Public Education and Redistribution when Talents are Mismatched*

Michele Bernasconi † Paola Profeta‡

Abstract

In an overlapping generations model with two social classes, rich and poor, parents of the different social classes vote on two issues: redistributive policies for them and public education for their kids. Public education is the engine for growth through its effect on human capital; but it is also the vehicle through which kids born from poor families may exchange their positions with kids born from rich families. This is because education reduces the probability of the mismatch, i.e. individuals with low talent but coming from rich families being placed in jobs which should be reserved to people with high talent (and viceversa). We find a political economy equilibrium of the voting game using probabilistic voting. When the poor are more politically influent, the economy is characterized by a higher level of education, growth and social mobility than under political regimes supported by the rich; pre-tax inequality is greater in the first case, but post-tax is lower.

Keywords: political economy, social mobility, probabilistic voting

JEL Classification: J62, J24, H10

*Previously circulated as: "Redistribution or education? The political economy of the social race".We thank participants to SIEP 2006 Annual Meeting, CEPR Conference on The Formation of Human Capital and Knowledge (Bergen 2007), and to seminars at Bocconi University, Luiss University, Universtiy of Tolouse and CESifo. We also thank Vincenzo Galasso, John Hassler, Raffaello Seri and Guido Tabellini for useful comments. All errors are ours. Paola Profeta acknowledges support from CES during her visit in February 2007.

†Dipartimento di Economia, Università dell'Insubria, Via Monte Generoso 71, I-21100 Varese (Italy). E-mail: bernasconi@uninsubria.it.

‡Istituto di Economia Politica, Università "L. Bocconi", Via Gobbi 5 - I-20136 Milano (Italy). E-mail: paola.profeta@uni-bocconi.it

1 Introduction

Rich and poor have different preferences towards redistribution and public education. Yet, the simple view that the poor support both programs and the rich oppose them is misleading. Even the poorest may choose partial redistribution over full expropriation to limit the distortionary effects of taxation (Meltzer and Richard 1981), while the rich may support public education, because it increases the overall level of human capital in the economy, or because they are able to redirect education expenditure in their favour. They may do this, for instance, by skewing the public component of education toward tertiary education (Hansen and Weisbrod 1969), or by making it only a partial subsidy (Fernández and Rogerson 1995). Generational conflicts may also challenge the basic view. Since low income young are more likely to prefer public education and low income elderly to choose income redistribution, the two programs may be supported by different coalitions between some of the poor and the rich, whose main objective is to minimize public spending (Levy 2005).

In this paper we consider an alternative channel that affects individuals' preferences over pure redistribution and public education. This channel passes through social mobility and is based on a mechanism that we call the "mismatch of talents". The formation of social classes in a society does not depend exclusively on people talent and ability, but also on family lines, social relations, neighborhood networks (Bowles and Gintis 2002). These factors tend to give high chances to kids from rich families of being allocated in good job positions, and hence of remaining rich, even when they have low talent, while at the same time they reduce the chances of the poor to upgrade their status, even when they have high talent. This "mismatch of talents" happens because talent is quite difficult to be directly observed and the rich are in a better position than the poor to gain from this imperfect information.

We argue that this "mismatch of talents" is sensitive to some public policies. Mainly, public education may increase the capacity of a society to correctly recognize the individuals' true talent, and thus to allocate them to the correct social class – thereby increases exchange social mobility. This effect increases the support for public education by the poor, who may even be willing to give up some redistributive transfers, to reduce the talents' mismatch and increase exchange mobility. Yet, it strenghtens the opposition to public education by the rich, who may even be willing to accept some redistribution, in order to stop exchange mobility by maintaining a high mismatch of talents.

In an overlapping generation model we study the political process leading poor and rich parents to choose redistributive policies for them and public education programs for their kids. Parents' preferences determine their voting behavior in a probabilistic voting game, where the level of a general tax rate proportional to income and the distribution of revenues between a pure redistributive program and public education are simultaneously decided. The political economy equilibrium depends on who, amongst the poor and the rich, have more political influence. We find that in the neutral case, in which the two income groups have the same political influence, the tax rate is maximum. When the rich have more political influence, they decrease the overall level of taxation and tilt the composition of public spending towards more redistribution and less public education, in order to reduce the threat of social mobility. Thus, both public education and social mobility are lower than in the neutral case. When instead the poor have more political influence, taxation is again maximum, but public education is higher and redistribution is lower than the neutral case.

The model delivers interesting implications on the economic and social dynamics. In particular, when the poor have more political influence, the economy is characterized by higher growth with less mismatch and more social mobility. The reason is twofold. More education implies more human capital and growth. Moreover, more education induces fewer mismatches of talents. If these mismatches cause an efficiency cost in production, then a new channel emerges for more education to increase growth, for any level of average human capital.

