

The Political Economy of Redistribution under Democracy*

Jess Benhabib and Adam Przeworski
Department of Economics Department of Politics
New York University

March 5, 2004

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. We find that democracies survive in wealthy societies. Conditional on the initial income distribution and the capacity of the poor and the wealthy to overthrow democracy, each country has a threshold of capital stock above which democracy survives. This threshold is lower when the distribution of initial endowments is more equal and when the revolutionary prowess of these groups is lower. Yet in poor unequal countries there exist no redistribution scheme which would be accepted both by the poor and the wealthy. Hence, democracy cannot survive. As endowments increase, redistribution schemes that satisfy both the poor and the wealthy emerge. Moreover, as capital stock grows the wealthy tolerate more and the poor less redistribution, so that the set of feasible redistributions becomes larger. Since the median voter prefers one such scheme to the dictatorship of either group, democracy survives.

*We would like to thank, Daron Acemoglu, Marco Bassetto, Alberto Bisin, V.V. Chari, Pat Kehoe, Onur Ozgur for very useful comments.

1 Introduction

We ask what redistributions of income and assets are feasible in a democracy, given the initial assets, their distribution, and some features of the political environment. The question is motivated by the possibility that if the redistribution is insufficient for the poor or excessive for the wealthy, they may turn against democracy. Moreover, if no redistribution simultaneously satisfies the poor and the wealthy and if either group has any chance to establish its dictatorship, 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 assume that decisions about redistribution are made in elections and show that no majority coalition of poor and wealthy leaves both better off than the decision of the median voter (see Theorem 2). Moreover, given a linear redistribution scheme, 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 path of future redistribution. To be accepted, however, the decision of the median voter must leave the poor and the wealthy at least as well off as they expect to be if they sought to establish their respective dictatorships, where they would choose their best redistribution scheme unilaterally. Hence, if democracy is to survive, any redistribution must satisfy the constraints originating from the possibility of rebellion. While we do not solve for the optimal redistribution scheme of the median voter, we investigate these constraints.¹

We find that democracies survive in wealthy countries. Conditional on the initial income distribution and the capacity of the poor and the wealthy to overthrow democracy, each country has a threshold of capital stock above which democracy survives. This threshold is lower when the distribution of initial endowments is more equal and when the revolutionary prowess of these groups is lower. In the extreme, democracy survives at any income if its distribution is sufficiently egalitarian or if neither group can establish dictatorship. Yet in poor unequal countries there exist no redistribution scheme which would be accepted both by the poor and the wealthy. Hence, democracy cannot survive. As endowments increase, redistribution schemes that satisfy both the poor and the wealthy emerge. Moreover, as capital stock grows the wealthy tolerate more and the poor less redistribution, so that the set of feasible redistributions becomes larger. Since the median voter prefers one such scheme to the dictatorship of either group, the outcome of electoral competition is obeyed by everyone and democracy survives.

These results are driven by an assumption about preferences. The cost of

¹See however Benhabib and Przeworski (2003) for a the solution of the unique optimal redistribution scheme under logarithmic utilities.

dictatorship is the loss of freedom. We follow the argument of Sen (1991) that people suffer disutility when they are not free to live the lives of their choosing. Specifically, even if we allow that the losers in the conflict over dictatorship suffer more, we also allow that everyone may dislike dictatorship to some extent. This preference against dictatorship (or for democracy) is independent of income: as Dasgupta (1993: 47) put it, the view that the poor do not care about 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). Yet since the marginal utility of income declines in income, while the dislike of dictatorship is independent of income, at a sufficiently high income the additional gain that would accrue from being able to dictate tax rates becomes too small to overcome the loss of freedom².

Explanations in terms of preferences are justifiably suspect. We are driven to it because of two facts discussed in the next section: democracy is more likely, indeed certain, to survive in wealthy countries, and no plausible rival hypothesis eliminates the role of income in sustaining democracy. Hence, income matters and income is not a proxy for something else.

In the remainder of the paper, we present a model of the economy, analyze the constraints on electoral outcomes originating from the possibility of rebellion, and derive the results. The proof that the median voter is decisive over the entire path of tax rates is presented in section 4.1.

2 Per Capita Income and the Survival of Democracy

The probability that a democracy³ would survive rises steeply in per capita income⁴. Between 1950 and 1999, the probability that a democracy would die during any year in countries with per capita income under \$1,000 (1985 PPP dollars) was 0.0845, so that one in twelve died. In countries with incomes between \$1,001 and \$3,000, this probability was 0.0362, for one in twenty-eight. Between \$3,001 and \$6,055, this probability was 0.0163, one in sixty-one. And no democracy ever fell in a country with per capita income higher than that of Argentina in 1975, \$6,055. This is a startling fact, given that throughout history about seventy democracies collapsed in poorer countries, while thirty-seven democracies spent over 1000 years in more developed countries and not

²For a general game-theoretic approach to the dependence of social conflict and cooperation on wealth, based on the non-homotheticity of preferences, see Benhabib and Rustichini (1996).

³Democracy is defined here as a regime in which incumbents lose elections and leave office if they lose, but alternative definitions generate the same conclusion. The data set we use to identify democracies can be found at <http://pantheon.yale.edu/~jac236/>.

⁴This is not true of regimes in general. Transitions from dictatorship to democracy become slightly more likely as income increases up to about \$5,000 and then become less likely. Moreover, they appear to be driven not by income but by other factors. The only reason for the association between the density of democracy and wealth is that democracy is an absorbing state in wealthy societies. On this topic, see Przeworski and Limongi (1997) and Przeworski (2004).

one died. These patterns are portrayed in Figure 1, which shows a loess smooth of transitions to democracy as a function of per capita income (the vertical bars are local 95% confidence intervals).

Is income a proxy for something else? Following the classical book of Lipset (1960), an enormous literature sought to explain the observed prevalence of democracies in the developed countries and their paucity in the less developed ones. This literature did not distinguish the factors that lead to the emergence of democracy from those that cause it to survive once established, and these factors are different (Przeworski and Limongi 1997, Przeworski et al. 2000). Yet several arguments offered in this literature apply to the role of factors other than per capita income in sustaining democracy.

In Table 1 we show probit regressions in which the dependent variable are deaths of democracies and the column headings specify the rival hypotheses.⁵ The conclusion is clear: While some of the rival factors do matter in the presence of income, none of them eliminates the crucial role of income in sustaining democracy.

The most obvious candidate for a rival explanation is education, about which Lipset (1960: 38-40) observed that

Many people have suggested that the higher the education level of a nation's population, the better the chances for democracy.... Education presumably broadens man's outlook, enables him to understand the need for norms of tolerance, restrains him from adhering to extremist doctrines, and increases his capacity to make rational electoral choices....The higher one's education, the more likely one is to believe in democratic values and support democratic practices.

We take the years of education of an average member of the labor force and learn that education plays some additional rule in sustaining democracy, but it does not reduce the importance of income.

Coser (1956) argued, and many sociologists following him agreed, that democracy is easier to sustain if a country has a complex social structure. Coser's argument was that when social structure is complex, cleavages overlap, rather than pit one large group against another. Coalitions need to be built and they entail compromises. Hence, conflict is decentralized and attenuated. The argument was put forth originally by Ross (1920): "Every species of social conflict interferes with every other species ... save only when lines of cleavage coincide; in which case they reinforce each other." The example given by Lipset (1960) was that the Quebecois are both Catholic and poorer, while the Anglos are Protestant and richer: in this case cleavages reinforce and accentuate, rather than attenuate, conflict. On the other hand, Eckstein (1966) explained the stability of Norwegian democracy in terms of cross-cutting cleavages.

⁵Note that because of data availability, the numbers of annual observations, as well as the numbers of transitions to dictatorship, differ across columns.

We test this argument by calculating labor force fractionalization, that is, the probability that two random members of the labor force do not work in the same of nine one-digit sectors (This variable is called COMPLEXITY). Complexity strongly reduces the probability that a democracy would die, but income still matters.⁶

Already J.S. Mill (1991: 230) argued that democracy is more difficult to sustain in countries ridden with ethno-linguistic divisions:

Free institutions are next to impossible in a country made up of different nationalities. Among a people without fellow-feeling, especially if they read and speak different languages, the united public opinion necessary to the working of representative government cannot exist.

With all the caveats about measuring ethnicity across cultures, we take the index of ethno-linguistic fractionalization, ELF60. Again, while democracy is less likely to die in the more homogeneous countries, the role of income continues to be important.

The theory of political obligation asserts that people feel duty to accept the results of a process in which they participated. Democracy, the argument goes, requires the "willingness to accept outcomes of as yet undetermined content" (Lamounier 1979: 13). This conception is widely criticized: as Coleman (1989: 197) puts it, "consenting to a process is not the same thing as consenting to the outcomes of the process." Yet if participation is a function of income, then income may be a proxy for participation. We find that electoral participation plays no role in addition to income, while income matters.

We reserved inequality for last, since income distribution plays a role in our model. Lipset (1960: 49) referred explicitly to mean-preserving spread:

The fact that the form which those political parties take in poorer countries is more extremist and radical than it is in wealthier ones is probably more related to the greater degree of inequality in such countries than to the fact that their poor are actually poorer in absolute terms.

Unfortunately, data on inequality are scarce, unreliable, and not easily compared across countries. All we can do is to take the high quality data from Deininger and Squire (1996) and extend them by attributing the same degree of inequality to two years before and after each observation. The resulting sample is still heavily biased in favor of wealthy countries, an additional reason to take the results with a grain of salt. With all these caveats, income distribution appears not to matter, while income continues to do so.⁷

⁶To measure complexity, we also used the standard deviation of the proportions working in the nine-digit sectors, with the same result. In addition, we tested the impact of the proportion of labor force in agriculture, which appears to play no role.

