}+\tau_{s+1})} = r(1-\lambda_s-\lambda_s x_{s+1})$. If $x_{s+1} \to 0$ so that $g_{s+1} = \frac{x_{s+1}(r(1-\tau_{s+1}))}{(x_{s+2}+\tau_{s+1})} \to 0$. However we also have $g_{s+1} \to r(1-\lambda_{s+1}) > 0$ since Assumptions 1 and 2 imply that λ_s is bounded away from zero and one, but this is a contradiction. If alternatively $x_{s+2} \to \infty$ so that $\frac{x_{s+1}(r(1-\tau_{s+1}))}{(x_{s+2}+\tau_{s+1})} \to 0$, then $g_{s+2} = \frac{x_{s+2}(r(1-\tau_{s+2}))}{(x_{s+3}+\tau_{s+2})} \geq 0$, but then $g_{s+2} \to r(1-\lambda_{s+2}-\lambda_{s+2}x_{s+2}) \to -\infty$, which again is a contradiction. Q.E.D.

Note that $\delta < 1$ is possible under our assumptions, so the economy can in fact shrink if the maximal tax rate $\tilde{\tau}$ is set to hold forever since then $g = (\beta r(1-\tilde{\tau}))^{\frac{1}{\sigma}}$, which can be less than unity.

3 Democracy: Median Voter Model

3.1 The Optimal Tax Sequence for the Median Voter

The value function for the median voter indexed as voter i is:

$$V\left(k_{t_0}, k_{t_0}^i, t_0\right) = \sum_{n=t_0}^{\infty} \beta^{n-t_0} \frac{\left(c_n^i\right)^{1-\sigma} - 1}{(1-\sigma)}$$

Iterating the Euler Equation for future consumption we obtain:

$$V\left(k_{t_0}, k_{t_0}^M, t_0\right) = (1-\sigma)^{-1}\left(c_{t_0}^i\right)^{1-\sigma}\left(1 + \sum_{n=t_0+1}^{\infty} \beta^{n-t_0}\left(\prod_{s=t_0+1}^{n} \left(\beta r\left(1-\tau_s\right)\right)^{\frac{1}{\sigma}}\right)^{1-\sigma}\right) - (1-\sigma)^{-1}(1-\beta)^{-1}$$

where $c_{t_0}^i$ is given by (6). Then we have:

$$V\left(k_{t_0}, k_{t_0}^M, t_0\right) = \frac{X_{t_0}^{1-\sigma} Z_{t_0}^{\sigma} - (1-\beta)^{-1}}{(1-\sigma)}$$

where

$$\lambda_{t_0} = \left(1 + \sum_{j=t_0+1}^{\infty} \prod_{s=t_0+1}^{j} \beta^{\frac{1}{\sigma}} \left(r\left(1-\tau_s\right)\right)^{\frac{1-\sigma}{\sigma}}\right)^{-1}$$

$$\begin{aligned} X_{t_0} &= \left(\left(1-\tau_{t_0}\right) v_{t_0}^M + n^{-1}\tau_{t_0} + n^{-1} \sum_{j=t_0+1}^{\infty} \tau_j \prod_{s=t_0+1}^{j} g_s (r\left(1-\tau_s\right))^{-1}\right) rk_{t_0} \\ &= \left(\left(1-\tau_{t_0}\right) v_{t_0}^M + n^{-1}\left(\tau_{t_0} + x_{t_0+1}\right)\right) rk_{t_0} \\ Y_{t_0} &= \left(1 + \sum_{n=t_0+1}^{\infty} \beta^{n-t_0} \left(\prod_{s=t_0+1}^{n} \beta r\left(1-\tau_s\right)\right)^{\frac{1}{\sigma}}\right)^{1-\sigma}\Bigg) \\ &= \left(1 + \sum_{n=t_0+1}^{\infty} \left(\prod_{s=t_0+1}^{n} \beta^{\frac{1}{\sigma}} (r\left(1-\tau_s\right))^{\frac{1-\sigma}{\sigma}}\right)\right) \\ Z_{t_0}^{\sigma} &= \lambda_{t_0}^{1-\sigma} Y_{t_0} = \left(1 + \sum_{j=t_0+1}^{\infty} \prod_{s=t_0+1}^{j} \beta^{\frac{1}{\sigma}} \left(r\left(1-\tau_s\right)\right)^{\frac{1-\sigma}{\sigma}}\right)^{\sigma} \end{aligned}$$

Theorem 2 *a)If preferences are logarithmic ($\sigma \to 1$), then the median voter chooses an optimal tax sequence such that $\lim_{m\to\infty} \tau_m = 0$, b) If initial redistribution is not restricted so that $0 \leq \tau_{t_0} \leq \tilde{\tau}_{t_0} = 1$, then the optimal tax sequence is $\{1, 0, 0...\}$, c) If $0 \leq \tau_{t_0} \leq \tilde{\tau}_{t_0} < 1$, then $\{\tau_{t_0}, 0, 0...\}$, is not an optimal tax sequence.*

Remark 1 *This is the standard optimal tax result under commitment. We use logarithmic preferences to simplify the proof, but we conjecture that the result holds more generally. Note that we allow the upper bound on the initial period tax rate to be less than 1, that is $\tilde{\tau}_{t_0} \leq 1$, which significantly complicates the proof because the poor median voter may not be able to equalize wealth in the initial period.*

Remark 2 *In contrast to existing literature (see Chamley (1986) or Chari and Kehoe (1999)) which gives necessary conditions that equilibria must satisfy, part (a) in Theorem 2 above as well as Corollary 1 that follows below, characterize the unique equilibrium. The key lies in the recursive analysis of first order conditions, which allows us to determine, using Kuhn-Tucker conditions, whether the tax will be set at the upper bound $\tilde{\tau}_s$ or at zero, irrespective of the feasible tax sequence that follows. An inductive argument then permits us to proceed and therefore our proof is not local. In the Corollary that follows, we obtain Chari and Kehoe's (1999) characterization of necessary conditions, requiring that* *taxes stay at the upper bound for a while and then drop to zero or to the lower bound, possibly following one period in which the tax rate may be interior. Our results, based on a constructive proof, yields the unique global optimum in the set of feasible tax sequences.*

Proof. a) Assume for a contradiction that $\lim_{s\to\infty}\sup\tau_s > 0$. We start by optimizing with respect to the initial period tax rates, τ_{t_0}:

$$\frac{\partial V\left(k_{t_0}, k_{t_0}^M, t_0\right)}{\partial \tau_{t_0}} = Z_{t_0}^{\sigma}\left(\left(1-\tau_{t_0}\right) r k_{t_0}^M + n^{-1}\left(\tau_{t_0} + x_{t_0+1}\right) r k_{t_0}\right)^{-\sigma} \times\left(-v_{t_0}^M + n^{-1}\right) r k_{t_0}$$

If the median voter is poor , that is if $\left(-v_{t_0}^M + n^{-1}\right) > 0$, and $0 \leq \tau_{t_0} \leq \tilde{\tau}_{t_0} < 1$, he sets $\tau_{t_0} = \tilde{\tau}_{t_0}$, independently of how future taxes are set. If he is rich, he sets $\tau_{t_0} = 0$.

For studying how future taxes are set, we consider:

$$\begin{aligned}\frac{\partial V\left(k_{t_0}, k_{t_0}^M, t_0\right)}{\partial \tau_m} &= Z_{t_0}^{\sigma} X_{t_0}^{-\sigma}\frac{\partial X_{t_0}}{\partial \tau_m} + \sigma\left(1-\sigma\right)^{-1} X_{t_0}^{1-\sigma} Z_{t_0}^{\sigma-1}\frac{\partial Z_{t_0}}{\partial \tau_m} \\ &= \left(Z_{t_0}^{\sigma} X_{t_0}^{1-\sigma}\right)\left(X_{t_0}^{-1}\frac{\partial X_{t_0}}{\partial \tau_m} + \sigma\left(1-\sigma\right)^{-1} Z_{t_0}^{-1}\frac{\partial Z_{t_0}}{\partial \tau_m}\right)\end{aligned}$$

Let $m > t_0$.

$$\begin{aligned}\sigma\left(1-\sigma\right)^{-1} Z_{t_0}^{-1}\frac{\partial Z_{t_0}}{\partial \tau_m} &= \left(1 + \sum_{n=t_0+1}^{\infty}\prod_{s=t_0+1}^{n}\beta^{\frac{1}{\sigma}}\left(r\left(1-\tau_s\right)\right)^{\frac{1-\sigma}{\sigma}}\right)^{-1} \\ &\times\left(-\beta^{\frac{1}{\sigma}} r\left(r\left(1-\tau_m\right)\right)^{\frac{1-2\sigma}{\sigma}}\right)\sum_{n=m}^{\infty}\left(\prod_{s=t_0+1, s\neq m}^{n}\left(\beta^{\frac{1}{\sigma}}\left(r\left(1-\tau_s\right)\right)^{\frac{1-\sigma}{\sigma}}\right)\right) \\ &= \left(1 + \sum_{n=t_0+1}^{\infty}\prod_{s=t_0+1}^{n}\beta^{\frac{1}{\sigma}}\left(r\left(1-\tau_s\right)\right)^{\frac{1-\sigma}{\sigma}}\right)^{-1} \\ &\times\left(-\left(1-\tau_m\right)\right)^{-1}\Big)\sum_{n=m}^{\infty}\left(\prod_{s=t_0+1}^{n}\left(\beta^{\frac{1}{\sigma}}\left(r\left(1-\tau_s\right)\right)^{\frac{1-\sigma}{\sigma}}\right)\right) < 0\end{aligned}$$

Letting $\sigma \to 1$ in evaluating the derivative,

$$\begin{aligned}\sigma\left(1-\sigma\right)^{-1} Z_{t_0}^{-1}\frac{\partial Z_{t_0}}{\partial \tau_m} &= \left(1-\beta\right)^{-1}\left(-\left(1-\tau_m\right)\right)^{-1}\Big)\beta^{m-t_0}\left(1-\beta\right) \\ &= -\beta^{m-t_0}\left(1-\tau_m\right)^{-1} < 0.\end{aligned}$$

Now consider:

$$X_{t_0}^{-1}\frac{\partial X_{t_0}}{\partial \tau_m} = \left(\left((1-\tau_{t_0})\, v_{t_0}^{M} + n^{-1}\tau_{t_0} + n^{-1}x_{t_0+1}\right) rk_{t_0}\right)^{-1} n^{-1}\frac{\partial x_{t_0+1}}{\partial \tau_m} rk_{t_0}$$

We have, first

$$x_{t_0+1} = \lim_{n\to\infty}\frac{A\,(2n+1)}{B\,(2n+1)} = \frac{\begin{bmatrix}1\,0\end{bmatrix}\left\{\prod_{s=t_0+1}^{n}\begin{bmatrix}1-\lambda_s\,(1-\tau_s) & \lambda_s\,(1-\tau_s)\\ \tau_s & 1-\tau_s\end{bmatrix}\right\}\begin{bmatrix}1-\lambda_{t_0}\\ 0\end{bmatrix}}{\begin{bmatrix}1\,0\end{bmatrix}\left\{\prod_{s=t_0+1}^{n}\begin{bmatrix}1-\lambda_s\,(1-\tau_s) & \lambda_s\,(1-\tau_s)\\ \tau_s & 1-\tau_s\end{bmatrix}\right\}\begin{bmatrix}\lambda_{t_0}\\ 1\end{bmatrix}}$$

$$\lim_{n\to\infty}\left(\frac{A\,(2n+1)}{B\,(2n+1)}\right) = \frac{\begin{bmatrix}1\,0\end{bmatrix}\begin{bmatrix}\prod_{s=m+1}^{\infty}\begin{bmatrix}1-\lambda_s\,(1-\tau_s) & \lambda_s\,(1-\tau_s)\\ \tau_s & 1-\tau_s\end{bmatrix}\\ \cdot\begin{bmatrix}1-\lambda_m\,(1-\tau_m) & \lambda_m\,(1-\tau_m)\\ \tau_m & 1-\tau_m\end{bmatrix}\\ \cdot\prod_{s=t_0+1}^{m-1}\begin{bmatrix}1-\lambda_s\,(1-\tau_s) & \lambda_s\,(1-\tau_s)\\ \tau_s & 1-\tau_s\end{bmatrix}\end{bmatrix}\begin{bmatrix}1-\lambda_{t_0}\\ 0\end{bmatrix}}{\begin{bmatrix}1\,0\end{bmatrix}\begin{bmatrix}\prod_{s=m+1}^{\infty}\begin{matrix}1-\lambda_s\,(1-\tau_s) & \lambda_s\,(1-\tau_s)\\ \tau_s & 1-\tau_s\end{matrix}\\ \cdot\begin{bmatrix}1-\lambda_m\,(1-\tau_m) & \lambda_m\,(1-\tau_m)\\ \tau_m & 1-\tau_m\end{bmatrix}\\ \cdot\prod_{s=t_0+1}^{m-1}\begin{bmatrix}1-\lambda_s\,(1-\tau_s) & \lambda_s\,(1-\tau_s)\\ \tau_s & 1-\tau_s\end{bmatrix}\end{bmatrix}\begin{bmatrix}\lambda_{t_0}\\ 1\end{bmatrix}}$$

where $m > t_0+1$, with the convention that if $m = t_0+1$, we have to define, with some abuse of notation, $\prod_{s=t_0+1}^{t_0}\begin{bmatrix}1-\lambda_s\,(1-\tau_s) & \lambda_s\,(1-\tau_s)\\ \tau_s & 1-\tau_s\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$. Note that we have restricted preferences to be logarithmic by letting $\sigma \to 1$, therefore assuring that $\lambda_s = \lambda = 1-\beta$, so that in evaluating the derivative we can set $\frac{\partial \lambda_s}{\partial \tau_m} = 0$, for all s, m.[3]

[3] The reason for the restriction is that in differentiating with respect to τ_m we want λ not to be affected. Otherwise, x_{t_0+1} depends on λ_t which depends on future taxes as well as τ_m. In principle if m is large, so that τ_m is far out relative to time t_0, it will not affect λ_{t_0} much, and by zero in the limit. But x_{t_0+1} depends on all future g's and x's, which contain concurrent λ_m. For tractability, we assume logarithmic preferences for the median voter's selection of τ_m. If we do this, we avoid the differentiation of λ_s that appears in all the product matrices above because they all depend on τ_m for $m \geq s$. Note however that under log preferences the optimal tax problem does not disappear, (just as it does not in the standard Chamley-Judd case) even though the savings rate is independent of the return. The difficulty remains in full force because future labor income discounted by the return net of taxes still depends on tax rates.

Since the product of stochastic matrices is stochastic,

$$= \lim_{n\to\infty} \frac{\partial A(2n+1)}{\partial \tau_m} = \begin{bmatrix} 1\, 0 \end{bmatrix} \begin{bmatrix} \begin{bmatrix} 1-a_m & a_m \\ b_m & 1-b_m \end{bmatrix} \\ \cdot \begin{bmatrix} \lambda & -\lambda \\ 1 & -1 \end{bmatrix} \\ \cdot \begin{bmatrix} c_m & 1-c_m \\ f_m & 1-f_m \end{bmatrix} \end{bmatrix} \begin{bmatrix} 1-\lambda \\ 0 \end{bmatrix}$$

$$= \lim_{n\to\infty} \frac{\partial B(2n+1)}{\partial \tau_m} = \begin{bmatrix} 1\, 0 \end{bmatrix} \begin{bmatrix} \begin{bmatrix} 1-a_m & a_m \\ b_m & 1-b_m \end{bmatrix} \\ \cdot \begin{bmatrix} \lambda & -\lambda \\ 1 & -1 \end{bmatrix} \\ \cdot \begin{bmatrix} c_m & 1-c_m \\ f_m & 1-f_m \end{bmatrix} \end{bmatrix} \begin{bmatrix} \lambda \\ 1 \end{bmatrix}$$

where

$$\begin{bmatrix} 1-a_m & a_m \\ b_m & 1-b_m \end{bmatrix} = \prod_{s=m+1}^{\infty} \begin{bmatrix} 1-\lambda(1-\tau_s) & \lambda(1-\tau_s) \\ \tau_s & 1-\tau_s \end{bmatrix} \geq 0$$

and

$$\begin{bmatrix} c_m & 1-c_m \\ f_m & 1-f_m \end{bmatrix} = \prod_{s=t_0+1}^{m-1} \begin{bmatrix} 1-\lambda(1-\tau_s) & \lambda(1-\tau_s) \\ \tau_s & 1-\tau_s \end{bmatrix} \geq 0.$$

Multiplying out, we obtain:

$$\lim_{n\to\infty} \frac{\partial A(2n+1)}{\partial \tau_m} = [\lambda(1-a_m)+a_m](c_m-f_m)(1-\lambda) = -\lim_{n\to\infty} \frac{\partial B(2n+1)}{\partial \tau_m}$$

$$\frac{\partial x_{t_0+1}}{\partial \tau_m} = \lim_{n\to\infty} \frac{\partial \left(\frac{A(2n+1)}{B(2n+1)}\right)}{\partial \tau_m} = \lim_{n\to\infty} \frac{\left(\frac{\partial A(2n+1)}{\partial \tau_m}\right) B(2n+1) - \left(\frac{\partial B(2n+1)}{\partial \tau_m}\right) A(2n+1)}{(B(2n+1))^2}$$
$$= \lim_{n\to\infty} \frac{\left(\frac{\partial A(2n+1)}{\partial \tau_m}\right)(B(2n+1)+A(2n+1))}{(B(2n+1))^2}$$

To evaluate this, note that using the property that the product of stochastic matrices is stochastic,

$$\lim_{n\to\infty} A(2n+1) = [1\ 0] \begin{bmatrix} 1-z_n & z_n \\ y_n & 1-y_n \end{bmatrix} \begin{bmatrix} 1-\lambda \\ 0 \end{bmatrix}$$

$$\lim_{n\to\infty} B(2n+1) = [1\ 0] \begin{bmatrix} 1-z_n & z_n \\ y_n & 1-y_n \end{bmatrix} \begin{bmatrix} \lambda \\ 1 \end{bmatrix}$$

$$
\begin{aligned}
\frac{\partial x_{t_0+1}}{\partial \tau_m} &= \lim_{n\to\infty} \frac{\left(\frac{\partial A(2n+1)}{\partial \tau_m}\right)\left(B\left(2n+1\right)+A\left(2n+1\right)\right)}{\left(B\left(2n+1\right)\right)^2} \\
&= \lim_{n\to\infty} \frac{\left(\frac{\partial A(2n+1)}{\partial \tau_m}\right)\left(1-z_n\left(1-\lambda\right)+z_n\left(1-\lambda\right)\right)}{\left(B\left(2n+1\right)\right)^2} \\
&= \lim_{n\to\infty} \frac{\left(\frac{\partial A(2n+1)}{\partial \tau_m}\right)}{\left(B\left(2n+1\right)\right)^2} \\
&= \frac{\left[\lambda\left(1-a_m\right)+a_m\right]\left(c_m-f_m\right)\left(1-\lambda\right)}{\lim_{n\to\infty}\left(B\left(2n+1\right)\right)^2}
\end{aligned}
$$

Since $\lim_{n\to\infty}\left(B\left(2n+1\right)+A\left(2n+1\right)\right)=1,\ \lim_{n\to\infty}\ \left(B\left(2n+1\right)\right)\left(1+\frac{A(2n+1)}{B(2n+1)}\right)=1$, $\lim_{n\to\infty}\left(B\left(2n+1\right)\right)^{-1}=1+x_{t_0+1}$. Therefore,

$$
\frac{\partial x_{t_0+1}}{\partial \tau_m} = \left[\lambda\left(1-a_m\right)+a_m\right]\left(c_m-f_m\right)\left(1-\lambda\right)\left(1+x_{t_0+1}\right)^2
$$

To evaluate $\left(c_m-f_m\right)$ let

$$
\begin{bmatrix} c_m^n & 1-c_m^n \\ f_m^n & 1-f_m^n \end{bmatrix} = \prod_{s=t_0+1}^{t_0+n} \begin{bmatrix} 1-\lambda\left(1-\tau_s\right) & \lambda\left(1-\tau_s\right) \\ \tau_s & 1-\tau_s \end{bmatrix}
$$

$$
\begin{bmatrix} c_m & 1-c_m \\ f_m & 1-f_m \end{bmatrix} = \prod_{s=t_0+1}^{m-1} \begin{bmatrix} 1-\lambda\left(1-\tau_s\right) & \lambda\left(1-\tau_s\right) \\ \tau_s & 1-\tau_s \end{bmatrix}
$$

Note that for $n=1$,

$$
\left(c_m^1-f_m^1\right) = 1-\lambda\left(1-\tau_{t_0+1}\right)-\tau_{t_0+1} = \left(1-\lambda\right)\left(1-\tau_{t_0+1}\right) > 0
$$

Then for $\left(c_m^{n-1}-f_m^{n-1}\right)$ given,

$$
\begin{aligned}
&\left(c_m^n-f_m^n\right) \\
&= \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 1-\lambda\left(1-\tau_{t_0+n}\right) & \lambda\left(1-\tau_{t_0+n}\right) \\ \tau_{t_0+n} & 1-\tau_{t_0+n} \end{bmatrix} \begin{bmatrix} c_m^{n-1} & 1-c_m^{n-1} \\ f_m^{n-1} & 1-f_m^{n-1} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}
\end{aligned}
$$

Evaluating

$$
\begin{aligned}
\left(c_m^n-f_m^n\right) &= \left(c_m^{n-1}-f_m^{n-1}\right)\left(1-\lambda\right)\left(1-\tau_{t_0+n}\right) > 0 \\
\left(c_m-f_m\right) &= \prod_{s=t_0+1}^{m-1}\left(1-\lambda\right)\left(1-\tau_s\right) = \beta^{m-1-t_0}\prod_{s=t_0+1}^{m-1}\left(1-\tau_s\right) \geq 0
\end{aligned}
$$

Therefore

$$\frac{\partial x_{t_0+1}}{\partial \tau_m} = \frac{\left[\lambda\left(1-a_m\right)+a_m\right]\left(1-\lambda\right)}{\left(\lim_{n\to\infty} B\left(2n+1\right)\right)^2}\beta^{m-1-t_0}\prod_{s=t_0+1}^{m-1}\left(1-\tau_s\right) \tag{12}$$

$$= \left[\lambda\left(1-a_m\right)+a_m\right]\left(1+x_{t_0+1}\right)^2\beta^{m-t_0}\prod_{s=t_0+1}^{m-1}\left(1-\tau_s\right) \geq 0$$

Thus

$$X_{t_0}^{-1}\frac{\partial X_{t_0}}{\partial \tau_m} = \left(\left(1-\tau_{t_0}\right)v_{t_0}^M + n^{-1}\tau_{t_0} + n^{-1}x_{t_0+1}\right)^{-1} \tag{13}$$

$$\times n^{-1}\left(\left[\lambda\left(1-a_m\right)+a_m\right]\left(1+x_{t_0+1}\right)^2\right)\beta^{m-t_0}\prod_{s=t_0+1}^{m-1}\left(1-\tau_s\right)$$

Let $Q_{t_0,m} : \{\tau_{t_0}, \tau_{t_0+1}...\} \to R_+$ be given by

$$Q_{t_0,m} = \left(\left(1-\tau_{t_0}\right)v_{t_0}^M + n^{-1}\left(\tau_{t_0}+x_{t_0+1}\right)\right)^{-1} n^{-1}\left(\left[\lambda\left(1-a_m\right)+a_m\right]\left(1+x_{t_0+1}\right)^2\right)$$

$$= \left(\frac{n^{-1}(1+x_{t_0+1})}{\left(1-\tau_{t_0}\right)v_{t_0}^M + n^{-1}\left(\tau_{t_0}+x_{t_0+1}\right)}\right)\left[\lambda\left(1-a_m\right)+a_m\right]\left(1+x_{t_0+1}\right)$$

This implies

$$\frac{\partial V\left(k_{t_0}, k_{t_0}^M, t_0\right)}{\partial \tau_m} = \left(Z_{t_0}^{\sigma}X_{t_0}^{1-\sigma}\right)\left(X_{t_0}^{-1}\frac{\partial X_{t_0}}{\partial \tau_m} + \sigma\left(1-\sigma\right)^{-1}Z_{t_0}^{-1}\frac{\partial Z_{t_0}}{\partial \tau_m}\right)$$

$$= Z_{t_0}\beta^{m-t_0}\left(1-\tau_m\right)^{-1}\left(Q_{t_0,m}\prod_{s=t_0+1}^{m}\left(1-\tau_s\right) - 1\right)$$

Since $\frac{\partial x_{t_0+1}}{\partial \tau_s} \geq 0$, $s > t_0$ and $0 \leq a_m < 1$, if the median voter is poor, that is $v_{t_0}^M \leq n^{-1}$, then $Q_{t_0,m}$ is bounded above:

$$Q_{t_0,m} < \left(\frac{n^{-1}(1+\overline{x}_{t_0+1})}{v_{t_0}^M + n^{-1}\overline{x}_{t_0+1}}\right)\left(1+\overline{x}_{t_0+1}\right)$$

where $\overline{x}_{t_0+1} = \sum_{j=t_0+1}^{\infty}\tilde{\tau}_j\prod_{s=t_0+1}^{j}g_s(r\left(1-\tilde{\tau}_s\right))^{-1}$. We also have $\left(1-\tau_s\right) \leq 1$ since $0 \leq \tau_s \leq \tilde{\tau}_s < 1$. If $\lim_{s\to\infty}\sup\tau_s > 0$, then there exists an infinite subsequence $\{\tau_{s_i}\}$ where $0 < \delta \leq \tau_{s_i}$. It follows that there exists a least $\tilde{m} \geq t_0 + 1$ such that $\left(Q_{t_0,m}\prod_{s=t_0+1}^{\tilde{m}}\left(1-\tau_s\right) - 1\right) < 0$. This implies that $\tau_m = 0$ for $m \geq \tilde{m}$, which is a contradiction. Hence, $\lim_{s\to\infty}\sup\tau_s = 0$ which proves part (a) of the Theorem.