A large literature has studied how public education promotes social mobility[1]

[1]This relation is closely tied to a notion of social mobility as equality of opportunity. See Shorrocks (1978) and Atkinson (1981) for classical papers on the connections between social mobility and various notions of equality of opportunity, Roemer (1998) for an interpretation which

(see reviews in Solon 1999, Breen and Jonsson 2005, Checchi 2006). In most of the previous contributions, however, the impact of public education on social mobility is due to capital market imperfections, which in a world where private education is possible, may prevent the poor from undertaking the same level of education investment as the rich (Becker and Tomes 1986, Maoz and Moav 1999, Restuccia and Urrutia 2004). In order to emphasize that our mechanism of the mismatch of talents is a new, alternative channel through which education increases social mobility, we choose to abstract away from capital market imperfections and private education. Obviously, this is a deliberate simplification of reality and the real world has to be thought as a combination of different forces at work. Our framework is useful to highlight that, in a world where education is mostly public as in most European countries, an equilibrium with high redistribution and low education expenditure may emerge as supported by rich individuals in their attempt to stop social mobility[2].

The emergence of the mismatch of talents in a context of imperfect information is itself a new contribution of the paper. More precisely, in our overlapping generations model, innate ability concurs with family and social backgrounds to determine the economic attainment of kids (as in Becker and Tomes 1979, Loury 1981 and, more recently, Bénabou 1996). However, unlike in most of the previous contributions, we explicitly model a genetic transmission of talent, using transition probabilities, which are assumed, together with talent, to be unobservable.

The idea of a trade-off between education and redistribution is not new. Recent contributions in the political economy literature (Bénabou and Ok 2001) have in particular emphasized the emergence of a trade-off between social mobility and redistribution: in socially mobile communities, since the poor have more chance of upward mobility, they may be induced not to support high levels of redistribution.

emphasizes the distinction between equality of "outcomes" and equality of "opportunity", Fields and Ok (1999) for a review of the literature on social mobility measurement.

[2]We emphasize the similarity between our argument and the large literature which has explained the origin and extension of redistributive programs, from the "Poor Law" to the Bismarckian welfare systems, as a strategy followed by the dominant class to stop the pressure of more radical social rearrangements. This focus of analysis also differentiates our contribution from some previous studies that explored the effects of the allocation of individuals on social mobility and growth *via* technological changes (as in Galor and Tsiddon 1997, and Hassler and Rodriguez Mora 2000).

As a result, the level of redistribution arising in a more mobile society is lower than the one arising in a less mobile society (Alesina and La Ferrara 2005). Though this relation remains valid in our context, we endogenously derive social mobility from the voting behaviour of the two social classes, rich and poor.

The paper is organized as follows. Section 2 presents the general features of the overlapping generations model. Section 3 explains the political institution and the political economy equilibrium. Section 4 studies the dynamics of the system. Section 5 considers additional efficiency costs of the talents mismatch. Section 6 concludes with a discussion of our results. All proofs are in the appendix.

2 The economic setting

In this section, we introduce an overlapping generations economy made up of a continuum of dynasties i, with unit measure $i = [0; 1]$. Individuals are heterogenous in their innate talent, which can be high or low. They live for two periods: in the first period young individuals accumulate human capital building on their innate talent; in the second period adult individuals receive an income, which depends on the social class they have been allocated to and take voting decisions. Each adult person becomes a parent and gives birth to one kid. He dies at the end of the second period. Notice that in this environment individuals take no economic, but only political decisions. Moreover, there is imperfect information, since the innate talent is unobservable, even to each individual, and cannot be inferred from the economic outcome.

2.1 Social classes, preferences and public policies

In every period, the society is divided into two social classes of equal size. Social classes correspond to job's types and hence incomes. Rich individuals of dynasty i have a high-paid job and they receive income $y_{t,i} = y_t^R$, while poor individuals have a low-paid job and receive income $y_{t,i} = y_t^P$, where the subscript index t with $t = 0, 1, 2, ...$ identifies individuals born at time t, the subscript index i identifies the specific dynasty i and the superscript index R or P identifies the social class, rich or poor. Incomes and social classes will be endogenously determined (see Section 2.3). All parents are employed and there is no flexibility in the amount of working hours.

The process of class transition from parents of generation t to parents of generation $t+1$ is represented by the following social mobility matrix:

	y^P_{t+1}	y^R_{t+1}
y^P_t	$\widetilde{p}_{t+1}$	$1-\widetilde{p}_{t+1}$
y^R_t	$1-\widetilde{p}_{t+1}$	$\widetilde{p}_{t+1}$

(1)

where $\widetilde{p}_{t+1} \in [0,1]$ is the fraction of parents and kids who belong to the same social class, and $(1-\widetilde{p}_{t+1})$ is the fraction of those belonging to different social classes.

We assume that only parents take political decisions. Hence, only parents' preferences will matter. Parents care about their consumption in their second period of their life $C_{t,i}$, and about the human capital of their kids $h_{t+1,i}$, according to the following Cobb-Douglas utility function:

$$V(C_{t,i}, h_{t+1,i}) = \ln(C_t, i) + E_{t,i} \ln(h_{t+1,i}) \tag{2}$$

where $E_{t,i}$ is the expectation operator for the parent's belief about his kid human capital. We explain how these beliefs are formed in section 2.5[3].