⁷Testing for endogeneity, or other more complex analyses, is not feasible with these data.

In the end, then, none of the rival hypotheses for which data exist eliminates the role of income. But why would income matter, independently of everything else? Lipset (1960: 51), who seemed to have thought that the more explanations the better, also offered a hypothesis which we find convincing:

The general income level of a nation also affects its receptivity to democratic norms. If there is enough wealth in the country so that it does make too much difference whether some redistribution takes place, it is easier to accept the idea that it does not matter greatly which side is in power. But if loss of office means serious losses for major groups, they will seek to retain office by any means available.

While, for simplicity, we use a model of electoral competition which implies that parties converge the same platform, we show that the spirit of this argument is consistent with the observed patterns. The eventual income gain that would accrue to a dictator matters a lot when incomes are low, but is not worth the loss of freedom when incomes are high.

Table 1: Transitions to dictatorship, as a function of per capita income and rival variables

	None	Education	Complexity	ELF	Participation	Inequality
Constant	-13066 (0.1161)	-0.7771 (0.2002)	2.5750 (1.1970)	-1.0137 (0.1528)	-0.7488 (0.4334)	-0.8037 (0.6409)
GDP/cap	-0.2262 (0.0426)	-0.1820 (0.0633)	-0.1959 (0.1103)	-0.1755 (0.0404)	-0.2273 (0.0950)	-0.2734 (0.0867)
Rival		-0.0816 (0.0504)	-5.5095 (1.7709)	-0.6373 (0.2518)	-0.7150 (0.7344)	-0.0050 (0.0140)
N	2423	1085	1201	2234	581	771
TDA	47	30	10	46	12	14

GDP/cap in thousands of 1985 PPP, from PWT, extended by Easterly, <http://nyu.edu/fas/institute/dri/Easterly>.

Rival: changes according to column headings, namely,

Education: Years of education of average member of labor force, from Bhalla (1994).

Complexity: Labor force fractionalization index, from Kim (2004)

ELF: Ethno-linguistic fractionalization, from Easterly and Levine (1997)

Participation: Voters as a proportion of adults, from IDEA.

Inequality: Gini index, from Deininger and Squire (1996)

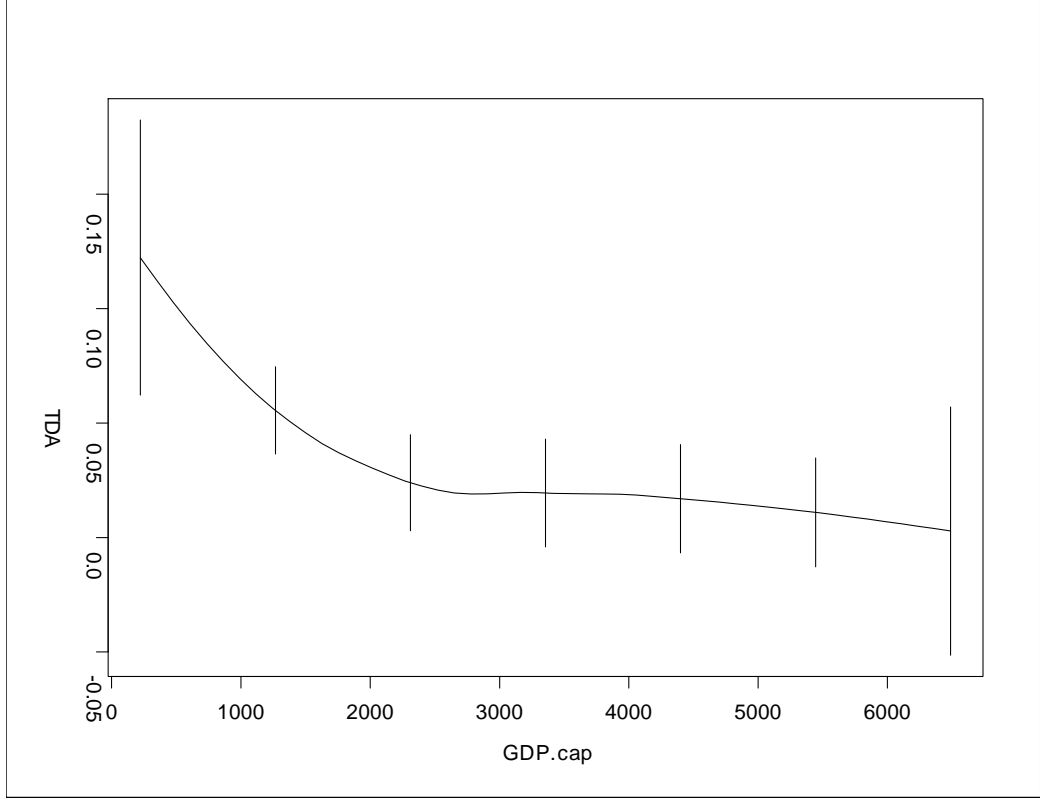


Figure 1: Transitions to dictatorship by per capita income

3 The Economy

The output y_t at time t is produced with capital k_t according to the linear production function

$$y_t = rk_t$$

with $r > 1$. There are n agents, indexed by i . In the initial period t_0 , they each own a share of the capital stock, $v_{t_0}^i$, with $\sum_{i=1}^n v_{t_0}^i = 1$. The shares of capital owned by agent i at time s are denoted by v_s^i , and the capital stock of this agent is $k_s^i = v_s^i k_s$. 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. Hence, the tax rate uniquely determines the extent of redistribution. The post-redistribution income is

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

where we assume that $\tau_s \in [0, \tilde{\tau}]$, where $\tilde{\tau} \leq 1$ for all s .

The standard median voter theorem applies when agents vote on a single issue, like the tax rate, and preferences over this issue satisfy the single-crossing property. In our context taxes are time varying, and the median voter chooses a sequence of taxes, not a single tax rate. Therefore, we have to show that the median voter theorem still holds, namely, that the tax sequence preferred by the median voter defeats any other tax sequence under majority voting. In particular, we have to rule out a tax sequence that would be preferred by a majority coalition of the poor and the rich, defeating the tax sequence preferred by the median voter. We address this issue in section 4.1, and show that there is no majority coalition that can outvote the preferred tax sequence of the median voter.

Redistribution has incentive effects on accumulation, so 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. They have CRRA preferences and the value function is

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

where q_s^i is the redistributive transfer agent i receives at time s . The constant term $(1-\sigma)^{-1} b$ is zero under a democratic regime and non-positive under a dictatorial regime. The utility function is discussed further in section 4.

We allow for the tax rates to vary across time. The first-order condition of the agent for an interior solution is:

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

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 $s = 1, 2, \dots$ (see Assumption 1 below), that:

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

From the no Ponzi and transversality conditions,

$$\lim_{t \rightarrow \infty} \left(\prod_{s=1}^t (r(1 - \tau_s))^{-1} \right) k_{t+1}^i = 0. \quad (3)$$

Iterating 1 forward, substituting into 2, using 3 and solving for c_0^i , we obtain:

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

Thus the agent consumes a fraction λ_t^i 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 $r(1-\tau_t)$. Note that $\lambda_t^i = \lambda_t$, so the propensity to consume is identical across agents. Note also that λ_t is the propensity to consume out of wealth because r is one plus the rate of return, net of depreciation.

The following assumptions assure that λ_t is bounded away from zero and one, that is $0 < \lambda^l \leq \lambda_t \leq \lambda^h < 1$ 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 $s = t_0 + 1, t_0 + 2, \dots$ where t_0 is the initial period.

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

The excluded case, $\tau_{\bar{t}} = 1$, 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 rule it out by Assumption 1.

3.0.1 Endogenizing transfers

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

$$k_t = \left(\prod_{s=t_0+1}^t g_s \right) k_{t_0}$$

Let the transfers be defined as:

$$\begin{aligned}
q_{t_0}^i &= n^{-1} \tau_{t_0} r k_{t_0} \\
q_t^i &= n^{-1} \tau_t r k_t = n^{-1} \tau_t r \left(\prod_{s=t_0+1}^t g_s \right) k_{t_0}
\end{aligned}$$

Using the definition of the growth rates and transfers,

$$c_t^i = \lambda_t \left((1 - \tau_t) r k_t^i + n^{-1} \left(\tau_t + \sum_{j=t+1}^{\infty} \tau_j \prod_{s=t+1}^j g_s (r(1 - \tau_s))^{-1} \right) r k_t \right) \quad (4)$$

Each agent's budget constraint implies

$$\begin{aligned} k_{t+1}^i &= (1 - \tau_t) r k_t^i + n^{-1} \tau_t r k_t - c_t^i \\ &= (1 - \tau_t) (1 - \lambda_t) r k_t^i \\ &\quad + \left(\tau_t (1 - \lambda_t) - \lambda_t \left(\sum_{j=t+1}^{\infty} \tau_j \prod_{s=t+1}^j g_s (r(1 - \tau_s))^{-1} \right) \right) n^{-1} r k_t \end{aligned} \quad (5)$$

Summing over agents,

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

Therefore, the equilibrium relation describing growth rates for our redistributive economy is:

$$g_{t+1} = r \left(1 - \lambda_t - \lambda_t \left(\sum_{j=t+1}^{\infty} \tau_j \prod_{s=t+1}^j g_s (r(1 - \tau_s))^{-1} \right) \right) \quad (6)$$

The solution of this equation, describing current growth rate as a function of future growth rates, expresses the growth rates in terms of an arbitrary sequence of current and future taxes and allows us to express the equilibrium of the economy for arbitrary tax sequences. Then a political system or mechanism chooses the equilibrium by selecting a particular tax sequence.⁸ 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 simply

$$g_s = r (1 - \lambda_s) (1 - \tau) \quad (7)$$

Note that λ_s and g_s are constant in this case since under taxes constant for $s > t_0$. We will not of course restrict our analysis to constant tax schemes.