We can further analyze $Q_{t_0,m}$ to determine how the tax sequence is related to the constraints on initial $0 \leq \tau_{t_0} \leq \tilde{\tau}_{t_0} \leq 1$. If $\tilde{\tau}_{t_0} = 1$, the poor median voter sets $\tau_{t_0} = 1$. In that case $\left(\frac{n^{-1}(1+x_{t_0+1})}{(1-\tau_{t_0})v_{t_0}^M + n^{-1}(\tau_{t_0}+x_{t_0+1})}\right) = 1$. In fact this implies that initial redistribution has achieved full equality. Further redistributions are unnecessary since shares will remain equal, and will only work to the detriment of growth since $\frac{\partial g_{t_0+1}}{\partial \tau_s} = -\lambda \frac{\partial x_{t_0+1}}{\partial \tau_s} \leq 0$. To prove this, we write, assuming $\tau_{t_0} = 1$,

$$
\begin{aligned}
\frac{\partial V\left(k_{t_0}, k_{t_0}^M, t_0\right)}{\partial \tau_{t_0+1}} &= Z_{t_0}\left(1-\tau_{t_0+1}\right)^{-1}\left(\frac{\left[\lambda\left(1-a_{t_0+1}\right)+a_{t_0+1}\right]}{\left(\lim_{n\to\infty} B\left(2n+1\right)\right)}\left(1-\tau_{t_0+1}\right)-1\right) \\
&= Z_{t_0}\left(1-\tau_{t_0+1}\right)^{-1} \\
&\times \left(\frac{\left[\lambda\left(1-a_{t_0+1}\right)+a_{t_0+1}\right]\left(1-\tau_{t_0+1}\right)}{\left[\left(1-a_{t_0+1}\right) a_{t_0+1}\right]\begin{bmatrix} 1-\lambda\left(1-\tau_{t_0+1}\right) & \lambda\left(1-\tau_{t_0+1}\right) \\ \tau_{t_0+1} & 1-\tau_{t_0+1} \end{bmatrix}\begin{bmatrix} \lambda \\ 1 \end{bmatrix}} - 1\right)
\end{aligned}
$$

$$
\begin{aligned}
&= Z_{t_0}\left(1-\tau_{t_0+1}\right)^{-1}\left(\frac{\left[\lambda\left(1-a_{t_0+1}\right)+a_{t_0+1}\right]\left(1-\tau_{t_0+1}\right)}{\left[\left(1-a_{t_0+1}\right) a_{t_0+1}\right]\begin{bmatrix} \alpha_0 \\ \omega_0 \end{bmatrix}} - 1\right) \\
&= Z_{t_0}\left(1-\tau_{t_0+1}\right)^{-1}\left(\frac{\left[\lambda\left(1-a_{t_0+1}\right)+a_{t_0+1}\right]\left(1-\tau_{t_0+1}\right)}{\left[\alpha_0\left(1-a_{t_0+1}\right)+\omega_0 a_{t_0+1}\right]} - 1\right) \qquad (14)
\end{aligned}
$$

where $1 > \alpha_0 > \lambda$, $1 > \omega_0 > \lambda$, and $\omega_0 = \lambda\tau_{t_0+1} + (1-\tau_{t_0+1}) > (1-\tau_{t_0+1})$. Thus, $\frac{\left[\lambda\left(1-a_{t_0+1}\right)+a_{t_0+1}\right]\left(1-\tau_{t_0+1}\right)}{\left[\alpha_0\left(1-a_{t_0+1}\right)+\omega_0 a_{t_0+1}\right]} < 1$, implying $\frac{\partial V\left(k_{t_0}, k_{t_0}^M, t_0\right)}{\partial \tau_{t_0+1}} < 0$; hence $\tau_{t_0+1} = 0$. Now, to proceed by induction, assume $\tau_s = 0$ for $s = t_0+1, \ldots, m-1$ where $m > t_0+1$, and consider

$$
\frac{\partial V\left(k_{t_0}, k_{t_0}^M, t_0\right)}{\partial \tau_m} = Z_{t_0}\beta^{m-t_0}\left(1-\tau_m\right)^{-1}\left(\frac{\left[\lambda\left(1-a_m\right)+a_m\right]}{\left(\lim_{n\to\infty} B\left(2n+1\right)\right)}\left(1-\tau_m\right)-1\right)
$$

Note that from the definition of $B(2n+1)$ and the assumption that $\tau_s = 0$ for $s = t_0+1, \ldots, m-1$, we can write the above as:

$$
\begin{aligned}
\frac{\partial V\left(k_{t_0}, k_{t_0}^M, t_0\right)}{\partial \tau_m} &= Z_{t_0}\beta^{m-t_0}\left(1-\tau_m\right)^{-1} \\
&\times \left(\frac{\left[\lambda\left(1-a_m\right)+a_m\right]\left(1-\tau_m\right)}{\left[1\, 0\right]\begin{bmatrix} 1-a_m & a_m \\ b_m & 1-b_m \end{bmatrix}\begin{bmatrix} 1-\lambda & \lambda \\ 0 & 1 \end{bmatrix}^{m-t_0-1}\begin{bmatrix} \lambda \\ 1 \end{bmatrix}} - 1\right)
\end{aligned}
$$

Now, we note that $\begin{bmatrix} 1-\lambda & \lambda \\ 0 & 1 \end{bmatrix}^{m-t_0-1} = \begin{bmatrix} 1-z & z \\ 0 & 1 \end{bmatrix}$, so that

$$\frac{\partial V\left(k_{t_0}, k_{t_0}^M, t_0\right)}{\partial \tau_m} = Z_{t_0}\beta^{m-t_0}\left(\frac{\lambda\left(1-a_m\right)+a_m}{\left(\lambda+\left(1-\lambda\right)z\right)\left(1-a_m\right)+a_m} - \left(1-\tau_m\right)^{-1}\right) < 0 \tag{15}$$

This implies that $\tau_m = 0$. This proves part (b) of the Theorem. [4] Note however that if $\tilde{\tau}_{t_0} < 1$, then if $\tau_s = 0$ for $s > t_0$, $\left(Q_{t_0}\prod_{s=t_0+1}^{m}\left(1-\tau_s\right)-1\right) > 0$, so that $\tau_m > 0$, which is a contradiction. Therefore if $\tilde{\tau}_{t_0} < 1$, $\tau_s = 0$ for $s > t_0$ is not an optimizing choice, as claimed in part (c).
Q.E.D.

We can further refine Theorem 2 to pin down the optimal tax sequence for the poor median voter.

Corollary 1 *If $\tilde{\tau}_{t_0} < 1$, there exists an $\tilde{m} > t_0$ with $\hat{\tau}_{\tilde{m}} \in [0, \tilde{\tau}_{\tilde{m}}]$ such that $\{\tilde{\tau}_{t_0}, \tilde{\tau}_{t_0+1}, ..\tilde{\tau}_{\tilde{m}-1}, \hat{\tau}_{\tilde{m}}, 0, 0, ...\}$ is the optimal tax sequence chosen by the median voter. Furthermore, $\tilde{m}$, which can be taken as a measure of the degree of redistribution, is non-increasing in $v_{t_0}^M$.*

Proof: To study the optimal tax sequence when $\tau_{t_0} = \tilde{\tau}_{t_0} < 1$, we note the following

$$Q_{t_0,t_0+1} = \left(\frac{n^{-1}(1+x_{t_0+1})}{\left(1-\tau_{t_0}\right)v_{t_0}^M + n^{-1}\left(\tau_{t_0}+x_{t_0+1}\right)}\right) \times \left(\frac{\left[\lambda\left(1-a_{t_0+1}\right)+a_{t_0+1}\right]}{\left[\left(1-a_{t_0+1}\right)a_{t_0+1}\right]\begin{bmatrix} 1-\lambda\left(1-\tau_{t_0+1}\right) & \lambda\left(1-\tau_{t_0+1}\right) \\ \tau_{t_0+1} & 1-\tau_{t_0+1} \end{bmatrix}\begin{bmatrix} \lambda \\ 1 \end{bmatrix}}\right)$$

[4] In the remarks following Theorem 1, equation (10), we showed that if $\tau_s = \tau$, $\begin{bmatrix} 1-\lambda\left(1-\tau\right) & \lambda\left(1-\tau\right) \\ \tau & 1-\tau \end{bmatrix}^k$ converges, as $k \to \infty$, to $\begin{bmatrix} \frac{\tau}{\lambda(1-\tau)+\tau} & \frac{\lambda(1-\tau)}{\lambda(1-\tau)+\tau} \\ \frac{\tau}{\lambda(1-\tau)+\tau} & \frac{\lambda(1-\tau)}{\lambda(1-\tau)+\tau} \end{bmatrix}$. It is easy to compute Q_{t_0} for constant $\tau_s = \tau > 0$ as well, since in that case $x_{t_0+1} = \tau\left(\lambda^{-1}-1\right)$, $a_m = \frac{\lambda(1-\tau)}{\lambda(1-\tau)+\tau}$. Then

$$Q_{t_0} = \left(\frac{n^{-1}(1+x_{t_0+1})}{\left(1-\tau_{t_0}\right)v_{t_0}^M + n^{-1}\left(\tau_{t_0}+x_{t_0+1}\right)}\right) \geq 1$$

because $\left[\lambda\left(1-a_m\right)+a_m\right]\left(1+x_{t_0+1}\right) = 1$. But then $\prod_{s=t_0+1}^{m}\left(1-\tau_s\right)-1 = \left(Q_{t_0}\left(1-\tau\right)^{m-t_0}-1\right) < 0$ for large m if $\tau > 0$. It is of course possible to compute a_m and x_{t_0+1} for policies which keep $\tau_s = \tau$ or $\tau_s = \tilde{\tau}_s$ for a finite number of periods from t_0+1 on and then set taxes to zero, as suggested by equation (11), and check for the optimality of this sequence.

$$= \left(\frac{n^{-1}(1+x_{t_0+1})}{(1-\tau_{t_0})\, v_{t_0}^M + n^{-1}(\tau_{t_0}+x_{t_0+1})} \right) \left(\frac{[\lambda(1-a_{t_0+1}) + a_{t_0+1}]}{\left[(1-a_{t_0+1})\, a_{t_0+1}\right] \begin{bmatrix} \alpha_0 \\ \omega_0 \end{bmatrix}} \right)$$

as in equation (14), where $1 > \alpha_0 > \lambda$, $1 > \omega_0 > \lambda$, and $\omega_0 = \lambda\tau_{t_0+1} + (1-\tau_{t_0+1}) > (1-\tau_{t_0+1})$. If $Q_{t_0,t_0+1} \le 1$, we set $\tau_s = 0$ for $s > t_0$. If $Q_{t_0,t_0+1} > 1$, we set $\tau_{t_0+1} = \hat{\tau}_{t_0+1}$ such that $Q_{t_0,t_0+1}(1-\hat{\tau}_{t_0+1}) = 1$, provided $\tau_{t_0+1} \le \tilde{\tau}_{t_0+1}$, and we set $\tau_s = 0$ for $s > t_0+1$. (Note from 15 that $\frac{\partial V\left(k_{t_0}, k_{t_0}^M, t_0\right)}{\partial \tau_m} < 0$ because a_m is independent of τ_m, implying that $\hat{\tau}_{t_0+1}$ is unique, although not necessarily less than $\tilde{\tau}_{t_0+1}$.) Otherwise if $\tau_{t_0+1} > \tilde{\tau}_{t_0+1}$, we set $\tau_{t_0+1} = \hat{\tau}_{t_0+1} = \tilde{\tau}_{t_0+1}$ and choose $\hat{\tau}_{t_0+2}$ so that $Q_{t_0,t_0+2}(1-\tilde{\tau}_{t_0+1})(1-\hat{\tau}_{t_0+2}) = 1$, provided $\hat{\tau}_{t_0+2} \le \tilde{\tau}_{t_0+2}$. If $\hat{\tau}_{t_0+2} > \tilde{\tau}_{t_0+2}$, we continue setting $\tau_s = \hat{\tau}_s = \tilde{\tau}_s,\ s > t_0+1$ until the first $s = s'$ where $Q_{t_0,t_0+s'}\left(\prod_{s=t_0+1}^{s'-1}(1-\hat{\tau}_s)\right)(1-\hat{\tau}_{s'}) = 1$ and $\tau_{s'} \le \tilde{\tau}_s$, and then we set $\tau_s = 0$ for $s > s'$. To prove that this is the optimal sequence of taxes, we have to show that $Q_{t_0,m}\prod_{s=t_0+1}^{m}(1-\hat{\tau}_s) - 1 < 0$ for $m > s'$ so that $\frac{\partial V\left(k_{t_0}, k_{t_0}^M, t_0\right)}{\partial \tau_m} < 0$ if $\tau_m > 0$, which is a contradiction. We will make heavy use of the averaging properties of stochastic matrices.

Let τ_n be the first non-zero tax rate where $n > s'$. Then

$$\begin{aligned} Q_{t_0,s'} \prod_{s=t_0+1}^{s'}(1-\tau_s) &= \left(\frac{n^{-1}(1+x_{t_0+1})}{(1-\tau_{t_0})\, v_{t_0}^M + n^{-1}(\tau_{t_0}+x_{t_0+1})} \right) [\lambda(1-a_{s'}) + a_{s'}] \\ &\quad \times (1+x_{t_0+1}) \prod_{s=t_0+1}^{s'}(1-\tau_s) \\ &= \left(\frac{n^{-1}(1+x_{t_0+1})}{(1-\tau_{t_0})\, v_{t_0}^M + n^{-1}(\tau_{t_0}+x_{t_0+1})} \right) \\ &\quad \times \left(\frac{(\lambda(1-a_{s'}) + a_{s'}) \prod_{s=t_0+1}^{s'}(1-\tau_s)}{\left[(1-a_{s'})\, a_{s'}\right] \prod_{s=t_0+1}^{s'} \begin{bmatrix} 1-\lambda(1-\hat{\tau}_s) & \lambda(1-\hat{\tau}_s) \\ \hat{\tau}_s & (1-\hat{\tau}_s) \end{bmatrix} \begin{bmatrix} \lambda \\ 1 \end{bmatrix}} \right) \end{aligned}$$

$$= \left(\frac{n^{-1}(1+x_{t_0+1})}{(1-\tau_{t_0})\, v_{t_0}^M + n^{-1}(\tau_{t_0}+x_{t_0+1})} \right) \left(\frac{(\lambda(1-a_{s'}) + a_{s'}) \prod_{s=t_0+1}^{s'}(1-\tau_s)}{\left[(1-a_{s'})\, a_{s'}\right] \begin{bmatrix} \alpha \\ \omega \end{bmatrix}} \right)$$

$$= \left(\frac{n^{-1}(1+x_{t_0+1})}{(1-\tau_{t_0})\, v_{t_0}^M + n^{-1}(\tau_{t_0}+x_{t_0+1})} \right) \left(\frac{(\lambda(1-a_{s'}) + a_{s'}) \prod_{s=t_0+1}^{s'}(1-\tau_s)}{\alpha(1-a_{s'}) + a_{s'}\omega} \right) = 1$$

with $1 > \omega > \lambda$ and $1 > \alpha > \lambda$. Now let

$$\begin{bmatrix} 1-z & z \\ y & 1-y \end{bmatrix} = \begin{bmatrix} 1-\lambda\left(1-\hat{\tau}_n\right) & \lambda\left(1-\hat{\tau}_n\right) \\ \hat{\tau}_n & \left(1-\hat{\tau}_n\right) \end{bmatrix} \begin{bmatrix} 1-\lambda & \lambda \\ 0 & 1 \end{bmatrix}^{n-1-s'}$$
$$= \begin{bmatrix} 1-\lambda\left(1-\hat{\tau}_n\right) & \lambda\left(1-\hat{\tau}_n\right) \\ \hat{\tau}_n & \left(1-\hat{\tau}_n\right) \end{bmatrix} \begin{bmatrix} 1-q & q \\ 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 1-z & z \\ \left(1-q\right)\hat{\tau}_n & q\hat{\tau}_n + \left(1-\hat{\tau}_n\right) \end{bmatrix}$$

so

$$Q_{t_0,s'} \prod_{s=t_0+1}^{s'} \left(1-\tau_s\right) = \left(\frac{n^{-1}(1+x_{t_0+1})}{\left(1-\tau_{t_0}\right) v_{t_0}^M + n^{-1}\left(\tau_{t_0} + x_{t_0+1}\right)}\right) \times \left(\frac{\left(\lambda\left(1-a_{s'}\right)+a_{s'}\right)\prod_{s=t_0+1}^{s'}\left(1-\tau_s\right)}{\alpha\left(1-a_{s'}\right)+a_{s'}\omega}\right)$$
$$= \left(\frac{n^{-1}(1+x_{t_0+1})}{\left(1-\tau_{t_0}\right) v_{t_0}^M + n^{-1}\left(\tau_{t_0} + x_{t_0+1}\right)}\right) \prod_{s=t_0+1}^{s'} \left(1-\tau_s\right) \times \left(\frac{\begin{bmatrix}\left(1-a_n\right) & a_n\end{bmatrix} \begin{bmatrix} 1-z & z \\ \left(1-q\right)\hat{\tau}_n & q\hat{\tau}_n + \left(1-\hat{\tau}_n\right) \end{bmatrix} \begin{bmatrix} \lambda \\ 1 \end{bmatrix}}{\begin{bmatrix}\left(1-a_n\right) & a_n\end{bmatrix} \begin{bmatrix} 1-z & z \\ \left(1-q\right)\hat{\tau}_n & q\hat{\tau}_n + \left(1-\hat{\tau}_n\right) \end{bmatrix} \begin{bmatrix} \alpha \\ \omega \end{bmatrix}}\right)$$

$$Q_{t_0,s'} \prod_{s=t_0+1}^{s'} \left(1-\tau_s\right) = \left(\frac{n^{-1}(1+x_{t_0+1})}{\left(1-\tau_{t_0}\right) v_{t_0}^M + n^{-1}\left(\tau_t + x_{t_0+1}\right)}\right) \prod_{s=t_0+1}^{s'} \left(1-\tau_s\right) \times \left(\frac{\begin{bmatrix}\left(1-a_n\right) & a_n\end{bmatrix} \begin{bmatrix} \lambda\left(1-z\right)+z \\ \lambda\left(1-q\right)\hat{\tau}_n + q\hat{\tau}_n + \left(1-\hat{\tau}_n\right) \end{bmatrix}}{\alpha'\left(1-a_n\right)+\omega' a_n}\right)$$
$$= \left(\frac{n^{-1}(1+x_{t_0+1})}{\left(1-\tau_{t_0}\right) v_{t_0}^M + n^{-1}\left(\tau_{t_0} + x_{t_0+1}\right)}\right) \prod_{s=t_0+1}^{s'} \left(1-\tau_s\right) \times \left(\frac{\left[\left(1-a_n\right)\left(\lambda\left(1-z\right)+z\right) + a_n\left(\lambda\left(1-q\right)\hat{\tau}_n + q\hat{\tau}_n + \left(1-\hat{\tau}_n\right)\right)\right]}{\alpha'\left(1-a_n\right)+\omega' a_n}\right)$$
$$= \left(\frac{n^{-1}(1+x_{t_0+1})}{\left(1-\tau_{t_0}\right) v_{t_0}^M + n^{-1}\left(\tau_{t_0} + x_{t_0+1}\right)}\right) \prod_{s=t_0+1}^{s'} \left(1-\tau_s\right) \times \left(\frac{\left[\left(1-a_n\right)\left(\lambda + z\left(1-\lambda\right)\right) + a_n\left(1-\hat{\tau}_n\right) + a_n\left(\lambda\left(1-q\right)\hat{\tau}_n + q\hat{\tau}_n\right)\right]}{\alpha'\left(1-a_n\right)+\omega' a_n}\right)$$

where

$$\begin{bmatrix} 1-z & z \\ (1-q)\,\hat{\tau}_n & q\hat{\tau}_n + (1-\hat{\tau}_n) \end{bmatrix} \begin{bmatrix} \alpha \\ \omega \end{bmatrix} = \begin{bmatrix} \alpha' \\ \omega' \end{bmatrix}$$

We also have

$$\begin{aligned} Q_{t_0,n} \prod_{s=t_0+1}^{n} (1-\tau_s) &= \left(\frac{n^{-1}(1+x_{t_0+1})}{(1-\tau_{t_0})\,v_{t_0}^M + n^{-1}(\tau_{t_0}+x_{t_0+1})} \right) \left(\prod_{s=s'+1}^{n} (1-\tau_s) \right) \\ & \left(\frac{[\lambda(1-a_n)+a_n] \prod_{s=t_0+1}^{s'} (1-\tau_s)}{\left[(1-a_n)\,a_n\right] \begin{bmatrix} 1-z & z \\ y & 1-y \end{bmatrix} \begin{bmatrix} \alpha \\ \omega \end{bmatrix}} \right) \\ &= \left(\frac{n^{-1}(1+x_{t_0+1})}{(1-\tau_{t_0})\,v_{t_0}^M + n^{-1}(\tau_{t_0}+x_{t_0+1})} \right) \\ & \left(\frac{[\lambda(1-a_n)+a_n]\,(1-\hat{\tau}_n) \prod_{s=t_0+1}^{s'} (1-\tau_s)}{\alpha'(1-a_n)+\omega' a_n} \right) \end{aligned}$$

Since

$$\begin{aligned} & [(1-a_n)(\lambda + z(1-\lambda)) + a_n(1-\hat{\tau}_n) + a_n(\lambda(1-q)\,\hat{\tau}_n + q\hat{\tau}_n)] \\ > & [\lambda(1-a_n)+a_n](1-\hat{\tau}_n) \end{aligned}$$

comparing $Q_{t_0,s'} \prod_{s=t_0+1}^{s'} (1-\tau_s)$ with $Q_{t_0,n} \prod_{s=t_0+1}^{n} (1-\tau_s)$ yields

$$1 = Q_{t_0,s'} \prod_{s=t_0+1}^{s'} (1-\tau_s) > Q_{t_0,n} \prod_{s=t_0+1}^{n} (1-\tau_s)$$

and $\frac{\partial V\left(k_t, k_t^i, t\right)}{\partial \tau_n} < 0$, contradicting $\tau_n > 0$, and therefore $\tau_s = 0$ for $s > s'$. Also from the definition of $Q_{t_0,m}$ and the fact that $\tau_{t_0} = \tilde{\tau}_{t_0}$ irrespective of $v_{t_0}^M$ if the median agent is poor, it follows that $\tilde{m}$ is decreasing in $v_{t_0}^M$. QED.

