The Evolution of Ideology, Fairness and Redistribution*

Alberto Alesina[†], Guido Cozzi[‡], and Noemi Mantovan[§]

November 3, 2011

Abstract

Ideas about what is "fair" influence preferences for redistribution. We study the dynamic evolution of different economies in which redistributive policies, perception of fairness, inequality and growth are jointly determined. We show how including beliefs about fairness can keep two otherwise identical countries in different development paths for a very long time. We show how different initial conditions regarding how "fair" is the same level of inequality can lead to two permanently different steady states. We also explore how bequest taxation can be an efficient way of redistributing wealth to correct "unfair" past accumulation of inequality.

1 Introduction

The poor want to tax the rich, but that is not all what determines redistributive policies. Views about what is "fair" and what is an acceptable level of inequality matters, above and beyond the individuals' position in the income ladder[1]. The same level of inequality may be more or less acceptable depending upon different beliefs about how wealth has been accumulated, namely by virtue of effort and ability or, on the contrary, by virtue of luck, connections or corruption. In one word, it matters for individual preferences whether different levels of income and wealth are "deserved" or not, as confirmed by extensive empirical evidence.[2] These views about inequality and justice (which we label "ideology") determine tax rates and the evolution of the distribution of income and wealth.

*We are grateful to participants in 2010 NBER Summer Institute, and especially Pierre Yared. The editor and an anonymous referee offered excellent comments.

† *Harvard University and IGIER Bocconi*

‡ *Durham University*

§ *Bangor University*

[1] See for instance the recent survey of preferences for redistribution by Alesina and Giuliano (2010) and the references cited therein. Alesina, Di Tella and McCulloch (2004) discuss different levels of inequality tolerance in various countries. Alesina and Glaeser (2004) focus on a comparison between Continental Europe and US. Persson and Tabellini (2000) provide an excellent overview of politico economic models of redistributive policies.

[2] See Alesina and Giuliano (2010) for a survey of the literature on this point.

The latter generates changes in the proportion of wealth inequality due to effort or to other factors including luck and government policies. As a result of these developments, individuals' preferences for redistribution evolve accordingly.

In this paper we provide a politico economic model that can trace over time the evolution of polices (income taxes, wealth taxes, and transfer schemes), the evolution of inequality, and of the preferences for redistribution, as a function of the changes in what individuals perceive as fair and unfair wealth differences. Generations of voters are linked by bequests; thus redistributive policies in the past, and past beliefs about what was "fair" inequality matter for today's initial conditions and therefore for today's' voters' preferences. We focus upon several issues. The first one is how different initial conditions regarding fairness lead to two different steady state. We consider two economies with the same initial level of inequality. In one all the inequality is "unfair", while in the other inequality is all "fair". We show that even though all the observable characteristics of the two countries are identical (including initial inequality), the two countries may converge to two different steady states. We can think of the initial period as the first one in which tax policy is decided democratically, and the initial conditions those determined by different pre-democratization histories. For instance a history of feudalism and rigid class divisions in Europe versus a more merit based and more recent American capitalism. This result formalizes the role of the perception of poverty and social justice as an explanation of differences in welfare polices in the two sides of the Atlantic as discussed informally in Alesina and Glaeser (2004). Second, we explore how a temporary shock to preferences, for instance a change in what is viewed as "fair" inequality, may lead to long lasting dynamic consequences. Third, we study how different forms of taxation, income taxation versus bequest taxation, interact with the perceptions about fairness. Under certain conditions, a high level of bequest taxation is superior to income taxation in order to reduce the perceived level of "unfairness" in inequality. This result is closely related to the issue of equalizing opportunities at birth.

This paper is related to the work of Alesina and Angeletos (2005a). However there are many differences. Fist of all, in the present paper we use a probabilistic voting model rather than median voter model as they use. This allows us to be more flexible and to analyze various types of redistributive schemes, including multidimensional policies. Second, and related to the previous point, we can study in depth bequest taxation in addition to income taxation. This is crucial since the issue of equalizing opportunity at birth can be central in a discussion about fairness . Third, given our rich dynamic specification we can address issues of "shocks" to preferences and to other parameters of the model, contrary to the static model by Alesina and Angeletos (2005a).[3] This allows us to make statements about the endogenous evolution of ideology and tax policy. In addition, our framework is sufficiently flexible that it could incorporate Piketty's (1995) intra-dynasties evolution of heterogeneous beliefs about

[3] For additional results along the same lines see Alesina and Angeletos (2005b). For a comment see Di Tella and Dubra (2010).

the incentive costs of redistribution. Another possible extension would be to incorporate Bénabou and Tirole's (2006) important point, namely allowing beliefs to be shaped not only by the actual data, but also by agents' psychological needs and objectives.

The present paper is organized as follows. Section 2 describes the model. Section 3 illustrates the dynamic evolution of the model and reports the most representative simulation results. Section 4 extends to wealth taxation as distinguished from income taxation. The last section concludes. The Matlab codes used in the present paper are available from the authors upon request.

2 The economy

We consider non overlapping generations of individuals, indexed by t. The size of the population is constant, there is one active individual per-family, and the total mass of families is normalized to one. Each individual, indexed by $i \in [0,1]$, lives for one period and has a certain level of endurance to effort, $\beta_i > 0$, luck, $\eta_i \in R$, and inner abilities, $A_i > 0$; average luck is zero, that is $\int_0^1 \eta_i di = 0$. These family-specific variables are assumed to be fully persistent over time but our results are robust if we allow for non persistent luck[4]. Each individual i cares about consumption, c_{it}, and how much wealth to bequeath to the next generation, k_{it} - which we label "capital". Effort, e_{it}, on the job enters negatively in the utility function. All choice variables are constrained to be non-negative. The "private" utility function is:

$$u_{it} = \frac{1}{(1-\alpha)^{1-\alpha}\alpha^{\alpha}} c_{it}^{1-\alpha} k_{it}^{\alpha} - \frac{1}{2\beta_i} e_{it}^2, \tag{1}$$

$0 < \alpha < 1$. The end of life gross wealth is:

$$z_{it} = A_i e_{it} + \eta_i + k_{it-1}. \tag{2}$$

For simplicity, capital yields zero rate of return. Each generation votes on the proportional tax rate, τ_t, which is applied to end-of-life gross wealth z_{it}; tax revenues are redistributed lump sum to all individuals, and the government budget is always balanced. For now we impose that income and initial wealth taxes are the same; in Section 4 we allow for different tax rates on income and wealth. End of life post-tax and transfer wealth is:

$$w_{it} = (1-\tau_t) z_{it} + G_t, \tag{3}$$

where $G_t = \tau_t \int_0^1 z_{it} di$ is the percapita transfer. Individual income is $y_{it} =$

[4] See the working paper version of this paper, Alesina, Cozzi, and Mantovan (2009).

$(A_i e_{it} + \eta_i)(1 - \tau_t) - \tau_t k_{it-1} + G_t$, and the aggregate income of generation t is

$$Y_t = \int_0^1 \left[(A_i e_{it} + \eta_i)(1 - \tau_t) - \tau_t k_{it-1} + G_t\right] di = \int_0^1 A_i e_{it} di,$$

which is identical to percapita income due to the population normalization.

The warm glow intergenerational altruism implies that fraction α of end of life wealth is bequeathed, as seen by maximizing u_{it} subject to $c_{it} + k_{it} = w_{it}$. Therefore, plugging the optimal consumption and bequest into the private utility function, we obtain:

$$u_{it} = w_{it} - \frac{e_{it}^2}{2\beta_i}. \tag{4}$$

Individuals vote on the tax rate at the beginning of life, before deciding on effort. Maximizing u_{it}, using (4), (2), and (3), gives

$$e_{it} = (1 - \tau_t) A_i \beta_i,$$

which shows the distortion on the supply of effort induced by expected taxation; effort increases with the individual work ability and decreases in the disutility of effort[5].

The definition of a period needs discussion. In the model the period is one generation and it is also the length of time for which the redistributive policy cannot be changed. The choice of a "tax rate" should not be interpreted as the day to day or year to year changes in fiscal policy, but the broad redistributive stand of a certain period in a certain country. For instance more redistribution in the US with the Great Society in the Sixties, or with the New Deal in the Thirties, less redistribution starting with the Eighties in the US and England. In Europe an increase in redistribution at the end of the Sixties, possibly a slowing down today, etc. In the numerical simulations, we can generalize the model on this point and allow a distinction between the time span of the life of a generation and that of the choice of a tax rate.[6]

2.1 Inequality and fairness

In addition to the standard utility function described above, individuals care also about some measure of inequality. In our benchmark case we assume that individuals tolerate inequality arising from innate ability and effort, but are averse to inequality arising from everything else, luck and government policies. More specifically, let us define "fair" utility and wealth as follows:

$$\begin{aligned} \widehat{u}_{it} &= \widehat{w}_{it} - \frac{e_{it}^2}{2\beta_i}, \\ \widehat{w}_{it} &= A_i e_{it} + \widehat{k}_{it-1}. \end{aligned}$$

[5] As in Heckman (2008), we could distinguish between cognitive abilities (here summarized by A_i) and non-cognitive abilities ($1/\beta_i$).

[6] See the working paper version of the present paper for an extensive discussion of this point.

Since each agent chooses $k_{it} = \alpha w_{it}$, where α represents the generosity towards the next generation, we define fair consumption, fair bequest, and fair disposable wealth as:

$$\widehat{c}_{it} = (1-\alpha)\widehat{z}_{it}, \qquad \widehat{k}_{it} = \alpha\widehat{z}_{it}, \qquad \widehat{z}_{it} = \widehat{w}_{it} = A_i e_{it} + \widehat{k}_{it-1}. \tag{5}$$

U_{it}, is then defined as:

$$U_{it} = u_{it} - \gamma\Omega_t, \tag{6}$$

where

$$\Omega_t = \int_0^1 (u_{jt} - \widehat{u}_{jt})^2 dj = \int_0^1 (w_{jt} - \widehat{w}_{jt})^2 dj. \tag{7}$$

and $\gamma > 0$ is the parameter which measures the importance of fairness . This representation of utility implies that individuals dislike deviations from a distribution of wealth/utility in which everybody gets only the benefits from effort and innate ability. Note that the difference between total wealth and fair wealth is due to luck and government intervention with taxes and transfers. The higher the tax rate, the lower the equilibrium choice of effort; therefore the larger is the percentage of individual income due to luck rather than effort, and the larger the proportion of differences across individuals due to luck rather than effort.

