Democracy and the Dynamics of Income Distribution

Ignacio Ortuño Ortín*
Departamento de Economía.
Universidad Carlos III

December, 2013

Abstract. We consider an overlapping generation model with income redistribution. The income of an agent depends on his level of human capital which depends on the educational resources their parents spent on him, on his parents human capital and on his innate abilities. The income tax rate at each generation is the result of a democratic political process. We analyze the dynamics of the after tax income and human capital distributions. The degree of inequality of those distributions might evolve in a non-monotonic way. Thus, democracy does not always bring monomotonic reductions in inequality over time.

1. Introduction

The dynamics of income distribution has always been a very important topic in economics. The interest on it has increased in recent years due to the raise in inequality experienced in several advanced economies during the past decades (Atkinson 2003, Atkinson and Voitchovsky 2011, Atkinson et al, 2011). Several explanations have been proposed to understand this phenomenon, among which are those based on economic globalization and technological changes. However, the possible role played by democratic institutions in the dynamics of the income distribution has not been sufficiently studied. Most work on democracy and income distribution focus on comparative studies between countries and the possible effect of different types of institutions (for example, proportional representation versus majority systems), and rarely discusses the dynamics across several generations.

The main objective of this paper is to develop a theoretical model that helps to understand how democratic processes determine the evolution of the distribution of income across several generations. The empirical observation that our model is trying to explain is the U-shape shown by inequality during the second half of the twentieth century in some democracies (see Atkinson 2003, Atkinson et al. 2011), or the N-shape if we also consider the first part of that century. So, we develop a model where the level of redistribution is determined democratically and the level of inequality across generations can have a non-monotonic behavior. Although we think that actually the dynamics of inequality across generations is a very complex issue and cannot be explained by a single factor (see Brandolini and Smeeding 2008), our model can bring new ideas to help us better understand this subject.

We consider an overlapping generation model with incomplete markets and only labor as production factor. The income of an agent depends on his level of human capital which depends on the educational resources their parents spent on him, on his parents human capital (a family externality) and on his innate abilities. There is pure redistribution and

*I wish to thank Salvador Ortigueira, John E. Roemer, Alejandro Corvalán and Carlos Ponce for helpful comments. I acknowledge the financial help from the Spanish MEC through grant ECO2010-19596.

the income tax rate at each generation is the result of a democratic political process. We analyze the dynamics of the after tax income and human capital distributions. We show that this type of "democratic" economic environment does not always bring non-monotonic reductions in inequality over time. Thus, we identified an additional reason to those in the existing literature that helps us understand the empirical evidence observed in some advanced economies. The intuition behind our result is simple: Suppose that in a first generation inequality is high. Unless the democratic system gives too much weight to the rich, the corresponding tax policy will be very redistributive, and as a consequence the distribution of resources spent on education quite equal. It follows that inequality in the next generation will be low, which makes for a less redistributive tax policy, and a more unequal distribution of resources spent on education. Thus, in the third generation inequality goes up which implies a more redistributive tax policy and so on. We show that this non-monotonic dynamics is more likely to occur the lower is the influence of the family externality and the greater the effect of the educational resources in the acquisition of human capital.

Related Literature: The closest works to ours are those of Roemer (2005) and Benabou (2002, 2005). Roemer (2005) analyzes the long-run income distribution under democracy using an economy similar in many aspects to the one in this paper[1].The fundamental differences with our approach are that i) he focuses on the long term, i.e. in the steady state, rather than in the dynamics and ii) he does not include "innate abilities" as an input in the formation of human capital. It is because of this second difference that his model cannot generate the type of non-monotonic income distribution dynamics that we obtain here.

Our economic model is very much based on Benabou (2005) from whom we draw the way to model the formation of human capital, the functional form of several distributions and the specific type of admissible income redistribution policies. We, however, make some simplifications and changes to be able to obtain results on the dynamics of the model. As in the case of Roemer (2005), the main difference with our work is that Benabou (2005) does not study the dynamics but only the steady state equilibrium. In particular he analyzes the possibility of several steady state equilibria each of them showing a different relation between inequality and redistribution (see also Alesina and Angeletos 2005).

Ichino et al. (2011) consider an overlapping generation economy in which, as in our model, in each period there are fathers and sons. Fathers choose the policy redistribution for their sons and choose investment to produce son's talent. As in the previous cited work they don not consider the short run dynamics, i.e. the dynamics of the first two or three generations, but only the steady state equilibrium.

Other papers where only the long run or the steady state equilibrium is analyzed are Hassler et al (2003), Krusell et al. (1997) and Azzimonti et al (2006). Hassler et al (2005), however, focus on the long-run level of redistribution as well as in the redistribution dynamics, as we do here. Their economic model is quite different from ours and in particular it does not consider fathers and sons and the process of human capital formation. Interestingly, they also get, like us, a non-monotonic dynamics of the income distribution. However, their result is generated by a mechanism completely different from ours. In their case the oscillations are generated by the economy, and the political system plays a moderating role of these oscillations, the opposite that in our case where it is

[1] Roemer (2001) develops a dynamic model of political cycles based on Hirschman(1982) where political moods change from a private orientation to a public orientation. However, political cycles in such model are generated in a very ad hoc way.

the political system which generates oscillations in the distribution of income. Moreover, in their model redistribution policies show persistence, i.e. redistribution today implies that more voters tomorrow will benefit from redistribution, just the opposite from the conclusion obtained here.

2. The basic model

The economy is modeled according to Benabou (2002, 2005) and Roemer (2005), albeit with certain modifications that allow us to obtain analytical results. We consider an overlapping generations economy with a continuum of agents at each period $t = 1, 2, 3, ...$ Each agent lives for two periods. In the first period the agent is a child and in the second period is an adult. As a child an agent neither works nor consumes. As an adult the agent works, consumes and invests in the education of his child. We identify the level of education of an agent with his human capital. It is assumed that a child born in generation t acquires a level of human capital h_{t+1}

$$h_{t+1} = k\varepsilon r^c h_t^b \tag{1}$$

where h_t is the human capital of his father, r is the amount of education inputs received, $c > 0, b > 0, k > 0$ are constants and ε is his level of innate ability. The distribution of innate ability among agents is and i.i.d. random variable[2] and $\log(\varepsilon)$ is distributed according to a normal $N(\frac{-s^2}{2}, s^2)$, so that $E[\varepsilon] = 1$. At time $t = 0$ the distribution of human capital h among adult agents is given by a lognormal with parameters m_0 and Δ_0, i.e. $\log(h) \sim N(m_0, \Delta_0)$. The realization of ε takes place at the beginning of the adult period. Thus, when an adult invest r in the education of his child he does not know the realization of ε. It will turn out that in all subsequent periods $t = 1, 2, ...,$the distribution of h will follow a lognormal as well. Thus, we write the distribution function of human capital at time t as $F(h; m_t, \Delta_t)$ where m_t denotes the mean and Δ_t the variance of $\log(h)$.