There are no capital markets. Government imposes a proportional tax rate τ_t on income. Per-head tax proceeds at all times $t = 0, 1, 2, ...$ are given by $\tau_t \overline{y}_t$, with $\overline{y}_t = 0.5y^P_t + 0.5y^R_t$ being the average gross income at time t. Tax proceeds can either finance a pure redistributive program or public education: let $\gamma_t \in [0,1]$ be the fraction of tax proceeds going into redistribution and $(1-\gamma_t)$ into education. The redistributive program provides a lump-sum transfer b_t to parents of generation t; public education finances a per-head amount e_t of education spending that will enter the human capital of each young individual of generation $t+1$ (kids), as specified below. All education is public and the overall government budget is balanced at

[3]Our assumption about preferences is standard in the literature (see Glomm and Ravikumar 1992, Galor and Zeira 1993, Bénabou 2000, and Cremer and Pestieau 2004).

every period t, so that $b_t = \gamma_t \tau_t \overline{y}_t$ and $e_t = (1 - \gamma_t)\tau_t \overline{y}_t$.

Net resources available for the second period consumption of each parent of generation t are thus $C_{t,i} = y_{t,i}(1 - \tau_t) + b_t$.

2.2 Human capitals and innate talent

Individuals form their human capital $h_{t,i}$ in the first period of life. A fundamental variable is innate ability or talent. Innate talent is a random shock hitting all individuals at the moment of their birth. For each individual i it can take a low value A^L, or a high value A^H. For the generation born at time $t = 0$, we assume that the two values, A^L and A^H, have equal probability. Starting with generation $t = 1$, there is a genetic mechanism of talent transmission, which follows a simple Markov process: with probability p an individual i has the same talent of his parent and with probability $1 - p$ he has the opposite talent. The law of large number holds, so that at all t half individuals are born with A^L and half with A^H.

Individuals form their human capitals according to a Cobb-Douglas learning technology, which builds on the average stock of human capital, transmitted to the new born individuals through education. For any individual i of generation $t + 1$ and innate talent A^j (where A^j can either be A^L or A^H), human capital is given by:

$$h_{t+1,i} = e_t{}^{\xi} \overline{H}_t^{\delta} A^j \tag{3}$$

where $\overline{H}_t$ is the average human capital at time t; e_t is the per-head level of public education decided by parents at time t for their kids; and ξ and δ are the parameters of the Cobb-Douglas, with both δ and $\xi \in (0, 1)$.

Notice that at time $t = 0$ society starts with no parents, thus the human capital accumulation for a young person with talent A^j is some primitive knowledge k_0 directly available to all individuals[4].

[4]Equation (3) is conceptually in line with the literature, in particular with the seminal work of Becker and Tomes (1979) and Loury (1981). These papers focus on human capital investment and innate ability as major sources of intergenerational earnings persistence. We combine this idea with a Cobb-Dogulas learning technology as in Glomm and Ravikumar (1992). Similar, more recent, contributions include Bénabou (1996), Fernández and Rogerson (1998), Davies, Zhang and Zeng (2005).

This process of human capital formation guarantees that at all $t = 0, 1, 2, ...,$ there are only two types of human capital in the society: $h^L_{t+1} = e_t{}^{\xi}\overline{H}^{\delta}_t A^L$ for all individuals with talent A^L ($h^L_0 = k_0 A^L$ at $t = 0$); and $h^H_{t+1} = e_t{}^{\xi}\overline{H}^{\delta}_t A^H$ for all individuals with A^H ($h^H_0 = k_0 A^H$ at $t = 0$). Further, since at all t half individuals born with A^L and half with A^H, both categories count for half of the population.

2.3 Imperfect information and social mismatch

An important feature of our setting is that we explicitly model the genetic transmission of talent, through the probability p. We assume, however, that the genetic probability p of talent transmission is not generally known. Several authors have provided estimates that innate ability of a child is positively correlated with innate ability of the parent (see Bowles and Gintis 2002, Sacerdote 2002, Plug and Vijverberg 2003, and references therein), which implies that p lies in some range of $(0.5, 1)$. A complete agreement on the precise value of p is however far from having being reached. This is because talent is very difficult to observe, even by the individuals themselves. Thus, we have an imperfect information context[5].

If talents were perfectly observed, given the process of formation of human capital at equation (3) at all $t = 0, 1, 2, ...$high human capital individuals would be allocated to high-paid jobs and viceversa. Assume that all jobs are paid according to the individual productivity measured by human capital. Thus, high human capital individuals would be rich, receiving income $y^R_t = h^H_t$, and low human capital individuals would be poor with income $y^P_t = h^L_t$. In this case the probability of class persistence of matrix (1) would always correspond to the probability of genetic

[5]In particular, the hypothesis of imperfect information implies that talent cannot be observed either *directly* through the individual's human capital, or even *indirectly* through the individual's productivity. Suppose, for example, that each individual i of generation $t+1$ produces an output $q_{t+1,i}$ according to the production function:

$$q_{t+1,i} = h_{t+1,i} + \varepsilon_{t+1,i} \tag{4}$$

where $h_{t+1,i}$ is the human capital of the individual, and $\varepsilon_{t+1,i}$ is an unobservable random shock in production (with $E(\varepsilon_{t+1,i}) = 0$). Since there are only two types of human capitals, an individual with talent A^H will produce $q_{t+1,i} = e_t{}^{\xi}\overline{H}^{\delta}_t A^H + \varepsilon_{t+1,i}$; while an individual with talent A^L will produce $q_{t+1,i} = e_t{}^{\xi}\overline{H}^{\delta}_t A^L + \varepsilon_{t+1,i}$. Thus, equation (4) implies that if talent cannot be directly recognized from human capital, given the white noise $\varepsilon_{t+1,i}$, it cannot either be perfectly known from $q_{t+1,i}$.

talent transmission, that is $\widetilde{p}_{t+1} = p$. However, given imperfect information, this allocation process is unfeasible. In particular, social classes will be formed with some fundamental mismatch, i.e. people with high and low talents are mixed up in both classes of "poor" and "rich"[6].