Finally, in case the median voter is rich, that is if $v_s^M > n^{-1}$, we have the following:

Corollary 2 *If the median voter is rich, he sets the tax sequence to* $\{0, 0, 0, \ldots\}$.

Proof. Intuitively this is obvious. Formally, for use in later theorems,

$$\frac{\partial V\left(k_{t_0},k_{t_0}^M,t_0\right)}{\partial\tau_m} = \left(Z_{t_0}^{\sigma}X_{t_0}^{1-\sigma}\right)\left(X_{t_0}^{-1}\frac{\partial X_{t_0}}{\partial\tau_m}+\sigma\left(1-\sigma\right)^{-1}Z_{t_0}^{-1}\frac{\partial Z_{t_0}}{\partial\tau_m}\right)$$

$$= Z_{t_0}\beta^{m-t_0}\left(\left(\frac{n^{-1}(1+x_{t_0+1})}{v_{t_0}^M+\left(n^{-1}-v_{t_0}^M\right)\tau_{t_0}+n^{-1}x_{t_0+1}}\right)\right.$$

$$\left.\times\left(\frac{\lambda\left(1-a_m\right)+a_m}{\left(\lambda+\left(1-\lambda\right)z\right)\left(1-a_m\right)+a_m}\right)-\left(1-\tau_m\right)^{-1}\right)<0$$

from (15) and from noting that $\left(\frac{n^{-1}(1+x_{t_0+1})}{v_{t_0}^M+\left(n^{-1}-v_{t_0}^M\right)\tau_{t_0}+n^{-1}x_{t_0+1}}\right)<1$ if $n^{-1}<v_{t_0}^M$ Q.E.D.

3.2 Political Constraints

We can now study the political constraints on the median voter that prevent him from implementing the preferred tax scheme: if the median voter is poor, he wants $\tau_{t_0}=1$ and $\tau_{t_0+s}=0$ for $s=1,2,\ldots$, whereas if the median voter is rich, he wants $\tau_{t_0+s}=0$ for $s=0,1,2,\ldots$

There is a pivotal agent w on the right whose share of initial capital, larger than the average share, is denoted by $v_{t_0}^w$. He prefers the tax scheme $\tau_{t_0+s}=0$ for $s=0,1,2,\ldots$. The pivotal agent on the left, p, has an initial share of capital smaller than or equal to the share of the median voter: $v_{t_0}^p\leq v_{t_0}^i$. This agent wants $\tau_{t_0}=1$ and $\tau_{t_0+s}=0$ for $s=1,2,\ldots$, a complete redistribution in the first period followed by zero taxes afterwards. If in any period, the pivotal agents receive less discounted utility under democracy than under their preferred authoritarian regime, they will institute an authoritarian regime.

Assumption 3 *Let t_a be the first period in which an authoritarian regime is established. Then $\tau_{t_a}\in[0,1]$, and $\tau_s\in[0,\tilde{\tau}_s]$, where $\tilde{\tau}_s<1$ for all $s>t_a$.*

This assumption allows the pivotal agent to reset initial taxes when she reverts to an authoritarian regime. We assume, for simplicity, that once established, an authoritarian regime lasts forever.

To sustain democracy, the median voter has to accommodate by setting taxes that will keep the pivotal agents on the right and left from establishing authoritarian regimes. Since the median voter is poor, we start by analyzing the tax sequence that the poor median voter must set to keep the rich pivotal agent from establishing right-wing authoritarianism. In the next section we will study the conditions under which the median voter can simultaneously prevent left and right authoritarian regimes.

Under log preferences the Euler equations are:

$$c_{t+1} = \beta r \left(1 - \tau_{t+1}\right) c_t$$

and the median voter's discounted utility is

$$V\left(k_{t_0}, k^i_{t_0}, t_0\right) = \sum_{t=t_0}^{\infty} \beta^{t-t_0} \log c_t$$

$$V\left(k_{t_0}, k^i_{t_0}, t_0\right) = \log c_{t_0} + \beta \log c_{t_0} \beta r \left(1 - \tau_{t_0+1}\right) + \beta^2 \log c_{t_0} \left(\beta r\right)^2 \left(1 - \tau_{t_0+1}\right)\left(1 - \tau_{t_0+2}\right) + ...$$

$$V\left(k_{t_0}, k^i_{t_0}, t_0\right) = \left(1 - \beta\right)^{-1} \log c^i_{t_0} + \beta \left(1 - \beta\right)^{-2} \log \beta r + \sum_{n=t_0+1}^{\infty} \beta^{n-t_0} \log \prod_{s=t_0+1}^{n} \left(1 - \tau_s\right)$$

Iterating the budget forward, as before,

$$c^i_{t_0} = \lambda_{t_0} \left(\left(1 - \tau_{t_0}\right) r k^i_{t_0} + n^{-1} \left(\tau_{t_0} + \sum_{j=t_0+1}^{\infty} \tau_j \prod_{s=t_0+1}^{j} g_s (r \left(1 - \tau_s\right))^{-1} \right) r k_{t_0} \right)$$

$$= \lambda_{t_0} \left(\left(1 - \tau_{t_0}\right) v^i_{t_0} + n^{-1} \left(\tau_{t_0} + \sum_{j=t_0+1}^{\infty} \tau_j \prod_{s=t_0+1}^{j} g_s (r \left(1 - \tau_s\right))^{-1} \right) \right) r k_{t_0}$$

where $\lambda_{t_0} = \lambda = 1 - \beta$, and $v^i_{t_0}$ is the initial share of the median voter. Note: If $\tau_s = 0, \ s > t_0$

$$V\left(k_{t_0}, k^i_{t_0}, t_0\right) = \left(1 - \beta\right)^{-1} \log c_{t_0} + \beta \left(1 - \beta\right)^{-2} \log \beta r$$

Thus under democracy median voter chooses taxes so that

$$V\left(k_{t_0}, k^i_{t_0}, t_0\right) = Max_{\{\tau\}^{\infty}_{t_0}} \ \left(1 - \beta\right)^{-1} \log c_{t_0} + \beta \log \left(1 - \beta\right)^{-2} \beta r + \sum_{n=t_0+1}^{\infty} \beta^{n-t_0} \log \prod_{s=t_0+1}^{n} \left(1 - \tau_s\right)$$

while agent with share $v^w_{t_0}$ gets

$$V^{wD}\left(k_{t_0}, k^w_{t_0}, t_0\right) = \left(1 - \beta\right)^{-1} \log c^w_{t_0} + \beta \left(1 - \beta\right)^{-2} \log \beta r + \sum_{n=t_0+1}^{\infty} \beta^{n-t_0} \log \prod_{s=t_0+1}^{n} \left(1 - \tau_s\right)$$

where

$$c^w_{t_0} = \lambda \left(\left(1 - \tau_{t_0}\right) r k^w_{t_0} + n^{-1} \left(\tau_{t_0} + \sum_{j=t_0+1}^{\infty} \tau_j \prod_{s=t_0+1}^{j} g_s (r \left(1 - \tau_s\right))^{-1} \right) r k_{t_0} \right)$$

$$= \lambda \left(\left(1 - \tau_{t_0}\right) v^w_{t_0} + n^{-1} \left(\tau_{t_0} + \sum_{j=t_0+1}^{\infty} \tau_j \prod_{s=t_0+1}^{j} g_s (r \left(1 - \tau_s\right))^{-1} \right) \right) r k_{t_0}$$

The alternative strategy of the rich pivotal agent is to switch to authoritarianism and implement zero taxes. We assume that discounted utility or value is scaled down under authoritarianism so that the discounted utility of the rich pivotal agent becomes:

$$V^{wA}\left(k_{t_0}, k_{t_0}^{w}, t_0\right) = \mu^{w}\left((1-\beta)^{-1}\log c_{t_0} + \beta(1-\beta)^{-2}\log\beta r\right)$$

$$\begin{aligned} c_{t_0} &= \lambda\left((1-\tau_{t_0})\, v_{t_0}^{w} + n^{-1}\tau_{t_0}\right) r k_{t_0} \\ &= \lambda v_{t_0}^{w} r k_{t_0} \end{aligned}$$

where

$$\begin{aligned} \mu^{w} < 1 \text{ if } (1-\beta)^{-1}\log\left(\lambda v_{t_0}^{s} r k_{t_0}\right) + \beta(1-\beta)^{-2}\log\beta r > 0 \\ \mu^{w} > 1 \text{ if } (1-\beta)^{-1}\log\left(\lambda v_{t_0}^{s} r k_{t_0}\right) + \beta(1-\beta)^{-2}\log\beta r < 0 \end{aligned}$$

$$\begin{aligned} \mu^{w} < 1 \text{ if } 1 &> \frac{(\beta r)^{-\beta(1-\beta)^{-1}}}{\lambda v_{t_0}^{s} r k_{t_0}} = \frac{k^{*}}{k_{t_0}} \\ \mu^{w} \geq 1 \text{ if } 1 &< \frac{(\beta r)^{-\beta(1-\beta)^{-1}}}{\lambda v_{t_0}^{s} r k_{t_0}} = \frac{k^{*}}{k_{t_0}} \\ \text{where } k^{*} &= \frac{(\beta r)^{-\beta(1-\beta)^{-1}}}{\lambda v_{t_0}^{s} r} \end{aligned}$$

$$\text{Therefore} \quad \left(\frac{k^{*}}{k_{t_0}}\right)^{1-\mu^{w}} < 1$$

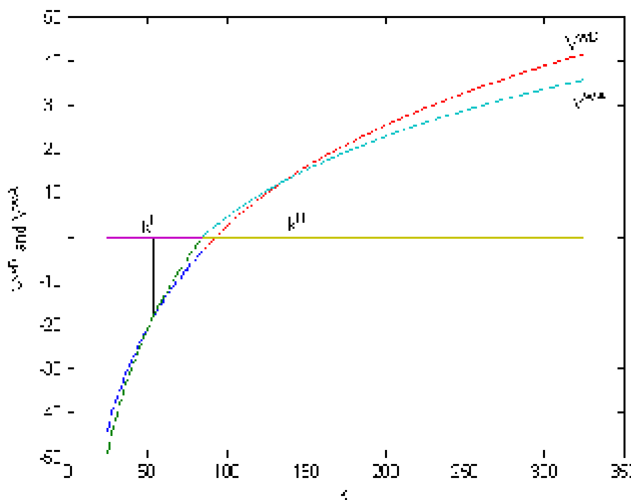
Before proceeding with a formal analysis, we will explore the features and content of our assumption about the costs of dictatorship, that is of μ^{w}.

First let us note a technical point: μ^{w} represents the cost of authoritarianism imposed on the discounted utility stream, not on period by period utility. If we had imposed it on period by period utility, the Euler equation would have been affected during a period where we cross from μ^{w} at or above one to below one: agents anticipating the changes in marginal utility across periods due to the switch in μ^{w} would change their savings. For simplicity therefore we impose the cost on lifetime utility. Note that if we constrained initial capital to a range where $\mu^{w} < 1$, and the economy grew, then period by period costs are the same as the lifetime utility costs: since μ^{w} remains constant, savings and the growth of consumption is unaffected by μ^{w}.Note also that μ^{w} jumps precisely when the value function crosses zero, so the value function remains continuous. To keep track, we define $\mu^w = \mu^{wL}$ if $\mu^w < 1$ and $\mu^w = \mu^{wH}$ if $\mu^w \geq 1$.

The specification with μ^w assures that, for any capital stock level and shares, utility is weakly lower under authoritarianism than it would be under democracy. Without costs of reverting to an authoritarian regime implicit in the μ^w, democracy may be unsustainable: the rich pivotal agent would always prefer to prevent a redistribution by the median voter and to implement zero taxes[5]. However, if $\mu^w < 1$, the pivotal agent may still find it advantageous to incur the costs of authoritarianism, if the redistribution of shares through taxation by the median voter proves sufficiently costly. To highlight the features of our specification of the costs of authoritarianism and compare it to alternatives, before a formal analysis we present a simplified graphical exposition. The value of sticking with democracy for the rich pivotal agent depends on the tax sequence chosen by the median voter, as well as on the stock of capital and its distribution across agents. By way of example, consider a fully redistributive tax sequence of the form $\{1, 0, 0...\}$ that the median voter would like to impose, and let us analyze whether it will be sustainable under democracy. In this illustrative example we ignore the incentive constraints of the poor pivotal agent, although we will incorporate them into our analysis later on. It is relatively straightforward to study whether a particular tax system is "unsustainable" in the initial period, but "sustainability" will require that the tax sequence remain sustainable not just initially but at all points in the future as well. We leave a formal study of "sustainability" for later, and proceed with an illustrative example for the initial period.

Let $n = 10,\ v^w = 0.11,\ \beta = 0.97,\ r = 1.07,\ \mu^{wH} = 1.2,\ \mu^{wL} = 0.8$. The value function of the rich pivotal agent under democracy where the median voter fully redistributes, V^{wD}, and under authoritarianism where taxes are zero, V^{wA}, can be plotted as follows:

[5]See however section **??** where a revolution succeeds with probability less than 1.



V^{wD} and V^{wA}

The rich pivotal agent will prefer to revert to an authoritarian regime if $k^L < k_{t_0} < k^H$. If we confine our attention to the first period only, and ignore the incentive constraints of the poor pivotal agent, we see that democracy can be sustained if $k_{t_0} \geq k^H$ or $k_{t_0} \leq k^L$. Of course democracy may become unsustainable along the growth path in the future, and agents who foresee this will take it into account in making their current decisions (see section 4.2 below). The figure above nevertheless illustrates the crucial feature of our particular specification of the costs of authoritarianism: whether a particular tax sequence, in this case $\{1, 0, 0...\}$, is sustainable or not, depends on the level of capital, that is, the "sustainability of democracy" is wealth dependent (see equation 19). By contrast, we may consider simple CRRA preferences, $U(c) = \mu^w (1-\sigma)^{-1} c^{(1-\sigma)}$, $\mu^w < (>) 1$, $\sigma < (>) 1$. In the CRRA case it is easy to show that the sustainability under democracy of a particular tax sequence depends on parameters and shares, but not on the stock of capital. The wealth dependence of "sustainability" which we are trying to capture would be absent in this case. To introduce wealth dependence into the CRRA case, we may modify preferences either as $U(c) = \mu^w (1-\sigma)^{-1} \left(c^{(1-\sigma)} - 1\right)$, which is the proper generalization of the logarithmic case, and amounts to introducing a differential fixed cost that favors utility under democracy over utility under authoritarianism. As in the logarithmic case, care must be given to set μ^w above or below unity as utility is positive or negative. An alternative specification of utility under authoritarianism that also delivers wealth dependence is

$U(c) = (1-\sigma)^{-1} c^{(1-\sigma)\mu^w}$. Here as well, if utility under democracy is to dominate utility under authoritarianism, μ^w must be set above or below unity, as consumption c is above or below unity. A common feature of all these non-homothetic specifications is that at some low level of consumption, utility levels under democracy and authoritarianism coincide (in the latter case as in the log case for example, at $c = 1$). At higher consumption or wealth levels (since consumption is proportional to wealth) agents receive higher utility under democracy than under authoritarianism because a marginal unit of consumption produces more utility in a democratic society than in an authoritarian one: the slope of utility is steeper under democracy. At lower levels of consumption, where the utility levels under democracy and dictatorship are close, a left dictatorship that redistributes in order to raise the endowment of the poor may be preferred by the left, and a right dictatorship that prevents the democratic implementation of redistributive taxes may be preferred by the right. At higher wealth levels however, beyond a threshold $\left(k^H\right)$, dictatorship is longer be attractive because it is not worth it.

An undesirable aspect of the specification of preferences is the following. If we set $\mu^w > 1$ for $1 < \frac{(\beta r)^{-\beta(1-\beta)^{-1}}}{\lambda v_{t_0}^s r k_{t_0}}$, as discounted utility becomes unbounded at high or low capital stocks, democracy always dominates: the benefits of blocking the median voter's redistribution and enforcing zero taxes will outweigh the costs of an authoritarian regime for the pivotal agent at intermediate levels of the capital, but not for sufficiently high or low levels. While it may be reasonable to think that the utility difference between repressive regimes and democratic ones grow in proportion to wealth and consumption, our logarithmic preference specification is unrealistic in its implication that the high costs of dictatorship necessarily overwhelm redistributive considerations at very low levels of wealth. Therefore under our specification of preferences, it may be sensible to confine attention to wealth levels above a certain threshold.

An alternative approach is to adopt a the specification that sets $\mu^w = 1$ for $1 < \frac{(\beta r)^{-\beta(1-\beta)^{-1}}}{\lambda v_{t_0}^s r k_{t_0}}$. Thus when wealth is low, the agents are indifferent to regime, but at higher wealth levels, democracy dominates dictatorship. The above arguments and reasoning still hold, and the figure above still has an intersection at k^H, but the lower intersection at k^L now disappears.[6] Indeed we can now set μ^w to switch from unity to

[6] Yet another approach is to introduce wealth dependendce into discounted utility through production where, for example, dictatorship becomes more detrimental to productive activity at higer levels of capital. We will not pursue this approach in this paper.

below unity at any level at or above k^*.[7] However, the assumption that the poor do not care at all about freedom is indeed hard to swallow. As Dasgupta (1993) puts it, the view that the poor do not care about the freedoms associated with democracy "is a piece of insolence that only those who don't suffer from their lack seem to entertain." (See also Sen 1994.)

While the above figure and analysis are illustrative, the sustainability or unsustainability of democracy cannot be studied in terms of a particular tax scheme. The median voter would seek to construct feasible tax sequences, and conditions for sustainability or unsustainability must be given in terms of initial aggregate wealth, the wealth shares and other parameters, that hold for all feasible tax sequences. This will be done in section 4.

For our analysis in subsequent sections we also define the discounted utility under the left wing dictatorship. The poor agent implements $\tau_{t_0} = 1$ to achieve full asset equality, and therefore has discounted utility:

$$V^{pA}\left(k_{t_0}, k_{t_0}^M, t_0\right) = \mu^p\left((1-\beta)^{-1}\log\left(\lambda n^{-1} r k_{t_0}\right) + \beta(1-\beta)^{-2}\log\beta r\right) \tag{16}$$

where μ^p is defined analogously to the case of rich pivotal agent, so that

$$\begin{aligned}\mu^p < 1 \text{ if } (1-\beta)^{-1}\log\left(\lambda n^{-1} r k_{t_0}\right) + \beta(1-\beta)^{-2}\log\beta r > 0 \\ \mu^p \geq 1 \text{ if } (1-\beta)^{-1}\log\left(\lambda n^{-1} r k_{t_0}\right) + \beta(1-\beta)^{-2}\log\beta r \leq 0\end{aligned}$$

or

$$\begin{aligned}\mu^p < 1 \text{ if } 1 > \frac{(\beta r)^{-\beta(1-\beta)^{-1}}}{\lambda n^{-1} r k_{t_0}} = \frac{k^{**}}{k_{t_0}} \\ \mu^p \geq 1 \text{ if } 1 \leq \frac{(\beta r)^{-\beta(1-\beta)^{-1}}}{\lambda n^{-1} r k_{t_0}} = \frac{k^{**}}{k_t} \\ \text{where } k^{**} = \frac{(\beta r)^{-\beta(1-\beta)^{-1}}}{\lambda n^{-1} r} > k^* = \frac{(\beta r)^{-\beta(1-\beta)^{-1}}}{\lambda v_{t_0}^w r}\end{aligned}$$

$$\text{Therefore } \left(\frac{k^{**}}{k_{t_0}}\right)^{1-\mu^p} < 1$$

In this section, we only consider the case where the right pivotal agent can overthrow democracy and establish an authoritarian regime.

[7] A switch below k^* would also be possible, but since discounted utility is negative, μ^w would first shift above unity and then to below unity after k^*, but this may produce multiple intersections.

The poor median voter M, to preserve democracy starting at time t_0, must set the sequence of taxes τ_{t_0+s}, $s = 0, 1, \ldots$ to maximize, for all $t = t_0 + s$,

$$Max_{\{\tau\}_{t_0}^{\infty}} \; (1-\beta)^{-1} \log \lambda \left((1-\tau_{t_0}) v_{t_0}^{M} + n^{-1} \left(\tau_{t_0} + \sum_{j=t_0+1}^{\infty} \tau_j \prod_{s=t_0+1}^{j} g_s (r(1-\tau_s))^{-1} \right) \right) r k_{t_0}$$

$$+ \beta (1-\beta)^{-2} \log \beta r + \sum_{n=t_0+1}^{\infty} \beta^{n-t_0} \log \prod_{s=t_0+1}^{n} (1-\tau_s)$$

subject to, for all $t \geq t_0$

$$\begin{aligned} (1-\beta)^{-1} \log \lambda \left((1-\tau_{t_0}) v_{t_0}^{w} + n^{-1} \left(\tau_{t_0} + \textstyle\sum_{j=t_0+1}^{\infty} \tau_j \prod_{s=t_0+1}^{j} g_s (r(1-\tau_s))^{-1} \right) \right) r k_{t_0} \\ + \beta (1-\beta)^{-2} \log \beta r + \textstyle\sum_{n=t_0+1}^{\infty} \beta^{n-t_0} \log \prod_{s=t_0+1}^{n} (1-\tau_s) \\ \geq \mu^{w} \left((1-\beta)^{-1} \log \left(\lambda v_{t_0}^{w} r k_{t_0} \right) + \beta (1-\beta)^{-2} \log \beta r \right) \end{aligned} \tag{17}$$

The dynamics of shares evolve as described before: if taxes are constant from $t_0 + 1$ onwards, the shares are also constant at $v_s^w = (1-\tau) v_t^w + n^{-1}\tau$ for all $s > t_0$. We will ignore the constraint by the poor pivotal agent for the moment, until the next section.

Theorem 3 *a. If preferences are logarithmic, under democracy the poor median voter will choose a tax sequence* $\lim_{t\to\infty} \tau_t = 0$.

b. If preferences are logarithmic and $\hat{\tau}_{t_0} = 1$, *then* $\{1, 0, 0...\}$ *is an optimal tax sequence for the poor median voter.*

c. Let $k^* = \frac{(\beta r)^{-\beta(1-\beta)^{-1}}}{\lambda v_{t_0}^{w} r}$. *Then,* $\hat{\tau}_{t_0} = 1$ *iff* $\left(\frac{v_{t_0}^{w}}{n^{-1}} \right) \left(\frac{k^*}{k_{t_0}} \right)^{1-\mu^w} \leq 1$.