The income obtained by an adult agent born in generation $t-1$ is equal to his level of human capital h_t[3]. An adult agent cares about his consumption, q_t, and the (expected) human capital of his child, h_{t+1}. We represent his preferences by

$$U(q_t, h_{t+1}) = q_t^{\theta} h_{t+1}^{1-\theta} \tag{2}$$

where $0 < \theta < 1$.

The consumption level q_t is equal to the after tax income minus the resources spent on the education of his child. We are assuming that there is only private education. Our results can be easily generalized to the case where there is also public education. Taxes are used to make income transfers among adults. At each generation t society has to choose the type of income transfer policy to be implemented that period. Let μ denote the mean income in society. We follow Benabou (2002, 2005) and assume that the after-tax income, $\widehat{h}$, for an (adult) agent with pre-tax income h is given by[4]

$$\widehat{h} = w(\tau)\widetilde{h}^{\tau} h^{1-\tau} \tag{3}$$

[2] Having this random variable is essential for our results. Romer(2005) does not include it and therefore can not generate the kind of dynamics that we obtain here.

[3] Our results remain true if, as in Benabou (2005), we assume that this income also depends on effort and on a random variable capturing market luck.

[4] Here we have removed the subscripts. Whenever it is clear we will not use subscripts.

where the break-even human capital level $\widetilde{h}$ is determined by

$$\int_0^\infty \widetilde{h}^\tau h^{1-\tau} dF(h; m, \Delta) = \mu \tag{4}$$

Thus, τ indicates the degree of progressiveness of the income redistribution scheme. If $\tau = 0$ there is no redistribution at all. When $\tau = 1$ all agents end up with the same after tax income. Taxation creates a distortion and the term $w(\tau)$ captures the economic waste associated to it. It is assumed that $0 \leq w(\tau) \leq 1$, $w(0) = 1$ and $w'(\tau) < 0$. This term $w(\tau)$ does not appear in Benabou (2002, 2005) where the distortionary effect of taxation is obtained by considering an elastic labor supply. We adopt this ad hoc specification for simplicity. In the Appendix we show that our numerical results remain true under a more general model with elastic labour supply as the one in Benabou(2005).

Given a redistribution policy τ an agent with human capital h spends the fraction γ of his after-tax income in the education of his child. Thus, consumption is $q = (1-\gamma)\widehat{h}$ and his optimal γ maximizes utility $U(q, h_s)$, where h_s denotes the human capital of his son. I.e.,γ is the solution to

$$\max_\gamma \left((1-\gamma)\widehat{h}\right)^\theta \left(E[\varepsilon]k\left(\gamma\widehat{h}\right)^c h^b\right)^{1-\theta} \tag{5}$$

It is easy to see that the solution to (5) does not depend on h and $E[\varepsilon]$ and is given by $\gamma^* = \frac{c(1-\theta)}{\theta + c(1-\theta)}$. It will be convenient to write $\theta + c(1-\theta) \equiv \eta$ and $\gamma^* = \frac{c(1-\theta)}{\eta}$ From now on, we will drop the asterisk from γ^*.

Notice that if $F(h; m, \Delta)$ is a lognormal distribution we have that

$$\int_0^\infty h^a dF(h; m, \Delta) = \exp(am + a^2\frac{\Delta^2}{2}) \tag{6}$$

From (4) and (6) we obtain

$$\widetilde{h}^\tau = \exp(m\tau + \frac{\Delta^2}{2}(1-(1-\tau)^2)) \tag{7}$$

Thus, using (7) and (3) the indirect utility function for an adult agent with human capital h can be written as,

$$\begin{aligned} u(\tau; h) \quad \equiv \quad & \left((1-\gamma)\, w(\tau) \exp(m\tau + \frac{\Delta^2}{2}(1-(1-\tau)^2))h^{1-\tau}\right)^\theta \\ & \left(\left(\gamma w(\tau) \exp(m\tau + \frac{\Delta^2}{2}(1-(1-\tau)^2))h^{1-\tau}\right)^c h^b\right)^{1-\theta} \end{aligned} \tag{8}$$

Dropping the terms that do not depend on τ and writing $y \equiv \log(h)$, we write the log of the indirect utility function (8) as

$$v(\tau; y) \equiv \log(w(\tau)) + m\tau - \frac{\Delta^2}{2}(1-\tau)^2 - \tau y \tag{9}$$

2.1. The Political System. The tax policy chosen at each period will depend on the distribution of human capital among adult agents and on the specific type of political institution adopted by that society. In this paper we deal with several political models that are commonly used by economists. We will consider first the "median voter" case. The median voter approach is justified when the political system consists of two parties competing for office. Parties propose redistribution policies and adult agents vote for the party that propose the policy that renders them with the highest after-tax income. The objective of each of the parties is to maximize the probability of winning the election. At equilibrium both parties propose the same policy which coincides with the optimal one for the agent with the median level of human capital. However, some scholars see such approach as a poor description of the political process that takes place in democratic societies. The median voter approach predicts that the amount of redistribution is increasing on the degree of inequality (Meltzer 1981). Empirical evidence generally does not support this approach (Alesina and Glaeser 2004, Benabou 1996). In recent years many theories have been proposed to understand this empirical evidence. For example, it might be the case that agents do not have fully selfish preferences (Alesina and Giulano 2009) or that not all groups have the same political influence (Karabourbonis 2010). Other authors defend the idea that social mobility, rather than the level of inequality, is what explains the level of redistribution (Benabou and Ok 2001), and other emphasize the role of ethnic diversity on redistribution (Alesina and Glaeser 2004, Desmet 2009) . The type of electoral rule might also be important, and countries with proportional representation present more redistribution (Person et al 2000, 2007). Some authors (see Roemer 2000) argue that political parties care about being in office as well as about the implemented policy. Thus, parties are ideologic, i.e. they have preferences on the set of possible policies. Under this alternative approach, parties do not need to propose the same policy. This makes for a much more complex and (perhaps) realistic political process. The price to pay, though, is the tractability of the model. In most of the paper we will adopt the median voter approach so that the main ideas can be develop in a very simple setting. The results, however, will be robust to the adoption of a political process with ideological parties. In any case, at this point we only need to assume that at each time t the implemented redistribution policy τ_t depends on $F(h; m_t, \Delta_t)$ and on the political institution, which is denoted by Ψ. Since the distribution of human capital will remain log-normal over time we can write $\tau_t = T(m_t, \Delta_t, \Psi)$ so that T denotes the mapping from the set of human capital distributions and political institutions into the set of redistribution policies.