2.2 The polity

We use a probabilistic voting model[7]. There are two parties - L , for "left", and R, for "right" - each of which simultaneously and credibly commits to a tax rate $\tau_P \in [0,1]$, $P = L$, R, at the beginning of each period. The individuals vote for a party at the beginning of their life, than they choose efforts. The party that obtained the majority of the votes is elected to office, and it implements the announced tax rate and redistributes accordingly. Finally, individuals choose their consumption and bequest. Individuals have heterogeneous degrees of party identification: the complete utility function is then:

$$\tilde{U}_{itP} = u_{it} - \gamma\Omega_t + (\sigma_{it} + \varepsilon_t)\chi_L(P), \text{ where } P = L, R.$$

P denotes the party that wins the election, and can be L or R. Indicator function $\chi_L(P)$ is 1 if $P = L$ and 0 if $P = R$. The random variable σ_{it} represents individual i's pro-party L ideological bias, while ε_t is an aggregate random variable capturing party L's popularity for time t. While we assumed (for simplicity) that individuals' pecuniary utility and luck shocks are fully persistent across

[7] Note that this voting model, due to Lindbeck and Weibull (1987,1993) does not require single peakness of preferences.

generations, political popularity may change from generation to generation both at the aggregate and at the family level. ε_t is uniformly distributed on support $\left[-\frac{1}{2\psi}, \frac{1}{2\psi}\right]$, and individual specific variables σ_{it} are uniformly distributed on support $\left[-\frac{1}{2\varphi_i}, \frac{1}{2\varphi_i}\right]$. All random variables are independent. Therefore, in the support of the corresponding distributions, the density function of aggregate popularity of party L is $\psi > 0$, and family-specific density functions are $\varphi_i > 0$, with the correlated (aggregate) component of the party identification assumed less variable than the individual components - that is $\psi > \varphi_i$, $\forall i \in [0,1]$. The two parties commit to their tax rates before they know the realization of the random variables ε_t and σ_{it}. They only care about winning the election, and hence choose their policies τ_t^L and τ_t^R by trying to maximize the probability of being elected, p_P, $P = L$, R.

The "popularity shocks" should not be viewed as the day ebbs and flows of electoral politics. Given our definition of a period as one generation these shocks should be seen as long term switches of one generation to the left (say the Sixties) or to the right, (say the Eighties in the US)[8].

2.3 Equilibrium

After simple substitutions, and momentarily neglecting the party L bias components, we obtain the indirect utility function of each individual in each generation. That function ultimately depends on exogenous parameters, on expected taxation, and on all the actual and fair wealth distribution of the previous generation:

$$
\begin{aligned}
U_{it} &= \left[\delta_i(1-\tau_t) + \eta_i + k_{it-1}\right](1-\tau_t) + \int_0^1 \left[\delta_j(1-\tau_t)\tau_t + \tau_t k_{jt-1}\right] dj - (1-\tau_t)^2 \frac{\delta_i}{2} \\
&- \gamma \int_0^1 \left[(\delta_s(1-\tau_t) + \eta_s + k_{st-1})(1-\tau_t) + \int_0^1 (\delta_j(1-\tau_t)\tau_t + \tau_t k_{jt-1}) dj - \delta_s(1-\tau_t) - \widehat{k}_{st-1}\right]^2 ds \qquad (8) \\
&\equiv \hat{U}_{it}(\tau_t).
\end{aligned}
$$

Where $\delta_i \equiv A_i^2 \beta_i$. The proof of Lemma 1 is in the Appendix.

Lemma 1. *In pairwise majority voting, there exists a unique equilibrium in which the two parties will select the same policy variable, $\tau_t^L = \tau_t^R \equiv \tau_t^*$, given by*

$$
\tau_t^* = \arg \max_{\tau_t \in [0,1]} \int_0^1 \varphi_i \hat{U}_{it}(\tau_t) di. \qquad (9)
$$

[8] See Song (2008) for an interesting model of political economy with persistent political ideology shocks.

The same equilibrium policy variable would also be chosen by a biased social planner who maximizes the following weighted aggregate welfare function:

$$W(\tau) \equiv \int_0^1 \varphi_i \hat{U}_{it}(\tau_t) di, \tag{10}$$

with each individual's indirect utility function (where effort, consumption, and bequest are all optimal) being weighted inversely to vulnerability, $1/\varphi_i$, to party-related attributes. In the special case in which individuals have the same densities $\varphi_i = \varphi$, Lemma 1 implies that $\tau_t^* = \arg\max_{\tau_t} W(\tau_t)$. Note that, from eq. (8), the equilibrium tax rate τ_t^* depends on generation $t-1$'s bequest distribution k_{t-1}, generation $t-1$'s fair bequest distribution $\widehat{k}_{t-1}$, and n the parameter vectors δ and η; that is $\tau_t^* = \tau^*(k_{t-1}, \widehat{k}_{t-1}, \delta, \eta)$.

Under the benchmark symmetry assumption, $\varphi_i = \varphi$, and normalizing by φ, we can simplify the relevant welfare function to:

$$\begin{aligned} W(\tau_t) &= \int_0^1 \hat{U}_{it}(\tau_t) di = \int_0^1 u_{it} di - \gamma \Omega_t = \frac{\bar{\delta}}{2}(1-\tau_t^2) + \bar{k}_{t-1} \\ &- \gamma \int_0^1 \left[-(\delta_s - \bar{\delta})(1-\tau_t)\tau_t + \eta_s(1-\tau_t) + (k_{st-1} - \bar{k}_{t-1})(1-\tau_t) - (\widehat{k}_{st-1} - \bar{k}_{t-1}) \right]^2 ds. \end{aligned}$$

The equilibrium tax rate τ_t^* determines the level of capital and fair capital for each family of the current generation. Therefore the link between different generations is summarized by the dynamics of k_{it} and $\widehat{k}_{it}$:[9]

$$k_{it} = \left[\delta_i(1-\tau_t) + \eta_i + k_{it-1}\right](1-\tau_t)\alpha + \alpha G_t \tag{11}$$

$$\widehat{k}_{it} = \alpha \delta_i (1-\tau_t) + \alpha \widehat{k}_{it-1}. \tag{12}$$

2.4 Discussion

Note that in eq. (12), "fair" bequests are obtained by removing from the parental end of life wealth, the effects of the "luck" variable, η_i, and of the taxes paid to and transfers received by the government. However, the indirect effect of tax rates on individual efforts is included in this definition of fairness. The reader may wonder why "$(1-\tau_t)$" should enter the "fair wealth": after all, it is an individually rational response to the distortion induced by taxation, and indeed $e_{it} = (1-\tau_t)A_i\beta_i$. If redistribution did not exist in the model, the individual would have exerted a first best effort level $e_{it}^F = A_i\beta_i$. We have also run

[9] Note that the distribution of δ_i should be high enough relative to the support of the distribution of η_i in order for end of life wealth never to be negative. See Lemma 2 in the Appendix for a sufficiency condition.

simulations under such a different view of fairness, based on "potential" rather than actual efforts, without much change in the results about the dynamics of k_{it}. By eq. (12), it simplifies the dynamics of $\widehat{k}_{it}$, which would tend to $\frac{\alpha\delta_i}{1-\alpha}$. However, the results of our computations do not change qualitatively.[10]

3 Intergenerational Dynamics

Starting from an initial vector of actual and fair wealth levels, $(k_{i0}, \hat{k}_{i0})_{i\in[0,1]}$, we use equations (8), problem (9), and eq.s (11) and (12), to iterate the model for an arbitrary number of generations, and calculate the sequence of equilibrium values of the endogenous variables of our dynamic economy for all parameter vectors, initial wealth distributions, and initial fair wealth distributions. By simulating the model for a sufficiently high number of generations, we can approximate the stable steady state value of the endogenous variables associated with each initial condition.

Generation t's pair of distributions $(k_{it}, \hat{k}_{it})_{i\in[0,1]}$ describe the interaction of real and "ideal" variables at time t. More precisely, the comparison between how society currently is - the actual distribution $(k_{it})_{i\in[0,1]}$ - and how society thinks it "should be" - the fair distribution $(\hat{k}_{it})_{i\in[0,1]}$ - sets the goals of the political action; together with the method of political competition - i.e. pairwise majority voting - this describes the political ideology prevailing for generation t in that economy. The resulting political equilibrium generates the evolution of $(k_{it+1}, \hat{k}_{it+1})_{i\in[0,1]}$, and therefore the political ideology (i.e. policy goals) prevailing in the next generation. Thus we trace the evolution of ideology, fairness and redistribution, as well as the aggregate GDP per capita.

3.1 US versus Europe: fair and unfair initial conditions

Define as time zero the first period in which tax policies are decided democratically with full participation.[11] Therefore, the initial distribution of "fair" and "unfair" wealth $(k_{i0}, \hat{k}_{i0})_{i\in[0,1]}$ is the result of previous history from $t = -\infty$ to 0. Before democratization, tax policy was chosen in political systems in which only a (wealthy) minority decided tax policy. In Europe, a history of feudalism and wealth related to nobility is clearly different from the history of the US,

[10] Another objection could be raised against purging additive luck η_i rather than both luck and ability A_i. Formally, luck enters additively while ability as the marginal product of effort: both could be viewed as "gifts of nature". Replacing A_i with $\bar{A} = \int_0^1 A_i di$ and using $\bar{A}e_{it} = \bar{A}(1-\tau_t)A_i\beta_i$ as the valued added component of the end-of-life wealth, however, would not change the qualitative results. In fact, individual ability, A_i, would still enter multiplicatively indirectly via optimal effort choice. Hence we can say that all the main qualitative results from the simulations are robust to the introduction of multiplicative luck, provided that also additive luck is present. Notice that, while in the previous case replacing A_i with its expected value in the direct abilities reduced the variance of δ_i (due to the elimination of the quadratic exponent on abilities), eliminating the variance of A_i completely could even increase the variance of δ_i.