Remark 3 *It is not surprising that the parts a and b of Theorem 2 also apply in the case of political constraints. If it is optimal to set taxes to zero in the limit without having to keep the rich pivotal agent in check, then it will be optimal to do so in the presence of a rich pivotal agent as well, since the rich pivotal agent prefers zero taxes. For t sufficiently large, the incentive constraints of the rich pivotal agent will be slack, so the the optimal tax policy is not time consistent. However, see Corollary 4 and Theorem 7 below.*

Remark 4 *The optimal policy of the median voter can be implemented as a stationary policy. Suppose the share of the median voter is* v_t^i. *After the first period, where the poor median voter sets* $\tau_{t_0} = \tilde{\tau}_{t_0}$, *the ratios of* *these shares will change. Consider the stationary policy where tax rates are set as a function of the ratio of shares:* $\tau_s = f\left(\frac{v_s^i}{1/n}\right)$*. If political system could commit to an optimal choice of a tax sequence, the trajectory of growth rates and shares would be determined, and the solution can then be expressed as a stationary tax policy that maps the equilibrium shares into the optimal tax rates independently of time, because agents making consumption decisions take shares as given.*

Proof: Consider rewriting the above maximization as a Lagrangian:

$$\mathcal{L}\left(k_{t_0}, k_{t_0}^M, k_{t_0}^w, t_{0\infty}\right) =$$

$$Max_{\{\tau\}_{t_0}^{\infty}}\ (1-\beta)^{-1}\log\lambda\left(\left(1-\tau_{t_0}\right)v_{t_0}^M + n^{-1}\left(\tau_{t_0} + \sum_{j=t_0+1}^{\infty}\tau_j\prod_{s=t_0+1}^{j} g_s(r\left(1-\tau_s\right))^{-1}\right)\right) rk_{t_0}$$

$$+\,\beta\left(1-\beta\right)^{-2}\log\beta r + \sum_{n=t_0+1}^{\infty}\beta^{n-t_0}\log\prod_{s=t_0+1}^{n}\left(1-\tau_s\right)$$

$$+\sum_{t=t_0}^{\infty}\phi_t\left(\begin{array}{c}(1-\beta)^{-1}\log\lambda\left(\left(1-\tau_{t_0}\right)v_{t_0}^w + n^{-1}\left(\tau_{t_0} + \sum_{j=t_0+1}^{\infty}\tau_j\prod_{s=t_0+1}^{j} g_s(r\left(1-\tau_s\right))^{-1}\right)\right) rk_t \\ +\beta\left(1-\beta\right)^{-2}\log\beta r + \sum_{n=t_0+1}^{\infty}\beta^{n-t_0}\log\prod_{s=t_0+1}^{n}\left(1-\tau_s\right) \\ -\mu^w\left(\left(1-\beta\right)^{-1}\log\left(\lambda v_{t_0}^w rk_{t_0}\right) + \beta\left(1-\beta\right)^{-2}\log\beta r\right)\end{array}\right)$$

$$+\,\theta_{1t}\tau_t + \theta_{2t}\left(1-\tau_t\right)$$

where ϕ_t are Lagrange multipliers on the incentive compatibility constraints and θ_{it}, $i = 1, 2$ are the multipliers for constraints $\tau_t \geq 0$ and $\tau_t \leq 1$. Note that for the above problem, setting the tax rate to zero forever is feasible, and therefore the feasible set of tax sequences is not empty.

The Lagrange multipliers $\{\phi_s\}_{s=t_0+1}^{\infty}$ must be appropriately chosen. The incentive constraints map the space of feasible tax sequences into the difference of the two value functions, a space that is not necessarily bounded since the difference depends on $\ell n\ k_t$. Lagrange multipliers must be in the dual of that space. First, we note that at each time t the coefficient of $\ell n\ k_t$ in the continuation value of the right pivotal agent in the constraint,

$$(1-\beta)^{-1}\log\lambda\left(\left(1-\tau_{t_0}\right)v_{t_0}^w + n^{-1}\left(\tau_{t_0} + \sum_{j=t_0+1}^{\infty}\tau_j\prod_{s=t_0+1}^{j} g_s(r\left(1-\tau_s\right))^{-1}\right)\right) r$$

$$> (1-\beta)^{-1}\log\lambda\left(v_{t_0}^w + n^{-1}\left(1+x_{t+1}\right)\right) r$$

remains is bounded: Since $g_{t+1} = r\left(1-\lambda_t-\lambda_t x_{t+1}\right)$ yields $x_{t+1} = \lambda^{-1}(1-\frac{g_{t+1}}{r}) - 1$, $g_{t+1} \leq r$, and $0 < \lambda$, by Assumptions 1 and 2, $x_{t+1}$is bounded. Therefore Lagrange multipliers $\{\phi_s\}_{s=t_0+1}^{\infty}$ must be chosen so that $\phi_s \ell n\ k_s = \phi_s \ell n\ k_{t_0} \prod_{t_0+1}^{s} g_n$ goes to zero appropriately. Note that from Theorem 1 and its Corollary we have $0 < \delta \leq g_s \leq r$ for $s > t_0$, so we need to set Lagrange multipliers $\{\phi_s\}_{s=t_0+1}^{\infty}$ to shrink at a rate faster than t^{-1}, unless constraints are slack $\left(\phi_{t_0} = 0\right)$ in which case Lagrange multipliers $\{\phi_s\}_{s=t_0+1}^{\infty}$ can all be set to zero. The latter case occurs if μ^w is sufficiently small and the median voter can assure that the right wing pivotal agent will not initiate an authoritarian regime even if $\tau_{t_0} = 1$. In that case $\theta_{2t_0} > 0$, and from (18) below, it can be set so that $\phi_{t_0} = 0$. This means that assets are fully equalized across agents in the first period, incentive constraints no longer bind, and tax rates are set to zero for all future dates. Note that we also have to resort to Assumption 2 to prevent all capital from being consumed at once.

Let $\hat{\tau}_{t_0}$ be the first period optimal tax. The first-order conditions with respect to τ_{t_0} are

$$\frac{d\mathcal{L}\left(k_t, k_t^M, t\right)}{d\tau_{t_0}} = \left(\left(-v_{t_0}^M + n^{-1}\right) rk_{t_0}\right) X_{t_0}^{-1} Z + \phi_{t_0} \left(X_{t_0}^w\right)^{-1} Z \left(-v_{t_0}^w + n^{-1}\right) rk_{t_0} + \theta_{1t_0} - \theta_{2t_0} = 0 \tag{18}$$

and

$$\phi_{t_0} = \frac{\left(\left(v_{t_0}^M - n^{-1}\right) k_{t_0}\right) X_{t_0}^{-1} Z - \theta_{1t} + \theta_{2t}}{\left(X_{t_0}^w\right)^{-1} Z \left(-v_{t_0}^w + n^{-1}\right) k_{t_0}}$$

where

$$Z = \lambda^{-1} = (1-\beta)^{-1},\ X_{t_0} = \left(\left(1-\tau_{t_0}\right) v_{t_0}^M + n^{-1} \left(\tau_{t_0} + \sum_{j=t_0+1}^{\infty} \tau_j \prod_{s=t_0+1}^{j} g_s (r\left(1-\tau_s\right))^{-1}\right)\right) rk_{t_0}$$

$$X_{t_0}^w = \left(\left(1-\tau_{t_0}\right) v_{t_0}^w + n^{-1} \left(\tau_{t_0} + \sum_{j=t_0+1}^{\infty} \tau_j \prod_{s=t_0+1}^{j} g_s (r\left(1-\tau_s\right))^{-1}\right)\right) rk_{t_0}$$

Note that if $\mu^w \neq 1$ and $\tau_{t_0} = 0$, the constraint on the utility of the wealthy pivotal agent would be slack, so we would have $\phi_{t_0} = 0$. As is clear from the above, in this case the poor median voter would redistribute so that $\theta_{2t} > 0$, which implies that if the median voter is poor, $\tau_{t_0} > 0$.

The first order conditions with respect to τ_m are

$$\frac{\partial \mathcal{L}\left(k_{t_0}, k_{t_0}^M, t_0\right)}{\partial \tau_m} = \frac{\partial V\left(k_{t_0}, k_{t_0}^i, t_0\right)}{\partial \tau_m} + \sum_{t=t_0}^{m} \phi_t \frac{\partial V^{wD}\left(k_t, k_t^w, t\right)}{\partial \tau_m} + \theta_{1m} - \theta_{2m} = 0$$

where $k_{t_0}^i$ is the median voter's initial wealth. In previous sections we had, under CRRA preferences,

$$\frac{dV\left(k_{t_0}, k_{t_0}^i, t_0\right)}{d\tau_m} = \left(Z_{t_0}^{\sigma} X_{t_0}^{1-\sigma}\right) \left(\left(X_{t_0}\right)^{-1} \frac{dX_{t_0}}{d\tau_m} + \sigma\left(1-\sigma\right)_t^{-1} \left(Z_{t_0}\right)^{-1} \frac{dZ_{t_0}}{d\tau_m}\right)$$

which becomes, for log preferences $(\sigma = 1)$,

$$\begin{aligned}\frac{dV\left(k_{t_0}, k_{t_0}^i, t_0\right)}{d\tau_m} &= Z\left((X_{t_0})^{-1}\frac{dX_{t_0}}{d\tau_m} - (1-\beta)(1-\tau_m)^{-1}\sum_{n=m}^{\infty}\beta^{n-t_0}\right)\\ &= Z\left((X_{t_0})^{-1}\frac{dX_{t_0}}{d\tau_m} - (1-\beta)(1-\tau_m)^{-1}\beta^{m-t_0}(1-\beta)^{-1}\right)\\ &= Z\left((X_{t_0})^{-1}\frac{dX_{t_0}}{d\tau_m} - (1-\tau_m)^{-1}\beta^{m-t_0}\right)\end{aligned}$$

This is as in section 3.1, where we let $\sigma \to 1$ to evaluate the derivative,

$$\begin{aligned}\sigma(1-\sigma)^{-1} Z_t^{-1}\frac{dZ_t}{d\tau_m} &= (1-\beta)^{-1}\left(-(1-\tau_m))^{-1}\right)\beta^{m-t}(1-\beta)\\ &= -\beta^{m-t}(1-\tau_m)^{-1} < 0\end{aligned}$$

Similarly, for $m > t$,

$$\begin{aligned}&\frac{dV^{wD}\left(k_t, k_t^w, t\right)}{d\tau_m}\\ &= \left(Z_t^{\sigma}(X_t^w)^{1-\sigma}\right)\left((X_t^w)^{-1}\frac{dX_t^w}{d\tau_m} + \sigma(1-\sigma)_{t_0}^{-1}(Z_{t_0})^{-1}\frac{dZ_{t_0}}{d\tau_m}\right)\\ &= Z\left((X_t^w)^{-1}\frac{dX_t^w}{d\tau_m} - (1-\tau_m)^{-1}\beta^{m-t_0}\right)\end{aligned}$$

where

$$X_t^{\omega} = \left((1-\tau_t)\, v_t^w + n^{-1}\left(\tau_t + \sum_{j=t+1}^{\infty}\tau_j\prod_{s=t+1}^{j} g_s(r(1-\tau_s))^{-1}\right)\right) rk_t$$

For $m = t$,using log preferences directly from above,

$$\begin{aligned}&\frac{dV^{wD}\left(k_t, k_t^w, t\right)}{d\tau_t}\\ &= \left(Z\,(X_t^w)^{-1}\right)\left(\left(-v_t^w + n^{-1}\right) rk_t\right)\end{aligned}$$

Thus

$$\begin{aligned}\frac{d\mathcal{L}\left(k_t, k_t^i, t\right)}{d\tau_m} &= Z\left((X_{t_0})^{-1}\frac{dX_{t_0}}{d\tau_m} - (1-\tau_m)^{-1}\beta^{m-t_0}\right)\\ &\quad + \left(Z\,(X_m^w)^{-1}\right)\left(\left(-v_m^w + n^{-1}\right) rk_m\right)\\ &\quad + \sum_{t=t_0+1}^{m-1}\phi_t Z\left((X_m^w)^{-1}\frac{dX_m^w}{d\tau_m} - (1-\tau_m)^{-1}\beta^{m-t}\right)\\ &\quad + \phi_{t_0} Z\left(\left(X_{t_0}^w\right)^{-1}\frac{dX_{t_0}^w}{d\tau_m} - (1-\tau_m)^{-1}\beta^{m-t_0}\right)\\ &+\theta_{1m} - \theta_{2m} = 0\end{aligned}$$

If we let $\sigma \to 1$, from the Corollary to Theorem 2,

$$\left(Z_s^{\sigma}\left(X_s^w\right)^{1-\sigma}\right)\left(\left(X_s^w\right)^{-1}\frac{\partial X_s^w}{\partial \tau_m}+\sigma\left(1-\sigma\right)^{-1}\left(Z_s\right)^{-1}\frac{\partial Z_s}{\partial \tau_m}\right)$$
$$=Z\left(\left(X_s^w\right)^{-1}\frac{\partial X_s^w}{\partial \tau_m}-\left(1-\tau_m\right)^{-1}\beta^{m-t_0}\right)<0$$

where $Z = 1/1-\beta$ and from the proof of Theorem 2, for $\sigma \to 1$,

$$\lim_{m\to\infty}\left[\left(Z_{t_0}^{\sigma}X_{t_0}^{1-\sigma}\right)\left(\left(X_{t_0}\right)^{-1}\frac{\partial X_{t_0}}{\partial \tau_m}+\sigma\left(1-\sigma\right)^{-1}\left(Z_{t_0}\right)^{-1}\frac{\partial Z_{t_0}}{\partial \tau_m}\right)\right]$$
$$=Z\left(\left(X_{t_0}\right)^{-1}\frac{\partial X_{t_0}}{\partial \tau_m}-\left(1-\tau_m\right)^{-1}\beta^{m-t_0}\right)<0$$

These imply that part of the expression for $\frac{\partial \mathcal{L}\left(k_{t_0},k_{t_0}^i,t_0\right)}{\partial \tau_m}$, apart from the term $\theta_{1m}-\theta_{2m}$ becomes negative. So, $\lim_{m\to\infty}\theta_{1m}-\theta_{2m}>0$ which implies that $\lim_{m\to\infty}\tau_m=0$. This proves part (a) of the Theorem.

If $\tau_{t_0}=1$, since from the Corollary to Theorem 2

$$Z\left(\left(X_m^w\right)^{-1}\frac{dX_m^w}{d\tau_m}-\left(1-\tau_m\right)^{-1}\beta^{m-t}\right)<0,$$

we can use the same inductive argument as part (b) of Theorem 2 to establish that $Z\left(\left(X_t\right)^{-1}\frac{\partial X_t}{\partial \tau_m}-\left(1-\tau_m\right)^{-1}\beta^{m-t_0}\right)<0$. It follows so that $\{1,0,0...\}$ is an optimal tax sequence, which proves part (b).

We now analyze the constraints on τ_t for various cases.[8] The constraint implies, for $t \geq t_0$:

$$\left(1-\beta\right)^{-1}\log\left(\lambda\left(\left(1-\tau_t\right)v_t^w+n^{-1}\left(\tau_t+x_{t+1}\right)\right)rk_t\right)$$
$$+\beta\left(1-\beta\right)^{-2}\log\beta r+\sum_{n=t+1}^{\infty}\beta^{n-t}\log\prod_{s=t+1}^{n}\left(1-\tau_s\right)$$
$$\geq\mu^w\left(\left(1-\beta\right)^{-1}\log\left(\lambda v_t^w rk_t\right)+\beta\left(1-\beta\right)^{-2}\log\beta r\right)$$

$$\log\left(\frac{\lambda rk_t\left(\left(1-\tau_t\right)v_t^w+n^{-1}\left(\tau_t+x_{t+1}\right)\right)}{\left(\lambda v_t^w rk_t\right)^{\mu^w}}\right)$$
$$\geq-\left(1-\mu^w\right)\beta\left(1-\beta\right)^{-1}\log\beta r-\left(1-\beta\right)\sum_{n=t+1}^{\infty}\beta^{n-t}\log\prod_{s=t+1}^{n}\left(1-\tau_s\right)$$

[8] Note that we do that for the log period-utility function. The log forms in the constraint are obtained after simplifying the value function using laws of the logarithm.

Taking exponential of both sides

$$\frac{(\lambda r k_t)^{1-\mu^w}\left((1-\tau_t)\,v_t^w + n^{-1}(\tau_t + x_{t+1})\right)}{(v_t^w)^{\mu^w}}$$
$$\geq (\beta r)^{-(1-\mu^w)\beta(1-\beta)^{-1}}\left(\prod_{n=t+1}^{\infty}\left(\prod_{s=t+1}^{n}(1-\tau_s)\right)^{(1-\beta)\beta^{n-t}}\right)^{-1}$$

and solving for τ_t

$$\tau_t \leq \frac{v_t^w + n^{-1}x_{t+1}}{(v_t^w - n^{-1})} - \frac{(v_t^w)^{\mu^w}(\beta r)^{-(1-\mu^w)\beta(1-\beta)^{-1}}K_t}{(v_t^w - n^{-1})(\lambda r k_t)^{1-\mu^w}}$$

where

$$1 \leq K_t = \left(\prod_{n=t+1}^{\infty}\left(\prod_{s=t+1}^{n}(1-\tau_s)\right)^{(1-\beta)\beta^{n-t}}\right)^{-1}$$

we obtain

$$\tau_t \leq \frac{v_t^w - n^{-1}\left(\frac{v_t^w}{n^{-1}}\left(\frac{k^*}{k_t}\right)^{1-\mu^w}K_t - x_{t+1}\right)}{v_t^w - n^{-1}} = \tilde{\tau}_t^R \tag{19}$$

Note that, since $x_{t+1} \geq 0$, as $k_{t_0} \to \infty$, so that $\mu^w < 1$ (or if $k_{t_0} \to 0$ when $\mu^w > 1$ for low k_{t_0}), the right side becomes larger than unity. Hence τ_{t_0} can be set to 1 . The right side will also exceed 1 if $\frac{v_t^w}{n^{-1}}\left(\frac{k^*}{k_t}\right)^{1-\mu^w}K_t - x_{t+1} < 1$. Since $\frac{\partial V(k_{t_0},k_{t_0}^i,t_0)}{\partial \tau_{t_0}} > 0$ as shown before, if we can show that the condition $\frac{v_t^w}{n^{-1}}\left(\frac{k^*}{k_t}\right)^{1-\mu^w} \leq 1$ implies that $\tau_{t_0} = 1$ is feasible, it will be implemented. Set $\tau_{t_0} = 1$. From part (b) above if $\tau_{t_0} = 1$, then $\{1, 0, 0...\}$ is necessarily the optimal tax sequence, in which case $K_t = 1$ and $x_{t+1} = 0$, so that $\tilde{\tau}_t^R = \frac{v_t^w - n^{-1}\left(\frac{v_t^w}{n^{-1}}\left(\frac{k^*}{k_t}\right)^{1-\mu^w}\right)}{v_t^w - n^{-1}} \geq 1$ and $\tau_{t_0} = 1$ is indeed feasible. This proves sufficiency for part (c). To prove necessity note that from the proof of Theorem 5 below, we have $K_t \geq 1 + x_{t+1}$. If $\frac{v_t^w}{n^{-1}}\left(\frac{k^*}{k_t}\right)^{1-\mu^w} > 1$ we have

$$\left(\frac{v_t^w}{n^{-1}}\left(\frac{k^*}{k_t}\right)^{1-\mu^w}K_t - x_{t+1}\right) > \left(\frac{v_t^w}{n^{-1}}\left(\frac{k^*}{k_t}\right)^{1-\mu^w}(x_{t+1}+1) - x_{t+1}\right)$$
$$= \left(\left(\frac{v_t^w}{n^{-1}}\left(\frac{k^*}{k_t}\right)^{1-\mu^w} - 1\right)x_{t+1} + \frac{v_t^w}{n^{-1}}\left(\frac{k^*}{k_t}\right)^{1-\mu^w}\right) > 1$$

and therefore $\tau_{t_0} \leq \tilde{\tau}_t^R < 1$. This proves part (c). QED.

In analogy to Corollary 1 of Theorem 2, we can establish a stronger result if we replace the exogenous constraints $\tilde{\tau}_s$ with the endogenous political incentive constraints of the rich pivotal agent $\tilde{\tau}_s^R$ that must of course hold period by period.

Corollary 3 *There exists an $m' < \infty$ such that $\{\hat{\tau}_{t_0}, \hat{\tau}_{t_0+1}, ..\hat{\tau}_{m'}, 0, 0, ...\}$ is the optimal tax sequence chosen by the median voter. Furthermore m', which can be taken as a measure of the degree of redistribution, is non-increasing in v_t^i.*

Proof: We modify Corollary 1 of Theorem 2 to account for the endogenous constraints, and first provide a constructive method for the solution. The median voter will choose $\hat{\tau}_{t_0}, \hat{\tau}_{t_0+1}, ..$ such that the rich pivotal agent's constraint is binding up to period before m', and after m' he will switch to zero taxes in order not to discourage savings in the early periods. If the constraint (17) at $t = t_0$ is satisfied for $\tau_{t_0} = 1$, we have the solution with full redistribution in the initial period, and future taxes are zero. Otherwise we start by setting $\{\hat{\tau}_{t_0}, 0, 0, ...\}$ and solving for the $\hat{\tau}_{t_0}$ from the constraint (17) at $t = t_0$ satisfied with equality. If $\hat{\tau}_{t_0} < 1$, the constraint is binding and from Theorem 4, $\{\hat{\tau}_{t_0}, 0, 0, ...\}$ is not optimal for the median voter. We then reset the tax sequence as $\{\hat{\tau}_{t_0}, \hat{\tau}_{t_0+1}, 0, 0, ...\}$ and solve for $\hat{\tau}_{t_0}$ in terms of $\hat{\tau}_{t_0+1}$ from the constraint (17) at $t = t_0$, satisfied with equality, and for τ_{t_0+1} such that $Q_{t_0,t_0+1}(1 - \tau_{t_0+1}) = 1$ (See Corollary 1 of Theorem 2). If the constraint (17) at $t = t_0+1$ is satisfied we have the solution, since the best solution for the rich pivotal agent is zero future taxes. If the constraint is not satisfied at $t = t_0 + 1$, we set $\{\hat{\tau}_{t_0}, \hat{\tau}_{t_0+1}, \hat{\tau}_{t_0+2}, 0, 0, ..\}$ and solve for $(\hat{\tau}_{t_0}, \hat{\tau}_{t_0+1}, \hat{\tau}_{t_0+2})$ from the constraints (17) at $t = t_0$, at $t = t_0 + 1$, and from $Q_{t_0,t_0+2}((1 - \tau_{t_0+1}))(1 - \tau_{t_0+2}) = 1$, with the proviso that if the solution to τ_{t_0+2} is negative, we set it to zero. Note that from the constraints (17) at $t = t_0$ and at $t = t_0 + 1$ we can solve recursively, first for $\hat{\tau}_{t_0+1}$ in terms of $\hat{\tau}_{t_0+2}$, then for $\hat{\tau}_{t_0}$ in terms of $\hat{\tau}_{t_0+1}$ and $\hat{\tau}_{t_0+2}$, and then for $\hat{\tau}_{t_0+2}$ that satisfies $Q_{t_0,t_0+2}((1 - \tau_{t_0+1}))(1 - \tau_{t_0+2}) = 1$. If there are multiple solutions, we pick the one that yields the highest utility to the median voter. If the constraint (17) at $t = t_0 + 2$ is satisfied, we have the solution. Otherwise we continue in this manner until $t_0 + m'$ where the constraint is satisfied and we have a solution, or until $Q_{t_0,t_0+m'}\left(\prod_{s=t_0+1}^{t_0+m'}(1 - \hat{\tau}_s)\right) \leq 1$, so that we can set $\tau_{t_0+m} = 0$ for $m \geq m'$ so that therefore the constraint (17) will be satisfied for $t_0 + m'$ and beyond.