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

To characterize the equilibrium dynamics of the economy, to be used in the next section, we first describe the evolution of asset shares from an initial distribution,

⁸For a general solution to equation 6 for arbitrary tax sequences, and an analysis of the optimal sequence for the median voter, see Benhabib and Przeworski (2003) who characterize a globally unique solution for the optimal redistributive tax sequence which is in agreement with the results of Chari and Kehoe (1999).

given the redistribution scheme. If we express asset shares as $v_t^i = \frac{k_t^i}{k_t}$, then equation 5 yields:

$$v_{t+1}^i = \left((g_{t+1})^{-1} r (1 - \lambda_t) (1 - \tau_t) \right) v_t^i + (g_{t+1})^{-1} r n^{-1} \left(\tau_t (1 - \lambda_t) - \lambda_t \sum_{j=t+1}^{\infty} \tau_j \prod_{s=t+1}^j g_s (r (1 - \tau_s))^{-1} \right) \quad (8)$$

where the law of motion for g_{t+1} is given by (6).

Note that the last term on the right-hand side of equation 8 describing the evolution of shares is independent of i , while the first term is proportional to the endowment v_t^i . Therefore even if shares change, their ordering is unaffected, and the median voter will be the same agent in each period. Note that if $\tau_s = \tau$, then $\lambda_s = \lambda$ and, using (7) and (8), we can solve for:

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

Therefore, shares v_{t+1}^i remain constant if taxes are constant and $g_s = r (1 - \tau) (1 - \lambda)$. That is, in this linear economy with redistribution, initial ownership ratios remain constant if taxes are constant. This is because agents always consume all income other than capital income, just as in Bertola's (1993) results, derived in a slightly different context.⁹

Suppose now that the tax rate is τ_{t_0} at t_0 , and τ afterwards. In this case the growth rate and $\lambda_s = \lambda$ are constant, and $g_s = r (1 - \lambda) (1 - \tau)$. Then the shares at $t_0 + 1$ are

$$v_{t_0+1}^i = (1 - \tau)^{-1} \left((1 - \tau_{t_0}) v_{t_0}^i + n^{-1} (\tau_{t_0} - \tau) \right) \quad (9)$$

and they are constant thereafter. Note that if $\tau_{t_0} > \tau$, $(v_{t_0+1}^j - v_{t_0}^j) (v_{t_0}^j - n^{-1}) < 0$, and if $\tau_{t_0} < \tau$, $(v_{t_0+1}^j - v_{t_0}^j) (v_{t_0}^j - n^{-1}) > 0$. If $\tau = 0$, $g_{t_0+1} = r (1 - \lambda_{t_0}) = (\beta r)^{\frac{1}{\sigma}}$. Therefore, if the initial share is $v_{t_0}^i$,

$$v_{t_0+1}^i = (1 - \tau_{t_0}) v_{t_0}^i + \tau_{t_0} n^{-1} \quad (10)$$

$$v_{t_0+s+1}^i = v_{t_0+s}^i \quad (11)$$

$\tau_t = 1$ produces complete non-distortionary equality in one step. That is $v_{t_0+s}^i = n^{-1}$ for all $s = 1, 2, \dots$

4 A political model of the sustainability of democracy

We study the political constraints on the median voter that prevent him from implementing his preferred tax scheme: if the median voter is poorer than the

⁹Bertola (1993) allows for differentially productive labor by introducing increasing returns to scale.

average, he prefers the tax sequence $\tau_{t_0} = 1$ and $\tau_{t_0+s} = 0$ for $s = 1, 2, \dots$, whereas if the median voter is richer than the average, he wants $\tau_{t_0+s} = 0$ for $s = 0, 1, 2, \dots$. We will proceed under the assumption that the median voter is poor.

There is a wealthy pivotal agent w , 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 $s = 0, 1, 2, \dots$. In turn, the poor pivotal agent, 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 $s = 1, 2, \dots$, a complete redistribution resulting in equal shares in the first period, followed by zero taxes afterwards. If in any period the pivotal agents receive less discounted utility under democracy than the expected value of a revolt aimed at instituting an authoritarian regime, they will revolt.

We assume the following:

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 who initiates a successful revolt to reset initial taxes when she reverts to an authoritarian regime. We assume, for simplicity, that once established, an authoritarian regime lasts forever. The success of a revolt, however, is probabilistic. If the median voter chooses a tax sequence under which the wealthy 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 and bear the costs of the autocratic regime: it is better to suffer autocracy under one's preferred tax sequence than under the tax sequence set by the other class. So we assume that if the tax sequence chosen by the median voter induces the right-wing wealthy agents to revolt, the revolution will succeed with probability π , but the left-wing poor agents will counter-revolt and may come to power with probability $1 - \pi$. 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 may come to power with probability 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 $\pi = \pi'$. Democracy is sustained if the median voter accommodates the right and the left by setting taxes that deter both of the pivotal agents from attempting to establish an authoritarian regime.

We also assume that the agents suffer a loss of utility under dictatorship. The utilities of the agents are given by $(1 - \sigma)^{-1} (c^{1-\sigma} - 1) + (1 - \sigma)^{-1} b^j$, where $j = s, u$, and $\sigma > 1$ ¹⁰. We set the parameters $b^u > 0$, $b^s \geq 0$ under dictatorship, and we set them to zero under democracy, where b^s is the per period utility cost to the group that successfully establishes a dictatorship, and b^u is the per period utility cost to the other group. We also assume the losers suffer at least as much under autocracy than the winners, and perhaps more.

¹⁰The results of this section also hold for $\sigma < 1$, in which case we must set $b^u < 0$, $b^s \leq 0$. We leave the derivations and proofs for this case to the reader.

Assumption 4 $b^u \geq b^s \geq 0$ and $b^u > 0$ under dictatorship, and $b^u = b^s = 0$ under democracy.

To preserve democracy starting at time t_0 , the median voter M must set the sequence of taxes τ_{t_0+s} , $s = 0, 1, \dots$ to maximize his value function, for all $t = t_0 + s$,

$$V(k_{t_0}, k_{t_0}^M, t_0) = \text{Max}_{\{\tau_t\}_{t=t_0}^\infty} \frac{1}{(1-\sigma)} \left(\lambda_{t_0}^{1-\sigma} \left((1-\tau_{t_0}) v_{t_0}^M + n^{-1} \tau_{t_0} + n^{-1} \sum_{j=t_0+1}^\infty \tau_j \prod_{s=t_0+1}^j g_s(r(1-\tau_s))^{-1} \right) r k_{t_0} \right)^{1-\sigma} \cdot \left(1 + \sum_{n=t_0+1}^\infty \beta^{n-t_0} \left(\prod_{s=t_0+1}^n (\beta r (1-\tau_s))^{\frac{1}{\sigma}} \right)^{1-\sigma} \right) - (1-\beta)^{-1} \right)$$

subject to, for all $t \geq t_0$,

$$0 \leq \left((1-\sigma)^{-1} \lambda_t^{1-\sigma} \left((1-\tau_t) r v_t^w + n^{-1} \tau_t + n^{-1} \sum_{j=t+1}^\infty \tau_j \prod_{s=t+1}^j g_s(r(1-\tau_s))^{-1} r k_t \right)^{1-\sigma} \cdot \left(1 + \sum_{n=t+1}^\infty \beta^{n-t} \left(\prod_{s=t+1}^n (\beta r (1-\tau_s))^{\frac{1}{\sigma}} \right)^{1-\sigma} \right) - (1-\sigma)^{-1} (1-\beta)^{-1} \right) - \pi (1-\sigma)^{-1} \left(\tilde{\lambda}_t^{-\sigma} (r v_t^w)^{1-\sigma} k_t^{1-\sigma} - (1-\beta)^{-1} + B^s \right) - (1-\pi) (1-\sigma)^{-1} \left(\tilde{\lambda}_t^{-\sigma} (r n^{-1})^{1-\sigma} k_t^{1-\sigma} - (1-\beta)^{-1} + B^u \right)$$

and

$$0 \leq \left((1-\sigma)^{-1} \lambda_t^{1-\sigma} \left((1-\tau_t) r v_t^p + n^{-1} \tau_t + n^{-1} \sum_{j=t+1}^\infty \tau_j \prod_{s=t+1}^j g_s(r(1-\tau_s))^{-1} r k_t \right)^{1-\sigma} \cdot \left(1 + \sum_{n=t+1}^\infty \beta^{n-t} \left(\prod_{s=t+1}^n (\beta r (1-\tau_s))^{\frac{1}{\sigma}} \right)^{1-\sigma} \right) - (1-\sigma)^{-1} (1-\beta)^{-1} \right) - (1-\pi') (1-\sigma)^{-1} \left(\tilde{\lambda}_t^{-\sigma} (r n^{-1})^{1-\sigma} k_t^{1-\sigma} - (1-\beta)^{-1} + B^s \right) - \pi' (1-\sigma)^{-1} \left(\tilde{\lambda}_t^{-\sigma} (r v_t^p)^{1-\sigma} k_t^{1-\sigma} - (1-\beta)^{-1} + B^u \right)$$