[11] Historically democratization has been a long and gradual process, but for simplicity we abstract from this here.

where modern capitalism developed without a long previous history of privilege and rigid class separations. Thus for given total level of inequality at time zero (full democratization), the degree of "unfair" inequality can be safely assumed to have been higher in Europe.[12]

The first experiment looks only at the effect of different initial conditions, namely different conditions at the time of democratization. After democratization the two countries become identical. The "United States" would be represented by our Country A", while "Europe" would be Country B. The two countries have the exact same level of inequality at time zero, but in the US it is all "fair", whereas in Europe it is all "unfair". In other words, the "ideal" level of inequality in the US is the one actually observed, in Europe it is totally different, namely Europe would like to have perfect equality since all the inherited one is unfair. Obviously this is an extreme example which generalizes to different ratios of fair versus unfair wealth in the two places. In the Appendix we describe how a different social welfare function before democratization, and/or a different pattern of "luck" versus effort, may have led to different distributions of "fair" and "unfair" wealth at the time of democratization, which we label time zero. Here we examine the post-democratization dynamics.

In order to isolate fully the effect of initial conditions we have assumed that after the democratization the two countries are totally identical. They have the same wealth distribution, the same technology, the same preferences and the same process determining luck. As can be seen from Figure 1, the effect of these initial differences in "fairness" lead to two different steady states. Country A converges to a steady state characterized by low unfairness, low redistribution, and high GDP, whereas country B to a steady state characterized by high unfairness, high redistribution, and low GDP.

[12] We do not consider here issues of slavery and race relations in the US as a determinant of inequality and preferences for redistribution. See Alesina and Glaeser (2004) for an extensive discussion.

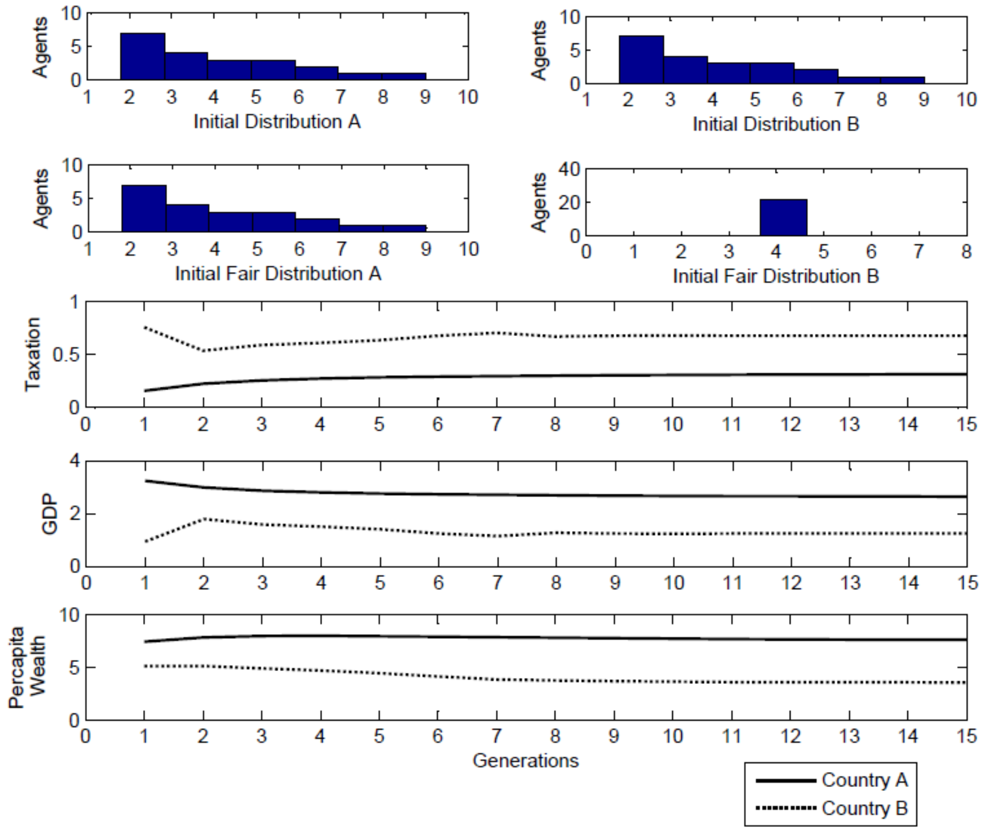

Figure 1

it

The intuition is simple. By striving to redistribute wealth in order to correct unfairness in the initial conditions generated by the past role of luck in individual success, Country B chooses higher tax rates, which discourages individual effort, thereby self-perpetuating luck as a source of income differences. Thus Country B reintroduces luck in an attempt to eliminate past accumulated "luck". As a consequence of the implied higher tax rates, its GDP never catches up with that of Country A. We have two steady sates only for some parameter values (but for a vast range of them). For other parameter values the two countries converge more or less slowly to the same steady state distribution of actual and fair wealth.

Higher values of the parameter γ, which represents the importance of fairness in society, and of the intergenerational generosity α, increase the possibility of having multiple steady states. In fact, if fairness is considered important, differences in the initial level of unfairness have stronger influence on tax policies: a country which starts with an unfair distribution of wealth implements higher tax rates. Similarly, the persistence of wealth and ideology depends highly on the intergenerational generosity parameter α. The higher parameter α the higher the influence of parental wealth and ideology on children. Therefore, different initial conditions are more likely to have long term influence, including to lead the economy to different steady states.

3.2 A temporary shock; the "Hippie" Generation

In this section, we study the effects of a one period unexpected ideological shock; we label it the "hippie" generation of the Sixties with its leftist turn. We can model this ideological shock in various ways, but one which seems appealing is as follows.[13] Imagine that the hippie generation becomes not only averse to "unfair inequality", as in our benchmark model, but becomes averse to inequality *per se*, as measured by the variance of end-of-life post-tax wealth. In other words the "hippies" turn against inequality even when generated by differences in hard work and ability. Thus they have preferences for redistribution in which:

$$\Omega_t = var(w_{it}).$$

The effects of this shock are shown in Figure 2:

[13] An alternative way would be an increase in γ. This can can be easily studied as well.

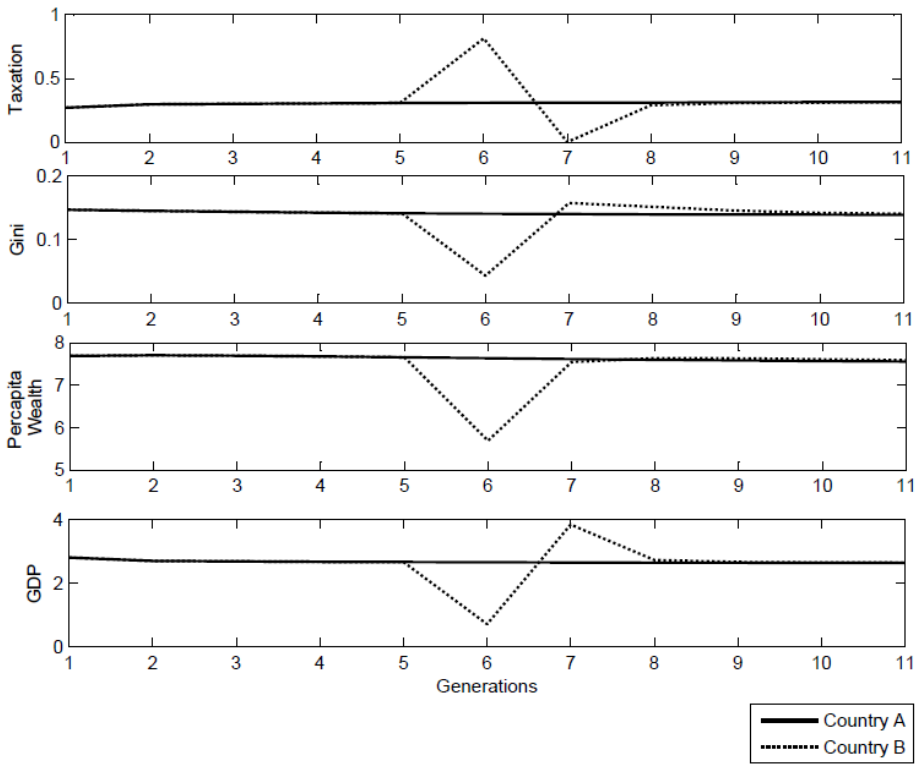

Figure 2

In order to evaluate the effect of this temporary shock, we consider two countries in the same initial steady state (defined by the model parameters). Country A illustrates the benchmark without the ideological shock, Country B has the shock. The "hippie" generation of Country B raises taxes, with a consequent immediate reduction of total inequality, measured by the Gini coefficient. This policy also reduces per capita GDP and wealth. What's interesting is that, though the egalitarian ideological shock lasts only one generation, the high level of redistribution reduces fairness due to the large government redistributive intervention which has not distinguished between luck and effort. When ideological preferences against unfairness revert back to normal, in order to correct the consequences of the previous wave of redistribution, taxes and transfers are heavily downsized, overshooting even the steady state.[14] Examples of oscillations like these my capture the Thatcher and Reagan reaction to the Sixties and Seventies, and then an adjustment by Blair and Clinton.

[14] This overshooting occurred in all the simulations we have tried exploring a wide range of parameter values.

3.3 Permanent difference: US versus Europe (again)

The case of permanent differences in ideology is simple. Imagine that at some point in country B (Europe) the aversion to inequality becomes stronger than in country A, and all inequality is viewed as unfair, so that, as above, $\Omega_t = var(w_{it})$ in country B. This is the case of a permanent shock, rather than a temporary one as studied above. The change relative to country A can be due to historical evolutions. For example the growth of Communist and Socialists parties in Europe after the two World Wars. This is depicted in Figure 3[15]. The two countries are equal in everything except their views on fairness, i.e. their Ω_t. They converge to two different steady states in which country A has lower tax rates, higher GDP, and higher percapita wealth, while country B has higher tax rates, and lower GDP and percapita wealth, but less inequality.

[15] In Figure 3 we have chosen a common value of α =0.1955 (which approximately reflects a combination of the American and the European propensity to save), and δ computed as the weighted average of the GDP divided by the weighted average of 1 minus the income tax rate. The values of γ and the distribution of δ are chosen to generate steady states that approximately match the average income tax rates in the US and in Europe, as well as real (base year, 2008) per-capita GDP.