We can show that $m' < \infty$. By construction, the constraint of the rich pivotal agent is binding for $m < m'$ so that $\hat{\tau}_{t_0+m} = \tilde{\tau}^R_{t_0+m}$ for $m < m'$. Since from Theorem 4, we know that $\lim_{m\to\infty} \tau_{t_0+m} = 0$, which implies $K_{t_0+m} \to 1$, $x_{t_0+m} \to 0$ and from (19) $\tilde{\tau}^R_{t_0+m} = \frac{v^w_{t_0+m} - n^{-1}\left(\frac{v^w_{t_0+m}}{n^{-1}}\left(\frac{k^*}{k_{t_0+m}}\right)^{1-\mu^w} K_{t_0+m} - x_{t_0+m+1}\right)}{v^w_{t_0+m} - n^{-1}} \to \frac{v^w_{t_0+m}\left(1 - \left(\frac{k^*}{k_{t_0+m}}\right)^{1-\mu^w}\right)}{v^w_{t_0+m} - n^{-1}} > 0$, unless $\frac{k^*}{k_{t_0+m}} \to 1$. Thus $\lim_{m\to\infty} \tilde{\tau}^R_{t_0+m} > 0$, so we would have $\lim_{m\to\infty} \tau_{t_0+m} > 0$. This yields a contradiction because then eventually $Q_{t_0,t_0+m}\left(\prod_{s=t_0+1}^{t_0+m}(1 - \hat{\tau}_s)\right) \le 1$ for large m. If $\frac{k^*}{k_{t_0+m}} \to 1$ lifetime utility is zero and since μ^w is multiplicative, there is no cost of dictatorship at $\frac{k^*}{k_{t_0+m}}$, and therefore taxes must be set to zero to sustain democracy. However, if $\lim_{m\to\infty} \tau_{t_0+m} = 0$, then $\lim_{m\to\infty} g_{t_0+m} = r(1-\lambda) > 1$, so $\frac{k^*}{k_{t_0+m}} \to 1$ is impossible.

As in Corollary 1 of Theorem 2, m' is non-increasing in $v^i_{t_0}$ since Q_{t_0,t_0+s} is increasing in $v^i_{t_0}$ for any s.

QED

The tax sequence in the above Theorem and Corollary is not time consistent, since the median voter will want to reset taxes at and beyond time $\tilde{m}$ once the incentive constraint of the rich pivotal agent is not binding. In the absence of legal commitment devices or reputational considerations associated with violation of campaign promises in terms of low future saving rates or electoral losses, the median voter will always set taxes so that the constraints of the rich pivotal agent is binding. Of course voters will always prefer to vote for someone who does have reputational and re-election concerns and faces costs associated with violations of campaign promises, since discounted utility will be higher if initially announced taxes are not reset. (See Theorem 7) But if such costs and concerns are absent, the only time consistent equilibrium is the one where the rich pivotal agent's constraint binds perpetually. If in addition, however, there are exogenous constraints on taxes such that $\tau_t \le \tilde{\tau}_t$, we have:

$$\tau_t = Min\left(\frac{v^w_t - n^{-1}\left(\frac{v^w_t}{n^{-1}}\left(\frac{k^*}{k_t}\right)^{1-\mu^w} K_t - x_{t+1}\right)}{v^w_t - n^{-1}}, \tilde{\tau}\right) \qquad (20)$$

Corollary 4 *The tax sequence along which 19 holds for all t is time consistent. Unless $g = r(1-\lambda)(1-\tilde{\tau}) > 1$, the economy will not grow.*

Proof: If k_t grew, then since $v^w_s \ge n^{-1}$, $\left(\frac{k^*}{k_t}\right)^{1-\mu^w} \to 0$ and

$\frac{v_t^w - n^{-1}\left(\frac{v_t^w}{n^{-1}}\left(\frac{k^*}{k_t}\right)^{1-\mu^w} K_t - x_{t+1}\right)}{v_t^w - n^{-1}}$ becomes greater than unity. Thus the incentive constraints of the rich pivotal agent do not bind, and eventually $\tau_s = \tilde{\tau}_s$. QED

Thus the consequences of time consistency in the absence of reputational considerations are quite drastic since it is a stretch to expect that the ad-hoc exogenous constraints on taxes will be low enough not to extinguish growth.

3.3 Coalitions of Rich and Poor

If the median voter is to be decisive at each time with regard to the entire path of future taxes, then at no time can a coalition of the poor and the rich make at least one better off and the other no worse off relative to the proposal of the median. Hence, we need to check whether the poor and rich pivotal agents can improve their utility by forming a stable coalition against the median voter under democracy. Here a stable coalition will require that the incentive constraint for the two pivotal agents hold period by period, and that the maximized discounted utility of the poor agent subject to the constraint that the rich agent's utility is at least as large as what he gets under the tax sequence chosen by the median voter under democracy exceeds the discounted utility the poor agent receives under tax sequence chosen by the median voter under democracy.

If we maximize the discounted utility of the poor agent subject to incentive constraints of the rich and with the additional constraint that the poor agent is no worse off than he was under the tax sequence implemented by the median voter, the poor agent will be forced to implement the same tax sequence as the median voter. The reason is that the poor pivotal agent is even more anxious to redistribute early on. However if he were to do so he would make the rich pivotal agent worse off, so the best he can do is to duplicate the median voter's tax scheme. Formally,

Theorem 4 *There is no tax sequence that satisfies all incentive constraints and that can make the poor pivotal agent better off than he is under the tax scheme chosen by the median voter, while making the rich pivotal agent no worse off.*

Proof: First we note that the value functions of the poor pivotal agent, the median voter and the rich pivotal agent are given by

$$
\begin{aligned}
V\left(k_{t_0}, k_{t_0}^i, t_0\right) &= \sum_{t=t_0}^{\infty} \beta^{t-t_0} \log c_t \\
V\left(k_{t_0}, k_{t_0}^i, t_0\right) &= \log c_{t_0}^i + \beta \log c_{t_0}^i \beta r \left(1-\tau_{t_0+1}\right) + \beta^2 \log c_{t_0}^i \left(\beta r\right)^2 \left(1-\tau_{t_0+1}\right)\left(1-\tau_{t_0+2}\right) + ... \\
V\left(k_{t_0}, k_{t_0}^i, t_0\right) &= (1-\beta)^{-1} \log c_{t_0}^i + \beta (1-\beta)^{-2} \log \beta r + \sum_{n=t_0+1}^{\infty} \beta^{n-t_0} \log \prod_{s=t_0+1}^{n} (1-\tau_s) \\
V\left(k_{t_0}, k_{t_0}^i, t_0\right) &= (1-\beta)^{-1} \log \left(\left(1-\tau_{t_0}\right) v_{t_0}^i + n^{-1}\tau_{t_0} + n^{-1} \sum_{j=t_0+1}^{\infty} \tau_j \prod_{s=t_0+1}^{j} g_s (r\left(1-\tau_s\right))^{-1} \right) \\
&\quad + (1-\beta)^{-1} \log \lambda r k_{t_0} + \beta (1-\beta)^{-2} \log \beta r + \sum_{n=t_0+1}^{\infty} \beta^{n-t_0} \log \prod_{s=t_0+1}^{n} (1-\tau_s) \\
&= \log \left(\begin{array}{c} \left(\left(1-\tau_{t_0}\right) v_{t_0}^i + n^{-1}\tau_{t_0} + n^{-1} \sum_{j=t_0+1}^{\infty} \tau_j \prod_{s=t_0+1}^{j} g_s (r\left(1-\tau_s\right))^{-1} \right)^{(1-\beta)^{-1}} \\ \cdot \prod_{n=t_0+1}^{\infty} \left(\prod_{s=t_0+1}^{n} (1-\tau_s) \right)^{\beta^{n-t_0}} \end{array} \right) \\
&\quad + (1-\beta)^{-1} \log \lambda r k_{t_0} + \beta (1-\beta)^{-2} \log \beta r
\end{aligned}
$$

where $i = p, M, w$ for the poor, median and rich agents. First we note that the value functions of the poor agent, the median voter and the rich agent are identical except with respect to the terms $\left(1-\tau_{t_0}\right) v_t^i$.Changing taxes for periods after t_0 affects consumptions of these agents in the same way, by changing first period consumption as well as its rate of growth. Changing τ_{t_0} however affects first period consumptions differentially. Since we have decreasing marginal utility, changes in consumption will have larger effects on the utility of the poor agent and smaller effect on the utility of the rich agent.

Assume that the poor agent implements a tax sequence $\left\{\tau_{t_0}^p, (\tau_s^p)_{s=t_0+1}^{\infty}\right\}$ that improves his payoff relative to the taxes chosen by the median voter$\left\{\tau_{t_0}^M, (\tau_s^M)_{s=t_0+1}^{\infty}\right\}$, respects all incentive constraints, and leaves the rich agent no worse off. If this change makes the median voter better off as well, we have a contradiction, since the median voter could have chosen those tax rates to start with. Suppose then that these tax rates, $\left\{\tau_{t_0}^p, (\tau_s^p)_{s=t_0}^{\infty}\right\}$, make the median voter worse off.

$$
\begin{aligned}
V\left(k_{t_0}, v_{t_0}^p, t_0; \left\{\tau_{t_0}^p, (\tau_s^p)_{s=t_0+1}^{\infty}\right\}\right) - V\left(k_{t_0}, v_{t_0}^p, t_0; \left\{\tau_{t_0}^M, (\tau_s^M)_{s=t_0+1}^{\infty}\right\}\right) &> 0 \\
V\left(k_{t_0}, v_{t_0}^M, t_0; \left\{\tau_{t_0}^p, (\tau_s^p)_{s=t_0+1}^{\infty}\right\}\right) - V\left(k_{t_0}, v_{t_0}^M, t_0; \left\{\tau_{t_0}^M, (\tau_s^M)_{s=t_0+1}^{\infty}\right\}\right) &< 0 \\
V\left(k_{t_0}, v_{t_0}^w, t_0; \left\{\tau_{t_0}^p, (\tau_s^p)_{s=t_0+1}^{\infty}\right\}\right) \geq V\left(k_{t_0}, v_{t_0}^w, t_0; \left\{\tau_{t_0}^M, (\tau_s^M)_{s=t_0+1}^{\infty}\right\}\right)
\end{aligned}
$$

Since the value functions are continuous in the shares $v_{t_0}^i$, from the Intermediate Value Theorem, there exist $v_{t_0}^{pM} \in \left(v_{t_0}^p, v_{t_0}^M\right)$, $v_{t_0}^{wM} \in \left(v_{t_0}^M, v_{t_0}^w\right)$

such that

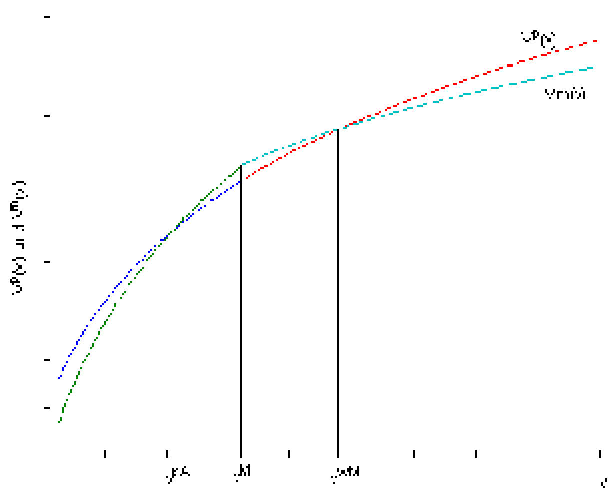
$$V\left(k_{t_0}, v_{t_0}^{pM}, t_0; \left\{\tau_{t_0}^{p}, (\tau_s^{p})_{s=t_0+1}^{\infty}\right\}\right) = V\left(k_{t_0}, v_{t_0}^{pM}, t_0; \left\{\tau_{t_0}^{M}, (\tau_s^{M})_{s=t_0+1}^{\infty}\right\}\right)$$

$$V\left(k_{t_0}, v_{t_0}^{M}, t_0; \left\{\tau_{t_0}^{p}, (\tau_s^{p})_{s=t_0+1}^{\infty}\right\}\right) < V\left(k_{t_0}, v_{t_0}^{M}, t_0; \left\{\tau_{t_0}^{M}, (\tau_s^{M})_{s=t_0+1}^{\infty}\right\}\right)$$

$$V\left(k_{t_0}, v_{t_0}^{wM}, t_0; \left\{\tau_{t_0}^{p}, (\tau_s^{p})_{s=t_0+1}^{\infty}\right\}\right) = V\left(k_{t_0}, v_{t_0}^{wM}, t_0; \left\{\tau_{t_0}^{M}, (\tau_s^{M})_{s=t_0+1}^{\infty}\right\}\right)$$

We take the smallest $v_{t_0}^{pM}$ and $v_{t_0}^{wM}$ for which the above holds.

For a graphical illustration, let $V^p(v)$ be the discounted utility value of the poor pivotal agent as a function of his share under his preferred taxes $\left\{\tau_{t_0}^{p}, (\tau_s^{p})_{s=t_0+1}^{\infty}\right\}$, and $V^m(v)$ be the same under the taxes $\left\{\tau_{t_0}^{M}, (\tau_s^{M})_{s=t_0+1}^{\infty}\right\}$ preferred by the median voter. Then the conditions above imply:



Double-Crossing

$V^m(v)$ and $V^p(v)$ must intersect at least cross, as drawn, to satisfy the conditions above, but of course they may cross more than twice.

From $V\left(k_{t_0}, v_{t_0}^{pM}, t_0; \left\{\tau_{t_0}^{p}, (\tau_s^{p})_{s=t_0+1}^{\infty}\right\}\right) = V\left(k_{t_0}, v_{t_0}^{pM}, t_0; \left\{\tau_{t_0}^{M}, (\tau_s^{M})_{s=t_0+1}^{\infty}\right\}\right)$, we have

$$\left((1-\tau_{t_0})\, v_{t_0}^{pM} + n^{-1}\tau_{t_0}^{p} + n^{-1} \sum_{j=t_0+1}^{\infty} \tau_j \prod_{s=t_0+1}^{j} g_s (r\,(1-\tau_s^{p}))^{-1} \right)^{(1-\beta)^{-1}}$$
$$\times \prod_{n=t_0+1}^{\infty} \left(\prod_{s=t_0+1}^{n} (1-\tau_s^{p}) \right)^{\beta^{n-t_0}}$$
$$=$$
$$\left((1-\tau_{t_0})\, v_{t_0}^{pM} + n^{-1}\tau_{t_0}^{M} + n^{-1} \sum_{j=t_0+1}^{\infty} \tau_j^{M} \prod_{s=t_0+1}^{j} g_s (r\left(1-\tau_s^{M}\right))^{-1} \right)^{(1-\beta)^{-1}}$$
$$\times \prod_{n=t_0+1}^{\infty} \left(\prod_{s=t_0+1}^{n} (1-\tau_s^{M}) \right)^{\beta^{n-t_0}}$$

and from $V\left(k_{t_0}, v_{t_0}^{wM}, t_0; \left\{\tau_{t_0}^{p}, (\tau_s^{p})_{s=t_0+1}^{\infty}\right\}\right) = V\left(k_{t_0}, v_{t_0}^{wM}, t_0; \left\{\tau_{t_0}^{M}, (\tau_s^{M})_{s=t_0+1}^{\infty}\right\}\right)$,

$$\left((1-\tau_{t_0})\, v_{t_0}^{wM} + n^{-1}\tau_{t_0}^{p} + n^{-1} \sum_{j=t_0+1}^{\infty} \tau_j \prod_{s=t_0+1}^{j} g_s (r\,(1-\tau_s^{p}))^{-1} \right)^{(1-\beta)^{-1}}$$
$$\times \prod_{n=t_0+1}^{\infty} \left(\prod_{s=t_0+1}^{n} (1-\tau_s^{p}) \right)^{\beta^{n-t_0}}$$
$$=$$
$$\left((1-\tau_{t_0})\, v_{t_0}^{wM} + n^{-1}\tau_{t_0}^{M} + n^{-1} \sum_{j=t_0+1}^{\infty} \tau_j \prod_{s=t_0+1}^{j} g_s (r\left(1-\tau_s^{M}\right))^{-1} \right)^{(1-\beta)^{-1}}$$
$$\times \prod_{n=t_0+1}^{\infty} \left(\prod_{s=t_0+1}^{n} (1-\tau_s^{M}) \right)^{\beta^{n-t_0}}$$

We also have

$$\frac{\partial V^i}{\partial v_{t_0}^i} = \left(\left((1-\tau_{t_0})\, v_{t_0}^{i} + n^{-1}\tau_{t_0} + n^{-1} \sum_{j=t_0+1}^{\infty} \tau_j \prod_{s=t_0+1}^{j} g_s (r\,(1-\tau_s))^{-1} \right) r k_{t_0} \right)^{-1}$$
$$\times (1-\beta)^{-1} (1-\tau_{t_0})$$

Since $V\left(k_{t_0}, v_{t_0}^{p}, t_0; \left\{\tau_{t_0}^{p}, (\tau_s^{p})_{s=t_0+1}^{\infty}\right\}\right) > V\left(k_{t_0}, v_{t_0}^{p}, t_0; \left\{\tau_{t_0}^{M}, (\tau_s^{M})_{s=t_0+1}^{\infty}\right\}\right)$, if the first intersection of value functions for $v_{t_0} \geq v_{t_0}^{p}$ is at $v_{t_0}^{pM}$, we have

$$\frac{\partial V\left(k_{t_0}, v_{t_0}^{pM}, t_0; \left\{\tau_{t_0}^{p}, (\tau_s^{p})_{s=t_0+1}^{\infty}\right\}\right)}{\partial v_{t_0}} < \frac{\partial V\left(k_{t_0}, v_{t_0}^{M}, t_0; \left\{\tau_{t_0}^{p}, (\tau_s^{p})_{s=t_0+1}^{\infty}\right\}\right)}{\partial v_{t_0}}$$

and if the first intersection of value functions for $v_{t_0} \geq v_{t_0}^M$ is at $v_{t_0}^{wM}$, we have

$$\frac{\partial V\left(k_{t_0}, v_{t_0}^{wM}, t_0; \left\{\tau_{t_0}^p, (\tau_s^p)_{s=t_0+1}^{\infty}\right\}\right)}{\partial v_{t_0}} > \frac{\partial V\left(k_{t_0}, v_{t_0}^{M}, t_0; \left\{\tau_{t_0}^p, (\tau_s^p)_{s=t_0+1}^{\infty}\right\}\right)}{\partial v_{t_0}}$$

But the choice of $v_{t_0}^{pM}$ implies

$$\begin{aligned}&\left(\left(\left(1-\tau_{t_0}^M\right) r v_{t_0}^{pM} + n^{-1}\tau_{t_0}^M + n^{-1}\sum_{j=t_0+1}^{\infty}\tau_j^M\prod_{s=t_0+1}^{j} g_s(r\left(1-\tau_s^M\right))^{-1}\right) r k_{t_0}\right)^{-1}\\ &\times(1-\beta)^{-1}\left(1-\tau_{t_0}^M\right)\\ &>\left(\left(\left(1-\tau_{t_0}^p\right) r v_{t_0}^{pM} + n^{-1}\tau_{t_0}^P + n^{-1}\sum_{j=t_0+1}^{\infty}\tau_j^p\prod_{s=t_0+1}^{j} g_s(r\left(1-\tau_s^p\right))^{-1}\right) r k_{t_0}\right)^{-1}\\ &\times(1-\beta)^{-1}\left(1-\tau_{t_0}^p\right)\end{aligned}$$

or

$$\begin{aligned}&\frac{\left(\left(\left(1-\tau_{t_0}^p\right) r v_{t_0}^{pM} + n^{-1}\tau_{t_0}^p + n^{-1}\sum_{j=t_0+1}^{\infty}\tau_j^p\prod_{s=t_0+1}^{j} g_s(r\left(1-\tau_s^p\right))^{-1}\right)\right)}{\left(\left(\left(1-\tau_{t_0}^M\right) r v_{t_0}^{pM} + n^{-1}\tau_{t_0}^M + n^{-1}\sum_{j=t_0+1}^{\infty}\tau_j^M\prod_{s=t_0+1}^{j} g_s(r\left(1-\tau_s^M\right))^{-1}\right)\right)}\\ &=\left(\frac{\prod_{n=t_0+1}^{\infty}\left(\prod_{s=t_0+1}^{n}\left(1-\tau_s^M\right)\right)^{\beta^{n-t_0}}}{\prod_{n=t_0+1}^{\infty}\left(\prod_{s=t_0+1}^{n}\left(1-\tau_s^p\right)\right)^{\beta^{n-t_0}}}\right)^{1-\beta} > \frac{\left(1-\tau_{t_0}^p\right)}{\left(1-\tau_{t_0}^M\right)}\end{aligned}$$

Similarly, the choice of $v_{t_0}^{wM}$ implies

$$\begin{aligned}&\frac{\left(\left(\left(1-\tau_{t_0}^p\right) r v_{t_0}^{wM} + n^{-1}\tau_{t_0} + n^{-1}\sum_{j=t_0+1}^{\infty}\tau_j^p\prod_{s=t_0+1}^{j} g_s(r\left(1-\tau_s^p\right))^{-1}\right)\right)}{\left(\left(\left(1-\tau_{t_0}^M\right) r v_{t_0}^{wM} + n^{-1}\tau_{t_0} + n^{-1}\sum_{j=t_0+1}^{\infty}\tau_j^M\prod_{s=t_0+1}^{j} g_s(r\left(1-\tau_s^M\right))^{-1}\right)\right)}\\ &=\left(\frac{\prod_{n=t_0+1}^{\infty}\left(\prod_{s=t_0+1}^{n}\left(1-\tau_s^M\right)\right)^{\beta^{n-t_0}}}{\prod_{n=t_0+1}^{\infty}\left(\prod_{s=t_0+1}^{n}\left(1-\tau_s^p\right)\right)^{\beta^{n-t_0}}}\right)^{1-\beta} < \frac{\left(1-\tau_{t_0}^p\right)}{\left(1-\tau_{t_0}^M\right)}\end{aligned}$$

which is a contradiction. Thus, it is not possible for the poor agent to implement a tax sequence $\left\{\tau_{t_0}^p, (\tau_s^p)_{s=t_0+1}^{\infty}\right\}$ that improves his payoff relative to the taxes chosen by the median voter $\left\{\tau_{t_0}^M, (\tau_s^M)_{s=t_0+1}^{\infty}\right\}$, respects all incentive constraints, and leaves the rich agent no worse off. *QED*

While we derived the result above for log preferences, it should be possible to establish the same result using the same argument for more general preferences.

4 Sustainability of Democracy

4.1 Unsustainability

In the Theorem below we study conditions for unsustainability, taking into account the incentive constraints of both the poor and the rich pivotal agent. Note that the conditions for unsustainability are not for particular tax sequences but are given in terms of initial stocks and parameters only. The harder issue of "sustainability" is taken up in the next section.

First however we must more fully explore the strategies of the rich and poor pivotal agents. If the median voter chooses a tax sequence under which the rich pivotal agent finds it optimal to revolt, the poor pivotal agent will also want to revolt, rather than passively accept a right-wing rule with zero taxes, since both pivotal agents will bear the costs of the autocratic regime: it is better to suffer autocracy under one's prefferred tax sequence than under the tax sequence set by the other class. So we assume that if the tax sequence chosen by the median induces the right to revolt, the revolution will succeed with probability π, but the left will counter-revolt and come to power with proability $1-\pi$. Similarly, if the tax sequence chosen by the median voter induces the left to revolt, the revolution will succeed with probability $1-\pi'$, but the right will counter-revolt and come to power with proability π'. Of course it may be reasonable to assume that it makes no difference whether the right or the left initiates the revolution, in which case we can set $\pi = \pi'$.