where $\tilde{\lambda}_t = \left(1 + \sum_{n=t+1}^\infty \beta^{n-t} \left(\prod_{s=t+1}^n (\beta r)^{\frac{1}{\sigma}} \right)^{1-\sigma} \right)^{-1} = 1 - \beta^{\frac{1}{\sigma}} r^{\frac{1-\sigma}{\sigma}} = \tilde{\lambda}$ because under dictatorship $\{0, 0, \dots\}$ is the preferred tax sequence of the rich and $\{1, 0, \dots\}$ is the preferred tax sequence of the poor. Also, $B^j = (1-\beta)^{-1} b^j$, $j = s, u$, : the group successfully establishing dictatorship loses $(1-\sigma)^{-1} B^s$ and the other group loses $(1-\sigma)^{-1} B^u$. The constraints restrict the taxes implemented by the median voter to yield utilities to the pivotal agents that exceed the expected utility that they would obtain by revolting. Since we consider the case where $\sigma > 1$, rearranging the constraints for the rich and the poor agent for $t \geq t_0$:

$$\begin{aligned}
0 \leq & (1-\sigma)^{-1} \lambda_t^{-\sigma} \left((1-\tau_t) r v_t^w + n^{-1} \tau_t + n^{-1} \sum_{j=t+1}^{\infty} \tau_j \prod_{s=t+1}^j g_s (r(1-\tau_s))^{-1} r k_t \right)^{1-\sigma} \\
& - (1-\sigma)^{-1} (1-\beta)^{-1} \\
& - (1-\sigma)^{-1} \left[\left(\pi (v_t^w)^{1-\sigma} + (1-\pi) (n^{-1})^{1-\sigma} \right) r^{1-\sigma} \tilde{\lambda}_t^{-\sigma} k_t^{1-\sigma} - (1-\beta)^{-1} + \pi B^s + (1-\pi) B^u \right]
\end{aligned} \tag{12}$$

$$\begin{aligned}
0 \leq & (1-\sigma)^{-1} \lambda_t^{-\sigma} \left((1-\tau_t) r v_t^p + n^{-1} \tau_t + n^{-1} \sum_{j=t+1}^{\infty} \tau_j \prod_{s=t+1}^j g_s (r(1-\tau_s))^{-1} r k_t \right)^{1-\sigma} \\
& - (1-\sigma)^{-1} (1-\beta)^{-1} \\
& - (1-\sigma)^{-1} \left[\left((1-\pi') (n^{-1})^{1-\sigma} + \pi' (v_t^p)^{1-\sigma} \right) r^{1-\sigma} \tilde{\lambda}_t^{-\sigma} k_t^{1-\sigma} - (1-\beta)^{-1} + (1-\pi') B^s + (1-\pi') B^u \right]
\end{aligned} \tag{13}$$

Assumption 5 $\beta r > 1$.

The assumption that $\beta r > 1$ assures that the growth rate for the tax sequence $\{\tau_t, 0, 0, \dots\}$, is positive: $g = (\beta r)^{\frac{1}{\sigma}} > 1$.

Theorem 1 *There exists $\hat{k}(\pi, \pi')$ and $\tilde{k}(\pi, \pi')$, $\tilde{k}(\pi, \pi') \geq \hat{k}(\pi, \pi') \geq 0$, such that democracy is sustainable for $k_{t_0} \geq \tilde{k}(\pi, \pi')$, and democracy is unsustainable for $k_{t_0} < \hat{k}(\pi, \pi')$.*

Proof. See Appendix ■

For a given an initial capital stock, democracy will be sustainable if wealth is sufficiently equally distributed, or if the probability of a successful revolution is sufficiently small:

Corollary 1 *Democracy is always sustainable (i) if v^w and v^p are sufficiently close to n^{-1} , that is, if income distribution is sufficiently equal, or (ii) if distribution is unequal, that is $v^w > v^p > n^{-1}$, but π and $1-\pi'$ are sufficiently small, that is, if the probability of a successful revolt for both of the pivotal agents is small.*

Proof. See Appendix ■

To inquire whether higher stocks of capital allow for higher levels of redistribution. We will say that the tax sequence $\{\tau'_{t_0}, \tau'_{t_0+1}, \tau'_{t_0+2}, \dots\}$ is “more redistributive” than $\{\tau_{t_0}, \tau_{t_0+1}, \tau_{t_0+2}, \dots\}$ if $1 \geq \tau'_{t_0+i} \geq \tau_{t_0+i}$, $i = 0, 1, \dots$. This of course is not a complete ranking of tax sequences, but sufficient for our purposes.

Corollary 2 *Let $\{\tau_{t_0}, \tau_{t_0+1}, \tau_{t_0+2}, \dots\}$ be a tax sequence with $\tau_{t_0} < 1$ for which democracy is sustainable from initial stock k_{t_0} , and let $\{\tau'_{t_0}, \tau_{t_0+1}, \tau_{t_0+2}, \dots\}$, $\tau'_{t_0} > \tau_{t_0}$, be a "more redistributive" sequence that is not sustainable from initial stock k_{t_0} . Then democracy is sustainable with the more redistributive tax sequence $\{\tau'_{t_0}, \tau_{t_0+1}, \tau_{t_0+2}, \dots\}$ for some $k'(\tau'_{t_0}) > k_{t_0}$.*

Proof. See Appendix. ■

Finally, we can ask whether the tax sequences in the feasible set for which democracy is sustainable contain sequences that are also time consistent from the perspective of the median voter. The set of time consistent tax sequences depends on the public's reaction to a deviation, which in equilibrium can deter deviations by the median voter. The public for example may lose confidence in the median voter and reduce savings, or refuse to re-elect a deviant government, which in turn can deter deviations from an announced tax sequence. (See for example Benhabib and Rustichini (1996).) However even without getting into such considerations, we can show that the feasible set of tax sequences contains a time consistent sequence. In proving the Theorem 1 above, we showed that for capital stocks that are sufficiently high, the sequence $\{\tau_{t_0}, 0, 0, \dots\}$ is always sustainable, and in particular $\{1, 0, 0, \dots\}$ is sustainable: if the stock is high enough, the rich agent will accept a full redistribution without revolt. This sequence is clearly time consistent, so even if the feasible set of tax sequences sustaining democracy are further constrained to time-consistent sequences, democracy is still sustainable for high enough initial capital.

To illustrate these results, consider an example where $\sigma = 2$, $\beta = 0.95$, $r = 1.08$, $\beta = .95$, $v^w = .2$, $n = 10$, $B = 2$ and where the tax sequence under democracy is constant: $\tau_s = \tau = .1$. In this case $\lambda_{t_0} = 1 - \beta^{\frac{1}{\sigma}} (r(1 - \tau))^{\frac{1-\sigma}{\sigma}}$, $\tilde{\lambda}_{t_0} = 1 - \beta^{\frac{1}{\sigma}} (r)^{\frac{1-\sigma}{\sigma}}$ and $g_s = r(1 - \tau)(1 - \lambda)$. The value functions for the wealthy agent under democracy and dictatorship, assuming a revolution by the wealthy succeeds with probability one are:

$$\begin{aligned} \text{Democracy} & : V = (1 - 2)^{-1} (7194.3) (k)^{-1} - (1 - 2)^{-1} (1 - 0.95)^{-1} \\ \text{Dictatorship} & : W = (1 - 2)^{-1} (1200) (k) - (1 - 2)^{-1} (1 - 0.95)^{-1} - 2 \end{aligned}$$

where V and W are the values and k is the capital stock. We can plot these as follows:

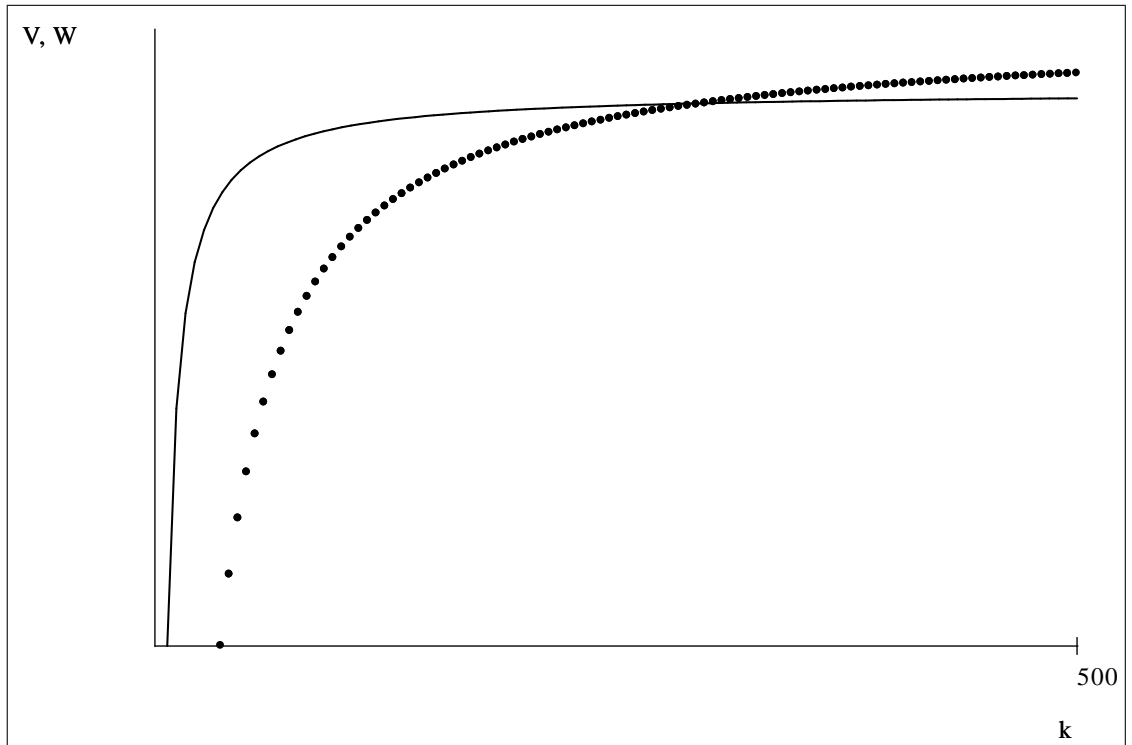


Figure 2: Value functions for democracy (points) and dictatorship (solid line)

At low values of the capital stock where consumption is low, redistributing the capital stock has a large impact on utility, since the marginal utility of consumption is big. This gain overwhelms the fixed utility loss of the rich associated with dictatorship. When the capital stock is high however, the marginal utility gain of a revolution is smaller than the utility loss associated with dictatorship, so the rich prefer not to revolt. If the probability of the success of the revolution becomes less than one, then the solid line shifts down, but the intersection of the two curves remain. The analysis of this example for the poor agent, say if $v^p = 0.1$, is analogous.