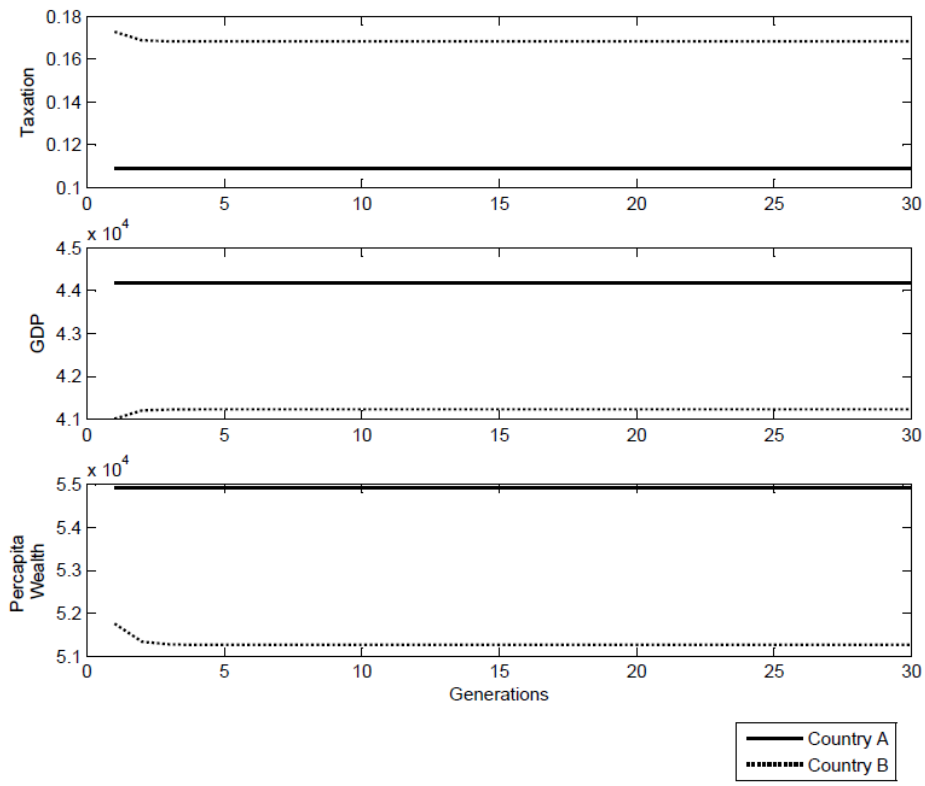

Figure 3

It is important to note that in this figure the different steady states are driven by different concepts of unfairness between the two countries, whereas the two steady states of Figure 1 are only driven by different initial views of the fair wealth distribution. That is, while in the case of Figure 1, characterized by multiple steady states, a one period shock to ideology has the potential to displace the country from the current steady state and make it converge to another, in the case of Figure 3 a one period shock cannot alter the long-term dynamics.

3.4 Extensions

In the working paper version of this paper we describe several additional simulation experiments and examples. For instance we have also investigated several alternative definitions of fairness. First we consider the case in which tax and transfers are considered part of fair wealth. Second, we have looked at cases in which the effect of A_i is part of luck. One may argue that being born smart is part of a sort of genetically induced "luck". Alternatively one may argue that intelligence is fostered by growing up in a rich family with more child care and investment in education. Finally we have considered the case in which individuals dislike inequality per se, namely any deviation of wealth and utility from equality for all at the average is costly. The latter would be an extreme definition of fairness in which any difference in wealth even if arising form harder work and more effort is unfair. Our results generalize nicely and intuitively. The more stringent is the definition of what is "fair" the more society will prefer to redistribute even and the cost growing less and the other way around. Finally, we looked at unexpected shocks (like the recent financial crisis) which may have led to sudden redistributions of wealth.

4 Inheritance (Bequest)Taxation

Our model, which focuses on "social justice", is especially useful to discuss inheritance taxation. Taxing wealth, or more precisely inheritance/ bequests, can be an especially useful and a "quick" way of reestablishing "fairness" without having to tax income for many generations. A bequest tax which redistributes at the beginning of individuals' life is also closely related to the issue of equal opportunities at birth. Our probabilistic voting structure allows multidimensional voting and so we can consider different wealth taxes and income taxes, which in the previous sections were constrained to be the same. In order to keep the analysis as simple as possible, we model taxation in two polar cases, stacking the deck in its favor first, (labelled "inheritance taxation") and against it in the second "(labeled bequest taxation".)

4.1 Non distortionary inheritance taxation

4.1.1 The economy

Each generation votes on the wealth tax rate, τ_{wt}, which is proportionally applied to beginning-of-life gross wealth k_{it-1} and on the income tax rate τ_{yt}; all tax revenues are to be redistributed lump sum to all individuals. The government budget is always balanced. Using the previous definitions, the end of life wealth of each individual is

$$w_{it} = (A_i e_{it} + \eta_i)(1 - \tau_{yt}) + (1 - \tau_{wt})k_{it-1} + G_t \tag{13}$$

where $G_t = \tau_{wt} \int_0^1 k_{jt-1} dj + \tau_{yt} \int_0^1 (A_j e_{jt} + \eta_j) \, dj$

We assume that each generation cares about gross inheritance. This does not take into account how the next generation will redistribute to implement initial wealth taxation. One way of rationalizing it is to say that parents leave a certain amount of wealth to the next generation allowing the latter to decide how to redistribute it.

Individual income is $y_{it} = (A_i e_{it} + \eta_i)(1 - \tau_{yt}) - \tau_{wt} k_{it-1} + G_t$, and the aggregate income of generation t is

$$Y_t = \int_0^1 \left[(A_i e_{it} + \eta_i)(1 - \tau_{yt}) - \tau_{wt} k_{it-1} + G_t \right] di = \int_0^1 A_i e_{it} di, \tag{14}$$

which is identical to percapita income due to the population normalization. Warm glow intergenerational altruism implies that fraction α of end of life wealth is bequeathed, as seen by maximizing u_{it} subject to $c_{it} + k_{it} = w_{it}$. Therefore, plugging the optimal consumption and bequest into the private utility function, we obtain:

$$u_{it} = w_{it} - \frac{e_{it}^2}{2\beta_i}. \tag{15}$$

Individuals vote on taxation at the beginning of their life, before deciding on effort. Maximizing u_{it}, using (13) and (15), gives

$$e_{it} = (1 - \tau_{yt}) A_i \beta_i, \tag{16}$$

which shows that individual effort gets discouraged only by expected income taxation, and is increasing in the individual work ability and decreasing in the disutility of effort. This suggests that our inheritance tax rate could be the best fiscal instrument to reduce unfairness, because it does not discourage individual effort. In fact, in light of equations (16) and (14), aggregate income becomes:

$$Y_t = \int_0^1 A_i (1 - \tau_{yt}) A_i \beta_i di = (1 - \tau_{yt}) \int_0^1 A_i^2 \beta_i di \equiv (1 - \tau_{yt}) \bar{\delta}. \tag{17}$$

Since inheritance taxation is non distortionary[16], in order to maximize GDP one should use only the inheritance tax and set the income tax to zero. This would apply to a standard model, but not in our set up, since our voters care both about consumption (thus GDP) and fairness. The next section will clarify this point. The law of motion of actual and fair wealth of each family i can now be written as:

$$k_{it} = \alpha \left[\delta_i (1 - \tau_{yt})^2 + \eta_i (1 - \tau_{yt}) + k_{it-1} (1 - \tau_{wt}) \right] + \alpha G_t \tag{18}$$

[16] Generation t cares about gross bequest, and the inheritance tax is a lump sum tax freely decided by the next generation.

and

$$\widehat{k}_{it} = \alpha\delta_i(1-\tau_{yt}) + \alpha\widehat{k}_{it-1}. \tag{19}$$

They are simply the equivalent of eq. (11) and eq. (12), after decomposing τ_t into τ_{yt} and τ_{wt}.As before, we use a probabilistic voting model, no changes in the structure of the polity. The equilibrium tax rates[17] $\left(\tau^*_{yt}, \tau^*_{wt}\right)$ are used to determine the level of capital k_{it} and fair capital $\widehat{k}_{it}$, whose distribution will be the state variable of period $t+1$..

4.1.2 Equilibrium Tax Rates

We can characterize the equilibrium inheritance tax rate as follows:

Proposition 1. *The equilibrium value of the wealth tax rate* $\tau^*_{wt} \in [0,1]$*, at all dates* t *is:*

$$\tau^*_{wt} = \min\left[\max\left(1 - \frac{(1-\tau_{yt})\left[cov(\delta_s, k_{st-1})\tau_{yt} - cov(\eta_s, k_{st-1})\right] + cov(\widehat{k}_{st-1}, k_{st-1})}{var(k_{st-1})}, 0\right), 1\right]. \tag{20}$$

Proof. See Appendix.

According to Proposition 1, τ_{wt} is high if luck plays a large part in explaining beginning of life wealth (i.e. $cov(\eta_s, k_{st-1})$ is high), and actual inheritances are not positively related with fair inheritances (i.e. $cov(\widehat{k}_{st-1}, k_{st-1})$ is low). Will voters always prefer a zero income tax rate in order to maximize aggregate income? As we shall see, this is not the case, unless people do not care about fairness: as long as $\gamma > 0$ the distortionary income taxation will be used to correct unfairness in period t. Indeed we have:

Corollary 1. The steady state equilibrium income tax rate τ^*_y is always positive if $\gamma > 0$.

Proof. See Appendix.

Numerical simulations show the following:

1. The transition to the steady state is generically faster than in the case of $\tau_{yt} = \tau_{wt} = \tau_t$;
2. The steady state $\left(\tau^*_{yt}, \tau^*_{wt}\right)$ is unique;
3. The wealth tax rate is often large and income tax small;
4. The use of both income and wealth taxes always delivers a superior macroeconomic performance, in terms of aggregate income and long run wealth, than the use of income tax only.

[17] See Lemma 3 in the Appendix for a more technical discussion.