Precisely, we assume that at any time each social class contains a fraction $\overline{\alpha}_t$ of individuals allocated in the "correct" social class (low talented people in the class of poor, and high talented people in the class of rich) and a fraction $(1 - \overline{\alpha}_t)$ of individuals allocated in the "wrong" social class. Thus, $(1 - \overline{\alpha}_t)$ represents the "mismatch of talents". Formally, this means that for people born at $t + 1$, the incomes of the rich and the poor are respectively given by:

$$y_{t+1}^P = \overline{\alpha}_{t+1}(e_t{}^{\xi}\overline{H}_t^{\delta}A^L) + (1 - \overline{\alpha}_{t+1})(e_t{}^{\xi}\overline{H}_t^{\delta}A^H) \tag{5}$$

$$y_{t+1}^R = \overline{\alpha}_{t+1}(e_t{}^{\xi}\overline{H}_t^{\delta}A^H) + (1 - \overline{\alpha}_{t+1})(e_t{}^{\xi}\overline{H}_t^{\delta}A^L) \tag{6}$$

For individuals born at $t = 0$, the two incomes are: $y_0^P = \overline{\alpha}_0 k_0 A^L + (1 - \overline{\alpha}_0)k_0 A^H$ and $y_0^R = \overline{\alpha}_0 k_0 A^H + (1 - \overline{\alpha}_0)k_0 A^L$.

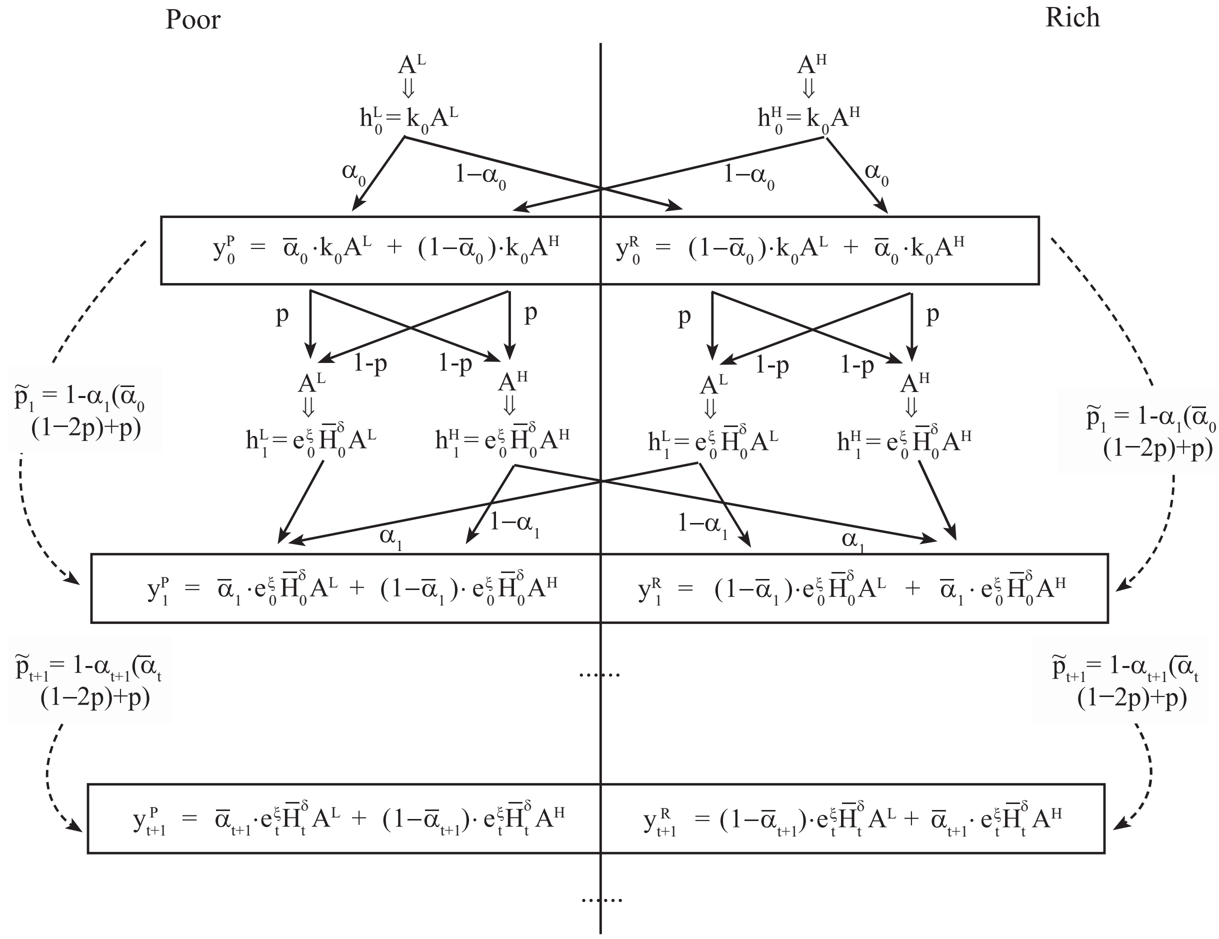
Fig. 1 shows how the mismatch between people talent and social classes can arise[7]. Each individual is assigned to a social class according to an allocation mechanism which, for all $t = 1, ...$, is based on two simple rules[8]: *i)* a low talent kid with poor parents is always assigned to the class of poor and a high talent kid with rich parents to the class of rich; *ii)* a high talent kid with poor parents is assigned with probability α_{t+1} to the class of rich and with probability $(1 - \alpha_{t+1})$ to the class of poor, and a low talent kid with rich parents is assigned with probability α_{t+1} to the class of poor and with probability $(1 - \alpha_{t+1})$ to the class of rich. In other words,