4.1 The median voter

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 majority coalition of the poor and the rich make at least one party better off and the other no worse off, relative to the preferred tax proposal of the median voter.¹¹ Hence we need to check whether

¹¹The standard median voter theorems apply when there is a single issue to be voted on. In our case there is an infinite sequence of tax rates, so the decisiveness of the median voter under simple assumptions is no longer assured. See however Gans and Smart (1996). We thank Marco Bassetto for this reference.

under democracy, the poor and rich pivotal agents, by proposing a tax sequence that draws all the voters poorer than the poor pivotal agent and those richer than the rich pivotal agent, can improve their utility by forming a majority coalition against the median voter. Note that the initial wealth share of a poor (rich) pivotal agent, other than being below (above) the median, is arbitrary, so that our Theorem 2 below implies that there is no tax sequence that a majority coalition of the rich and poor could propose to Pareto improve the utilities of the coalition over the tax sequence proposed by the median voter. Here a 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, exceeds the discounted utility the poor agent receives under tax sequence chosen by the median voter.

Intuitively, if we maximize the discounted utility of the poor agent subject to incentive constraints of the rich 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 a lot 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. Of course if both the poor and the rich pivotal agents prefer democracy to revolution under the tax sequence chosen median voter, the median voter will also prefer democracy to revolution under his own tax scheme.

The value functions of the poor pivotal agent, the median voter, and the rich pivotal agent are given by

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

where $i = p, M, w$ for the poor, median, and rich agents. Note that the value functions of the poor agent, the median voter, and the rich agent are identical except with respect to the terms $(1-\tau_{t_0}) v_{t_0}^i$. Changing taxes for periods after t_0 affects consumptions of these agents identically, by changing first period consumption as well as its rate of growth but changing τ_{t_0} 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 $\{\tau_{t_0}^p, (\tau_s^p)_{s=t_0+1}^\infty\}$ that improves his payoff relative to the taxes chosen by the median voter $\{\tau_{t_0}^M, (\tau_s^M)_{s=t_0+1}^\infty\}$, that 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, $\{\tau_{t_0}^p, (\tau_s^p)_{s=t_0+1}^\infty\}$, make the median voter worse off. Then, it follows that:

$$V(k_{t_0}, v_{t_0}^p, t_0; \{\tau_{t_0}^p, (\tau_s^p)_{s=t_0+1}^\infty\}) - V(k_{t_0}, v_{t_0}^p, t_0; \{\tau_{t_0}^M, (\tau_s^M)_{s=t_0+1}^\infty\}) > 0$$

$$V(k_{t_0}, v_{t_0}^M, t_0; \{\tau_{t_0}^p, (\tau_s^p)_{s=t_0+1}^\infty\}) - V(k_{t_0}, v_{t_0}^M, t_0; \{\tau_{t_0}^M, (\tau_s^M)_{s=t_0+1}^\infty\}) < 0$$

$$V(k_{t_0}, v_{t_0}^w, t_0; \{\tau_{t_0}^p, (\tau_s^p)_{s=t_0+1}^\infty\}) - V(k_{t_0}, v_{t_0}^w, t_0; \{\tau_{t_0}^M, (\tau_s^M)_{s=t_0+1}^\infty\}) \geq 0$$

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

$$V(k_{t_0}, v_{t_0}^{pM}, t_0; \{\tau_{t_0}^p, (\tau_s^p)_{s=t_0+1}^\infty\}) = V(k_{t_0}, v_{t_0}^{pM}, t_0; \{\tau_{t_0}^M, (\tau_s^M)_{s=t_0+1}^\infty\}) \quad (14)$$

$$V(k_{t_0}, v_{t_0}^M, t_0; \{\tau_{t_0}^p, (\tau_s^p)_{s=t_0+1}^\infty\}) < V(k_{t_0}, v_{t_0}^M, t_0; \{\tau_{t_0}^M, (\tau_s^M)_{s=t_0+1}^\infty\}) \quad (15)$$

$$V(k_{t_0}, v_{t_0}^{wM}, t_0; \{\tau_{t_0}^p, (\tau_s^p)_{s=t_0+1}^\infty\}) = V(k_{t_0}, v_{t_0}^{wM}, t_0; \{\tau_{t_0}^M, (\tau_s^M)_{s=t_0+1}^\infty\}) \quad (16)$$

Let $V^p(v)$ be the discounted utility value of the poor pivotal agent as a function of his share under his preferred taxes $\{\tau_{t_0}^p, (\tau_s^p)_{s=t_0+1}^\infty\}$, and $V^m(v)$ be the same under the taxes $\{\tau_{t_0}^M, (\tau_s^M)_{s=t_0+1}^\infty\}$ preferred by the median voter. Then the conditions above can be graphically illustrated in Figure 3:

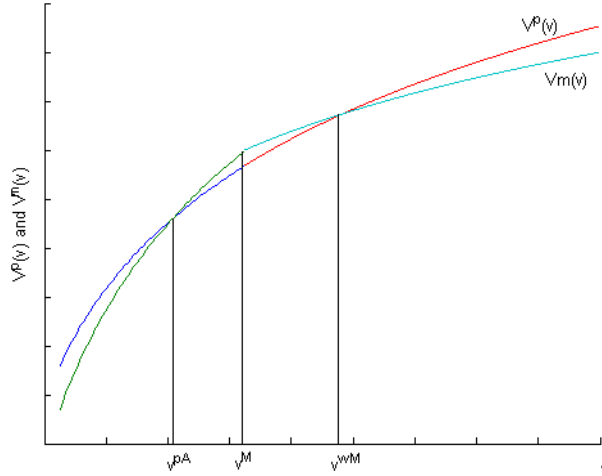


Figure 3: Double-Crossing

$V^m(v)$ and $V^p(v)$ must intersect at least twice, as drawn, to satisfy the conditions above, but of course they may cross more than twice. If there are more than two intersections, we take the smallest $v_{t_0}^{pM}$ and $v_{t_0}^{wM}$ for which the above holds.¹² If the agents below $v_{t_0}^{pM}$ together with those above $v_{t_0}^{wM}$ can form a majority, then the median voter's proposed tax sequence would be defeated. Theorem 2 proves that under democracy, there is no coalition of the rich and poor that can Pareto improve their utility by proposing an alternative tax sequence to defeat the median voter, as depicted in Figure 3 above.

Theorem 2 *There is no feasible tax sequence that a majority coalition of the rich and poor could propose to Pareto improve the utilities of the coalition over the tax sequence proposed by the median voter.*

Proof. See Appendix. ■

5 Appendix

Theorem 1: There exists $\hat{k}(\pi, \pi')$ and $\tilde{k}(\pi, \pi')$, $\tilde{k}(\pi, \pi') \geq \hat{k}(\pi, \pi') \geq 0$, such that democracy is sustainable for $k_{t_0} \geq \tilde{k}(\pi, \pi')$, and democracy is unsustainable for $k_{t_0} < \hat{k}(\pi, \pi')$

Proof. Let $G^p = (1 - \pi')B^s + \pi'B^u$ and $G^w = \pi B^s + (1 - \pi)B^u$. Consider the tax sequence $\{\tau_{t_0}, \tau, \tau, \dots\}$. The constraints 12 and 13 for period t_0 in the optimization problem of the median voter become (assuming $\sigma > 1$):

$$G^w (rk_{t_0})^{\sigma-1} \geq \lambda_{t_0}^{-\sigma} ((1 - \tau_{t_0})v_{t_0}^w + n^{-1}\tau_{t_0} + n^{-1}(\lambda^{-1} - 1)\tau)^{1-\sigma} - \tilde{\lambda}^{-\sigma} \left(\pi (v_{t_0}^w)^{1-\sigma} + (1 - \pi)(n^{-1})^{1-\sigma} \right) \quad (17)$$

$$G^p (rk_{t_0})^{\sigma-1} \geq \lambda_{t_0}^{-\sigma} ((1 - \tau_{t_0})v_{t_0}^p + n^{-1}\tau_{t_0} + n^{-1}(\lambda^{-1} - 1)\tau)^{1-\sigma} - \tilde{\lambda}^{-\sigma} \left((1 - \pi')(n^{-1})^{1-\sigma} + \pi'(v_{t_0}^p)^{1-\sigma} \right) \quad (18)$$

To show that democracy is sustainable, we must show that the feasible set satisfying the constraints is not empty. We have to show that the inequalities above will hold for a feasible tax sequence at every $t_0 + s$, $s = 0, 1, \dots$ for k_{t_0} sufficiently large. For the sequence of $\{\tau_{t_0}, \tau, \tau, \dots\}$, $\lambda_{t_0+s} = \lambda$, $s = 0, 1, \dots$, and the growth rate is $g = r(1 - \lambda)(1 - \tau)$. Since taxes are constant after the initial period, from equation 8 the shares $v_{t_0}^j$, $j = w, p$ become, for $s = 1, 2, \dots$,

$$v_{t_0+1+s}^j = v_{t_0+1}^j = (1 - \tau)^{-1} \left((1 - \tau_{t_0})v_{t_0}^j + n^{-1}(\tau_{t_0} - \tau) \right) \quad (19)$$