We can find an explicit solution analytically for an important special case. Most notably:

Corollary 2. *If all the inequality derives from luck, i.e. if* $\sigma_\eta^2 > 0$ *and* $\sigma_\delta^2 = 0$*, then the steady state levels of* τ_w^* *and* τ_y^* *are:*

$$\tau_w^* = 1, \text{ and}$$
$$\tau_y^* = \frac{2\gamma\sigma_\eta^2}{\bar{\delta} + 2\gamma\sigma_\eta^2}.$$

Proof. See Appendix.

4.1.3 What if Inheritance is viewed as Luck?

So far we have assumed that the agents consider bequest as a fair component of wealth, for the part that depends on the ability and effort of their parents. In the opposite scenario, people would think that being born in a rich family is just luck. Hence, assuming that the individuals consider any difference in the inherited capital endowment as undeserved, and hence unfair, it is relatively straightforward to prove[18] that if the variance of "luck", σ_η^2, is not less than 25% of the variance of the combined ability σ_δ^2, the steady state inheritance tax rate becomes 100%. Moreover, if ability and luck are not persistent, but instead are independently and identically distributed over the generations, then[19] $\tau_{wt}^* = 1$, for all t.

4.2 Distortionary bequest taxation

We now return to a wealth tax levied at the end of life, but the difference from the benchmark case of Section 2 is that now we allow it to be different from the income tax. The sum of both taxes would be a tax on total end-of-life wealth. Since bequests are taxed before being transmitted to the next generation, the parent's saving gets taxed while they are still alive, with the proceeds being redistributed within their generation. This enhances the distortionary character of bequest taxation, with the potential to strongly discourage individual effort, and therefore lead to much lower equilibrium values of the bequest tax rate than in the previous sections' inheritance tax rate.

Within the stylized demographics of our model, this structure is reminiscent of the *Unified Gift and Estate Tax* system used until recently in the United States, in which taxation is levied on the estate transferred, regardless of the transfer happening after death (by a will) or before (as a gift). A similar estate tax system is currently in existence in the United Kingdom. In this section we

[18] See Proposition 2 in the Appendix.
[19] See Corollary 3 in the Appendix.

will therefore assume that, when evaluating end-of-life utility according to eq. (1), the parents take into account net bequests, that is:

$$k_{it} = (1 - \tau_{bt})b_{it}.$$

Intergenerational altruism implies that fraction α of her end of life wealth is bequeathed, as seen by maximizing u_{it} subject to $c_{it} + b_{it} = w_{it}$. Therefore, plugging the optimal consumption and bequest into eq. (1), we obtain:

$$u_{it} = w_{it}(1 - \tau_{bt})^{\alpha} - \frac{e_{it}^2}{2\beta_{it}}, \tag{21}$$

and hence:

$$k_{it} = (1 - \tau_{bt})\alpha w_{it}.$$

As before, fiscal policy platforms are voted by each generation before exerting their effort choices. Income tax rate, τ_{yt}, is proportionally applied to end of life incomes. All tax revenues are to be redistributed lump sum to all individuals. Hence,

$$w_{it} = (1 - \tau_{yt})\left(A_{it}e_{it} + \eta_{it}\right) + k_{it-1} + G_t. \tag{22}$$

Government budget is always balanced, and, after rearraging, can be written as:

$$G_t = \frac{[\tau_{yt} + \tau_{bt}\alpha(1 - \tau_{yt})]\int_0^1 A_{jt}e_{jt}dj + \tau_{bt}\alpha\int_0^1 k_{jt-1}dj}{1 - \tau_{bt}\alpha}.$$

Since taxation is known at the beginning of life, before the effort choice is taken, maximizing u_{it}, using (21), (1), and (22), gives optimal effort choice

$$e_{it} = (1 - \tau_{yt})(1 - \tau_{bt})^{\alpha} A_{it}\beta_{it},$$

which shows that individual effort will be discouraged by expected taxation, and is increasing in the individual work ability and decreasing in the disutility of effort.

Hence equilibrium lump sum transfers are:

$$G_t = \frac{1}{1 - \tau_{bt}\alpha}\left\{[\tau_{yt} + \tau_{bt}\alpha(1 - \tau_{yt})]\int_0^1 \delta_{jt}(1 - \tau_{yt})(1 - \tau_b)^{\alpha}dj + \tau_{bt}\alpha\int_0^1 k_{jt-1}dj\right\}.$$

Consequently, the reduced form private utility is:

$$\begin{aligned} u_{it} &= [(1 - \tau_{yt})((1 - \tau_{yt})(1 - \tau_{bt})^{\alpha}\delta_{it} + \eta_{it}) + k_{it-1} + G_t]\,(1 - \tau_{bt})^{\alpha} - \frac{(1 - \tau_{yt})^2(1 - \tau_{bt})^{2\alpha}\delta_{it}}{2} = \\ &= \left[(1 - \tau_{yt})\left((1 - \tau_{yt})(1 - \tau_{bt})^{\alpha}\frac{\delta_{it}}{2} + \eta_{it}\right) + k_{it-1} + G_t\right](1 - \tau_{bt})^{\alpha}. \end{aligned} \tag{23}$$

The generation t individual i utility, U_{it}, after fairness considerations are included and before including the political party bias, is:

$$\hat{U}_{it}(\tau_{yt}, \tau_{bt}) = u_{it} - \gamma \Omega_t, \tag{24}$$

where

$$\begin{aligned}\Omega_t \quad = \quad & \int_0^1 (u_{jt} - \widehat{u}_{jt})^2 dj = \int_0^1 \left[w_{jt}(1-\tau_{bt})^{\alpha} - \widehat{w}_{jt}\right]^2 dj = \qquad (25) \\ & \int_0^1 \left[\begin{array}{c} \left[(1-\tau_{yt})(\left(1-\tau_{yt})(1-\tau_{bt})^{\alpha}\delta_{jt} + \eta_{jt}\right) + k_{jt-1} + G_t\right](1-\tau_{bt})^{\alpha} \\ -[(1-\tau_{yt})(1-\tau_{bt})^{\alpha}\delta_{jt} + \widehat{k}_{jt-1}] \end{array} \right]^2 dj\end{aligned}$$

The law of motion of actual and fair wealth of each family i now become:

$$k_{it} = \alpha \left[\delta_i (1-\tau_{yt})^2 (1-\tau_{bt})^{\alpha} + \eta_i (1-\tau_{yt}) + k_{it-1}\right] + \alpha G_t \tag{26}$$

and

$$\widehat{k}_{it} = \alpha \delta_i (1-\tau_{yt})(1-\tau_{bt})^{\alpha} + \alpha \widehat{k}_{it-1}. \tag{27}$$

Like in the case of the previous section, with perfectly symmetric political bias the resulting probabilistic voting equilibrium will maximize the utilitarian welfare functional, that is

$$(\tau_{yt}^*, \tau_{bt}^*) = \arg \max_{(\tau_{yt}, \tau_{bt}) \in [0,1]^2} \int_0^1 \hat{U}_{it}(\tau_{yt}, \tau_{bt}) di. \tag{28}$$

In the following section we illustrate some numerical simulations on inheritance and bequest taxation

4.3 Numerical Simulations

We have run several simulations of the different scenarios in the presence of multidimensional policy, and tracked the evolution of ideology, fiscal policy, GDP and inequality in the presence of taxes on inheritance/ bequests. The inheritance tax along the lines of Section 4.1 has strong beneficial effects on the economy, whereas the bequest taxation of Section 4.2 does not add to percapita income anything substantial, especially when simulated using data calibrated to match some empirical moments.

In particular, with data calibrated to the United Kingdom, the increase in steady state percapita GDP from having the two tax instruments, compared to the only presence of an income tax, is a scanty 1%. The economic intuition from this result is easily given: as eq. (23) shows, the bequest tax rate distorts effort no less than does the income tax rate, but it does not selectively correct the effects of luck, as instead partially achieved by the income tax rate.

Instead the inheritance taxation of section 4.1 can be very beneficial. In what follows, we allow a switch to the non-distortionary inheritance tax regime sketched in Section 4.1. That is first we assume that the country initially adopts a distortionary bequest tax regime stylized in Section 4.2, then we make the country switch to a non-distortionary inheritance tax regime voted by the inheritance recipients. In both cases, voters also vote for an income tax rate. We assume the economy to be initially in a steady state and, considering the average individual, we find values of parameters[20] α, $\overline{\overline{\delta}}$, and γ that allow to best match the following moments of the British economy: the net national percapita income, final consumption expenditure as a percentage of GDP, and taxes on income, as well as bequest and gift taxes, as a percentage of GDP (2008 data).

Figure 4 depicts, in a highly stylized way, the current approximated situation[21] in solid line (Country A) and the new fiscal regime (Country B), in dashed lines.

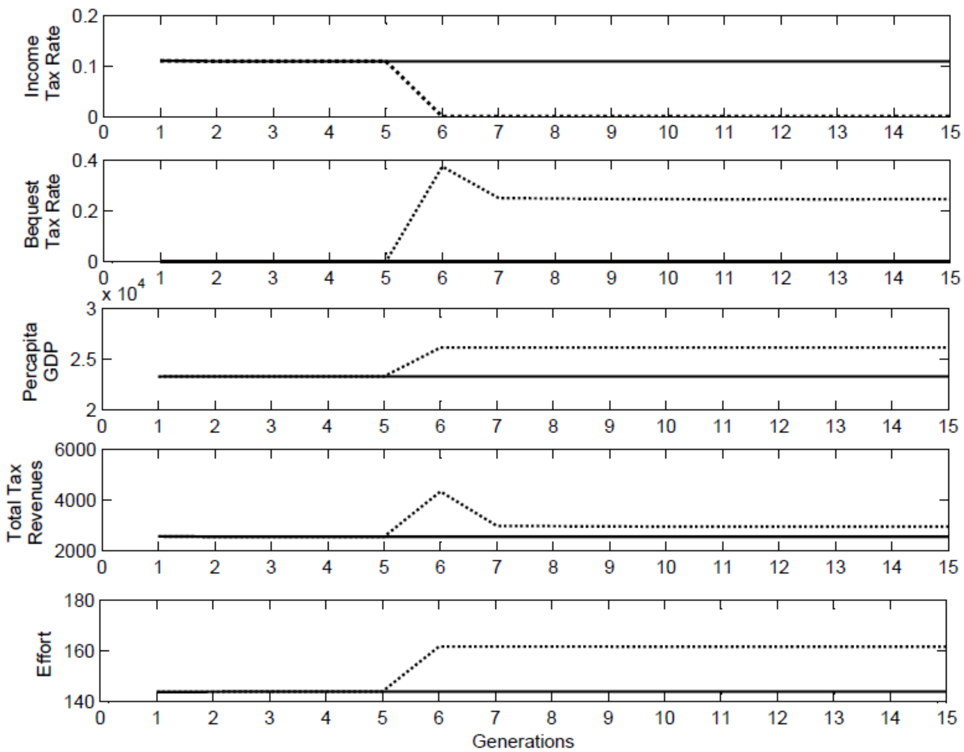

Figure 4

We assume that the fiscal reform takes place in period 5, just to facilitate the illustration of the initial steady state. As the simulated dashed lines show,

[20] Namely, α =0.3135, $\overline{\overline{\delta}} = 26,152$, and $\gamma = 3.553$.