[6]Other papers have studied the role of imperfect information on innate talent (e.g. Maoz and Moav 1999, Checchi, Ichino and Rustichini 1999). In these studies, however, uncertainty is completely resolved with the acquisition of education (for at least those individuals who acquire it). In our paper, instead, some uncertainty about innate talent remains even after education causing the mismatch.

[7]From equations (5) and (6) also notice that an incentive compatibility constraint in society requires that $y_{t+1}^R > y_{t+1}^P$, i.e. that $\overline{\alpha}_{t+1} > 0.5$. In the next section we give the condition for this constraint to be satisfied.

[8]For the first generation born at $t = 0$ the fractions $\overline{\alpha}_0$ and $(1-\overline{\alpha}_0)$ of people allocated respectively in the right and in the wrong social classes are determined randomly, with $\overline{\alpha}_0$ representing the initial condition of the system.

Figure 1: The formation of social classes with the mismatch of talents



α_{t+1} is the probability that the society correctly recognizes the talents of individuals who should be assigned to a different class than their parents'; while $(1-\alpha_{t+1})$ is the probability of having mistakes or errors in the allocation of these individuals.

This process of class formation suggests that, while in general there are little problems in putting both poor kids with low talent in the lower class and rich kids with high talent in the upper class[9], it is more difficult to upgrade kids with high talents from poor parents and to downgrade kids from rich parents but with low talent[10].

[9]Still, the setting can be easily generalized to include random errors even for the allocation of low-talented kids born from poor parents and high-talented kids born from rich parents.

[10]Given that social classes correspond to jobs' types, the allocation process of individuals in social classes may replicate quite realistic stories. For example, kids of a rich family have better opportunities to find an initial better-paid job (say, a stage), independently on their talent, because of family background, social connections, neighborhood networks etc. Their on-the-job performance

The process of class formation hence depends crucially on the probability α_{t+1}. In the next section we will assume that α_t is endogenously determined based on the level of public education in the society, and we will discuss various factors which may affect α_{t+1}. Notice now that α_{t+1} enters both the determination of the fraction $\overline{\alpha}_{t+1}$ of people with the correct talent in each social class and the probability of class persistence $\widetilde{p}_{t+1}$ in society. Iterating from the example of Fig. 1 with $t = 1$, the precise proportion $\overline{\alpha}_{t+1}$ follows this law of motion:

$$\overline{\alpha}_{t+1} = 1 + [\overline{\alpha}_t(2p - 1) - p](1 - \alpha_{t+1}) \tag{7}$$

Similarly, in the society the probability of class persistence $\widetilde{p}_{t+1}$ for kids of poor parents to remain poor and for kids of rich parents to remain rich evolves according to the following equation (see dotted lines in Fig. 1):

$$\widetilde{p}_{t+1} = 1 - \alpha_{t+1}(\overline{\alpha}_t(1 - 2p) + p) \tag{8}$$

The latter equation shows that $\widetilde{p}_{t+1}$ is equal to p only when both α_{t+1} and $\overline{\alpha}_t$ are 1; while $\widetilde{p}_{t+1} > p$ whenever either α_{t+1} or $\overline{\alpha}_t$ (or both) are less than 1. Thus, when imperfect information generates a mismatch of talents, class persistence is greater than the genetic probability of talent transmission.

2.4 Education and class transitions

In this section we argue that the probability α_{t+1} of correctly recognizing individuals' talents is affected by the level of public education. A society with a higher level of education is more able to correctly allocate individuals in their appropriate job or social class. In particular, although family background, social connections, neighborhood networks and all similar factors still remain at the origin of the social mismatch illustrated in Fig. 1, education better allows the society to disentangle

may then reveal the true talent of this person. If he has a high talent, it is reasonable that he will keep the job, while if he has a low talent, with probability α_{t+1} he is recognized and he has to quit. Instead, the kids of a poor family are on their own. Low talent kids will mainly find a low-paid job. Yet, if they have high talent, with (the same) probability α_{t+1}, they may be recognized and upgraded to high-paid jobs.

the impact of family background from innative talent. Education thus increases the equality of opportunity. Yet, the ability of the society to separate these two effects, family background and innate talent, may be reduced by the size of the group of individuals (low talent kids from rich parents and high talent kids from poor parents) who has to be evaluated.