We are assuming that the policy τ is chosen by the adult agents at the beginning of each period and remains constant during the length of the period. In our model a period coincides with the life span of an adult agent. Since in most democracies general elections take place every four or five years the assumption of an election per period is an strong one. The idea is that it takes time to change the tax system in a significant way. Moreover, we can see the outcome of an election in our model as the average outcome of the several elections that take place in a generation.

2.2. Dynamics. In this section we analyze the dynamics of the distribution of human capital. It is assumed that at each generation t, a tax policy $\tau_t = T(m_t, \Delta_t, \Psi)$ is chosen and implemented. Recall that the share of (after-tax) income spent in the education of a child is given by $\gamma = \frac{c(1-\theta)}{\eta}$. Notice that γ does not depend on h and it remains constant over time. However, the level of after-tax income does depend on the implemented tax policy and, as a consequence, the level of resources spent on the education of a child also depends on the tax policy. Thus, the distribution of human capital among agents in a given period depends on their parents' human capital distribution and on the tax policy implemented in their parent's generation.

Using (3) and the fact that $r = \gamma\widehat{h}_t$ the law of motion given by (1) becomes

$$h_{t+1} = k\varepsilon\,(\gamma w(\tau_t)\,\widetilde{h}_t^{\tau_t})^c h_t^{(1-\tau_t)c+b} \tag{10}$$

and substituting (7) in it we obtain

$$h_{t+1} = k\varepsilon\;\;(\gamma w(\tau_t)\;\;\exp(m\tau_t + \frac{\Delta_t^2}{2}(1-(1-\tau_t)^2))^{\tau_t})^c\;\,h_t^{(1-\tau_t)c+b} \tag{11}$$

that can be written as

$$h_{t+1} = C\varepsilon h_t^{(1-\tau_t)c+b} \tag{12}$$

where the term $C = k\,(\gamma w(\tau_t)\exp(m\tau_t + \frac{\Delta_t^2}{2}(1-(1-\tau_t)^2))^{\tau_t})^c$ is the same for all agents. Taking the log of both sides of equation (12) we have

$$y_{t+1} = \log C + \log\varepsilon + ((1-\tau_t)\,c + b)y_t \tag{13}$$

Since the original distribution of human capital, $F(h; m_0, \Delta_0)$, and the distribution of ε are log-normal, and h and ε are independent random variables, equation (13) implies that for any period t the distribution of human capital remains log-normal and Δ^2, the variance of $y \equiv \log(h)$, follows the dynamic process

$$\Delta_{t+1}^2 = ((1-\tau_t)\,c + b)^2\,\Delta_t^2 + s^2 \tag{14}$$

In a similar way, one can easily show that the mean of y follows the process

$$m_{t+1} = (c+b)\,m_t + \log\left(k\left(\frac{c(1-\theta)}{\eta}\right)^c\right) + c\log\left(w(\tau_t)\right) + c\frac{\Delta_t^2}{2}\left(1-(1-\tau_t)^2\right) - \frac{s^2}{2} \tag{15}$$

Since we are interested on distributional issues an index of inequality has to be specified. The coefficient of variation is quite often used to measure income or wealth inequality. In our case all the distributions are log-normal and such coefficient is $CV = \sqrt{\exp(\Delta^2) - 1}$. Thus, income inequality is measure by Δ^2 and our objective is to analyze the dynamic process of Δ_t^2 (or Δ_t) given by the system of simultaneous equations (15) and (14). Notice that since the tax policy τ_t is given by $T(m_t, \Delta_t, \Psi)$, the evolution of Δ_{t+1}^2 depends, in principle, on Δ_t^2 and on m_t. The implications of the chosen policy τ on the rate of economic growth and on the evolution of the income distribution can be easily analyze using the previous equations. The relationship between inequality and growth is not the objective of our paper and we remit the reader to Benabou (2002, 2005) who provides an extensive analysis on it, even though in a more general framework.

From equation (14) we can see that the inequality in human capital in a period, measured by the variance, is a proportion of the previous period inequality plus the "innate" inequality represented by ε. Thus, the tern $(1-\tau_t)\,c + b$ measures the intergenerational persistence of human capital and, as indicated by Benabou, it can be seen as an inverse measure of social mobility. If the society adopts a leassez fair policy we have $\tau_t = 0$ so that $\Delta_{t+1}^2 = (c+b)^2\,\Delta_t^2 + s^2$. In the opposite case, the full redistribution policy $\tau_t = \frac{c+b}{c}$ is implemented and we have $\Delta_{t+1}^2 = s^2$ so that inequality is exclusively due to the differences in the innate ability term.

The inequality level associated to a steady state of this dynamic process is given by $\Delta^2 = \frac{s^2}{1-((1-\tau)c+b)^2}$. Of course, nothing guarantees yet that the process converges to a steady state solution. To obtain convergence we will need, in addition to restrictions on the parameters b and c, that the function $T(m_t, \Delta_t, \Psi)$ behaves nicely. At this point, however, there is no much we can say about the function $T(m_t, \Delta_t, \Psi)$,which depends on the political institution Ψ. In any case, we should mention here that the goal in this paper, and contrary to Benabou (2002, 2005), is not the study of the steady state solutions. It might take several generations to reach such steady state and therefore the dynamic path could be as important as the final state.

3. THE MEDIAN VOTER SOLUTION

It has already being explained under which assumptions on the political process the median voter solution can be seen as a good prediction of the implemented redistribution policy. In this case, at each period the median voter's optimal policy is implemented. We first need to show that such a median voter is well defined in our model. Thus, we need to check that the preferences of the agents are single-peaked. We will work with the specific case of tax distortion given by

$$w(\tau) = 1 - z\tau, \tag{16}$$

where $0 < z \leq 1$.