[21] In our model, the approximated income tax/GDP is 10.92% and the approximated bequest tax/GDP is 0.006%.

convergence to the new steady state is achieved in just two periods. The income tax rate converges to a very small but positive level (.02%) in one period, while the initial wealth tax rate, after an initial overshooting, tends in two periods to its new steady state level of about 24.69%.

As clear from Figure 4, our simulation confirms the strong potential gains from introducing a tax targeted to inheritance-related wealth, confirming what we had expected based on our previous analytical results: the new fiscal policy would allow to drastically reduce income taxation and stimulate a huge increase in GDP. Percapita GDP increases from £23,293 to £26,148, with an 12.26% growth. The transition to the new steady state seems quite rapid. Interestingly, due to this change in the tax composition, the total tax collection increases along with a permanent increase in percapita GDP. Notice that percapita GDP, negatively depending on the income tax rate as per eq. (17), converges monotonically to its new steady state, insensitive to the inheritance tax's non-monotonic pattern.

5 Conclusions

We have shown how the evolution of the political ideology regarding the fairness of the constellation of income and wealth in society can generate economic and political persistence in inequality, redistribution, and growth. We have shown how the perception about what is "fair" or not may generate very different policy choices even in countries which would otherwise be identical. We have also shown how temporary shock to preferences/ideology may have long lasting effects. Finally we have explored in detail the issue of inheritance taxation. We have shown under which conditions inheritance taxes can be used to reduce more quickly than otherwise the accumulated stock of unfair inequality embodied in bequests. This result is closely connected with the discussion of equalization of opportunities at birth.

6 Appendix

6.1 Proofs

Lemma 1. *In pairwise majority voting, there will exist a unique equilibrium in which the two parties will select the same policy variable, $\tau_t^L = \tau_t^R \equiv \tau_t^*$, given by*

$$\tau_t^* = \arg\max_{\tau_t \in [0,1]} \int_0^1 \varphi_i \hat{U}_{it}(\tau_t) di. \tag{29}$$

Proof. In fact, individual i of generation t will vote for party R if $\hat{U}_{it}(\tau_t^R) > \hat{U}_{it}(\tau_t^L) + \sigma_{it} + \varepsilon_t$, that is if $\sigma_{it} < \hat{U}_{it}(\tau_t^R) - \hat{U}_{it}(\tau_t^L) - \varepsilon_t$. Given our assumption on σ_{it}, this event happens with probability $\left[\hat{U}_{it}(\tau_t^R) - \hat{U}_{it}(\tau_t^L) - \varepsilon_t\right]\varphi_i + \frac{1}{2}$. Aggregating over all individuals and using the law of large numbers, the fraction of votes that goes to party R is: $\pi_R = \int_0^1 \left\{\left[\hat{U}_{it}(\tau_t^R) - \hat{U}_{it}(\tau_t^L) - \varepsilon_t\right]\varphi_i + \frac{1}{2}\right\} di = \int_0^1 \left[\hat{U}_{it}(\tau_t^R) - \hat{U}_{it}(\tau_t^L)\right]\varphi_i di - \varphi\varepsilon_t + \frac{1}{2}$, where $\varphi \equiv \int_0^1 \varphi_i di$ is the average of the individual ideological densities. Party R wins if $\pi_R > \frac{1}{2}$, which happens if and only if $\varepsilon_t < \frac{\int_0^1 \left[\hat{U}_{it}(\tau_t^R) - \hat{U}_{it}(\tau_t^L)\right]\varphi_i di}{\varphi}$. From our assumptions on ε_t, this happens with probability $\left(\frac{\int_0^1 \left[\hat{U}_{it}(\tau_t^R) - \hat{U}_{it}(\tau_t^L)\right]\varphi_i di}{\varphi} - \left[-\frac{1}{2\psi}\right]\right)\psi = \psi\frac{\int_0^1 \left[\hat{U}_{it}(\tau_t^R) - \hat{U}_{it}(\tau_t^L)\right]\varphi_i di}{\varphi} + \frac{1}{2} \equiv p_R$. Party R therefore chooses $\tau_t^* = \arg\max p_R = \arg\max_{\tau_t^R} \int_0^1 \hat{U}_{it}(\tau_t^R)\varphi_i di$. Swapping notations, party L chooses $\tau_t^* = \arg\max p_L = \arg\max_{\tau_t^L} \int_0^1 \hat{U}_{it}(\tau_t^L)\varphi_i di$. By Weierstrass theorem a maximum certainly exists. Moreover, it is generically unique. Q.E.D.

Lemma 2. *Let us assume that the distribution of abilities and luck are such that* $\inf\left\{A_i^2\beta_i : i \in [0,1]\right\} > -\inf\left\{\eta_i : i \in [0,1]\right\}$. *Then* $w_{it} \geq 0$ *for all* $i \in [0,1]$,*and* $t = 1, 2, ...,$ *for every non-negative initial capital vector* k_{i0}, $i \in [0,1]$, *and for every tax rate sequence* $\tau_t \in [0,1]$.

Proof. First notice that the above stated condition implies:

$$\int_0^1 A_j^2\beta_j dj > -\inf\left\{\eta_i : i \in [0,1]\right\} \equiv |\eta^{\inf}|. \tag{30}$$

Let us consider the worst possible scenario, in which $k_{it-1} = 0$ for all $i \in [0,1]$: if we are able to prove that $k_{it} = \alpha w_{it} \geq 0$ in this case, then $k_{it} \geq 0$ will hold in all other cases.

From the definition of end-of-life post-tax wealth, optimal effort choice,and government transfer, it easily follows that

$$w_{it} = (1-\tau_t)z_{it} + G_t = (1-\tau_t)^2 A_i^2\beta_i + (1-\tau_t)\eta_i + \tau_t(1-\tau_t)\int_0^1 A_j^2\beta_j dj, \tag{31}$$

which expresses w_{it} as a quadratic function of τ_t. Hence, $w_{it} = 0$ if and only if $\tau_t = 1$ and

$$\tau_t = -\frac{\eta_i + A_i^2\beta_i}{\int_0^1 A_j^2\beta_j dj - A_i^2\beta_i}. \tag{32}$$

Let us first focus on the $\tau_t = 1$ root. Since in $\tau_t = 1$, $w_{it} = 0$, as τ_t becomes lower than 1, we need to make sure that w_{it} does not immediately become negative: that is we want w_{it} to be locally a decreasing function of τ_t. Taking the derivative of w_{it} with respect to τ_t we get:

$$\frac{dw_{it}}{d\tau_t} = -2(1-\tau_t)A_i^2\beta_i - \eta_i + (1-\tau_t)\int_0^1 A_j^2\beta_j dj - \tau_t\int_0^1 A_j^2\beta_j dj < 0 \quad (33)$$

if and only if:

$$\eta_i > -2(1-\tau_t)A_i^2\beta_i + (1-2\tau_t)\int_0^1 A_j^2\beta_j dj. \quad (34)$$

Notice that if $\tau_t = 1$ inequality (34) holds true if

$$\eta_i > -\int_0^1 A_j^2\beta_j dj, \quad (35)$$

holds, which is a consequence of inequality (30).Clearly, this guarantees only that $w_{it} > 0$ for τ_t slightly less than 1.

Setting $\tau_t = 0$ in (31), it becomes:

$$w_{it} = A_i^2\beta_i + \eta_i, \quad (36)$$

which is positive if $A_i^2\beta_i > -\eta_i$, which holds under the condition in the statement. Hence, being wealth (31) quadratic in τ_t, the second root of $w_{it} = 0$ - given by eq. (32) - has to be negative if the corresponding parabola is concave or larger than 1 if it is convex[22]. In both cases, $w_{it} > 0$ for all $\tau_t \in [0,1]$. QED

Lemma 3. *In pairwise majority voting, there exists a unique equilibrium in which the two parties select the same policy variable,* $\left(\tau_{yt}^L, \tau_{wt}^L\right) = \left(\tau_{yt}^R, \tau_{wt}^R\right) = \left(\tau_{yt}^*, \tau_{wt}^*\right)$*, given by*

$$\left(\tau_{yt}^*, \tau_{wt}^*\right) = \arg\max_{(\tau_{yt},\tau_{wt})\in[0,1]^2} \int_0^1 \varphi_i \hat{U}_{it}\left(\tau_{yt}, \tau_{wt}\right) di. \quad (37)$$

Proof. Using (7), (13), and (15), gives

$$e_{it} = (1-\tau_{yt})A_i\beta_i, \quad (38)$$

Momentarily neglecting the party L bias components, we obtain the indirect utility function of each individual in each generation. That function ultimately depends on exogenous parameters, on expected taxation, and on the actual and fair wealth distribution of the previous generation:

[22] Simple graphing shows that any parabola $y = ax^2 + bx + c$ sloping down at $x = 1$ and positive at $x = 0$, will be positive for all $0 \leq x < 1$.

$$U_{it} = \left[\delta_i(1-\tau_{yt})+\eta_i\right](1-\tau_{yt}) + k_{it-1}(1-\tau_{wt}) + \int_0^1 \left[\delta_j(1-\tau_{yt})\tau_{yt}+\tau_{wt}k_{jt-1}\right]dj - (1-\tau_{yt})^2\frac{\delta_i}{2}$$

$$-\gamma\int_0^1 \left[\begin{array}{c} \left[\delta_s(1-\tau_{yt})+\eta_s\right](1-\tau_{yt}) + k_{st-1}(1-\tau_{wt}) + \\ \int_0^1 \left[\delta_j(1-\tau_{yt})\tau_{yt}+\tau_{wt}k_{jt-1}\right]dj - \delta_s(1-\tau_{yt}) - \widehat{k}_{st-1} \end{array}\right]^2 ds \tag{39}$$

$$\equiv \hat{U}_{it}\left(\tau_{yt},\tau_{wt}\right).$$

For the rest of the proof, simply follow the same steps as in the proof of Lemma 1, with obvious modifications. QED.