Formally, we assume the following relation:

$$\alpha_{t+1} = \frac{1 - c + d\frac{e_t}{\overline{y}_t}}{(1-\overline{\alpha}_t)p + \overline{\alpha}_t(1-p)} \tag{9}$$

where $1-c$ represents the general degree of openness in society, that is the extent to which society offers equal chances to all individuals independently of the level of public education; the parameter d is the degree to which openness responds to the per-head education expenditure[11] ; and the denominator expresses for each social class the number of kids at time $t+1$ who, if correctly allocated according to their talent, should change their social position with respect to that of their parents.

We generally expect c greater than 0.5 (and lower than 1) and $d \in (0, 0.5)$. We also assume that the following condition is satisfied:

$$c - d \geq p. \tag{10}$$

Condition (10) guarantees that α_{t+1} is a genuine probability belonging to the interval $(0,1)$ at all t. Specifically, when $\frac{e_t}{\overline{y}_t} = 0$ and $c = 1$, society is completely closed and immobile ($\alpha_{t+1} = 0$ and $\widetilde{p}_{t+1} = 1$; see equation 8); when instead $\frac{e_t}{\overline{y}_t} = 1$ and $c - d = p$, the numerator of (9) reaches its maximum value $1-p$, which is also the lower bound of the denominator, when in particular $\overline{\alpha}_t = 1$.

More specifically, using equation (9), both the proportion of individuals correctly allocated in the society $\overline{\alpha}_{t+1}$ (equation 7) and the probability of class persistence

[11] Also notice that although in this paper we only consider the effect of public education on the mismatch of talents and mobility, expression (9) can be easily extended to include the impact of other factors and public policies, amongst which health, security, or liberalization policies.

$\widetilde{p}_{t+1}$ (equation 8), can be recomputed as follows:

$$\overline{\alpha}_{t+1} = 2 - (c - d\frac{e_t}{\overline{y}_t}) - p + \overline{\alpha}_t(2p-1) \tag{11}$$

and

$$\widetilde{p}_{t+1} = c - d\frac{e_t}{\overline{y}_t}. \tag{12}$$

Equation (11) describes how the mismatch evolves in the society at any time $t > 0$ (starting from some initial condition $\overline{\alpha}_0$). It also shows that condition (10) is necessary and sufficient for $\overline{\alpha}_{t+1} \in [0.5, 1]$ at all t, that is for society to respect the incentive compatibility constraint $y^P_{t+1} < y^R_{t+1}$ at all $t = 0, 1, 2, ...$ (see footnote 7). Furthermore, it is worthwhile noticing that, for some time-invariant $\frac{e_t}{\overline{y}_t} \in [0, 1]$, equation (11) implies a unique steady state $\overline{\alpha}_v$, obtained as:

$$\overline{\alpha}_v = \frac{2 - (c - d\frac{e_t}{\overline{y}_t}) - p}{2(1-p)}. \tag{13}$$

The latter entails $\overline{\alpha}_v = 1$, namely that society can end-up in a state without mismatch, if and only if $c - d = p$ and $\frac{e_t}{\overline{y}_t} = 1$.[12]

Equation (12) is crucial in what follows: it makes clear that public education increases exchange mobility, by reducing the mismatch of talents. It also shows that while in our model ability may be genetically transmitted, the genetic probability p of talent transmission plays no role in the mechanism of inheritance of economic inequality[13]. Moreover, in our world of imperfect information, it represents the only relationships that individuals can observe, as explained in the following section.

2.5 Individuals' information set

Remember that, according to their preferences at equation (2), parents care about their kids' human capital. Given however the imperfect information context and the way in which social classes are formed, what matters for parents is the human

[12] Convergence is monotonic with an increasing trajectory if $\overline{\alpha}_v > \overline{\alpha}_0$ and a decreasing trajectory if $\overline{\alpha}_0 > \overline{\alpha}_v$.

[13] This result is indeed consistent with various empirical evidence (see e.g. Bowles and Gintis 2002, in particular their section on "the role of genetic inheritance of cognitive skill", pp. 10-13).

capital that the society will recognize to their kids, which determines their social status. Thus, although parents may not know all the process leading to equation (12), they have information on the ex-post realizations of $\widetilde{p}_t$ and e_t and use them to form beliefs on the allocation of their kids in social classes.