Proposition 1. *At any period the preferences of the (adult) agents over redistribution policies are single-peaked. The optimal redistribution policy τ for the "median voter", i.e. the agent with human capital $y = m$, is given by*

$$\Gamma(\Delta) \equiv \frac{\Delta(1+z) - \sqrt{\Delta^2(1-z)^2 + 4z^2}}{2\Delta z}$$

The proof of this and following propositions are provided in the Appendix.

Thus, the optimal redistribution policy τ for the median voter only depends on the variance Δ and not on the mean m. In this case, equation (14) is enough to understand the evolution of the human capital inequality when politics is captured by the median voter solution**.** To analyze such dynamics of the inequality index Δ_t we substitute the expression $\Gamma(\Delta)$ in Proposition 1 into equation (14) to obtain

$$\begin{aligned} \Delta_{t+1} &= \sqrt{\left((1-\Gamma(\Delta_t))\,c+b\right)^2 \Delta_t^2 + s^2} \\ &= \sqrt{\left(\left(1 - \frac{\Delta_t(1+z) - \sqrt{\Delta_t^2(1-z)^2 + 4z^2}}{2\Delta_t z}\right)c + b\right)^2 \Delta_t^2 + s^2} \end{aligned} \tag{17}$$

We will restrict ourselves to cases for which there exists an stable steady state. Since we want to show that the standard view under which democracy brings a monotonic convergency to a steady state is not always appropriate, this is not an strong assumption. If the dynamic process described by (17) does not converge to a steady state the inequality index Δ_t is either always increasing without a bound or increasing in some periods and decreasing in others.

We say that the dynamics of inequality is *monotonic* if we have either $\Delta_t \leq \Delta_{t+1} \leq \Delta_{t+2} \leq ...$ or $\Delta_t \geq \Delta_{t+1} \geq \Delta_{t+2} \geq ...$It is not difficult to see that a sufficient condition to observe a non-monotonic dynamics of Δ_t is that the function $g(\Delta) = \left((1-\Gamma(\Delta))\,c+b\right)^2 \Delta^2 + s^2$ has negative slope, see Figure 1.

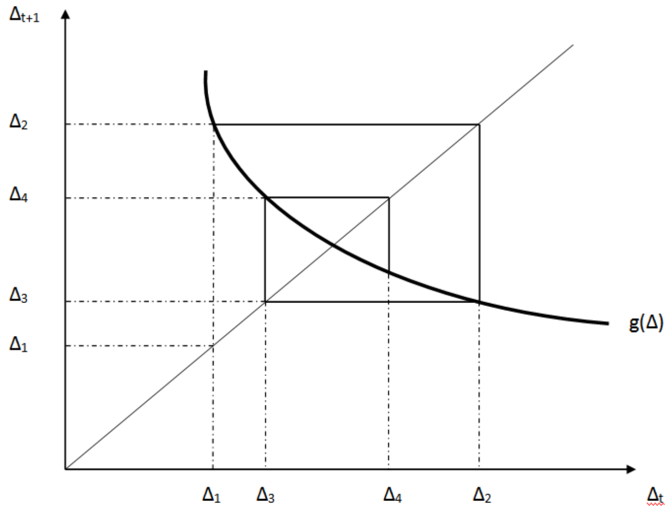

Figure 1: Example of g'(Δ)<0 and non-monotonic dynamics of Δ_t

Proposition 2. *If for all t we have $b < \frac{c}{2z}\left(\frac{-\Delta_t(1-z)^2}{\sqrt{\Delta_t^2(1-z)^2+4z^2}} + 1 - z\right)$ then $g'(\Delta_t) < 0$, and the dynamics of Δ_t is non-monotonic.*

It is clear that our restrictions on the parameters b, c and z do not exclude the possibility that the inequality in the proposition holds. Such inequality is more likely to hold the greater the parameter c is and the smaller the parameter b is[5]. Recall that b measures the externality of the parents human capital and c measures the efficiency of the education resources on the son's human capital production function. When $g'(\Delta_t) < 0$ the dynamics of Δ_t might be characterized by "over-shooting" in the inequality index. For small values of Δ_t the median voter's optimal tax rate is low and as a consequence the value of Δ_{t+1} is high. On the other hand, when Δ_t is high the median voter's optimal tax policy is very redistributive and the inequality in the next generation, Δ_{t+1} , will be low. Notice that a progressive redistribution policy implies a more equal distribution of consumption and a more equal distribution of resources spent on the education of children. Thus, it is intuitive that "over-shooting" is more likely to happen when b is small and c is large.

Table 1 shows an example of "over-shooting" for realistic values of the parameters[6]. The last column gives, V, the variance of the log after-tax income. Since the share of income invested in education is the same for all families the variance V also indicates the variance in log consumption. The index of inverse measure of social mobility, $b + c(1-\tau)$, goes from 0.39 in the first generation to 0.61 in the second generation, which are values within the range of the empirical estimations[7] (see Benabou 2002, Black and Devereux

[5] In particular, if $b = 0$ we always have $g'(\Delta_t) < 0$.

[6] The Mathematica program used to obtain this example and the ones in the following sections is available upon request.

[7] Interestingly, this index of social mobility does not change monotonically. See Nybom and Stuhler (2013) for a model where this also occurs, albeit for different reasons unrelated to the democratic process analyzed here.

2011).

Median Voter Solution

	Δ	τ	V
$t=1$	1.2	0.629	0.445
$t=2$	0.847	0.351	0.549
$t=3$	0.875	0.382	0.540
$t=4$	0.872	0.379	0.541
steady state	0.87195	0.37968	0.54088
$z=0.4, b=0.1, c=0.8, s=0.7$			

Table 1

In the first generation the inequality is high, $\Delta_1 = 1.2$, the corresponding tax policy is very progressive, $\tau_1 = 0.629$, and as a consequence the distribution of resources spent on education is quite equal. It follows that the inequality in the next generation is quite low, $\Delta_2 = 0.84\dot{7}$, which makes for a less progressive tax policy, $\tau_2 = 0.351$ and a more unequal distribution of resources spent on education. Thus, in the third generation inequality goes up to $\Delta_3 = 0.875$ which implies a more progressive tax policy and so on.

The previous example is robust to changes of the values of the parameters as long as b remains relative low and c high. If, for example, we take $b = 0.2$ there is no longer over-shooting and the variance converges monotonically to the steady state value. However, the aim of the paper is just to show the logic behind the possibility of over-shooting, i.e. of a non-monotonic dynamic.