Theorem 5 *Let* $k^{**} = \frac{(\beta r)^{-\beta(1-\beta)^{-1}}}{\lambda n^{-1} r}$ *and* $k^{*} = \frac{(\beta r)^{-\beta(1-\beta)^{-1}}}{\lambda v_{t_0}^{w} r}$. *If preferences are logarithmic, democracy is not sustainable if*

$$\left(\frac{k_{t_0}}{k^{**}}\right) < (1-\alpha)\left(\frac{k_{t_0}}{k^{**}}\right)^{\mu^p}\left(\frac{v_{t_0}^{p}}{n^{-1}}\right)^{\pi'\mu^p} + \alpha\left(\frac{k_{t_0}}{k^{**}}\right)^{\mu^w}\left(\frac{v_{t_0}^{w}}{n^{-1}}\right)^{\mu^w \pi}$$

where $\alpha = \frac{n^{-1}-v_{t_0}^{p}}{\left(v_{t_0}^{w}-v_{t_0}^{p}\right)}$. *For* $\alpha \in (0,1)$ *this condition holds if the probability of a successful revolution by the initiating group is close to* 1 $(\pi' \to 0, \pi \to 1)$, *and either* k_{t_0} *is close to* k^{**}, *or if both* μ^p *and* μ^w *are close to* 1. *It fails to hold if* α *is sufficiently close to* 1 *or* 0.

Remark 5 *The term* α *is a measure of the skewness of the income distribution: its denominator measures the range of the distribution while the numerator measures how far the poor is from the average. If* $\alpha = 0$, *the poor agent is at the mean and the tax sequence* $(0, 0, ...)$ *is acceptable* *to all, while if* $1-\alpha = \frac{v^w_{t_0} - n^{-1}}{\left(v^w_{t_0} - v^p_{t_0}\right)} = 0$*, the rich agent is at the mean and the tax sequence* $(1, 0, ...)$ *is acceptable to all.*

Proof: As seen above, in the log case we have that for the rich pivotal agent

$$\begin{array}{l} \mu^w < 1 \text{ if } (1-\beta)^{-1} \log\left(\lambda v^w_{t_0} r k_{t_0}\right) + \beta (1-\beta)^{-2} \log \beta r > 0 \\ \mu^w \geq 1 \text{ if } (1-\beta)^{-1} \log\left(\lambda v^w_{t_0} r k_{t_0}\right) + \beta (1-\beta)^{-2} \log \beta r \leq 0 \end{array}$$

that is

$$\mu^w < 1 \text{ if } 1 > \frac{(\beta r)^{-\beta(1-\beta)^{-1}}}{\lambda v^w_{t_0} r k_{t_0}} = \frac{k^*}{k_{t_0}}$$

$$\mu^w \geq 1 \text{ if } 1 \leq \frac{(\beta r)^{-\beta(1-\beta)^{-1}}}{\lambda v^w_{t_0} r k_{t_0}} = \frac{k^*}{k_{t_0}}$$

$$\text{where } k^* = \frac{(\beta r)^{-\beta(1-\beta)^{-1}}}{\lambda v^w_{t_0} r}$$

Therefore, $\left(\frac{k^*}{k_{t_0}}\right)^{1-\mu^w} \leq 1$ in both cases.

Similarly for the poor pivotal agent in the log case, we have

$$\begin{array}{l} \mu^p < 1 \text{ if } (1-\beta)^{-1} \log\left(\lambda n^{-1} r k_{t_0}\right) + \beta (1-\beta)^{-2} \log \beta r > 0 \\ \mu^p \geq 1 \text{ if } (1-\beta)^{-1} \log\left(\lambda n^{-1} r k_{t_0}\right) + \beta (1-\beta)^{-2} \log \beta r \leq 0 \end{array}$$

$$\mu^p < 1 \text{ if } 1 > \frac{(\beta r)^{-\beta(1-\beta)^{-1}}}{\lambda n^{-1} r k_{t_0}} = \frac{k^{**}}{k_{t_0}}$$

$$\mu^p \geq 1 \text{ if } 1 \leq \frac{(\beta r)^{-\beta(1-\beta)^{-1}}}{\lambda n^{-1} r k_{t_0}} = \frac{k^{**}}{k_t}$$

$$\text{where } k^{**} = \frac{(\beta r)^{-\beta(1-\beta)^{-1}}}{\lambda n^{-1} r} > k^* = \frac{(\beta r)^{-\beta(1-\beta)^{-1}}}{\lambda v^w_{t_0} r}$$

$$\text{Therefore } \left(\frac{k^{**}}{k_{t_0}}\right)^{1-\mu^p} \leq 1$$

The constraint for the poor pivotal agent is

$$\left(\begin{array}{c} (1-\beta)^{-1} \log\left(\lambda\left((1-\tau_t) v^p_t + n^{-1}(\tau_t + x_{t+1})\right) r k_t\right) \\ +\beta (1-\beta)^{-2} \log \beta r + \sum_{n=t+1}^{\infty} \beta^{n-t} \log \prod_{s=t+1}^{n} (1-\tau_s) \\ \geq (1-\pi') \left(\mu^p \left((1-\beta)^{-1} \log\left(\lambda n^{-1} r k_t\right) + \beta (1-\beta)^{-2} \log \beta r\right)\right) \\ +\pi' \left(\mu^p \left((1-\beta)^{-1} \log\left(\lambda v^p_t r k_t\right) + \beta (1-\beta)^{-2} \log \beta r\right)\right) \end{array}\right)$$

$$(1-\beta)^{-1}\log\left(\frac{\lambda r k_t\left((1-\tau_t)\,v_t^p+n^{-1}\left(\tau_t+x_{t+1}\right)\right)}{(\lambda r k_t)^{\mu^p}\left(\left(n^{-1}\right)^{1-\pi'}\left(v_t^p\right)^{\pi'}\right)^{\mu^p}}\right)$$
$$\geq -\left(1-\mu^p\right)\beta\left(1-\beta\right)^{-2}\log\beta r-\sum_{n=t+1}^{\infty}\beta^{n-t}\log\prod_{s=t+1}^{n}\left(1-\tau_s\right)$$

$$\log\left(\frac{\lambda r k_t\left((1-\tau_t)\,v_t^p+n^{-1}\left(\tau_t+x_{t+1}\right)\right)}{(\lambda r k_t)^{\mu^p}\left(\left(n^{-1}\right)^{1-\pi'}\left(v_t^p\right)^{\pi'}\right)^{\mu^p}\left(n^{-1}\right)^{1-\mu^p}}\left(n^{-1}\right)^{1-\mu^p}\right)$$
$$\geq -\left(1-\mu^p\right)\beta\left(1-\beta\right)^{-1}\log\beta r-\left(1-\beta\right)\sum_{n=t+1}^{\infty}\beta^{n-t}\log\prod_{s=t+1}^{n}\left(1-\tau_s\right)$$

or

$$\left(\frac{\lambda r k_t\left((1-\tau_t)\,v_t^p+n^{-1}\left(\tau_t+x_{t+1}\right)\right)}{(\lambda r k_t)^{\mu^p}\left(\left(n^{-1}\right)^{1-\pi'}\left(v_t^p\right)^{\pi'}\right)^{\mu^p}\left(n^{-1}\right)^{1-\mu^p}}\left(n^{-1}\right)^{1-\mu^p}\right)$$
$$\geq (\beta r)^{-(1-\mu^p)\beta(1-\beta)^{-1}}\left(\prod_{n=t+1}^{\infty}\left(\prod_{s=t+1}^{n}\left(1-\tau_s\right)\right)^{(1-\beta)\beta^{n-t}}\right)^{-1}$$

Manipulating the constraint algebraically, as we did for the rich agent's constraint in section 3.2, we get,

$$\tau_t \geq \frac{\left(\left(n^{-1}\right)^{1-\pi'}\left(v_t^p\right)^{\pi'}\right)^{\mu^p}\left(n^{-1}\right)^{1-\mu^p}K_t}{\left(n^{-1}-v_t^p\right)}\left(\frac{k^{**}}{k_t}\right)^{1-\mu^p}-\frac{v_t^p+n^{-1}x_{t+1}}{\left(n^{-1}-v_t^p\right)}$$

$$\tau_t \geq \frac{\left(\left(\frac{v_t^p}{n^{-1}}\right)^{\pi'\mu^p}n^{-1}\right)K_t}{\left(n^{-1}-v_t^p\right)}\left(\frac{k^{**}}{k_t}\right)^{1-\mu^p}-\frac{v_t^p+n^{-1}x_{t+1}}{\left(n^{-1}-v_t^p\right)} \tag{21}$$

$$\tau_t \geq \frac{\left(n^{-1}\left(K_t\left(\frac{k^{**}}{k_t}\right)^{1-\mu^p}\left(\frac{v_t^p}{n^{-1}}\right)^{\pi'\mu^p}-x_{t+1}\right)-v_t^p\right)}{\left(n^{-1}-v_t^p\right)}=\tilde{\tau}_t^L \tag{22}$$

where

$$1\leq K_t=\left(\prod_{n=t+1}^{\infty}\left(\prod_{s=t+1}^{n}\left(1-\tau_s\right)\right)^{(1-\beta)\beta^{n-t}}\right)^{-1}$$

As $k_{t_0}\to 0$ (so that $\mu^p>1$) or $k_{t_0}\to\infty$ (so that $\mu^p<1$), the right side becomes negative, so the incentive constraint of the poor is always satisfied.

The rich pivotal agent's constraint also needs modification. We get

$$\begin{aligned}
&(1-\beta)^{-1}\log\left(\lambda\left((1-\tau_t)\,v_t^w+n^{-1}\left(\tau_t+x_{t+1}\right)\right)rk_t\right)\\
&+\beta\left(1-\beta\right)^{-2}\log\beta r+\sum_{n=t+1}^{\infty}\beta^{n-t}\log\prod_{s=t+1}^{n}\left(1-\tau_s\right)\\
\geq{}&\pi\left(\mu^w\left((1-\beta)^{-1}\log\left(\lambda v_t^w rk_t\right)+\beta\left(1-\beta\right)^{-2}\log\beta r\right)\right)\\
&+\left(1-\pi\right)\left(\mu^w\left((1-\beta)^{-1}\log\left(\lambda n^{-1}rk_t\right)+\beta\left(1-\beta\right)^{-2}\log\beta r\right)\right)
\end{aligned}$$

$$\begin{aligned}
&\log\left(\frac{\lambda rk_t\left((1-\tau_t)\,v_t^w+n^{-1}\left(\tau_t+x_{t+1}\right)\right)}{\left(\left(v_t^w\right)^{\pi}\left(n^{-1}\right)^{(1-\pi)}\right)^{\mu^w}\left(\left(\lambda rk_t\right)^{\mu^w}\right)}\right)\\
\geq{}&-\left(1-\mu^w\right)\beta\left(1-\beta\right)^{-1}\log\beta r-\left(1-\beta\right)\sum_{n=t+1}^{\infty}\beta^{n-t}\log\prod_{s=t+1}^{n}\left(1-\tau_s\right)
\end{aligned}$$

Taking exponential of both sides

$$\begin{aligned}
&\frac{\left(\lambda rk_t\right)^{1-\mu^w}\left((1-\tau_t)\,v_t^w+n^{-1}\left(\tau_t+x_{t+1}\right)\right)}{\left(\left(v_t^w\right)^{\pi}\left(n^{-1}\right)^{(1-\pi)}\right)^{\mu^w}}\\
&\geq\left(\beta r\right)^{-(1-\mu^w)\beta(1-\beta)^{-1}}\left(\prod_{n=t+1}^{\infty}\left(\prod_{s=t+1}^{n}\left(1-\tau_s\right)\right)^{(1-\beta)\beta^{n-t}}\right)^{-1}
\end{aligned}$$

and solving for τ_t

$$\tau_t\leq\frac{v_t^w+n^{-1}x_{t+1}}{\left(v_t^w-n^{-1}\right)}-\frac{\left(v_t^w\right)^{\pi\mu^w}\left(n^{-1}\right)^{(1-\pi)\mu^w}\left(\beta r\right)^{-(1-\mu^w)\beta(1-\beta)^{-1}}K_t}{\left(v_t^w-n^{-1}\right)\left(\lambda rk_t\right)^{1-\mu^w}}$$

$$\tau_t\leq\frac{v_t^w+n^{-1}x_{t+1}}{\left(v_t^w-n^{-1}\right)}-\frac{\left(v_t^w\right)^{\pi\mu^w+1-\mu^w}\left(n^{-1}\right)^{(1-\pi)\mu^w}\left(\beta r\right)^{-(1-\mu^w)\beta(1-\beta)^{-1}}K_t}{\left(v_t^w-n^{-1}\right)\left(\lambda rv_t^w k_t\right)^{1-\mu^w}}$$

we obtain

$$\tau_t\leq\frac{v_t^w-n^{-1}\left(\left(\frac{v_t^w}{n^{-1}}\right)^{1-\mu^w(1-\pi)}\left(\frac{k^*}{k_t}\right)^{1-\mu^w}K_t-x_{t+1}\right)}{v_t^w-n^{-1}}=\tilde{\tau}_t^R \qquad (23)$$

Combining the poor agents's incentive constraint with the constraint (23) for the rich agent, democracy is unsustainable if $\tilde{\tau}_t^R<\tilde{\tau}_t^L$, which implies

$$\frac{v_t^w + n^{-1}x_{t+1}}{v_t^w - n^{-1}} + \frac{v_t^p + n^{-1}x_{t+1}}{(n^{-1} - v_t^p)}$$
$$< K_t \left(\frac{(n^{-1}) \left(\frac{v_t^p}{n^{-1}}\right)^{\pi' \mu^p}}{(n^{-1} - v_t^p)} \left(\frac{k^{**}}{k_t}\right)^{1-\mu^p} + \frac{n^{-1}\left(\frac{v_t^w}{n^{-1}}\right)^{1-\mu^w(1-\pi)}}{v_t^w - n^{-1}} \left(\frac{k^*}{k_t}\right)^{1-\mu^w} \right) \tag{24}$$

The above inequality implies that the feasible interval for τ_{t_0} under democracy, $\left[\tilde{\tau}_{t_0}^L, \tilde{\tau}_{t_0}^R\right] \cap [0,1]$, is empty: the highest acceptable tax to the right is lower than the lowest acceptable tax to the left. Note that this inequality cannot hold if $k_{t_0} \to 0$ (so that $\mu^w > 1$) or $k_{t_0} \to \infty$ (so that $\mu^w < 1$), since the right side goes to zero while the left is positive. From (24), we obtain

$$(1 + x_{t+1})\,\kappa_t < \frac{\left(\left(\frac{v_t^p}{n^{-1}}\right)^{\pi'\mu^p} \left(\frac{k^{**}}{k_t}\right)^{1-\mu^p} (v_t^w - n^{-1}) + \left(\frac{v_t^w}{n^{-1}}\right)^{1-\mu^w(1-\pi)} \left(\frac{k^*}{k_t}\right)^{1-\mu^w} (n^{-1} - v_t^p) \right)}{(v_t^w - v_t^p)} \tag{25}$$

where

$$\kappa_t = \left(\prod_{n=t+1}^{\infty} \left(\prod_{s=t+1}^{n} (1 - \tau_s) \right)^{(1-\beta)\beta^{n-t}} \right) = K_t^{-1}$$

If this condition holds, $\tilde{\tau}_{t_0}^R < \tilde{\tau}_{t_0}^L$ so the feasible set is empty. Setting $(1 + x_{t+1})\,\kappa_t = \frac{(1+x_{t+1})}{K_t}$ we can interpret the numerator as the one plus the benefit from transfers, and the denominator as the cost of tax distortions. If $\tau_m = 0$ for $m > t$, it becomes equal to 1. We now show that $(1 + x_{t+1})\,\kappa_t \leq 1$. We have

$$\left(\frac{(1 + x_{t+1})}{K_t}\right) < \frac{\left(\left(\frac{v_t^p}{n^{-1}}\right)^{\pi'\mu^p} \left(\frac{k^{**}}{k_t}\right)^{1-\mu^p} (v_t^w - n^{-1}) + \left(\frac{v_t^w}{n^{-1}}\right)^{1-\mu^w(1-\pi)} \left(\frac{k^*}{k_t}\right)^{1-\mu^w} (n^{-1} - v_t^p) \right)}{(v_t^w - v_t^p)} \tag{26}$$

and have to show that

$$0 < \left(\frac{(1 + x_{t+1})}{K_t}\right) = (1 + x_{t+1})\,\kappa_t < 1$$

Note that for the sequence $\tau_s = 0, s > t_0,\ x_{t_0+1} = 0,\ \kappa_{t_0} = 1$, so in this case $(1 + x_{t_0+1})\,\kappa_{t_0} = 1$. We will maximize $(1 + x_{t_0+1})\,\kappa_{t_0}$ with respect to tax sequences τ_s, $s > t_0$, looking at

$$\frac{\partial\left(\log\left(1 + x_{t_0+1}\right) + \log \kappa_{t_0}\right)}{\partial \tau_m} = (1 + x_{t_0+1})^{-1} \frac{\partial x_{t_0+1}}{\partial \tau_m} + \frac{\partial \log \kappa_{t_0}}{\partial \tau_m}$$

We know from section 3.1 that

$$\frac{\partial x_{t_0+1}}{\partial \tau_m} = [\lambda(1-a_m)+a_m](1+x_{t_0+1})^2 \beta^{m-t_0} \prod_{s=t_0+1}^{m-1} (1-\tau_s) \geq 0$$

Hence

$$(1+x_{t_0+1})^{-1} \frac{\partial x_{t_0+1}}{\partial \tau_m} = [\lambda(1-a_m)+a_m](1+x_{t_0+1}) \beta^{m-t_0} (1-\tau_m)^{-1} \prod_{s=t_0+1}^{m} (1-\tau_s)$$

$$\begin{aligned}
\frac{\partial \log \kappa_{t_0}}{\partial \tau_m} &= \frac{\partial \left[\sum_{n=m}^{\infty} \log\left(\prod_{s=t_0+1}^{n} (1-\tau_s)\right)^{(1-\beta)\beta^{n-t_0}}\right]}{\partial \tau_m} \\
&= \frac{(1-\beta)\,\partial\left[\sum_{n=m}^{\infty} \beta^{n-t_0} \sum_{t_0+1}^{n} \log(1-\tau_s)\right]}{\partial \tau_m} \\
&= -(1-\beta)\left[\sum_{n=m}^{\infty} \beta^{n-t_0} (1-\tau_m)^{-1}\right] \\
&= -(1-\beta)(1-\tau_m)^{-1} \beta^{m-t_0} [1-\beta]^{-1}
\end{aligned}$$

So

$$\frac{\partial\left(\log(1+x_{t_0+1}) + \log \kappa_{t_0}\right)}{\partial \tau_m}$$
$$= \left((1-\tau_m)^{-1} \beta^{m-t_0}\right)\left([\lambda(1-a_m)+a_m](1+x_{t_0+1}) \prod_{s=t_0+1}^{m} (1-\tau_s) - 1\right)$$

Now we note, analogously to the proof of Theorem 2, (since products of stochastic matrices are stochastic),

$$\begin{aligned}
[\lambda(1-a_m)+a_m] &= \begin{bmatrix} 1 & 0 \end{bmatrix} \prod_{s=m+1}^{\infty} \begin{bmatrix} 1-\lambda(1-\tau_s) & \lambda(1-\tau_s) \\ \tau_s & 1-\tau_s \end{bmatrix} \begin{bmatrix} \lambda \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} (1-a_m) & a_m \end{bmatrix} \begin{bmatrix} \lambda \\ 1 \end{bmatrix}
\end{aligned}$$

while

$$\begin{aligned}
(1+x_{t_0+1})^{-1} &= \begin{bmatrix} 1 & 0 \end{bmatrix} \prod_{s=t_0+1}^{\infty} \begin{bmatrix} 1-\lambda(1-\tau_s) & \lambda(1-\tau_s) \\ \tau_s & 1-\tau_s \end{bmatrix} \begin{bmatrix} \lambda \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} (1-a_m) & a_m \end{bmatrix} \prod_{s=t_0+1}^{m-1} \begin{bmatrix} 1-\lambda(1-\tau_s) & \lambda(1-\tau_s) \\ \tau_s & 1-\tau_s \end{bmatrix} \begin{bmatrix} \lambda \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} (1-a_m) & a_m \end{bmatrix} \begin{bmatrix} c_m & 1-c_m \\ f_m & 1-f_m \end{bmatrix} \begin{bmatrix} \lambda \\ 1 \end{bmatrix}
\end{aligned}$$

Thus,

$$\frac{\partial\left(\log\left(1+x_{t_0+1}\right)+\log\kappa_{t_0}\right)}{\partial\tau_m}$$

$$=\left[\lambda\left(1-a_m\right)+a_m\right]\left(1+x_{t_0+1}\right)\beta^{m-t_0}\left(1-\tau_m\right)^{-1}\prod_{s=t_0+1}^{m}\left(1-\tau_s\right)-\left(1-\tau_m\right)^{-1}\beta^{m-t_0}$$

$$=\left(\left(1-\tau_m\right)^{-1}\beta^{m-t_0}\right)\left(\frac{\left[\left(1-a_m\right)a_m\right]\begin{bmatrix}\lambda\\1\end{bmatrix}\left(\prod_{s=t_0+1}^{m}\left(1-\tau_s\right)-1\right)}{\left[\left(1-a_m\right)a_m\right]\prod_{s=t_0+1}^{m-1}\begin{bmatrix}1-\lambda\left(1-\tau_s\right)&\lambda\left(1-\tau_s\right)\\\tau_s&1-\tau_s\end{bmatrix}\begin{bmatrix}\lambda\\1\end{bmatrix}}\right)$$

Now set $m=t_0+1$, and using the convention we used in the proof of Theorem 2, that $\prod_{s=t_0+1}^{t}\begin{bmatrix}1-\lambda\left(1-\tau_s\right)&\lambda\left(1-\tau_s\right)\\\tau_s&1-\tau_s\end{bmatrix}=\begin{bmatrix}1&0\\0&1\end{bmatrix}$,

$$\frac{\partial\left(\log\left(1+x_{t_0+1}\right)+\log\kappa_{t_0}\right)}{\partial\tau_{t_0+1}}$$

$$=\left[\lambda\left(1-a_{t_0+1}\right)+a_{t_0+1}\right]\left(1+x_{t_0+1}\right)\beta\left(1-\tau_{t_0+1}\right)^{-1}\prod_{s=t_0+1}^{m}\left(1-\tau_s\right)-\left(1-\tau_{t_0+1}\right)^{-1}\beta$$

$$=\left(\left(1-\tau_{t_0+1}\right)^{-1}\beta\right)\left(\left(\frac{\left[\left(1-a_{t_0+1}\right)a_{t_0+1}\right]\begin{bmatrix}\lambda\\1\end{bmatrix}}{\left[\left(1-a_{t_0+1}\right)a_{t_0+1}\right]\begin{bmatrix}\lambda\\1\end{bmatrix}}\right)\prod_{s=t_0+1}^{t_0+1}\left(1-\tau_s\right)-1\right)<0$$

Thus to maximize $\left(1+x_{t_0+1}\right)\kappa_{t_0}$ we have to set $\tau_{t+1}=0$, irrespective of future τ_s, $s>t_0+1$. Now we proceed by induction as we did in the proof of Theorem 2. Set $\tau_s=0$ for $s=t_0+1,\ldots,m-1$. This yields:

$$\frac{\partial\left(\log\left(1+x_{t_0+1}\right)+\log\kappa_{t_0}\right)}{\partial\tau_m}$$

$$=\left[\lambda\left(1-a_m\right)+a_m\right]\left(1+x_{t_0+1}\right)\beta^{m-t_0}\left(1-\tau_m\right)^{-1}\prod_{s=t_0+1}^{m}\left(1-\tau_s\right)-\left(1-\tau_m\right)^{-1}\beta^{m-t_0}$$

$$=\left(\left(1-\tau_m\right)^{-1}\beta^{m-t_0}\right)\left(\frac{\left[\left(1-a_m\right)a_m\right]\begin{bmatrix}\lambda\\1\end{bmatrix}}{\left[\left(1-a_m\right)a_m\right]\begin{bmatrix}1-\lambda&\lambda\\0&1\end{bmatrix}^{m-t_0}\begin{bmatrix}\lambda\\1\end{bmatrix}}\left(1-\tau_m\right)-1\right)$$