In particular, a poor parent will assign probability $\widetilde{p}_{t+1}$, as defined in equation (12), that his kid will be recognized by the society to have low human capital ${e_t}^{\xi}\overline{H}_t^{\delta}A^L$, and will assign probability $(1-\widetilde{p}_{t+1})$ that he will be recognized to have high human capital ${e_t}^{\xi}\overline{H}_t^{\delta}A^H$; the converse holds for the rich. Using this hypothesis in equation (2), the utility functions for the poor and the rich parents become, respectively:

$$V(C_t^P, h_{t+1}^j) = \ln(C_t^P) + \widetilde{p}_{t+1}\ln({e_t}^{\xi}\overline{H}_t^{\delta}A^L) + (1-\widetilde{p}_{t+1})\ln({e_t}^{\xi}\overline{H}_t^{\delta}A^H) \quad (14)$$

$$V(C_t^R, h_{t+1}^j) = \ln(C_t^R) + \widetilde{p}_{t+1}\ln({e_t}^{\xi}\overline{H}_t^{\delta}A^H) + (1-\widetilde{p}_{t+1})\ln({e_t}^{\xi}\overline{H}_t^{\delta}A^L) \quad (15)$$

These equations show two critical effects of education: on the one side, it increases human capital accumulation for all individuals, and hence growth; on the other side, it affects differently preferences of rich and poor. While the poor have an incentive to prefer more public education to increase the chance for their kids to have a recognized high talent and become rich, the rich have the opposite incentive to reduce public education to avoid that kids with poor parents will have recognized a high talent and take the good jobs at their place. In the remaining of the paper we will show how this effect of public education can affect substantially the political behavior of the two social classes and its consequences for the macroeconomy.

3 The political institution

At time t, based on their preferences at equations (14) and (15), poor and rich parents vote on both the overall tax rate τ_t and on the fraction γ_t of tax proceeds going into pure redistribution $b_t = \gamma_t\tau_t\overline{y}_t$ (with the fraction $(1-\gamma_t)$ financing the per-head public education expenditure $e_t = (1-\gamma_t)\tau_t\overline{y}_t$).

In this section we introduce a probabilistic voting model to determine the equilibrium levels of τ_t and γ_t. These will determine the GDP shares going into redistribution and public education, i.e. the ratios $\frac{b_t}{\overline{y}_t} = \gamma_t \tau_t$ and $\frac{e_t}{\overline{y}_t} = (1-\gamma_t)\tau_t$, respectively. In Section 4 we will analyze the dynamics of the model and the evolution of incomes growth, inequality and social mobility.

3.1 The political economy equilibrium

Models of probabilistic voting (in the tradition of Lindbeck and Weibull 1987, building in turn on Coughlin and Nitzan 1981a and 1981b), are used to solve for political equilibria in situations in which political platforms include more than one issue[14].

Consider two parties, or candidates. Before the election takes place, the parties commit to a policy platform. They act simultaneously and do not cooperate. Each party chooses the platform which maximizes its expected number of votes, or, equivalently, the probability of winning the election. Platforms are chosen when the election outcome is still uncertain. The two parties differ along some other dimension relevant to the voters than the announced policies and which may reflect ideological elements. Ideology may twist voters' preferences away from strict economic interest. In particular, when there is an ideological twist, it pays candidates to propose policy mix more attractive to more mobile voters, also called the "swing" voters.

In probabilistic voting models, there is a unique political equilibrium in which the two candidates propose the same policy. This policy maximizes a social welfare function weighting all voters' utility, with weights depending on the size of the "swing" voters in each class. If the number of swing voters is the same, all groups get equal weight in the candidate's decision. However, if the groups differ in how easily their votes can be swayed, the group containing more swing voters is more responsive to policy and gets a higher weight in the party's objective.

In our set-up, there are only two classes, the poor and the rich, with utility

[14]See Persson and Tabellini (2000) for alternative solutions to political economy models involving multiple issues. Probabilistic voting is particularly appropriate in our context, because it allows dealing with "ideological" components, which may differentiate the political influence of the two competing social classes, rich and poor.

functions given in equations (14) and (15), respectively. Let $\omega_t \in (0, +\infty)$ denote the weight measuring the proportion of "swing" voters in the class of "rich" relative to the proportion of "swing" voters in the class of "poor".

Definition. *A probabilistic voting equilibrium at time t is a pair (τ_t, γ_t) for $\tau_t \in [0,1]$ and $\gamma_t \in [0,1]$, which maximizes a policy maker's objective function given by:*

$$\max_{\gamma_t,\tau_t} W = \omega_t[\ln(C_t^P) + \widetilde{p}_{t+1}\ln(e_t{}^{\xi}\overline{H}_t^{\delta}A^L) + (1-\widetilde{p}_{t+1})\ln(e_t{}^{\xi}\overline{H}_t^{\delta}A^H)] \\ + \ln(C_t^R) + \widetilde{p}_{t+1}\ln(e_t{}^{\xi}\overline{H}_t^{\delta}A^H) + (1-\widetilde{p}_{t+1})\ln(e_t{}^{\xi}\overline{H}_t^{\delta}A^L) \tag{16}$$

where: i) $\omega_t > 0$*;* ii) $C_t^i = y_t^i(1-\tau_t) + \gamma_t\tau_t\overline{y}_t$ *for* $i = R, P$*;* iii) $\widetilde{p}_{t+1} = c - d \cdot \frac{e_t}{\overline{y}_t}$*, with* $e_t = (1-\gamma_t)\tau_t\overline{y}_t$ *and* $\overline{y}_t = 0.5(y_t^P + y_t^R)$*;* iv) *and with* y_t^P*,* y_t^R*,* $\overline{H}_t$ *all given at time* t*.*

Given the definition, when $\omega_t \in (0,1)$ the bias due to "swing" voters is in favor of the policy mix preferred by the rich; when $\omega_t \in (1, +\infty)$ the bias is in favor of the policy mix preferred by the poor; while when it is exactly $\omega_t = 1$, there is no ideological bias and all preferences count equally.