Providing a more general model would require adding new features as, for example, elastic labor supply which would make the analysis less transparent. The economic model in Benabou (2005) provides a more general technologic specification, in particular, the income, m, of an adult in generation t is given by

$$m_t = \sigma_t \; h_t^{\gamma} \; l_t^{\delta} \tag{18}$$

where l_t is effort, σ_t is a productivity shock and h_t denotes, as in this paper, the level of human capital. Providing effort is costly and at the moment of choosing the tax policy τ the agent does not know the realization of the productivity shock σ_t. Thus, the tax policy plays a twofold role in this model; as a redistribution mechanism and as an insurance policy. The cost of adopting this more general approach is that we cannot obtain simple analytical results, like the one in Proposition 2, which clarify the conditions for a non-monotonic dynamics of inequality. However it is important to know whether our results depend critically on the restrictive assumptions used in our model. To this end we provide in the Appendix a numerical example of non-monotonic dynamics for the technology specified in Benabou(2005). Thus, our result is not due to the simplicity of the model used here, and the same kind of result is obtained with a more general model.

4. Ideological Party Competition

One might wonder whether the over-shooting result obtained in the previous section depends critically on the median-voter approach adopted there. Thus, our next task is to show that a similar result can be obtained for more realistic political institutions. In this section we analyze a two ideological party competition model so that the political equilibrium will be different from the one in the median voter model. There are several possible modelling choices with respect to the following four main features of party competition (see Roemer 2000): i) the way party ideology is determined; ii) the information parties have about voters' preferences; iii) the possibility that parties make credible commitments on policy and iv) whether the policy implemented coincides with the platform announced by the winning party or not. We will assume that the ideology of the parties is exogenously given, parties have perfect information about the distribution of voters' preferences but can't make credible policy commitments, and the implemented policy is a weighted combination of the two platforms announced by the parties. All these assumptions have been used in other papers on political competition (see for example Alesina Rosenthal 1995, Gomberg et al 2000) and here we adopted them for simplicity. The main technical problem arising when the median voter approach is abandoned is the difficulty of proving existence of equilibrium of the game played by parties and the almost impossibility of obtaining explicit solutions[8] (see Roemer 2001). Hence, we are trying to model ideological party competition in the most simple way but still yielding an equilibrium outcome different from the one in the median voter approach. In any case, the same type of results provided in this section can be obtained with more complex models. For example, the ideology of the parties can be made endogenous as in Baron (1993) and Gomberg at al. (2004), the parties might have imperfect information about voters so that the election outcome is not deterministic anymore and the implemented policy coincides with the (credible) platform announced by the winning party[9].

We describe now the details of the political institution analyzed in this section. There are two parties L and R. Parties have preferences on the set of possible tax policies. The preferences of party L (M) coincide with the preferences of an agent with human capital level in the p^l (p^r) percentile of the distribution. This percentile remains the same over time and it is known by all agents. Recall that at each period t the distribution of y is lognormal and denoted by $F(h; m_t, \Delta_t)$. This implies that the distribution of $y = \log h$ is normal and we denote it by $F_y(y; m_t, \Delta_t)$. Thus, at period t the preferences of L are given by the preferences of agent y_t^l where

$$\begin{aligned} p^l &= F_y(y_t^l; m_t, \Delta_t) \\ &= \Phi(\frac{y_t^l - m_t}{\Delta_t}) \end{aligned} \tag{19}$$

and where Φ denotes the standard Normal distribution. Let $M^l = \Phi^{-1}(p^l)$, then from (19) we have

$$\frac{y_t^l - m_t}{\Delta_t} = M^l \tag{20}$$

In a similar way for party R we have

$$\frac{y_t^r - m_t}{\Delta_t} = M^r \tag{21}$$

[8] In the probabilistic voting model of Coughlin (1992) and Lindbeck and Weibull (1987) equilibrium exists under very general conditions, but the two parties propose the same policy, and in this section we want to consider cases where that does not happen.

[9] Numerical results for this type of model with endogenous formation of ideologies are available upon request.

Thus, once the distribution of y has been standardized, M^i is the value of the agent represented by party i.

At the beginning of each period t party L proposes policy τ_t^l and party R proposes policy τ_t^r. Agents vote for the policy they like the most according to their preferences $v(\tau; y)$. The outcome of the election will determine, in a way explained later on, the implemented policy for that period. Since agents know the ideology of the parties and there is not possibility of making credible policy commitments we assume that the proposals of parties coincide with their ideal policy. From the first order conditions $\frac{\partial v(\tau;y^l)}{\partial \tau} = 0$ we can compute the ideal policy for party i as the solution to

$$\frac{w'(\tau)}{w(\tau)} + m + \Delta^2(1-\tau) - y^i = 0$$

Continuing with the specific function $w(\tau) = 1 - z\tau$ and using (20) this first order condition can be written as,

$$-z + \Delta^2(1-\tau)(1-z\tau) - M^i\Delta(1-z\tau) = 0 \tag{22}$$

Notice that the solution to (22) does not depend on m, so that the optimal policy depends only on Δ. We write the solution to (22) as $\Gamma(\Delta; M^i)$ and it is easy to check that

$$\Gamma(\Delta; M^i) = \frac{1+z}{2z} + \frac{-M^i z - \sqrt{\left(\Delta\left(1 - z + zM^i\right)^2 + z^2 4\right)}}{2\Delta z}, i = l, r \tag{23}$$

The median agent has $M = 0$ and in that case the policy $\Gamma(\Delta; 0)$ coincides with the ideal policy of the median agent calculated in the previous section and given by $\Gamma(\Delta)$ in (39).

Thus, at period t, party L proposes $\tau_t^l = \Gamma(\Delta_t; M^l)$ and party R proposes $\tau_t^r = \Gamma(\Delta_t; M^r)$. A voter of type y votes for policy τ_t^l whenever $v(\tau_t^l; y) \geq v(\tau_t^r; y)$. Let the fraction of agents voting for L at period t be denoted by V_t. We assume that the implemented policy at t is given by

$$\tau_t = V_t\tau_t^l + (1 - V_t)\tau_t^r \tag{24}$$

Thus, the implemented policy can be seen as a political compromise between the two proposals. The justification for this approach can be found in, for example, Alesina-Rosenthal (1995), and Ortuño-Ortín (1997). It is important to note that our results do not depend on this assumption on the implemented policy. We could alternatively assume that parties can announce credible policy proposals -so that τ_t^i doesn't need to be equal to $\Gamma(\Delta; M^i)$- and the party that obtains most votes implements its announced policy. However, both parties will propose, at equilibrium, the ideal policy of the median voter. In order to obtain a result different from the median voter solution it would be needed to introduce uncertainty about the distribution of voters. In that case, the outcome of the election is not deterministic and the implemented policy in one period would be the average or expected winning policy. Since we want to keep the model as simple as possible the "political compromise approach" of equation (24) seems more convenient[10].