Now, we note that $\begin{bmatrix} 1-\lambda & \lambda \\ 0 & 1 \end{bmatrix}^{m-t_0} = \begin{bmatrix} 1-z & z \\ 0 & 1 \end{bmatrix}$, so that

$$\frac{\partial \left(\log\left(1+x_{t_0+1}\right)+\log\kappa_{t_0}\right)}{\partial\tau_m}$$

$$=\left[\lambda\left(1-a_m\right)+a_m\right]\left(1+x_{t_0+1}\right)\beta^{m-t_0}\left(1-\tau_m\right)^{-1}\prod_{s=t_0+1}^{m}\left(1-\tau_s\right)-\left(1-\tau_m\right)^{-1}\beta^{m-t_0}$$

$$=\left(\left(1-\tau_m\right)^{-1}\beta^{m-t_0}\right)\left(\frac{\lambda\left(1-a_m\right)+a_m}{\left(\lambda+\left(1-\lambda\right)z\right)\left(1-a_m\right)+a_m}\left(1-\tau_m\right)-1\right)<0$$

So $\tau_m=0$ for $m>t_0$. This means that $\left(1+x_{t_0+1}\right)\kappa_{t_0}$ is maximized at $\tau_m=0, m>t_0$, and therefore the maximum value is indeed $\left(1+x_{t_0+1}\right)\kappa_{t_0}=1$. Now we can claim, using the condition (25), that democracy is unsustainable if

$$1<\frac{\left(\left(\frac{v_{t_0}^p}{n^{-1}}\right)^{\pi'\mu^p}\left(\frac{k^{**}}{k_{t_0}}\right)^{1-\mu^p}\left(v_{t_0}^w-n^{-1}\right)+\left(\frac{v_{t_0}^w}{n^{-1}}\right)^{1-\mu^w(1-\pi)}\left(\frac{k^*}{k_{t_0}}\right)^{1-\mu^w}\left(n^{-1}-v_{t_0}^p\right)\right)}{\left(v_{t_0}^w-v_{t_0}^p\right)} \tag{27}$$

or

$$\left(\frac{k_{t_0}}{k^{**}}\right)<(1-\alpha)\left(\frac{v_{t_0}^p}{n^{-1}}\right)^{\pi'\mu^p}\left(\frac{k_{t_0}}{k^{**}}\right)^{\mu^p}+\alpha\left(\frac{v_{t_0}^w}{n^{-1}}\right)^{\mu^w(\pi-1)}\left(\frac{k_{t_0}}{k^*}\right)^{\mu^w}$$

$$=(1-\alpha)\left(\frac{v_{t_0}^p}{n^{-1}}\right)^{\pi'\mu^p}\left(\frac{k_{t_0}}{k^{**}}\right)^{\mu^p}+\alpha\left(\frac{v_{t_0}^w}{n^{-1}}\right)^{\mu^w\pi}\left(\frac{k_{t_0}}{k^{**}}\right)^{\mu^w}$$

where $\alpha=\frac{n^{-1}-v_{t_0}^p}{\left(v_{t_0}^w-v_{t_0}^p\right)}$. This condition holds if the probability of a successful revolution by the initiating group is close to 1 $\left(\pi'\to 0,\pi\to 1\right)$, and at the same time either k_{t_0} is close to k^{**}, or μ^p and μ^w are close to 1.(Note that $k^*<k^{**}$). It fails if α is close to 1 or 0 because μ^p, μ^w, $\left(\frac{v_{t_0}^p}{n^{-1}}\right)^{\pi'\mu^p}$ are less than unity and $\left(\frac{v_{t_0}^w}{n^{-1}}\right)=1$ when $\alpha=1$. QED

4.2 Sustainability

If the feasible set satisfying both the poor and rich agents' incentive constraints is not empty, democracy is always sustainable. The Theorem below gives conditions for democracy to be sustainable for the tax sequence $\{\tau_{t_0},0,0...\}$ in terms of parameters and the initial capital stock only.

Theorem 6 *Let $k_{t_0}>k^{**}$. If*

$$\left(\frac{k^{**}}{r\beta k_{t_0}}\right)^{1-\mu^p}<\left(\frac{v_{t_0}^p}{n^{-1}}\right)^{1-\pi'\mu^p} \tag{28}$$

and

$$(1-\alpha)\left(\frac{k^{**}}{k_{t_0}}\right)^{1-\mu^p}\left(\frac{\upsilon^p_{t_0}}{n^{-1}}\right)^{\pi'\mu^p}+\alpha\left(\frac{k^{**}}{k_{t_0}}\right)^{1-\mu^w}\left(\frac{\upsilon^w_{t_0}}{n^{-1}}\right)^{\mu^w\pi}\leq 1 \quad (29)$$

where $1-\alpha=\frac{\upsilon^w_{t_0}-n^{-1}}{\left(\upsilon^w_{t_0}-\upsilon^p_{t_0}\right)}$, *then* $\tilde{\tau}^L_{t_0},<\tilde{\tau}^R_{t_0}$ *and for* $\tau_{t_0}\in\left[\tilde{\tau}^L_{t_0},\tilde{\tau}^R_{t_0}\right]\cap[0,1]$ *democracy is sustainable for the tax sequence* $\{\tau_{t_0},0,0...\}$*. The above inequalities hold for* k_{t_0} *sufficiently large. 29 will hold for* α *close to* 1 *or* 0*, or* μ^w *and* μ^p *close to zero, or for* π *close to zero, or* π' *close to* 1*.*

Remark 6 *When* $\alpha=0,\ n^{-1}=\upsilon^p_{t_0}$*, which implies that the poor pivotal agent would tolerate* $\{\tau_{t_0},0,0...\}$ *with* $\tau_{t_0}=0$*, while if* $\alpha=1,\ \upsilon^w_{t_0}=n^{-1}$ *which implies that the rich pivotal agent would tolerate* $\{\tau_{t_0},0,0...\}$ *with* $\tau_{t_0}=1$*. If the probability of success for a right wing revolution is zero* $(\pi=0)$ *the tax sequence* $\{1,0,0...\}$ *is sustainable under democracy. Similarly, if the probability of success for a left wing revolution is zero* $(\pi'=1)$*, the tax sequence* $\{0,0,0...\}$ *is sustainable under democracy.*

Proof: From the 26 we have that $\tau_{t_0}\in\left[\tilde{\tau}^L_{t_0},\tilde{\tau}^R_{t_0}\right]$ will satisfy incentive constraints if

$$\left(\frac{(1+x_{t_0+1})}{K_{t_0}}\right)\geq\left((1-\alpha)\left(\frac{\upsilon^p_t}{n^{-1}}\right)^{\pi'\mu^p}\left(\frac{k^{**}}{k_t}\right)^{1-\mu^p}+\alpha\left(\frac{\upsilon^w_t}{n^{-1}}\right)^{\mu^w\pi}\left(\frac{k^{**}}{k_t}\right)^{1-\mu^w}\right)$$

$\left(\frac{(1+x_{t_0+1})}{K_t}\right)$ attains a maximum of 1 for $\tau_{t_0+s}=0,\ s\geq 1$ and the growth rate is constant: $g=r(1-\lambda)=r\beta>1$. At time t_0, the incentive constraints of both the poor and rich pivotal agents are satisfied by hypothesis. Since taxes are zero from t_0+1 on, the incentive constraints of the rich agent will continue to be satisfied, so we turn attention to the constraints of the poor pivotal agent. Note that $k^{**}=\frac{(\beta r)^{-\beta(1-\beta)^{-1}}}{\lambda n^{-1}r}$ is constant. From (22), the constraint for the poor pivotal agent is satisfied at $t\geq t_0$ if

$$\tau_t\geq\frac{\left(n^{-1}\left(K_t\left(\frac{k^{**}}{k_t}\right)^{1-\mu^p}\left(\frac{\upsilon^p_t}{n^{-1}}\right)^{\pi'\mu^p}-x_{t+1}\right)-\upsilon^p_t\right)}{(n^{-1}-\upsilon^p_t)} \quad (30)$$

If $\tau_{t_0+s}=0,\ s\geq 1$, we have $K_{t_0+n}=1,\ x_{t_0+n}=0$ for $n\geq 0$. Then, the right side of the above is non-positive if $\left(\frac{k^{**}}{k_t}\right)^{1-\mu^p}\leq\left(\frac{\upsilon^p_t}{n^{-1}}\right)^{1-\pi'\mu^p}$.

We have to check that the incentive constraint for the poor continues to hold for $t > t_0$. To see this, note that $v^p_{t_0+1} = (1 - \tau_{t_0}) v^p_{t_0} + \tau_{t_0} n^{-1}$ and shares stays constant from $t_0 + 1$ on, while capital grows at $r\beta$. Thus at $t_0 + 1$, by hypothesis,

$$\left(\frac{k^{**}}{k_{t_0+1}}\right)^{1-\mu^p} = \left(\frac{k^{**}}{r\beta k_{t_0}}\right)^{1-\mu^p} \le \left(\frac{v^p_{t_0}}{n^{-1}}\right)^{1-\pi'\mu^p} \le \left(\frac{v^p_{t_0}}{n^{-1}} + \tau_{t_0}\left(1 - \frac{v^p_{t_0}}{n^{-1}}\right)\right)^{1-\pi'\mu^p} = \left(\frac{v^p_{t_0+1}}{n^{-1}}\right)^{1-\pi'\mu^p}$$

It follows that the right side of 30 is non-positive for $s \ge t_0 + 1$, so that setting $\tau_s = 0$ remains feasible.

For the case $\alpha = 1$ we have $v^w_{t_0} = n^{-1}$. To show that 29 holds as $\alpha \to 1$, we write $\alpha \left(\frac{v^w_t}{n^{-1}}\right)^{\mu^w \pi} \left(\frac{k^{**}}{k_t}\right)^{1-\mu^w}$ as $\alpha \left(\frac{v^w_t}{n^{-1}}\right)^{\mu^w \pi} \left(\frac{k^*}{k_t}\right)^{1-\mu^w} \left(\frac{v^w_t}{n^{-1}}\right)^{1-\mu^w} = \alpha \left(\frac{k^*}{k_t}\right)^{1-\mu^w} \left(\frac{v^w_t}{n^{-1}}\right)^{1-\mu^{w(1-\pi)}}$. Since $k^* < k^{**}$, $\left(\frac{k^*}{k_t}\right)^{1-\mu^w} < 1$, and $\left(\frac{v^w_t}{n^{-1}}\right)^{1-\mu^{w(1-\pi)}} \to 1$ as $\alpha \to 1$. Similarly as $\alpha \to 0$, $\left(\frac{v^p_{t_0}}{n^{-1}}\right) \to 1$ and $\left(\frac{k^{**}}{k_{t_0}}\right)^{1-\mu^p} \le 1$ so 29 holds.

If $\pi = 0$, 29 holds because $(1-\alpha)\left(\frac{k^{**}}{k_{t_0}}\right)^{1-\mu^p} \left(\frac{v^p_{t_0}}{n^{-1}}\right)^{\pi'\mu^p} + \alpha \left(\frac{k^{**}}{k_{t_0}}\right)^{1-\mu^w} < 1$. Finally, if $\pi' = 1$ we check that $\tilde{\tau}^L_{t_0}, < \tilde{\tau}^R_{t_0}$ from 22 and 19 for a sequence tax sequence $\{\tau_{t_0}, 0, 0...\}$ for which $K_t = 1$ and $x_{t+1} = 0$ for $t > t_0$:

$$\frac{\left(n^{-1}\left(\left(\frac{k^{**}}{k_t}\right)^{1-\mu^p} \left(\frac{v^p_t}{n^{-1}}\right)^{\pi'\mu^p}\right) - v^p_t\right)}{(n^{-1} - v^p_t)} = \tilde{\tau}^L_{t_0} < \tilde{\tau}^R_{t_0} = \frac{v^w_t - n^{-1}\left(\left(\frac{v^w_t}{n^{-1}}\right)^{1-\mu^w(1-\pi)} \left(\frac{k^*}{k_t}\right)^{1-\mu^w}\right)}{v^w_t - n^{-1}}$$

If $\pi' = 1$,

$$\frac{\left(n^{-1}\left(\left(\frac{k^{**}}{k_t}\right)^{1-\mu^p} \left(\frac{v^p_t}{n^{-1}}\right)^{\mu^p}\right) - v^p_t\right)}{(n^{-1} - v^p_t)} = \tilde{\tau}^L_{t_0} < \tilde{\tau}^R_{t_0} = \frac{v^w_t - n^{-1}\left(\left(\frac{v^w_t}{n^{-1}}\right)^{1-\mu^w} \left(\frac{k^*}{k_t}\right)^{1-\mu^w}\right)\left(\frac{v^w_t}{n^{-1}}\right)^{\mu^w \pi}}{v^w_t - n^{-1}}$$

$$\frac{n^{-1}\left(\left(\frac{k^{**}}{k_t}\left(\frac{v^p_t}{n^{-1}}\right)\right)^{1-\mu^p} \left(\frac{v^p_t}{n^{-1}}\right) - \frac{v^p_t}{n^{-1}}\right)}{(n^{-1} - v^p_t)} = \tilde{\tau}^L_{t_0} < \tilde{\tau}^R_{t_0} = \frac{v^w_t\left(1 - \left(\frac{k^{**}}{k_t}\right)^{1-\mu^w} \left(\frac{v^w_t}{n^{-1}}\right)^{\mu^w \pi - 1}\right)}{v^w_t - n^{-1}}$$

Note now that the right side is positive since both $\left(\frac{k^{**}}{k_t}\right)^{1-\mu^w}$ and $\left(\frac{v^w_t}{n^{-1}}\right)^{\mu^w \pi - 1}$ are smaller than 1, and the left hand side is smaller than 0 if $\frac{k^{**}}{k_t} < 1 < \left(\frac{v^p_t}{n^{-1}}\right)^{-1}$, as hypothesized, so that the tax sequence $\{0, 0...\}$ is implementable under democracy.

QED

Remark 7 *The theorem above shows that democracy can be sustainable for the tax sequence $\{\tau_{t_0}, 0, 0...\}$, but of course the median voter may do much better while still sustaining democracy: generally other feasible tax sequences not constant after the second period will be preferred by the median voter. The result above only shows that the feasible set of tax sequences that sustain democracy is not empty.*

Remark 8 *We should note also that the condition $\left(\frac{k^{**}}{k_{t_0}}\right)^{1-\mu^p} < \left(\frac{v^p_{t_0}}{n^{-1}}\right)^{1-\pi'\mu^p}$ is more likely to be satisfied, and therefore democracy more likely to be feasible, the closer the wealth share of the poor agent is to the average, and the wealthier the economy is at the start $\left(\frac{k^{**}}{k_{t_0}} \text{is small}\right)$.*

We may slightly generalize the above result to the case of $\{\tau_{t_0}, \tau, \tau...\}$ where the constant sequence after the initial period is is τ rather than 0. In that case shares still remain constant after $t_0 + 1$ at

$$\begin{aligned} v^i_{t_0+1} &= g^{-1} r \left[(1-\lambda)(1-\tau_{t_0}) v^M_{t_0} + \tau_{t_0}(1-\lambda) n^{-1}\right] \\ &= \left[(1-\tau)^{-1}\left((1-\tau_{t_0}) v^M_{t_0} + \tau_{t_0} n^{-1}\right)\right] \end{aligned}$$

Capital grows if $g = r(1-\lambda)(1-\tau) = (\beta r)(1-\tau) > 1$. For a constant sequence, we also have $K_{t_0} = (1-\tau)^{\frac{-\beta}{1-\beta}}$, $x_{t_0} = \tau\left(\lambda^{-1} - 1\right) = \tau\left(\frac{\beta}{1-\beta}\right)$, and $\frac{1+x_{t_0+1}}{K_{t_0}} = \frac{1+\tau\left(\frac{\beta}{1-\beta}\right)}{(1-\tau)^{\frac{-\beta}{1-\beta}}}$. Thus for $t_0 + 1$ the constraint (27) may be written as

$$\frac{\left(\left(\frac{v^p_{t_0}}{n^{-1}}\right)^{\pi'\mu^p}\left(\frac{k^{**}}{k_{t_0}}\right)^{1-\mu^p}\left(v^w_{t_0} - n^{-1}\right) + \left(\frac{v^w_{t_0}}{n^{-1}}\right)^{1-\mu^w(1-\pi)}\left(\frac{k^*}{k_{t_0}}\right)^{1-\mu^w}\left(n^{-1} - v^p_{t_0}\right)\right)}{\left(v^w_{t_0} - v^p_{t_0}\right)}$$

$$\frac{1+\tau\left(\frac{\beta}{1-\beta}\right)}{(1-\tau)^{\frac{-\beta}{1-\beta}}} \geq \frac{\left(\begin{array}{c} \left(\frac{k^{**}}{\beta r(1-\tau)k_{t_0}}\right)^{1-\mu^p}\left(\frac{(1-\tau)^{-1}\left((1-\tau_{t_0})v^p_{t_0}+n^{-1}\tau_{t_0}\right)}{n^{-1}}\right)^{\pi'\mu^p} \\ \cdot\left((1-\tau)^{-1}\left((1-\tau_{t_0})v^w_{t_0} + n^{-1}\tau_{t_0}\right) - n^{-1}\right) \\ + \left(\frac{(\beta r)^{-\beta(1-\beta)^{-1}}}{\lambda r \beta r(1-\tau)k_{t_0}}\right)^{1-\mu^w}\left(\frac{(1-\tau)^{-1}\left(v^w_{t_0}\left(1-\tau_{t_0}\right)+\tau_{t_0}n^{-1}\right)}{n^{-1}}\right)^{1-\mu^w(1-\pi)} \\ \cdot\left(n^{-1} - (1-\tau)^{-1}\left(v^p_{t_0}\left(1-\tau_{t_0}\right) + n^{-1}\tau_{t_0}\right)\right) \end{array}\right)}{\left((1-\tau)^{-1}\left((1-\tau_{t_0})v^w_{t_0} + n^{-1}\tau_{t_0}\right) - (1-\tau)^{-1}\left((1-\tau_{t_0})v^p_{t_0} + n^{-1}\tau_{t_0}\right)\right)}$$

or

$$\frac{1+\tau\left(\frac{\beta}{1-\beta}\right)}{(1-\tau)^{\frac{-\beta}{1-\beta}}} \geq \frac{\left(\begin{array}{c}\left(\frac{k^{**}}{\beta r(1-\tau)k_{t_0}}\right)^{1-\mu^p}\left(\frac{(1-\tau)^{-1}\left(\left(1-\tau_{t_0}\right)v^p_{t_0}+n^{-1}\tau_{t_0}\right)}{n^{-1}}\right)^{\pi'\mu^p} \\ \cdot\left((1-\tau)^{-1}\left(\left(1-\tau_{t_0}\right)v^w_{t_0}+n^{-1}\tau_{t_0}\right)-n^{-1}\right) \\ +\left(\frac{(\beta r)^{-\beta(1-\beta)^{-1}}}{\lambda r\beta r(1-\tau)k_{t_0}}\right)^{1-\mu^w}\left(\frac{(1-\tau)^{-1}\left(v^w_{t_0}\left(1-\tau_{t_0}\right)+\tau_{t_0}n^{-1}\right)}{n^{-1}}\right)^{1-\mu^w(1-\pi)} \\ \cdot\left(n^{-1}-(1-\tau)^{-1}\left(v^p_{t_0}\left(1-\tau_{t_0}\right)+n^{-1}\tau_{t_0}\right)\right)\end{array}\right)}{\frac{\left(1-\tau_{t_0}\right)}{(1-\tau)}n^{-1}\left(v^w_{t_0}-v^p_{t_0}\right)}$$

If $\{\tau_{t_0},\tau,\tau...\}$ satisfies the constraint at t_0+1, it will continue to satisfy it beyond t_0+1 since shares will be constant and capital will grow, so the right side of the above will decline. Thus we have the following:

Corollary 5 *If $k_{t_0}\geq k^{**}$, for a tax sequence $\{\tau_{t_0},\tau,\tau,\ldots\}$ satisfying*

$$(1-\alpha)\left(\frac{k^{**}}{k_{t_0}}\right)^{1-\mu^p}\left(\frac{v^p_{t_0}}{n^{-1}}\right)^{\pi'\mu^p}+\alpha\left(\frac{k^{**}}{k_{t_0}}\right)^{1-\mu^w}\left(\frac{v^w_{t_0}}{n^{-1}}\right)^{-\mu^w(1-\pi)}\leq 1$$

and

$$\frac{\left(\begin{array}{c}\left(\frac{k^{**}}{\beta r(1-\tau)k_{t_0}}\right)^{1-\mu^p}\left(\frac{(1-\tau)^{-1}\left(\left(1-\tau_{t_0}\right)v^p_{t_0}+n^{-1}\tau_{t_0}\right)}{n^{-1}}\right)^{\pi'\mu^p} \\ \cdot\left((1-\tau)^{-1}\left(\left(1-\tau_{t_0}\right)v^w_{t_0}+n^{-1}\tau_{t_0}\right)-n^{-1}\right) \\ +\left(\frac{(\beta r)^{-\beta(1-\beta)^{-1}}}{\lambda r\beta r(1-\tau)k_{t_0}}\right)^{1-\mu^w}\left(\frac{(1-\tau)^{-1}\left(v^w_{t_0}\left(1-\tau_{t_0}\right)+\tau_{t_0}n^{-1}\right)}{n^{-1}}\right)^{1-\mu^w(1-\pi)} \\ \cdot\left(n^{-1}-(1-\tau)^{-1}\left(v^p_{t_0}\left(1-\tau_{t_0}\right)+n^{-1}\tau_{t_0}\right)\right)\end{array}\right)}{\frac{(1-\tau_t)}{(1-\tau)}n^{-1}\left(v^w_t-v^p_t\right)}$$
$$\leq\left((1-\tau)^{\frac{\beta}{1-\beta}}\left(1+\tau\left(\frac{\beta}{1-\beta}\right)\right)\right)$$

democracy is sustainable provided the constant growth rate $g=r(1-\lambda)(1-\tau)\geq 1$.

4.3 Characterization

In any period, the incentive constraints of both the median voter and the poor pivotal agent must also hold. Clearly, if $v^i_t>v^p_t$, and since this order is preserved through time, if the poor pivotal agent's incentive constraint is satisfied, so is the poor median voter's. The converse however is not true. Unconstrained by the poor pivotal agent and under commitment at time t, the poor median voter selects a sequence $\{\hat{\tau}_t, \hat{\tau}_{t+1}, ..\hat{\tau}_{m'}, 0, 0, ...\}$ However, such a sequence can violate the incentive constraints of the poor pivotal agent. To study this case, we write the Lagrangian for the median voter, to include the incentive constraints of the poor agent.

$$\mathcal{L} =$$

$$\frac{1}{1-\sigma}\left(\begin{array}{c}\lambda_{t_0}^{1-\sigma}\left(\left(\begin{array}{c}\left(1-\tau_{t_0}\right)v_{t_0}^i + n^{-1}\tau_{t_0} \\ +\sum_{j=t_0+1}^{\infty}\tau_j\prod_{s=t_0+1}^{j} g_s(r\left(1-\tau_s\right))^{-1}\end{array}\right) rk_{t_0}\right)^{1-\sigma} \\ \cdot\left(1+\sum_{n=t_0+1}^{\infty}\beta^{n-t_0}\left(\prod_{s=t_0+1}^{n}(\beta r\left(1-\tau_s\right))^{\frac{1}{\sigma}}\right)^{1-\sigma}\right) - (1-\beta)^{-1}\end{array}\right)$$

$$+\sum_{t=t_0}^{\infty}\phi_t\left(V^{wD}\left(k_t, k_t^w, t\right) - V^{wA}\left(k_t, k_t^w, t\right)\right)$$

$$+\sum_{t=t_0}^{\infty}\zeta_t\left(V^{pD}\left(k_t, k_t^p, t\right) - V^{pA}\left(k_t, k_t^p, t\right)\right)$$

$$+\theta_{1t}\tau_t + \theta_{2t}\left(1-\tau_t\right)$$

where ζ_t are the multipliers for the poor pivotal agent's constraints, and $V^{wA}\left(k_t, k_t^w, t\right)$ and. $V^{pA}\left(k_t, k_t^p, t\right)$ are the values that the rich and poor pivotal agents would receive by reverting to authoritarian regimes, incorporating the probabilistic outcomes of simultaneous resurrections dicussed in the previous section.