Note that if $\tau_{t_0} > \tau$, $(v_{t_0+1}^j - v_{t_0}^j)(v_{t_0}^j - n^{-1}) < 0$, and if $\tau_{t_0} < \tau$, $(v_{t_0+1}^j - v_{t_0}^j)(v_{t_0}^j - n^{-1}) > 0$. Using 19, the incentive constraints at $t_0 + 1$ are:

$$G^w (r^2(1 - \lambda)(1 - \tau)k_{t_0})^{\sigma-1} \geq \lambda_{t_0}^{-\sigma} \left((1 - \tau_{t_0})v_{t_0}^w + n^{-1}\tau_{t_0} + n^{-1}\tau(\lambda^{-1} - 1) \right)^{1-\sigma} - \tilde{\lambda}^{-\sigma} \left(\pi (v_{t_0+1}^w)^{1-\sigma} + (1 - \pi)(n^{-1})^{1-\sigma} \right) \quad (20)$$

¹²If the smallest $v_{t_0}^{pM}$ belongs to an interval over which 14 holds, we can redefine $v_{t_0}^{pM}$ as the sup v_{t_0} over that interval.

$$G^p (r^2 (1 - \lambda) (1 - \tau) k_{t_0})^{\sigma-1} \geq \lambda_{t_0}^{-\sigma} \left((1 - \tau_{t_0}) v_{t_0}^p + n^{-1} \tau + n^{-1} \tau (\lambda^{-1} - 1) \right)^{1-\sigma} - \tilde{\lambda}^{-\sigma} \left((1 - \pi') (n^{-1})^{1-\sigma} + \pi' (v_{t_0+1}^p)^{1-\sigma} \right) \quad (21)$$

Let $\tau_{t_0} > \tau$. Then, from 19 it follows that $v_{t_0+1}^p > v_{t_0}^p$ and $v_{t_0+1}^w < v_{t_0}^w$. To assure that the incentive constraints hold from t_0 on, we define the inequalities:

$$G^w (rk_{t_0})^{\sigma-1} \geq \lambda_{t_0}^{-\sigma} \left((1 - \tau_{t_0}) v_{t_0}^w + n^{-1} \tau_{t_0} + n^{-1} \tau (\lambda^{-1} - 1) \right)^{1-\sigma} - \tilde{\lambda}^{-\sigma} \left(\pi (v_{t_0}^w)^{1-\sigma} + (1 - \pi) (n^{-1})^{1-\sigma} \right) \quad (22)$$

$$G^p (rk_{t_0})^{\sigma-1} \geq \lambda_{t_0}^{-\sigma} \left((1 - \tau_{t_0}) v_{t_0}^p + n^{-1} \tau_{t_0} + n^{-1} \tau (\lambda^{-1} - 1) \right)^{1-\sigma} - \tilde{\lambda}^{-\sigma} \left((1 - \pi') (n^{-1})^{1-\sigma} + \pi' (v_{t_0+1}^p)^{1-\sigma} \right) \quad (23)$$

Under assumptions 1 and 2, $0 < \lambda^m \leq \lambda_{t_0} \leq \lambda^M < 1$, and the right hand sides of 22 and 23 are bounded. Note that the right side of 22 is identical to the right side of 17, and since $\sigma > 1$ and $v_{t_0+1}^p > v_{t_0}^p$, the right side of 23 is at least as large the right side of 18. Furthermore, the right side of 23 is identical to the right side of 21 and 23 is at least as large 20. Therefore, if 22 and 22 hold, the incentive constraints will be satisfied from time t_0 onwards. Let $\varsigma = \{\tau : g = r(1 - \lambda)(1 - \tau) > 1 \text{ and } \tau < \tau_{t_0}\}$. Note that $0 \in \varsigma$. Given π , for $\tau \in \varsigma$, there is a $\tilde{k}^w(\pi)$ such that for $k_{t_0} \geq \tilde{k}^w(\pi)$ 22 is satisfied with equality, and given π' , there is a $\tilde{k}^p(\pi')$ such that for $k_{t_0} \geq \tilde{k}^p(\pi')$ 23 is satisfied with equality. Let $\tilde{k}(\pi, \pi') = \text{Min} \left(\text{Max} \left(\tilde{k}^p(\pi'), 0 \right), \text{Max} \left(\tilde{k}^w(\pi), 0 \right) \right)$. Then if $\tilde{k}(\pi, \pi') < k_{t_0}$ democracy is sustainable for the feasible sequence $\{\tau_{t_0}, \tau, \tau, \dots\}$, $\tau \in \varsigma$. Of course there may be other feasible tax sequences that the median voter would prefer, but this establishes that democracy is sustainable by showing that the feasible set that satisfies the constraints and the assumptions is not empty.

To characterize unsustainability we show that for k_{t_0} sufficiently small, the feasible set is empty. Let $k \rightarrow 0$, so that $(rk_{t_0})^{\sigma-1} \rightarrow 0$. Then the constraints 17 and 18 become:

$$0 \geq \lambda_{t_0}^{-\sigma} \left((1 - \tau_{t_0}) v_{t_0}^w + n^{-1} \tau_{t_0} + n^{-1} x_{t_0} \right)^{1-\sigma} - \tilde{\lambda}_{t_0}^{-\sigma} \left(\pi (v_{t_0}^w)^{1-\sigma} + (1 - \pi) (n^{-1})^{1-\sigma} \right)$$

$$0 \geq \lambda_{t_0}^{-\sigma} \left((1 - \tau_{t_0}) v_{t_0}^p + n^{-1} \tau_{t_0} + n^{-1} x_{t_0} \right)^{1-\sigma} - \tilde{\lambda}_{t_0}^{-\sigma} \left((1 - \pi') (n^{-1})^{1-\sigma} + \pi' (v_{t_0}^p)^{1-\sigma} \right)$$

where $x_{t_0} = \sum_{j=t_0+1}^{\infty} \tau_j \prod_{s=t_0+1}^j g_s (r(1 - \tau_s))^{-1}$. Note that for $k_{t_0} \rightarrow 0$, we have $G^j (rk_{t_0})^{\sigma-1} \rightarrow 0$, $j = w, p$, which is equivalent to the case where $G^w = G^p = 0$, that is where there is no cost to an autocratic regime. In that case however, for any given tax sequence other than $\{\tau\} = \{0, 0, \dots\}$, there exists

$\tilde{\pi} \leq 1$ such that for $\pi > \tilde{\pi}$,

$$0 > \left((1-\sigma)^{-1} \lambda_{t_0}^{1-\sigma} \left((1-\tau_{t_0}) v_{t_0}^w + n^{-1} \tau_{t_0} + n^{-1} \sum_{j=t_0+1}^{\infty} \tau_j \prod_{s=t_0+1}^j g_s (r(1-\tau_s))^{-1} r k_{t_0} \right)^{1-\sigma} \right) \\ \cdot \left(1 + \sum_{n=t_0+1}^{\infty} \beta^{n-t_0} \left(\prod_{s=t_0+1}^n (\beta r (1-\tau_s))^{\frac{1}{\sigma}} \right)^{1-\sigma} \right) - (1-\sigma)^{-1} (1-\beta)^{-1} \\ - (1-\sigma)^{-1} \left[\tilde{\lambda}_{t_0}^{-\sigma} \left(\pi (v_{t_0}^w)^{1-\sigma} + (1-\pi) (n^{-1})^{1-\sigma} \right) r^{1-\sigma} k_{t_0}^{1-\sigma} - (1-\beta)^{-1} \right]$$

or

$$0 < \lambda_{t_0}^{-\sigma} \left((1-\tau_{t_0}) v_{t_0}^w + n^{-1} \tau_{t_0} + n^{-1} x_{t_0} \right)^{1-\sigma} - \tilde{\lambda}_{t_0}^{-\sigma} \left(\pi (v_{t_0}^w)^{1-\sigma} + (1-\pi) (n^{-1})^{1-\sigma} \right) \quad (24)$$

This follows because, when $\pi = 1$, the wealthy agent always prefers $\{\tau\} =$

$$\{0, 0, \dots\}, \text{ which is the same as the expected utility } (1-\sigma)^{-1} \left[\tilde{\lambda}_{t_0}^{-\sigma} (v_{t_0}^w)^{1-\sigma} r^{1-\sigma} k_{t_0}^{1-\sigma} - (1-\beta)^{-1} \right]$$

of revolting. Thus for $\pi = 1$, there is no tax sequence that satisfies the constraint for the rich for $G^w(k_{t_0})^{\sigma-1} = 0$, other than the sequence $\{\tau\} = \{0, 0, \dots\}$. Furthermore, the wealthy agent prefers any tax sequence to $\{\tau\} = \{1, 0, 0, \dots\}$ which establishes full wealth equality in the first period, so that the inequality 24 is reversed for $\pi = 0$. Therefore, since the right side of 24 is increasing in π , for any tax sequence there will be a $\tilde{\pi} \in (0, 1)$, which under defection gives a convex combination of utilities under $\{\tau\} = \{0, 0, \dots\}$ and $\{\tau\} = \{1, 0, 0, \dots\}$, such that 24 is satisfied with equality. Consider now

$$G^w(rk_{t_0})^{\sigma-1} = \lambda_{t_0}^{-\sigma} \left((1-\tau_{t_0}) v_{t_0}^w + n^{-1} \tau_{t_0} + n^{-1} x_{t_0} \right)^{1-\sigma} - \tilde{\lambda}_{t_0}^{-\sigma} \left(\pi (v_{t_0}^w)^{1-\sigma} + (1-\pi) (n^{-1})^{1-\sigma} \right) \quad (25)$$

and define \bar{k}^w as the value of k_{t_0} that satisfies 25 when $\pi = 1$. We can, since $\frac{G^w(rk_{t_0})^{\sigma-1}}{dk} > 0$, define $\hat{k}^w(\pi) : [0, 1] \rightarrow [0, \bar{k}]$ such that 25 is satisfied, where we have $\bar{k}^w = \hat{k}^w(1)$. Note that if $b^u > b^s$ so that $\frac{dG^w}{d\pi} < 0$, it also follows that $\frac{d\hat{k}(\pi)}{d\pi} > 0$. Therefore, given π there exists a unique $\hat{k}^w(\pi)$ such that revolt is preferred for $k < \hat{k}^w(\pi)$ because 17 cannot hold.