Remark. As in other probabilistic voting models, the same equilibrium policy variable would also be chosen by a biased social planner who maximizes the following weighted aggregate welfare functional:

$$W\left(\tau_{yt},\tau_{wt}\right) \equiv \int_0^1 \varphi_i \hat{U}_{it}\left(\tau_{yt},\tau_{wt}\right)di,$$

with each individual's indirect utility function (where effort, consumption, and bequest are all optimal) being weighted inversely to vulnerability, $1/\varphi_i$, to party-related attributes. In the special case of individuals who have the same densities $\varphi_i = \varphi$, Lemma 3 implies that $\left(\tau^*_{yt},\tau^*_{wt}\right) = \arg\max_{\left(\tau_{yt},\tau_{wt}\right)} W\left(\tau_{yt},\tau_{wt}\right)$ would coincide with the tax rate chosen by a social planner who adopts a utilitarian welfare functional. Notice that, from eq. (39), the equilibrium tax rates $\left(\tau^*_{yt},\tau^*_{wt}\right)$ will depend on generation $t-1$'s bequest distribution k_{t-1}, generation $t-1$'s fair bequest distribution $\widehat{k}_{t-1}$, and of course on the parameter vectors δ and η; that is $\left(\tau^*_{yt},\tau^*_{wt}\right) = \Upsilon(k_{t-1},\widehat{k}_{t-1},\delta,\eta)$.

Proposition 1. *The equilibrium value of the wealth tax rate* $\tau^*_{wt} \in [0,1]$, *follows a simple rule:*

$$\tau^*_{wt} = \min\left[\max\left(1 - \frac{(1-\tau_{yt})\left[cov(\delta_s,k_{st-1})\tau_{yt} - cov(\eta_s,k_{st-1})\right] + cov(\widehat{k}_{st-1},k_{st-1})}{var(k_{st-1})},0\right),1\right]. \tag{40}$$

Proof. Since equal weights are assumed, according to Lemma 5 the probabilistic voting equilibrium will be such that the following is maximized:

$$\int_0^1 \hat{U}_{it}\left(\tau_{yt},\tau_{wt}\right)di = \tag{41}$$

$$= \int_0^1 \left\{ \begin{array}{c} \left[\delta_i(1-\tau_{yt})+\eta_i\right](1-\tau_{yt})+k_{it-1}(1-\tau_{wt})+ \\ \int_0^1 \left[\delta_j(1-\tau_{yt})\tau_{yt}+\tau_{wt}k_{jt-1}\right]dj-(1-\tau_{yt})^2\frac{\delta_i}{2}- \\ \gamma\int_0^1 \left[\begin{array}{c} \left[\delta_s(1-\tau_{yt})+\eta_s\right](1-\tau_{yt})+k_{st-1}(1-\tau_{wt})+ \\ \int_0^1 \left[\delta_j(1-\tau_{yt})\tau_{yt}+\tau_{wt}k_{jt-1}\right]dj-\delta_s(1-\tau_{yt})-\widehat{k}_{st-1} \end{array} \right]^2 ds \end{array} \right\} di$$

Taking the first derivative of (41) with respect to the wealth tax rate, τ_{wt}, we obtain:

$$\frac{\partial\left(\int_0^1 \hat{U}_{it}\left(\tau_{yt},\tau_{wt}\right)di\right)}{\partial\tau_{wt}} = 0-\gamma\frac{\partial\Omega_t}{\partial\tau_{wt}}, \tag{42}$$

where

$$\begin{aligned} \Omega_t &= \\ & \int_0^1 \left\{ \begin{array}{c} \left[\delta_s(1-\tau_{yt})+\eta_s\right](1-\tau_{yt})+k_{st-1}(1-\tau_{wt})+ \\ \int_0^1 \left[\delta_j(1-\tau_{yt})\tau_{yt}+\tau_{wt}k_{jt-1}\right]dj-\delta_s(1-\tau_{yt})-\widehat{k}_{st-1} \end{array} \right\}^2 ds \\ &= \int_0^1 \left[-\delta_s(1-\tau_{yt})\tau_{yt}+\eta_s(1-\tau_{yt})+k_{st-1}(1-\tau_{wt})+\bar{\delta}(1-\tau_{yt})\tau_{yt}+\tau_{wt}\bar{k}_{t-1}-\widehat{k}_{st-1}\right]^2 ds \end{aligned}$$

Hence, $-\gamma\frac{\partial\Omega_t}{\partial\tau_{wt}} =$

$$\begin{aligned} &= 2\gamma\int_0^1 \left[\left(\bar{\delta}-\delta_s\right)(1-\tau_{yt})\tau_{yt}+\eta_s(1-\tau_{yt})+\left(k_{st-1}-\bar{k}_{t-1}\right)(1-\tau_{wt})+\bar{k}_{t-1}-\widehat{k}_{st-1}\right]\left(k_{st-1}-\bar{k}_{t-1}\right)ds \\ &= 2\gamma\left[var(k_s)(1-\tau_{wt})-cov(\delta_s,k_{st-1})(1-\tau_{yt})\tau_{yt}+(1-\tau_{yt})cov(\eta_s,k_{st-1})-cov(\widehat{k}_{st-1},k_{st-1})\right] \end{aligned}$$

which satisfies Kuhn-Tucker conditions only if τ_{wt} is as in the stated expression. QED.

Corollary 1. The steady state equilibrium income tax rate τ_y^* is always positive.

Proof. To characterize the equilibrium income tax rate, we differentiate (41) with respect to τ_{yt} getting:

$$\frac{\partial\left(\int_0^1 \hat{U}_{it}\left(\tau_{yt},\tau_{wt}\right)di\right)}{\partial\tau_{yt}} = -\tau_{yt}\bar{\delta}-\gamma\frac{\partial\Omega_t}{\partial\tau_{yt}}, \tag{43}$$

but, recalling $cov(\delta_s, \eta_s) = 0$, the last part becomes:

$$\begin{aligned}\frac{\partial \Omega_t}{\partial \tau_{yt}} &= 2\int_0^1 \left[\begin{array}{c} \left(\bar{\delta} - \delta_s\right)(1-\tau_{yt})\tau_{yt} + \eta_s(1-\tau_{yt}) + \\ \left(k_{st-1} - \overline{k}_{t-1}\right)(1-\tau_{wt}) + \overline{k}_{t-1} - \widehat{k}_{st-1} \end{array} \right] \left[\left(\bar{\delta} - \delta_s\right)(1-2\tau_{yt}) - \eta_s\right] ds \\ &= 2[\sigma_\delta^2(1-2\tau_{yt})(1-\tau_{yt})\tau_{yt} - (1-\tau_{yt})\sigma_\eta^2 \\ &\quad -cov(k_{st-1}, \delta_s)(1-\tau_{wt})(1-2\tau_{yt}) - cov(k_{st-1}, \eta_s)(1-\tau_{wt}) - \int_0^1 \widehat{k}_{st-1}\left(\bar{\delta} - \delta_s\right)(1-2\tau_{yt})ds].\end{aligned}$$

Since in the steady state $\widehat{k}_{s\infty} = \frac{\alpha\delta_s(1-\tau_y)}{1-\alpha}$, the last integral is

$$\int_0^1 \widehat{k}_{st-1}\left(\bar{\delta} - \delta_s\right)(1-2\tau_{yt})ds = (1-2\tau_{yt})\frac{\alpha(1-\tau_y)}{1-\alpha}\sigma_\delta^2, \tag{44}$$

and hence:

$$\frac{\partial \Omega_t}{\partial \tau_{yt}} = 2\left[\begin{array}{c} \sigma_\delta^2(1-2\tau_{yt})(1-\tau_{yt})\left(\tau_{yt} - \frac{\alpha}{1-\alpha}\right) - (1-\tau_{yt})\sigma_\eta^2 \\ -cov(k_{s\infty}, \delta_s)(1-\tau_{wt})(1-2\tau_{yt}) - cov(k_{s\infty}, \eta_s)(1-\tau_{wt}) \end{array} \right].$$

Noting that $cov(k_{s\infty}, \delta_s) > 0$ and $cov(k_{s\infty}, \eta_s) > 0$, it follows that $\frac{d\Omega_t}{d\tau_{yt}} < 0$ at $\tau_{yt} = 0$, which implies that $\tau_{yt} = 0$ is a local minimum for $\int_0^1 \hat{U}_{it}(\tau_{yt}, \tau_{wt})\, di$. QED.

Corollary 2. *If all the inequality derives from luck, i.e. if* $\sigma_\eta^2 > 0$ *and* $\sigma_\delta^2 = 0$*, then the steady state levels of* τ_w^* *and* τ_y^* *are:*

$$\begin{aligned}\tau_w^* &= 1, \text{ and} \\ \tau_y^* &= \frac{2\gamma\sigma_\eta^2}{\bar{\delta} + 2\gamma\sigma_\eta^2}.\end{aligned}$$

Proof. Plug $\sigma_\eta^2 > 0$ and $\sigma_\delta^2 = 0$ into the expression for (42), notice that $cov(\eta_s, k_s) > 0$, to obtain $\tau_w^* = 1$. Then use this result in (43) and solve for τ_y^*.

Proposition 2. *If* $\sigma_\eta^2 > \frac{\sigma_\delta^2}{4}$*, then* $\tau_w^* = 1$ *and* $\tau_y^* \in (0,1)$.