In the previous section we showed that agents have single peaked preferences. Hence, given the two proposals τ_t^l and τ_t^r there will be a type y_t' who is indifferent between those

[10]Since in our model a period coincides with the life span of an adult agent, actually many elections take place during such a period. Say that in each of those election party L wins with a probability given by its expected vote share V. Then, equation (24) could be seen as an average of the implemented tax policies during those periods.

two proposals, and all the agents with $y < y_t'$ prefer τ_t^l and all the agents with $y > y_t'$ prefer τ_t^r. Thus, the indifferent type is such that $v(\tau_t^l; y_t') = v(\tau_t^r; y_t')$, and using (9) we can solve for y_t'

$$y_t' = m_t + \frac{\log\left[\frac{w(\tau_t^r)}{w(\tau_t^l)}\right] + \frac{\Delta_t^2}{2}\left(\left(1-\tau_t^l\right)^2 - (1-\tau_t^r)^2\right)}{\tau_t^r - \tau_t^l} \tag{25}$$

The share of agents voting for party L is $V_t = F_y\left(y_t'; m_{t,}\Delta_t\right)$. Since F_y is a normal distribution we have $F_y\left(y_t'; m_{t,}\Delta_t\right) = \Phi(\frac{y_t'-m_t}{\Delta_t})$, and by (25) we have

$$\begin{aligned} V_t &= \Phi(\frac{y_t' - m_t}{\Delta_t}) \\ &= \Phi\left(\frac{\log\left[\frac{w(\tau_t^r)}{w(\tau_t^l)}\right] + \frac{\Delta_t^2}{2}\left(\left(1-\tau_t^l\right)^2 - (1-\tau_t^r)^2\right)}{\Delta_t(\tau_t^r - \tau_t^l)}\right) \end{aligned} \tag{26}$$

and substituting τ_t^l and τ_t^r in (26) for the expressions in (23) we get

$$V_t = V_t(\Delta_t) \equiv \Phi\left(\frac{\log\left[\frac{w(\Gamma(\Delta_t;M^l))}{w(\Gamma(\Delta_t;M^r))}\right] + \frac{\Delta_t^2}{2}\left(\left(1-\Gamma(\Delta_t;M^l)\right)^2 - (1-\Gamma(\Delta_t;M^r))^2\right)}{\Delta_t(\Gamma(\Delta_t;M^r) - \Gamma(\Delta_t;M^l))}\right) \tag{27}$$

Notice that the only variable in (27) is the variance Δ_t , so that the share of votes obtained by party L, $V_t(\Delta_t)$, is a function of just Δ_t .Hence, the implemented policy at period t depends on Δ_t and is given by

$$\tau_t = \tau(\Delta_t) \equiv V_t(\Delta_t)\Gamma(\Delta_t; M^l) + (1 - V_t(\Delta_t))\Gamma(\Delta_t; M^r) \tag{28}$$

From equation (14) and (28) the variance of y follows now the process

$$\Delta_{t+1} = \sqrt{\left((1-\tau(\Delta_t))\, c + b\right)^2 \Delta_t^2 + s^2} \tag{29}$$

A sufficient condition to obtain over-shooting in the dynamics of Δ_t is that the slope of $\sqrt{\left((1-\tau(\Delta_t))\, c + b\right)^2 \Delta_t^2 + s^2}$ be negative. Contrary to the previous section, here we cannot get an analytical expression for $\tau(\Delta_t)$ so that it is not possible to sign such slope. Thus, we have to rely on numerical simulations. It turns out to be quite easy to generate realistic and robust examples showing over-shooting. Table 2 shows the result for the same economy as the one in Table 1 and with the Left party representing and agent with human capital in the 40th percentile and the Right party representing an agent in the 60th percentile. As in the median voter case, we can only provide this type of result if the parameter b is low and c is large.

Ideological Party Competition

	Δ	τ	V
$t = 1$	1.2	0.658	0.408
$t = 2$	0.830	0.295	0.585
$t = 3$	0.891	0.376	0.555
$t = 4$	0.880	0.362	0.561
steady state	0.882	0.365	0.560
$z = 0.4, b = 0.1, c = 0.8, s = 0.7$			
$p^l = 40\%, p^r = 60\%$			

Table 2

Thus, the possibility of non-monotonic income distribution dynamics is not exclusive of the median voter approach and it can be observed under more general models of democracy.

5. Final comments

We have developed a simple overlapping generation model where the income and the human capital of an agent depend on the educational resources their parents spent on him, on his parents human capital and on his innate abilities. There is redistribution at each generation which is the result of a democratic political process. We show that democracy does not always bring non-monotonic reductions in inequality over time. We left for future work the development of a more complete model that includes capital and a more realistic demography in which coexist simultaneously many generations, not just parents of the same age and children of the same age. It would also be important to analyze the model's predictions on the growth rate, and if the empirical evidence is consistent with these predictions.

References

[1] Alesina, A. and H. Rosenthal (1995), *Partisan Politics, Divided Government, and the Economy*, New York: Cambridge University Press.

[2] Alesina, A. and H. Rosenthal (1996), "A theory of divided government," *Econometrica* **64**, 1311-1343.

[3] Alesina, A., and E. L. Glaeser (2004). *Fighting Poverty in the U.S. and Europe: A World of Difference*. New York: Oxford University Press.

[4] Alesina, A. and G. Angeletos, (2005), "Fairness and Redistribution," *American Economic Review*, **95**, 960-980.

[5] Azzimonti, M., de Francisco, E. and P. Krusell, (2006),"Median-voter Equilibria in the Neoclassical Growth Model under Aggregation," *Scand. J. of Economics*, **108**, 587-606.

[6] Atkinson, A.B. (2003), "Income Inequality in OECD countries: Data and Explanations," *CESifo* WP 881.

[7] Atkinson, A.B. and S. Voitchovsky, (2011), "The Distribution of Top Earnings in the UK since the Second War World," *Economica*, **78**, 440-459.

[8] Atkinson, A.B., Piketty, T., and E. Saez, (2011), "Top Incomes in the Long Run of History," *Journal of Economic Literature*, 49(1), 3-71.

[9] Baron, D.P., (1993), "Government formation and endogenous parties," *American Political Science Review*, **87**, 34-47.