Now consider the poor agent. Again, if there is no cost to dictatorship ($G^p(rk_{t_0})^{\sigma-1} = 0$), and the probability that the revolt fails is zero ($\pi' = 0$), the poor pivotal agent will revolt because he can implement his preferred tax scheme, $\{1, 0, 0, \dots\}$ costlessly and for sure. Under these assumptions, for any other tax scheme than $\{1, 0, 0, \dots\}$ put forth by the median voter, we have

$$0 > \left((1-\sigma)^{-1} \lambda_{t_0}^{1-\sigma} \left((1-\tau_{t_0}) r v_{t_0}^p + n^{-1} \tau_{t_0} + n^{-1} \sum_{j=t_0+1}^{\infty} \tau_j \prod_{s=t_0+1}^j g_s (r(1-\tau_s))^{-1} r k_{t_0} \right)^{1-\sigma} \right) \\ \cdot \left(1 + \sum_{n=t_0+1}^{\infty} \beta^{n-t_0} \left(\prod_{s=t_0+1}^n (\beta r (1-\tau_s))^{\frac{1}{\sigma}} \right)^{1-\sigma} \right) - (1-\sigma)^{-1} (1-\beta)^{-1} \\ - (1-\sigma)^{-1} \left[\left((1-\pi') (n^{-1})^{1-\sigma} + \pi' (v_{t_0}^p)^{1-\sigma} \right) r^{1-\sigma} \lambda_{t_0}^{-\sigma} k_{t_0}^{1-\sigma} - (1-\beta)^{-1} + G^p \right]$$

or, using $G^p (rk_{t_0})^{\sigma-1} = 0$,

$$0 < \lambda_{t_0}^{-\sigma} \left((1 - \tau_{t_0}) v_{t_0}^p + n^{-1} \tau_{t_0} + n^{-1} x_{t_0} \right)^{1-\sigma} - \tilde{\lambda}_{t_0}^{-\sigma} \left((1 - \pi') (n^{-1})^{1-\sigma} + \pi' (v_{t_0}^p)^{1-\sigma} \right) \quad (26)$$

Furthermore, the poor agent prefers any tax sequence to $\{\tau\} = \{0, 0, \dots\}$ which establishes full wealth equality in the first period, so that the inequality 26 is reversed for $\pi' = 1$. Therefore, since the right side of 26 is decreasing in π' , for any tax sequence $\{\tau\}$ there will be a $\hat{\pi} > 0$, which under defection gives a convex combination of utilities under $\{\tau\} = \{0, 0, \dots\}$ and $\{\tau\} = \{1, 0, 0, \dots\}$, such that 26, is satisfied with equality. Consider now

$$G^p (rk_{t_0})^{\sigma-1} = \lambda_{t_0}^{-\sigma} \left((1 - \tau_{t_0}) v_{t_0}^p + n^{-1} \tau_{t_0} + n^{-1} x_{t_0} \right)^{1-\sigma} - \tilde{\lambda}_{t_0}^{-\sigma} \left((1 - \pi') (n^{-1})^{1-\sigma} + \pi' (v_{t_0}^p)^{1-\sigma} \right) \quad (27)$$

and define \bar{k}^p as the value of k_{t_0} that satisfies 27 when $\pi' = 0$. We can, since $\frac{G^p (rk_{t_0})^{\sigma-1}}{dk} > 0$, define $\hat{k}^p (\pi) : [0, 1] \rightarrow [0, \bar{k}_{t_0}^p]$ such that 27 is satisfied, where we have $\bar{k}^p = \hat{k}^p (0)$. If $b^u > b^s$ so that $\frac{dG^p}{d\pi'} > 0$, it also follows that $\frac{d\hat{k}^p (\pi)}{d\pi'} < 0$. Therefore, given π' there exists a unique $\hat{k}^p (\pi')$ such that revolt is preferred for $k < \hat{k}^p (\pi')$ because 18 cannot hold.

For the sustainability of democracy both 17 and 18 must hold, so we define $\hat{k} (\pi, \pi') = \text{Max} \left(\hat{k}^p (\pi'), \hat{k}^w (\pi) \right)$, Democracy then will not be sustainable if $k_{t_0} < \hat{k} (\pi)$. Finally, note that it must be true that $\tilde{k} (\pi, \pi') \geq \hat{k} (\pi, \pi')$, for otherwise there would be values of k_{t_0} which would be both sustainable and unsustainable, which is a contradiction. ■

Corollary 1: Democracy is always sustainable (i) if v^w and v^p are sufficiently close to n^{-1} , that is if income distribution is sufficiently equal, or (ii) if distribution is unequal, that is $v^w > v^p > n^{-1}$, but π and $1 - \pi'$ are sufficiently small, that is if the probability of a successful revolt for either of the pivotal agents is small.

Proof. For $\{\tau\} = \{\tau_{t_0}, \tau, \tau, \dots\}$, the growth rate is $g = r (1 - \lambda) (1 - \tau)$ for all t_0 , and $\sum_{j=t_0+1}^{\infty} \tau_j \prod_{s=t_0+1}^j g_s (r (1 - \tau_s))^{-1} = \tau \sum_{j=t_0+1}^{\infty} (1 - \lambda)^{j-t_0} = \tau \lambda^{-1}$. Then the incentive constraints are:

$$G^w (rk_{t_0})^{\sigma-1} \geq \lambda^{-\sigma} \left((1 - \tau_{t_0}) v_{t_0}^w + n^{-1} \tau_{t_0} + \lambda^{-1} (1 - \lambda) \tau \right)^{1-\sigma} - \tilde{\lambda}^{-\sigma} \left(\pi (v_{t_0}^w)^{1-\sigma} + (1 - \pi) (n^{-1})^{1-\sigma} \right) \quad (28)$$

$$G^p (rk_{t_0})^{\sigma-1} \geq \lambda^{-\sigma} \left((1 - \tau_{t_0}) v_{t_0}^p + n^{-1} \tau_{t_0} + \lambda^{-1} (1 - \lambda) \tau \right)^{1-\sigma} - \tilde{\lambda}^{-\sigma} \left((1 - \pi') (n^{-1})^{1-\sigma} + \pi' (v_{t_0}^p)^{1-\sigma} \right) \quad (29)$$

Note that for $\{\tau\} = \{\tau_{t_0}, 0, 0, \dots\}$, $\tilde{\lambda} = \lambda = 1 - \beta^{\frac{1}{\sigma}} r^{\frac{1-\sigma}{\sigma}} < 1$. If $v^w = v^p = n^{-1}$, the tax sequence $\{\tau\} = \{\tau_{t_0}, 0, 0, \dots\}$ is sustainable if

$$G^w (rk_{t_0})^{\sigma-1} \geq \lambda^{-\sigma} \left((n^{-1})^{1-\sigma} - 1 \right) (v^w)^{1-\sigma}$$

$$G^p (rk_{t_0})^{\sigma-1} \geq \lambda^{-\sigma} \left((n^{-1})^{1-\sigma} - 1 \right) (v^p)^{1-\sigma}$$

But $\left((n^{-1})^{1-\sigma} - 1 \right) < 0$ if $\sigma > 1$, so the inequalities above hold strictly for any $k_{t_0} \geq 0$. It follows by continuity that the constraints hold if v^w and v^p are sufficiently close to n^{-1} .

If $\pi = 1 - \pi' = 0$, that is if a revolt never succeeds, for $\{\tau\} = \{\tau_{t_0}, 0, 0, \dots\}$

$$G^w (rk_{t_0})^{\sigma-1} \geq \lambda^{-\sigma} \left(((1 - \tau_{t_0}) v_{t_0}^w + n^{-1} \tau_{t_0})^{1-\sigma} - (n^{-1})^{1-\sigma} \right) \quad (30)$$

$$G^p (rk_{t_0})^{\sigma-1} \geq \lambda^{-\sigma} \left(((1 - \tau_{t_0}) v_{t_0}^p + n^{-1} \tau_{t_0})^{1-\sigma} - (v_{t_0}^p)^{1-\sigma} \right) \quad (31)$$

For $\sigma > 1$, the inequalities above will hold for any k_{t_0} , since the right sides of 30 and 31 are negative if $v^w > v^p > n^{-1}$. Furthermore the inequalities will continue to hold after period t_0 because the capital stock grows and the shares remain constant at $v^j = v_{t_0+s}^j = (1 - \tau_{t_0}) v_{t_0}^j + \tau_{t_0} n^{-1}$ for $s = 1, 2, \dots$ and $j = w, p$, with $v^p \leq n^{-1}$, and $v^w \geq n^{-1}$, so that the right sides of 30 and 31, with v^w and v^p replacing $v_{t_0}^w$ and $v_{t_0}^p$ are still negative. Thus democracy will be sustainable for any k_{t_0} under taxes $\{\tau\} = \{\tau_{t_0}, 0, 0, \dots\}$. Therefore, if π and $1 - \pi'$ are sufficiently close to 0, the set of tax sequences for which democracy is sustainable is not empty. Of course the median voter will then choose from the feasible set the tax sequence that maximizes his utility. ■

Corollary 2: Let $\{\tau_{t_0}, \tau_{t_0+1}, \tau_{t_0+2}, \dots\}$ be a tax sequence with $\tau_{t_0} < 1$ for which democracy is sustainable from initial stock k_{t_0} , and let $\{\tau'_{t_0}, \tau'_{t_0+1}, \tau'_{t_0+2}, \dots\}$, $\tau'_{t_0} > \tau_{t_0}$, be a "more redistributive" sequence that is not sustainable from initial stock k_{t_0} . Then democracy is sustainable with the more redistributive tax sequence $\{\tau'_{t_0}, \tau'_{t_0+1}, \tau'_{t_0+2}, \dots\}$ for some $k'(\tau'_{t_0}) > k_{t_0}$.