Proof. Notice that in this case $cov(\widehat{k}_{st-1}, k_{st-1}) = 0$. From the proof of Proposition 1, we get

$$\begin{aligned}\frac{\partial\left(\int_0^1 \hat{U}_{it}(\tau_{yt}, \tau_{wt})\, di\right)}{\partial \tau_{wt}} &= 0 - \gamma\frac{\partial \Omega_t}{\partial \tau_{wt}} = \\ &= 2\gamma\left[var(k_s)(1-\tau_{wt}) - cov(\delta_s, k_{st-1})(1-\tau_{yt})\tau_{yt} + (1-\tau_{yt})cov(\eta_s, k_{st-1})\right]\end{aligned}$$

In a steady state, $-cov(\delta_s, k_{st-1})(1-\tau_y)\tau_y + (1-\tau_y)cov(\eta_s, k_{st-1}) =$

$$\frac{\alpha(1-\tau_y)\left[-\sigma_\delta^2(1-\tau_y)\tau_y + \sigma_\eta^2\right]}{1-\alpha(1-\tau_w)}. \tag{45}$$

Expression (45) is positive if $\sigma_\eta^2 > \frac{\sigma_\delta^2}{4}$, which yields $\tau_w^* = 1$. Using this and simplifying, we can write $\frac{\partial\left(\int_0^1 \hat{U}_{it}(\tau_{yt},\tau_{wt})di\right)}{\partial\tau_{yt}}$ in a steady state, as

$$\begin{aligned} & -\tau_y\bar{\delta} - \gamma\frac{\partial\Omega}{\partial\tau_y} \\ = \quad & -\tau_y\bar{\delta} + \gamma 2(1-\tau_y)\left[-\sigma_\delta^2(1-2\tau_y)\tau_y + \sigma_\eta^2\right], \end{aligned}$$

which is a third degree polynomial which - being $\sigma_\eta^2 > \frac{\sigma_\delta^2}{8}$ - admits only one interior root $\tau_y \in (0,1)$. To see this, after excluding $\tau_y^* = 0$ and $\tau_y^* = 1$, rewrite

$$-\tau_y\bar{\delta} + \gamma 2(1-\tau_y)\left[-\sigma_\delta^2(1-2\tau_y)\tau_y + \sigma_\eta^2\right] \geq 0$$

as

$$\frac{\tau_y\bar{\delta}}{\gamma 2(1-\tau_y)} \leq \sigma_\delta^2(1-2\tau_y)\tau_y - \sigma_\eta^2$$

Since the right hand side is a concave parabola with roots between 0 and 0.5, the left hand side hyperbolic-like curve crosses it from below at most in one point[23].QED.

Corollary 3. If ability and luck are independently and identically distributed over the generations, then $\tau_{wt}^* = 1$, for all t. The steady state level of the income tax rate is: $\tau_y^* \in (0,1)$ if $\sigma_\eta^2 > \frac{\sigma_\delta^2}{8}$, and $\tau_y^* = 0$ if $\sigma_\eta^2 \leq \frac{\sigma_\delta^2}{8}$.

Proof. In this case $cov(\delta_s, k_{st-1}) = 0 = cov(\eta_s, k_{st-1})$, and we immediately get

$$\begin{aligned} \frac{\partial\left(\int_0^1 \hat{U}_{it}\left(\tau_{yt},\tau_{wt}\right)di\right)}{\partial\tau_{wt}} &= 0 - \gamma\frac{\partial\Omega_t}{\partial\tau_{wt}} = \\ &= 2\gamma var(k_s)(1-\tau_{wt}) > 0 \end{aligned}$$

which implies $\tau_{wt}^* = 1$. If $\sigma_\eta^2 \leq \frac{\sigma_\delta^2}{8}$, then $-\tau_y\bar{\delta} - \gamma\frac{\partial\Omega}{\partial\tau_y} < 0$ which implies that $\tau_y^* = 0$ is the only steady state value of τ_y. Instead if $\sigma_\eta^2 > \frac{\sigma_\delta^2}{8}$, following the related argument in the proof of Proposition 2 gives the result. QED.

[23] The intersection from above is a minimizer of the social welfare functional, and hence it will not be an equilibrium.

6.2 US versus Europe

In this section of the Appendix, we give an example of the case of Section 3.1, in which the initial conditions of the two Countries, A and B, are obtained endogenously from a pre-democratization dynamics characterized by a higher importance of luck in Country B than in Country A. In Figure 5 we have provided a representative simulation of two countries, A and B, representative of stylized US and European cases, characterized, by the following features:

1. Both countries share the same parameters, except for the luck distribution: Country B's luck distribution has a higher variance;
2. Both countries share the same political system, biased in favour of the wealthiest.

For brevity, we only show 5 generations, in which Countries A and B are in their steady state before democratization (time 6).

From generation 6 (democratization) the following fundamental changes take place:

1. Country B's luck distribution becomes identical to Country A's;
2. Both countries fully democratize, with different social groups sharing the same weight in power.

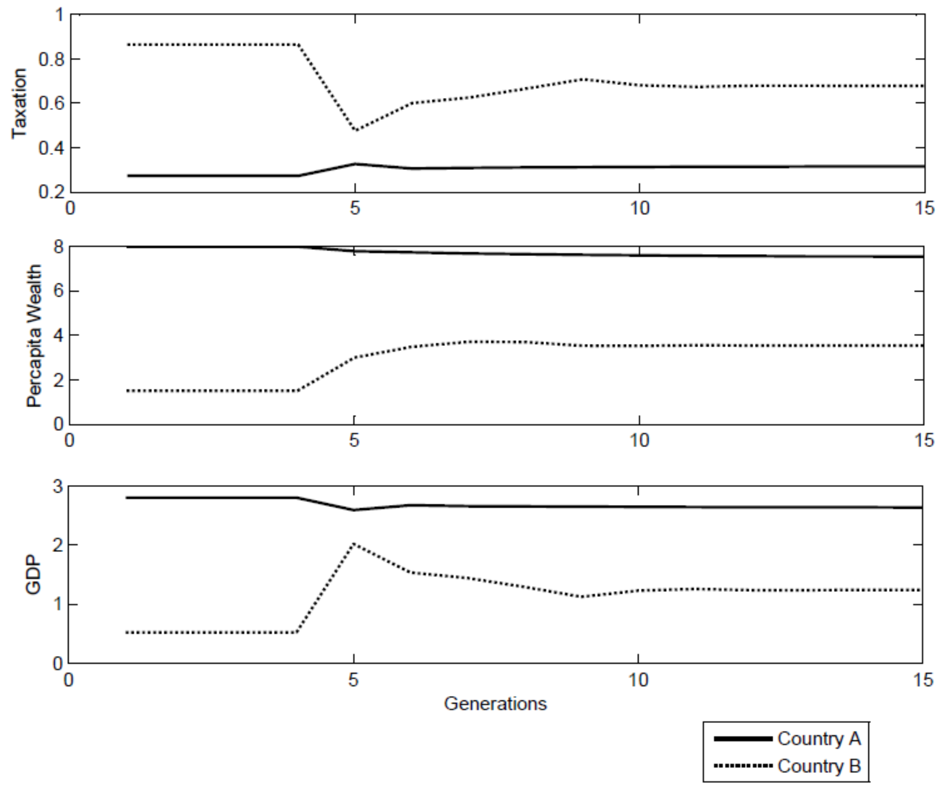

Figure 5

As the reader can see, due to their different initial conditions, the long term fate of these two hypothetical countries diverge: Country B converges to a different steady states than Country A, in which taxation is higher, and percapita wealth and percapita GDP are lower.

References

Alesina, A. and G.M. Angeletos, (2005a), "Fairness and Redistribution," *American Economic Review,* vol. 95(4), pp. 960-80.

Alesina A. and G. M. Angeletos (2005b), "Corruption, Inequality and Fairness" *Journal of Monetary Economics, 52, 1227-44.*

Alesina A., G. Cozzi, and N. Mantovan (2009), "The Evolution of Ideology Fairness and Redistribution", *NBER* Working Paper 15587.

Alesina, A. and E.L. Glaeser, (2004), "Fighting Poverty in the US and Europe: A World of Differences," *New York: Oxford University Press.*

Alesina, A. and P. Giuliano, (2010), "Preferences for Redistribution," in A. Bisin and J. Benhabib eds. *Handbook of Social Economics,* North Holland Amsterdam.

Alesina A., R. Di Tella, and R. McCulloch, (2004) , "Inequality and Happiness: are Europeans and Americans Different? " *Journal of Public Economics,* vol. 88 (9-10), pp. 2009-2042.

Bénabou, R., (1996), "Inequality and Growth," in Ben S. Bernanke and J. J. Rotemberg, eds., *National Bureau of Economic Research Macro Annual,* vol. 11, pp. 11-74.

Bénabou, R. and J. Tirole, (2006), "Beliefs in a Just World and Redistributive Politics," *Quarterly Journal of Economics*, vol. 121(2), pp. 699-746.

Di Tella, R. and R. McCulloch, (2009), "Why Doesn't Capitalism Flow to Poor Countries?," *Brookings Papers on Economic Activity*, pp. 285-321.

Di Tella, R., and J. Dubra, (2010), "A Note on Fairness and Redistribution." Harvard Business School Working Paper, No. 11-059.

Heckman, J.J., (2008), "Schools, Skills, and Synapses," *Economic Inquiry*, Vol. 46(3), 289-324.

Lindbeck A. and J. W. Weibull, (1987), "Balanced-Budget Redistribution as the Outcome of Political Competition," *Public Choice,* vol. 52(3), pp. 273–297.

Lindbeck A. and J. W. Weibull, (1993), "A Model of Political Equilibrium in a Representative Democracy," *Journal of Public Economics,* vol. 51(2), pp. 195-209.

Meltzer, A.H. and S. Richard, (1981), "A Rational Theory of the Size of Government," *Journal of Political Economy*, vol. 89(5), pp. 914-27.

Persson, T. and G. Tabellini, (1994), "Is inequality harmful for growth?" *American Economic Review,* vol. 84(3), pp. 600–621.

Piketty, T., (1995), "Social Mobility and Redistributive Politics," *The Quarterly Journal of Economics*, vol. 110(3), pp. 551-584.

Song, Z., (2008), "Persistent Ideology, and the Determination of Public Policies over Time", *International Economic Review*, forthcoming.