[10] Benabou, R. (1996), “Inequality and Growth”, *NBER Macroeconomics Annual* , B. Bernanke and J. Rotemberg, eds., 11-74.

[11] Benabou, R. and E. A. Ok, (2001), “Social Mobility and the Demand for Redistribution: The POUM Hypothesis”, *Quarterly Journal of Economics*, **116**, 447-487.

[12] Benabou, R. (2002), "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?", *Econometrica* **70** (2), 481-517.

[13] Benabou, R. (2005), "Inequality, Technology and the Social Contract," in *Handbook of Economic Growth*, vol1 B, edited by P. Aghion and S. Durlafau, Elsevier B.V. DOI: 10.1016/S1574-0684(05)01025-7

[14] Black, S.E. , and P. Devereux, (2011), "Recent Developments in Intergenerational Mobility," in *Handbook of Labor Economics*, ed. by O. Ashenfelter and D. Card, vol 4A. Elsevier.

[15] Coughlin, Peter J., (1992), *Probabilisitc Voting Theory*, New York: Cambridge University press.

[16] Desmet, K., Ortuño-Ortín, I. and S. Weber, (2009), "Linguisitc Diversity and Redistribution", *Journal of the European Economic Association*, **7**(6): 1291-1318.

[17] Gomberg, A., Marhuenda, F. and I. Ortuño-Ortín, (2004), "A Model of Endogenous Political Platforms," *Economic Theory* **24**, 373-394.

[18] Hassler, J., J.V: Rodríguez-Mora, Storesletten, K. and F. Zilibotti, (2003),"The Survival of the Welfare State," *American Economic Review*, **93**, 87-112.

[19] Hassler, J., Krusell, P., Storesletten, K. and F. Zilibotti, (2005), "The Dynamics of Government," *Journal of Monetary Economics,* **52**, 1331-1358.

[20] Hirshman, A. (1982), *Shifting Involvements: Private Interest and Public Action*, Princeton: Princeton University Press.

[21] Ichino, A., Karabarbounis, L. and E. Moretti, (2011), "The Political Economy of Intergenerational Income Mobility," *Economic Inquiry*, **49 (1)**, 47-69.

[22] Krusell, P., Quadrini, V. and J. Rios-Rull, (1997), *Journal of Economic Dynamic and Control*, **21**, 243-272.

[23] Lindbeck, A. and J. Weibull, (1987), "Balanced Budget Redistribution as the Outcome of Political Competition," *Public Choice* 52, 273-297.

[24] Meltzer, A., and S. F. Richard (1981), “A Rational Theory of the Size of the Government,” *Journal of Political Economy* **89**, 914 927.

[25] Nybom, M. and J. Stuhler, (2013), "Interpreting Trends in Intergenerational Mobility", manuscript, UCL.

[26] Ortuño-Ortin, I. (1997), "A spatial model of political competition and proportional representation," *Social Choice and Welfare* **14**, 427-38.

[27] Persson, T., Roland, G., and G. Tabellini, (2000), "Comparative Politics and Public Finance", *Journal of Political Economy* **108**, 1121-1161.

[28] Persson, T., Roland, G. and G. Tabellini (2007). "Electoral Rules and Government Spending in Parliamentary Democracies." *Quarterly Journal of Political Science* **2** (2): 155–88.

[29] Roemer, J.E. (2001), *Political Competition: A Theory with Applications to the Distribution of Income*, Cambridge: Harvard University Press.

[30] Roemer, J.E. (2005), *Democracy, Education, and Inequality.* Econometric Society Monographs; n. 40, Cambridge University Press, New York.

[31] Wittman, D.A. (1977), "Candidates with Policy Preferences: A Dynamic Model," *Journal of Economic Theory* 14, 180-189.

6. Appendix

6.1. Proof Proposition1. We first prove that agents have single-peaked preferences. From (2) we have

$$\frac{\partial v(\tau;y)}{\partial \tau} = \frac{w'(\tau)}{w(\tau)} + m + \Delta^2(1-\tau) - y \tag{30}$$

and

$$\frac{\partial\left(\frac{\partial v(\tau;y)}{\partial \tau}\right)}{\partial \tau} = \frac{d\left(\frac{w'(\tau)}{w(\tau)}\right)}{d\tau} - \Delta^2 \tag{31}$$

If the expression in (31) is negative the preferences are single peaked. Hence, a sufficient condition for single-peaked preferences, and the existence of the median voter, is that

$$\frac{d\left(\frac{w'(\tau)}{w(\tau)}\right)}{d\tau} \leq 0 \tag{32}$$

We have that $w(\tau) = 1 - z\tau$, where $0 < z \leq 1$. In this case $\frac{w'(\tau)}{w(\tau)} = \frac{-z}{1-z\tau}$ and

$$\frac{d\left(\frac{w'(\tau)}{w(\tau)}\right)}{d\tau} = \frac{-z^2}{()^2} < 0$$

so that (32) is satisfied.

We next compute the optimal policy for the median voter. Notice that the median voter is identified by the median value of the log human capital, so that he is the agent with $y = m$. The optimal policy for an agent with human capital y must satisfy the first order condition $\frac{\partial v(\tau;y)}{\partial \tau} = 0$, i.e.

$$\frac{w'(\tau)}{w(\tau)} + m + \Delta^2\ (1-\tau) - y = 0 \tag{33}$$

for the median type agent $y = m$ and this condition becomes

$$-\frac{w'(\tau)}{w(\tau)} = \Delta^2\ (1-\tau) \tag{34}$$

Notice that (34) doesn't depend on m. Hence, the optimal policy for the median voter is the solution to

$$\frac{z}{1-z\ \tau} = \Delta^2(1-\tau)$$

which is equivalent to

$$z = (1-z\ \tau)(\Delta^2 - \Delta^2\tau)$$

or

$$\begin{aligned} z &= \Delta^2 - z\ \tau\ \Delta^2 - \Delta^2\tau + z\ \Delta^2\tau^2 \Longrightarrow \\ 0 &= \tau^2\Delta^2\ z + \tau(-z\ \Delta^2 - \Delta^2) + \Delta^2 - z \end{aligned}$$

or

$$\tau^2\Delta^2 z - \tau\ \Delta^2(1+z) + \Delta^2 - z = 0 \tag{35}$$

And the solution to (35) is

$$\tau = \frac{1}{2z}\frac{\Delta + \Delta\ z \pm \sqrt{(\Delta^2 - 2\Delta^2\ z + \Delta^2\ z^2 + 4z^2)}}{\Delta} \tag{36}$$