Proof. By hypothesis for $\{\tau_{t_0}, \tau_{t_0+1}, \tau_{t_0+2}, \dots\}$, the incentive constraints for period t_0 ,

$$\begin{aligned} G^w (rk_{t_0+s})^{\sigma-1} &\geq \lambda_{t_0+s}^{-\sigma} \left((1 - \tau_{t_0+s}) v_{t_0+s}^w + n^{-1} \tau_{t_0+s} + n^{-1} (\lambda_{t_0+s}^{-1} - 1) \tau_{t_0+s} \right)^{1-\sigma} \\ &\quad - \tilde{\lambda}^{-\sigma} \left(\pi (v_{t_0+s}^w)^{1-\sigma} + (1 - \pi) (n^{-1})^{1-\sigma} \right) \end{aligned} \quad (32)$$

$$\begin{aligned} G^p (rk_{t_0+s})^{\sigma-1} &\geq \lambda_{t_0+s}^{-\sigma} \left((1 - \tau_{t_0+s}) v_{t_0+s}^p + n^{-1} \tau_{t_0+s} + n^{-1} (\lambda_{t_0+s}^{-1} - 1) \tau_{t_0+s} \right)^{1-\sigma} \\ &\quad - \tilde{\lambda}^{-\sigma} \left((1 - \pi') (n^{-1})^{1-\sigma} + \pi' (v_{t_0+s}^p)^{1-\sigma} \right) \end{aligned} \quad (33)$$

are satisfied for $s = 0, 1, \dots$. Under assumptions 1 and 2, $0 < \lambda^m \leq \lambda_{t_0} \leq \lambda^M < 1$, and the right hand sides of 32 and 33 are bounded. Suppose the period t_0 incentive constraint is not satisfied for $\{\tau'_{t_0}, \tau'_{t_0+1}, \tau'_{t_0+2}, \dots\}$. Since $G^p (rk_{t_0})^{\sigma-1}$ is increasing in k_{t_0} , there is a $k'_{t_0}(\tau'_{t_0}) > k_{t_0}$ such that the period t_0 constraint holds. Suppose instead that the period $t_0 + s$ incentive constraint is not satisfied for $\{\tau'_{t_0}, \tau'_{t_0+1}, \tau'_{t_0+2}, \dots\}$. Since λ_s and g_s depend only on $\{\tau_{t_0+1}, \tau_{t_0+2}, \dots\}$,

they are unaffected by τ_{t_0} . Thus $k_{t_0+s} = \left(\prod_{j=t_0+1}^{\tau+s-1} g_j \right) k_{t_0}$ is strictly increasing in

k_{t_0} . Then again there is a $k_{t_0}^{s'}(\tau'_{t_0}) > k_{t_0}$ such that the period $t_0 + s$ constraint is satisfied. Let $S = \{s \mid \text{the period } t_0 + s \text{ constraint is not satisfied for } \{\tau'_{t_0}, \tau_{t_0+1}, \tau_{t_0+2}, \dots\}\}$. Let $k_{t_0}^{s'}(\tau') = \sup_{s \in S} \{k_{t_0}^{s'}\}$. Then all the incentive constraints are satisfied for initial stock $k_{t_0}^{s'}(\tau')$. ■

Theorem 2: 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. If there is a coalition of the rich and poor that can make the poor better off than under the tax scheme chosen by the median voter, while making the rich no worse off, then 14-16 holds, and we show that this leads to a contradiction.

From 14,

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

and from 16,

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

Rearranging these yields:

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

and

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

so that:

$$\begin{aligned}
& \left(\frac{\left(\left((1 - \tau_{t_0}) v_{t_0}^{pM} + n^{-1} \tau_{t_0}^p + n^{-1} \sum_{j=t_0+1}^{\infty} \tau_j^p \prod_{s=t_0+1}^j g_s(r(1 - \tau_s^p))^{-1} \right) r k_{t_0} \right)}{\left(\left((1 - \tau_{t_0}) v_{t_0}^{pM} + n^{-1} \tau_{t_0}^M + n^{-1} \sum_{j=t_0+1}^{\infty} \tau_j^M \prod_{s=t_0+1}^j g_s(r(1 - \tau_s^M))^{-1} \right) r k_{t_0} \right)} \right) \\
&= \left(\frac{\left(\left((1 - \tau_{t_0}) v_{t_0}^{wM} + n^{-1} \tau_{t_0}^p + n^{-1} \sum_{j=t_0+1}^{\infty} \tau_j^p \prod_{s=t_0+1}^j g_s(r(1 - \tau_s^p))^{-1} \right) r k_{t_0} \right)}{\left(\left((1 - \tau_{t_0}) v_{t_0}^{wM} + n^{-1} \tau_{t_0}^M + n^{-1} \sum_{j=t_0+1}^{\infty} \tau_j^M \prod_{s=t_0+1}^j g_s(r(1 - \tau_s^M))^{-1} \right) r k_{t_0} \right)} \right) \tag{34}
\end{aligned}$$

We also have:

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

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

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

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

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

Evaluating the derivatives at $v_{t_0}^{pM}$ implies:

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

or

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

Similarly, evaluating the derivatives at $v_{t_0}^{wM}$ implies

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

or

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

This is only possible if

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

But, from equation 34,

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

which leads to a contradiction. ■

6 References

- Benhabib, Jess and Rustichini, Aldo. "Social Conflict and Growth." *Journal of Economic Growth*, 1996, 1, pp.25-142.
- Benhabib, Jess. and Rustichini, Aldo, "Optimal Taxes Without Commitment," , *Journal of Economic Theory*, 77, 1997, 231-259.
- Benhabib, Jess and Przeworski, Adam. "The Political Economy of Redistribution under Democracy," working paper, 2003, <http://www.econ.nyu.edu/user/benhabib/dg103.pdf>
- Bertola, Giuseppe. "Factor Shares and Savings in Endogenous Growth." *American Economic Review*, 1993, 83(5), pp. 1184-1198.
- Besley, Timothy and Stephen Coate. "Sources of Inefficiency in a Representative Democracy: A Dynamic Analysis." *American Economic Review*, 1998, 88, pp. 139-156.
- Bhalla, Surjit S. "Freedom and Economic Growth: A Virtuous Circle?" Paper presented at the Nobel Symposium Democracy's Victory and Crisis, Uppsala University, August 27-30, 1994.
- Chari, V. V. and Kehoe, Patrick. "Optimal Monetary and Fiscal Policy." *Handbook for Macroeconomics*, eds. John Taylor and Michael Woodford. New York: North-Holland, 1999, volume 1C, pp. 1671-1745.
- Coleman, Jules. "Rationality and the Justification of Democracy," in Geoffrey Brennan and Loren E. Lomasky, eds., *Politics and Process*, Cambridge: Cambridge University Press, 1989.
- Coser, Lewis. *The Functions of Social Conflict*. Glencoe, IL: Free Press, 1956.
- Dasgupta, Partha, *An Inquiry into Well-Being and Destitution*, Oxford: Oxford University Press, 1993.
- Deininger, Klaus and Lyn Squire. "A New Data Set Measuring Income Inequality." *World Bank Economic Review*, 1996, 10(3), pp. 565-91.
- Easterly, William and Ross Levine. "Africa's Growth Tragedy: Policies and Ethnic Divisions." *Quarterly Journal of Economics*, 1997, 112(November), pp. 1203-50.
- Eckstein, Harry. *Division and Cohesion in Democracy: A Study of Norway*. Princeton, NJ: Princeton University Press, 1966.
- Gans, Joshua S. and Smart, Michael. "Majority Voting with Single-Crossing Preferences." *Journal of Public Economics*, 1996, 59, pp. 219-237.
- Lamounier, Bolivar. "Notes on the Study of Re-democratization," Latin American Working Paper No. 58, Wilson Center, Washington, DC, 1979.
- Lipset, Seymour Martin. *Political Man: The Social Bases of Politics*. Garden City, NJ: Doubleday, 1960.
- Kim, Wonik. *The Political Economy of Unemployment Insurance*. Ph.D. Dissertation, Department of Politics, New York University, 2004.
- Mill, John Stuart. *Representations on Representative Government*. Buffalo, NY: Prometheus, 1991 [18610].
- Przeworski, Adam and Limongi, Fernando. "Modernization: Theories and Facts." *World Politics*, 1977, 49, pp.

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

Przeworski, Adam. "Economic Development and Transitions to Democracy." Ms. Department of Politics, New York University, 2004.

Ross, Edward A. *The Principles of Sociology*. New York, 1920.

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

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