The first-order conditions with respect to τ_{t_0} are[9]

$$\frac{\partial\mathcal{L}\left(k_{t_0}, k_{t_0}^M, t_0\right)}{\partial\tau_{t_0}} = \left(\left(-v_{t_0}^i + n^{-1}\right)rk_{t_0}\right)X_{t_0}^{-\sigma}Z_{t_0}^{\sigma} + \theta_{1t_0} - \theta_{2t_0}$$

$$+\zeta_{t_0}\left(X_{t_0}^p\right)^{-\sigma}Z_{t_0}^{\sigma}\left(-v_{t_0}^p + n^{-1}\right)rk_{t_0} \tag{31}$$

$$+\phi_{t_0}\left(X_{t_0}^w\right)^{-\sigma}Z_{t_0}^{\sigma}\left(-v_{t_0}^w + n^{-1}\right)rk_{t_0} \tag{32}$$

Note that the median voter as well as the poor pivotal agent would like to set τ_{t_0} to 1 unless the rich pivotal agent's constraint is binding, since $\left(-v_{t_0}^i + n^{-1}\right) < 0$ and $\left(-v_{t_0}^p + n^{-1}\right) < 0$. If τ_{t_0} is interior, that is if $\tau_{t_0} \in (0,1)$, the rich pivotal agent's constraint is binding, and we may proceed with the analysis as in the case of Theorem 4, provided the poor pivotal agent's constraints are slack. However, as we saw in the previous section, the feasible set may be empty under certain conditions, and the median voter may not be able to satisfy the constraints imposed by the poor and rich agents simultaneously. If this is not the case the

[9]Note that derivatives of $V^{wA}\left(k_t, k_t^w, t\right)$ and $V^{pA}\left(k_t, k_t^p, t\right)$ do not show up in the first order conditions with respect to taxes, since they do do not involve taxes.

median voter will set τ_{t_0} either at 1 or to satisfy the rich pivotal agent's constraints.

We now turn to the analysis of τ_m.

$$\begin{aligned}\frac{\partial \mathcal{L}\left(k_{t_0}, k_{t_0}^i, t_0\right)}{\partial \tau_m} =& \left(Z_{t_0}^{\sigma} X_{t_0}^{1-\sigma}\right)\left(\left(X_{t_0}\right)^{-1} \frac{\partial X_{t_0}}{\partial \tau_m}+\sigma(1-\sigma)^{-1}\left(Z_{t_0}\right)^{-1} \frac{\partial Z_{t_0}}{\partial \tau_m}\right)+ \\ & \left(Z_m^{\sigma}\left(X_m^w\right)^{-\sigma}\right) \phi_m\left(-v_m^w+n^{-1}\right) r k_m \\ & +\sum_{s=t_0+1}^{m-1} \phi_s\left(Z_s^{\sigma}\left(X_s^w\right)^{1-\sigma}\right)\left(\left(X_s^w\right)^{-1} \frac{\partial X_s^w}{\partial \tau_m}+\sigma(1-\sigma)^{-1}\left(Z_s\right)^{-1} \frac{\partial Z_s}{\partial \tau_m}\right) \\ & +\phi_{t_0}\left(Z_{t_0}^{\sigma}\left(X_{t_0}^w\right)^{1-\sigma}\right)\left(\left(X_{t_0}^w\right)^{-1} \frac{\partial X_{t_0}^w}{\partial \tau_m}+\sigma(1-\sigma)_{t_0}^{-1}\left(Z_{t_0}\right)^{-1} \frac{\partial Z_{t_0}}{\partial \tau_m}\right) \\ & \left(Z_m^{\sigma}\left(X_m^p\right)^{-\sigma}\right) \phi_m\left(-v_m^p+n^{-1}\right) r k_m \\ & +\sum_{s=t_0+1}^{m-1} \zeta_s\left(Z_s^{\sigma}\left(X_s^p\right)^{1-\sigma}\right)\left(\left(X_s^p\right)^{-1} \frac{\partial X_s^p}{\partial \tau_m}+\sigma(1-\sigma)^{-1}\left(Z_s\right)^{-1} \frac{\partial Z_s}{\partial \tau_m}\right) \\ & +\theta_{1m}-\theta_{2m}=0\end{aligned}$$

As before letting $\sigma \rightarrow 1$, from the Corollary 2 to Theorem 2

$$\begin{aligned}&\left(Z_s^{\sigma}\left(X_s^w\right)^{1-\sigma}\right)\left(\left(X_s^w\right)^{-1} \frac{d X_s^w}{d \tau_m}+\sigma(1-\sigma)^{-1}\left(Z_s\right)^{-1} \frac{d Z_s}{d \tau_m}\right) \qquad (33)\\ &=Z\left(\left(X_s^w\right)^{-1} \frac{d X_s^w}{d \tau_m}-\left(1-\tau_m\right)^{-1} \beta^{m-t_0}\right)<0\end{aligned}$$

where $Z=1/1-\beta$ and from the proof of Theorem 2, and

$$\begin{aligned}&\lim_{m \rightarrow \infty}\left[\left(Z_{t_0}^{\sigma} X_{t_0}^{1-\sigma}\right)\left(\left(X_{t_0}\right)^{-1} \frac{d X_{t_0}}{d \tau_m}+\sigma(1-\sigma)^{-1}\left(Z_{t_0}\right)^{-1} \frac{d Z_{t_0}}{d \tau_m}\right)\right] (34)\\ &=Z\left(\left(X_{t_0}\right)^{-1} \frac{d X_{t_0}}{d \tau_m}-\left(1-\tau_m\right)^{-1} \beta^{m-t_0}\right)<0\end{aligned}$$

Furthermore

$$\begin{aligned}&\lim_{m \rightarrow \infty}\left[\left(Z_s^{\sigma}\left(X_s^p\right)^{1-\sigma}\right)\left(\left(X_s^p\right)^{-1} \frac{d X_s^p}{d \tau_m}+\sigma(1-\sigma)^{-1}\left(Z_s\right)^{-1} \frac{d Z_s}{d \tau_m}\right)\right] (35)\\ &=Z\left(\left(X_s^p\right)^{-1} \frac{d X_s^p}{d \tau_m}-\left(1-\tau_m\right)^{-1} \beta^{m-t_0}\right)<0\end{aligned}$$

since the same analysis that applies to the median agent also applies to the poor median agent, as they differ only in their initial v_t's. However, we cannot claim that $\lim_{m \rightarrow \infty}\left[\frac{\partial \mathcal{L}\left(k_{t_0}, k_{t_0}^i, t_0\right)}{\partial \tau_m}-\left(\theta_{1m}-\theta_{2m}\right)\right]<0$ and therefore that $\lim_{m\to\infty} \tau_m = 0$: in this case we require the poor pivotal agent's incentive constraints to be satisfied at all times, not just in the limit from the perspective of time t_0. Unlike the rich pivotal agent who prefers zero taxes, the constraints of the poor pivotal agent may not allow the sequence of taxes $\{\hat{\tau}_{t_0}, \hat{\tau}_{t_0+1}, \ldots, \hat{\tau}_{m'}, 0, 0, ...\}$ given by Theorem 4 since such a sequence can violate the incentive constraints of the poor pivotal agent. Note the incentive constraint of the poor pivotal agent at time t_0, (and indeed at each time t as well):

$$\tau_{t_0} \geq \frac{\left(n^{-1}\left(K_{t_0}\left(\frac{k^{**}}{k_{t_0}}\right)^{1-\mu^p}\left(\frac{v^p_{t_0}}{n^{-1}}\right)^{\pi'\mu^p} - x_{t_0+1}\right) - v^p_{t_0}\right)}{\left(n^{-1} - v^p_{t_0}\right)} = \tau^L_{t_0} \qquad (36)$$

In either a growing or a shrinking economy if $\left(\frac{k^{**}}{k_t}\right)^{1-\mu^p} \to 0$, the constraint will be satisfied eventually and taxes will be set to zero eventually, but if $\frac{k^{**}}{k_t}$ is close to one, the constraint may not hold if taxes are prematurely close to zero, because the numerator is positive. To see this note that for $\frac{k^{**}}{k_t} = 1$ a negative numerator requires:

$$\left(n^{-1}\left(K_t\left(\frac{v^p_t}{n^{-1}}\right)^{\pi'\mu^p} - x_{t+1}\right) - v^p_t\right) < 0$$

$$K_t\left(\frac{v^p_t}{n^{-1}}\right)^{\pi'\mu^p} - (1 + x_{t+1}) < \frac{v^p_t}{n^{-1}} - 1 < 0$$

In the proof of Theorem 5 we showed that $K_t - (1 + x_{t+1}) \geq 0$, so for $\frac{k^{**}}{k_t}$ close to one, the constraint will be violated if $\left(\frac{v^p_t}{n^{-1}}\right)^{\pi'\mu^p}$ is close to one, for example if $\pi' = 0$ so that an insurrection of the poor suceeds with probabily one. Thus to keep the poor pivotal agent in check, the median voter may have to delay setting taxes to zero beyond m', but then would have to reduce earlier taxes to still satisfy the constraints of the rich pivotal agent.

Finally, note that if the poor pivotal agent's constraint were to be satisfied, the poor pivotal agent would have no incentive to implement a left wing authoritarian regime at any point in time. This implies however that the median voter, with $v^i_s > v^p_s$ and therefore higher discounted utility than the poor pivotal agent, also does not have an incentive to reset the tax sequence and deviate from the original sequence, provided such a deviation involves a deviation cost that scales down lifetime utility by μ^M where $\mid \mu^M - 1 \mid \geq \mid \mu^P - 1 \mid$: even an immediate full redistribution would not yield higher utility to the median voter if it does not do so for the poor pivotal agent. This implies that the tax sequence chosen by the median voter that respects the poor pivotal agent's incentive constraints in each period is time consistent. We have then the following Theorem.

Theorem 7 *When democracy is sustainable, the tax sequence chosen by the median voter is time consistent if by resetting taxes the median voter's lifetime utility is scaled down by* μ^M, *where* $\mid \mu^M - 1 \mid \geq \mid \mu^P - 1 \mid$.

5 Appendix

Proof of Theorem 1 The recurrence relation for the solution of continued fractions is standard (see [**?**], p. 8-9) and can be written as a first order system as

$$\begin{bmatrix} A(n) \\ C(n) \end{bmatrix} = \begin{bmatrix} b_n & a_n \\ 1 & 0 \end{bmatrix} \begin{bmatrix} A(n-1) \\ C(n-1) \end{bmatrix}, \quad C(n) = A(n-1)$$

$$\begin{bmatrix} B(n) \\ D(n) \end{bmatrix} = \begin{bmatrix} b_n & a_n \\ 1 & 0 \end{bmatrix} \begin{bmatrix} B(n-1) \\ D(n-1) \end{bmatrix}, \quad D(n) = B(n-1)$$

which in our framework reduces to stochastic matrices:

$$\begin{bmatrix} b_{2n+1} & a_{2n+1} \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \lambda_{t+n} & (1-\lambda_{t+n}) \\ 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} b_{2n} & a_{2n} \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \tau_{t+n} & (1-\tau_{t+n}) \\ 1 & 0 \end{bmatrix}$$

Furthermore

$$\begin{bmatrix} b_{2n+1} & a_{2n+1} \\ 1 & 0 \end{bmatrix} \begin{bmatrix} b_{2n} & a_{2n} \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1-\lambda_{t+n}(1-\tau_{t+n}) & \lambda_{t+n}(1-\tau_{t+n}) \\ \tau_{t+n} & 1-\tau_{t+n} \end{bmatrix}$$

with

$$A(1) = a_1, \ A(0) = 0, \ A(-1) = 1, \ B(1) = b_1, \ B(0) = 1, \ B(-1) = 0$$

We note that the product of stochastic matrices is stochastic. Then

$$\begin{bmatrix} A(2n+1) \\ C(2n+1) \end{bmatrix} = \begin{bmatrix} 1-\lambda_{t+n}(1-\tau_{t+n}) & \lambda_{t+n}(1-\tau_{t+n}) \\ \tau_{t+n} & 1-\tau_{t+n} \end{bmatrix} \begin{bmatrix} A(2n-1) \\ C(2n-1) \end{bmatrix} \quad (37)$$

$$\begin{bmatrix} B(2n+1) \\ D(2n+1) \end{bmatrix} = \begin{bmatrix} 1-\lambda_{t+n}(1-\tau_{t+n}) & \lambda_{t+n}(1-\tau_{t+n}) \\ \tau_{t+n} & 1-\tau_{t+n} \end{bmatrix} \begin{bmatrix} B(2n-1) \\ D(2n-1) \end{bmatrix} \quad (38)$$

For $n = 1, \ldots$, iteration gives

$$\begin{bmatrix} A(2n+1) \\ C(2n+1) \end{bmatrix} = \left\{ \prod_{s=1}^{n} \begin{bmatrix} 1-\lambda_{t+s}(1-\tau_{t+s}) & \lambda_{t+s}(1-\tau_{t+s}) \\ \tau_{t+s} & 1-\tau_{t+s} \end{bmatrix} \right\} \begin{bmatrix} a_1 \\ 0 \end{bmatrix}$$
$$= \left\{ \prod_{s=1}^{n} \begin{bmatrix} 1-\lambda_{t+s}(1-\tau_{t+s-1}) & \lambda_{t+s}(1-\tau_{t+s}) \\ \tau_{t+s} & 1-\tau_{t+s} \end{bmatrix} \right\} \begin{bmatrix} 1-\lambda_t \\ 0 \end{bmatrix}$$
$$\begin{bmatrix} B(2n+1) \\ D(2n+1) \end{bmatrix} = \left\{ \prod_{s=1}^{n} \begin{bmatrix} 1-\lambda_{t+s}(1-\tau_{t+s}) & \lambda_{t+s}(1-\tau_{t+s}) \\ \tau_{t+s} & 1-\tau_{t+s} \end{bmatrix} \right\} \begin{bmatrix} b_1 \\ 1 \end{bmatrix}$$
$$= \left\{ \prod_{s=1}^{n} \begin{bmatrix} 1-\lambda_{t+s}(1-\tau_{t+s}) & \lambda_{t+s}(1-\tau_{t+s}) \\ \tau_{t+s} & 1-\tau_{t+s} \end{bmatrix} \right\} \begin{bmatrix} \lambda_t \\ 1 \end{bmatrix}$$

and

$$\frac{A(2n+1)}{B(2n+1)} = \frac{\begin{bmatrix} 1\,0 \end{bmatrix} \left\{ \prod_{s=1}^{n} \begin{bmatrix} 1-\lambda_{t+s}(1-\tau_{t+s}) & \lambda_{t+s}(1-\tau_{t+s}) \\ \tau_{t+s} & 1-\tau_{t+s} \end{bmatrix} \right\} \begin{bmatrix} 1-\lambda_t \\ 0 \end{bmatrix}}{\begin{bmatrix} 1\,0 \end{bmatrix} \left\{ \prod_{s=1}^{n} \begin{bmatrix} 1-\lambda_{t+s-1}(1-\tau_{t+s}) & \lambda_{t+s}(1-\tau_{t+s}) \\ \tau_{t+s} & 1-\tau_{t+s} \end{bmatrix} \right\} \begin{bmatrix} \lambda_t \\ 1 \end{bmatrix}} \tag{39}$$

We have to show that $\lim_{n\to\infty} \frac{A(2n+1)}{B(2n+1)}$ exists, that is that the product of the stochastic matrices in (39) converges. since

$$J_{t+s} = \begin{bmatrix} 1-\lambda_{t+s}(1-\tau_{t+s}) & \lambda_{t+s}(1-\tau_{t+s}) \\ \tau_{t+s} & 1-\tau_{t+s} \end{bmatrix}$$

is stochastic, a theorem in [?] guarantees that

$$\lim_{n\to\infty} (J_n J_{n-1} \cdots J_{t+1})$$

converges to a stochastic matrix G_∞, if J_s is fully indecomposable, which in the case of 2×2 matrices requires $J_s > 0$. This will hold as long as $\tau_{t+s} \geq e > 0$ for $s \geq 1$. Note that Assumptions 1 and 2 imply $\tau_{t+s} < 1$ for all $s \geq 1$. (Note initial $\tau_t \in [0,1]$.) So now consider the case where $\lim_{s\to\infty} \sup \tau_{t+s} > 0$. Note that if $\tau_{t+s} > 0,\ J_s > 0$ and if $\tau_{s-1} = 0,\ J_{s-1} \geq 0$, with $J_s J_{s-1} > 0$. So we can relabel the sequence $G_n = J_n J_{n-1} \cdots J_{t+1}$ as sequence $G_n = H_n R_n H_{n-1} R_{n-2} \cdots H_{t+2} R_{t+2} H_{t+1} R_{t+1}$ where each H_s is the product of fully indecomposable positive matrices $J_s > 0$ with consecutive indices and R_s is the product of matrices with $J_s \geq 0$ with consecutive indices, with the convention that $R_{t+1} = I$ if $J_{t+1} > 0$, and is part of H_{t+1}. Let $F_n = H_n R_n$. The infinite product $\lim_{n\to\infty} F_n F_{n-1} ... F_{t+1}$ is the product of positive stochastic 2×2 matrices since each F_n is a positive matrix, and therefore converges to a stochastic matrix $G_\infty \geq 0$. Then

$$\lim_{n\to\infty} A(2n+1) = \lim_{n\to\infty} \begin{bmatrix} 1\,0 \end{bmatrix} [G_\infty] \begin{bmatrix} 1-\lambda_t \\ 0 \end{bmatrix} \geq 0$$

For $x_{t+1} = \lim_{n\to\infty} \frac{A(2n+1)}{B(2n+1)} < \infty$, we need $\lim_{n\to\infty} B(2n+1) > 0$. Since $\lambda_t > 0$, and G_∞ is a non-negative stochastic matrix with row sums of unity, and the product of stochastic matrices is stochastic,.

$$\lim_{n\to\infty} B(2n+1) = \begin{bmatrix} 1 \, 0 \end{bmatrix} [G_\infty] \begin{bmatrix} \lambda_t \\ 1 \end{bmatrix} > 0$$

By construction we had $g_{s+1} = \frac{x_{s+1}(r(1-\tau_{s+1}))}{(x_{s+2}+\tau_{s+1})} \geq 0$ and $g_{s+1} = r(1-\lambda_s - \lambda_s x_{s+1}) < r$; so $0 \leq g_s < r$ for all $s \geq t$.

Now consider the case where $\lim_{s\to\infty} \sup \tau_{t+s} = 0$. Thus

$$\lim_{s\to\infty} J_{t+s} = \begin{bmatrix} 1-\lambda & \lambda \\ 0 & 1 \end{bmatrix}$$

since $\lambda = (1 - \beta^{\frac{1}{\sigma}} r^{\frac{1-\sigma}{\sigma}})$ if $\tau_s = 0$. But the product

$$\lim_{n\to\infty} (J_n J_{n-1} \cdots J_{t+1})$$

converges since the tail of the product converges to $\lim_{n\to\infty} \begin{bmatrix} 1-\lambda & \lambda \\ 0 & 1 \end{bmatrix}^n = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$ if $\lambda < 1$.[10] **Q.E.D.**

[10] More generally, see also Seneta, Theorem 4.14, page 150, and Exercise 4.38, p. 158, and in particular Chatterjee and Seneta, Corollary, p. 93. Note in particular that the index 2 is the single essential class of indices of $\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$, which is aperiodic.

6 References

Barro, Robert J. 1989. "Economic Growth in a Cross Section of Countries." *NBER Working Paper no. 3120.* Cambridge, MA: National Bureau of Economic Research.

Benhabib, Jess. and Velasco, Andres.; (1996). On the Optimal and Best Sustainable Taxes in an Open Economy. *European Economic Review,* **40**, 135-154.

Benhabib, Jess. and Rustichini, Aldo, "Optimal Taxes Without Commitment," , Journal of Economic Theory, 77 (1997), 231-259.

Benhabib, Jess, Rustichini, Aldo and Andres Velasco (2001) "Public Capital and Optimal Taxes Without Commitment," Review of Economic Design, 6, Festschrift in honor of Roy Radner, 371-396.

Bertola,Giuseppe (1993), "Factor Shares and Savings in Endogenous Growth," *American Economic Review*, Vol. 83, No. 5. , 1184-1198.

Besley, Timothy, and Stephen Coate. 1998. "Sources of Inefficiency in a Representative Democracy: A Dynamic Analysis." *American Economic Review 88:* 139-156.

Chamley, C. (1985). Efficient Taxation in a Stylized Model of Intertemporal General Equilibrium. *International Economic Review*, **26, 2**, 451-468.

Chamley, Christophe (1985). Efficient Taxation in a Stylized Model of Intertemporal General Equilibrium. *International Economic Review*, **26, 2**, 451-468.

Chari, V. V., and Kehoe, Patrick,"Optimal Monetary and Fiscal Policy," Handbook for Macroeconomics, eds. John Taylor and Michael Woodford, North-Holland, New York, volume 1C, 1671-1745, , 1999.

Chatterjee, S. and Seneta, E., (1977), "Towards Consensus: Some Convergence Theorems on Repeated Averaging," *J. Appl. Prob.*, 14, 89-97.

Dasgupta, Partha, (1993), *An Inquiry into Well-Being and Destitution*, Oxford: Oxford University Press, p. 47.

Hartfiel, D. J., (1974), "On Infinite Products of Nonnegative Matrices," *SIAM Journal of Applied Mathematics*, 26, 297-301.

Helliwell, John. 1994. "Empirical Linkages Between Democracy and Economic Growth." *British Journal of Political Science 24*: 225-48.

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

Li, Hongyi, Lyn Squire, and Heng-fu Zou. 1997. "Explaining International and Intertemporal Variations in Income Inequality." *Economic Journal 108:* 1-18.

Meltzer, Allan H. and Scott F. Richard. 1981. "A Rational Theory of the Size of Government." *Journal of Political Economy 89:* 914-927.

Mukherjee, Dilip and Debraj Ray, "Is Equality Stable?" American Economic Review 92 (Papers and Proceedings), 253–259 (2002).

Piketty, Thomas.2000. "Income Inequality in France, 1901-1998." Ms.

Przeworski, Adam, Michael E. Alvarez, José Antonio Cheibub, and Fernando Limongi. 2000. *Democracy and Development: Political Institutions and Well-Being in the World, 1950-1990.* New York: Cambridge University Press.

Przeworski, Adam and Fernando Limongi 1977. "Modernization: Theories and Facts." *World Politics 49*:

Sen, Amartya, (1991), "Welfare, Preference and Freedom." *Journal of Econometrics 50*: 15-29.

Sen, Amartya, (1994), "Freedom and Needs." *The New Republic*, January 10, p. 7.

Seneta, E., (1980), *Non-Negative Matrices and Markov Chains*, Springer-Verlag, New York.

Sorensen, Lisa and Waadeland, Haakon, (1992), *Continued Fractions with Applications*, North Holland, New York.