We take the negative root. To see that this is correct consider the second order condition

$$2\tau\ \Delta^2\ z - (1+z)\ \Delta^2 < 0 \Longleftrightarrow \tau < \frac{1+z}{2\ z} \tag{37}$$

If in (36) we take the positive root we have

$$\begin{aligned} \tau &= \frac{1}{2z}\frac{\Delta + \Delta\ z + \sqrt{(\Delta^2 - 2\ \Delta^2\ z + \Delta^2\ z^2 + 4z^2)}}{\Delta} \\ &> \frac{\Delta(1+z) + \sqrt{\Delta^2\ (1-z)^2}}{2\ \Delta\ z} = \frac{1}{z} \end{aligned} \tag{38}$$

from equations (38) and (37) follows that

$$\frac{1}{z} < \frac{1+z}{2z}$$

which is true iff $2z < z + z^2$. This inequality holds iff $z < z^2$, i.e. it holds when $z > 1$ which is a contradiction with the assumption that $z \leq 1$.

Summing up, the optimal policy for the median voter is given by the negative solution in (36), which we can rewrite as

$$\Gamma(\Delta) \equiv \frac{\Delta(1+z) - \sqrt{\Delta^2(1-z)^2 + 4z^2}}{2\ \Delta\ z} \tag{39}$$

6.2. Proof of Proposition 2. It follows from equation (17) that a sufficient condition to observe a non-monotonic (convergence to the steady state) is that the function $g(\Delta) = \left((1-\Gamma(\Delta))\,c+b\right)^2\Delta^2+s^2$ has negative slope. We have

$$g'(\Delta) = -2((1-\Gamma(\Delta))\ c+b)\ c\ \Gamma'(\Delta)\ \Delta^2 + ((1-\Gamma(\Delta))\ c+b)^2\ 2\ \Delta < 0$$

which is equivalent to

$$g'(\Delta)<0 \Longleftrightarrow ((1-\Gamma(\Delta))\ c+b)^2 < ((1-\Gamma(\Delta))\ c+b)\ c\ \Gamma'(\Delta)\ \Delta$$

or

$$g'(\Delta)<0 \Longleftrightarrow (1-\Gamma(\Delta))\ \ c+b < c\ \Delta\ \Gamma'(\Delta) \tag{40}$$

Recall that $\Gamma(\Delta) = \frac{\Delta(1+z)-\sqrt{\Delta^2(1-z)^2+4z^2}}{2\Delta z}$. It is easy to check that

$$\Gamma'(\Delta) = \frac{2\ z}{\Delta^2\sqrt{\Delta^2(1-z)^2+4z^2}}$$

and substituting for $\Gamma(\Delta)$ and $\Gamma'(\Delta)$ in (40) we get

$$\begin{aligned} g'(\Delta) &< 0 \\ &\Leftrightarrow \\ (1-\frac{\Delta\ (1+z)-\sqrt{\Delta^2(1-z)^2+4z^2}}{2\ \Delta\ z})\ c+b &< c\frac{\Delta\ 2\ z}{\Delta^2\sqrt{\Delta^2(1-z)^2+4z^2}} \end{aligned} \tag{41}$$

The last inequality in (41) can be written as

$$b<c\ (\frac{2\ z}{\Delta\ \sqrt{\Delta^2(1-z)^2+4z^2}} - 1 + \frac{\Delta\ (1+z)-\sqrt{\Delta^2\ (1-z)^2+4z^2}}{2\ \Delta\ z} \tag{42}$$

which is equivalent to

$$b<c\ (-\frac{\Delta^2(1-z)^2+4z^2+\Delta(-1+z)\sqrt{\Delta^2(1-z)^2+4z^2}-4z^2}{2\ \Delta\ z\ \sqrt{\Delta^2(1-z)^2+4z^2}})$$

and simplifying

$$g'(\Delta)<0 \Leftrightarrow b < \frac{c}{2z}\left(\frac{-\Delta(1-z)^2}{\sqrt{\Delta^2(1-z)^2+4z^2}}+1-z\right) \tag{43}$$

6.3. Numerical example.. Here we provide a numerical example of non-monotonic dynamics for the economic specification in Benabou (2005). The income, m, of an adult agent in generation t is given by

$$m_t = \sigma_t h_t^{\gamma} l_t^{\delta} \tag{44}$$

where l_t is effort, σ_t is a productivity shock and h_t denotes, as in this paper, his level of human capital. The productive shock σ is a i.i.d. random variable and $\log(\sigma) \sim N(\frac{-v^2}{2}, v^2)$. The human capital h_t is given by equation (1) in the text. In a similar way to our equation (14) disposable income is given by $\widehat{m} = \widetilde{m}^{\tau} m^{1-\tau}$. The budget constraint is $\widehat{m}_t = c_t + r_t$ where c_t is consumption and r_t is investment in the education of the child. An agent chooses his effort and consumption (when he knows his productivity) to maximize the function $(1-\rho)(\ln c_t - l_t^{\eta}) + \rho \ln E_t[h_{t+1}]$. The ex-ante preferences (before he knows his productivity) is given by $U_t = \ln(E_t[V_t^{1-a} \mid h_t]^{1/(1-a)})$. Proposition 2 in Benabou (2005) shows that the evolution of Δ_t across generations is given by

$$\Delta_{t+1}^2 = \left((1-\tau_t)\, c\gamma + b\right)^2 \Delta_t^2 + c^2(1-\tau_t)v^2 + s^2$$

where τ_t is the rate of redistribution chosen by the decisive voter in the political process. Benabou (2005) assumes that such decisive voter is located at the $100 \times p$ percentile of the human capital distribution, where $p > 1/2$. Thus, the chosen rate of redistribution is the optimal one for an agent richer than the median. The numerical example here is for the case $p = 1/2$ so that we continue with the same "median voter" approach as in the text. However, similar type of examples can be easily generated for $p > 1/2$.

The following table summarizes the result of our specific numerical example. This non-monotonic dynamics is robust to small changes of any parameter value.

Economic model as in Benabou

	Δ	τ
$t = 1$	1.2	0.569
$t = 2$	0.836	0.462
$t = 3$	0.877	0.476
$t = 4$	0.870	0.474
steady state	0.872	0.474

$\gamma = 0.5; \delta = 0.5; b = 0.01; c = 0.9; v = 1;$
$w = 0.7; \eta = 4; \rho = 0.1; a = 0.015$