The following proposition characterizes the political economy equilibrium under these three different political conditions.

Proposition 1 *In the above economy, depending on ω_t, the political equilibria are as follow:*

- *For $\omega_t = 1$ (all voters count equally), $\tau_t = 1$ and $\gamma_t = \frac{1}{1+\xi}$. Hence, the GDP shares going into pure redistribution and into public education are, in the order: $\tau_t\gamma_t = \frac{1}{1+\xi}$ and $\tau_t(1-\gamma_t) = \frac{\xi}{1+\xi}$ at all $t = 0, 1, 2, ...$;*
- *For $\omega_t > 1$ (bias favors poor), $\tau_t = 1$ and $\gamma_t < \frac{1}{1+\xi}$ for all $t = 0, 1, 2, ...$ Hence, the GDP shares are: $\tau_t\gamma_t < \frac{1}{1+\xi}$ and $\tau_t(1-\gamma_t) > \frac{\xi}{1+\xi}$ at all $t = 0, 1, 2, ...$; further, for a time-invariant ω_t (that is, constant for all*

$t = 0,1,2,...$), γ_t is time-invariant, so that the two shares are also time-invariant;

- *For $\omega_t < 1$ (bias favors rich), τ_t and γ_t are elaborate functions of the parameters (their exact values are given in Appendix); the more interesting GDP shares are: $\gamma_t\tau_t < \frac{1}{1+\xi}$ all $t = 0,1,2,...$ with $\gamma_t\tau_t = 0$ for any $\omega_t \leq \frac{y_t^P}{y_t^R}$ and with $\frac{\partial\gamma_t\tau_t}{\partial(y_t^P/y_t^R)} > 0$ when $\omega_t \in (\frac{y_t^P}{y_t^R}, 1)$; $(1-\gamma_t)\tau_t < \frac{\xi}{1+\xi}$ at all $t = 0,1,2,...$; further, for a time-invariant ω_t, $(1-\gamma_t)\tau_t$ is time-invariant.*

Moreover, the shares of GDP going into redistribution $\gamma_t\tau_t$ and into public expenditure $\tau_t(1-\gamma_t)$, as functions of the various parameters of the political decision problem are characterized as depicted in Fig. 2, which is integral part of the Proposition.

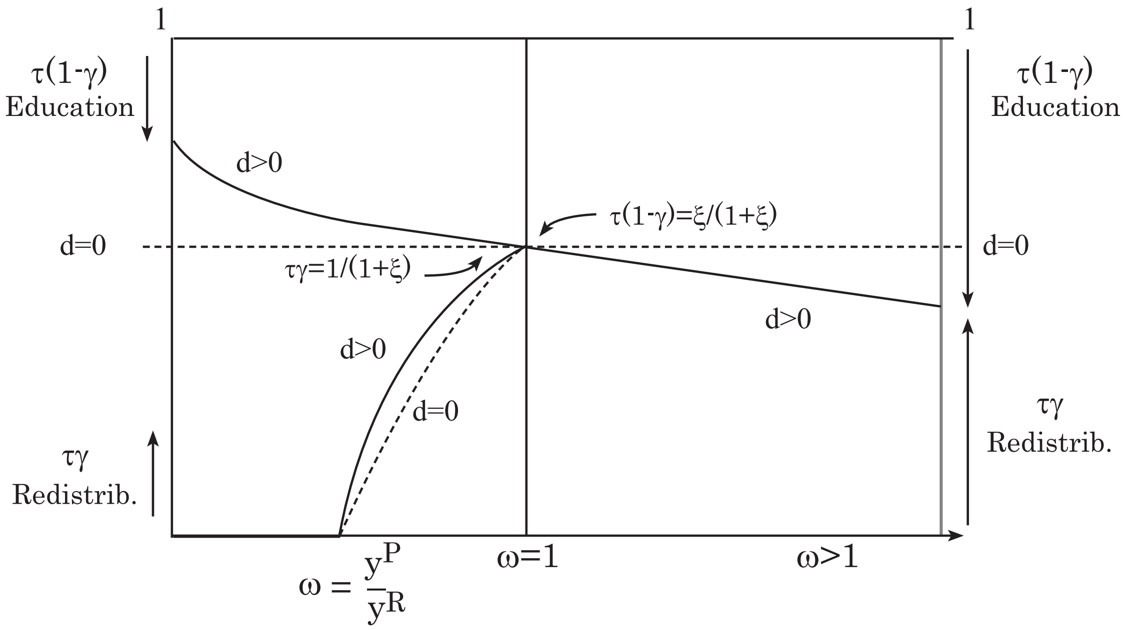


FIGURE 2: Political equilibrium for education and redistribution

The Proposition delivers the political economy equilibrium, as illustrated in Fig. 2. When $\omega = 1$ there is no political bias towards any of the two social classes and the policy maker optimally chooses between education and redistribution ignoring any effect of education on social mobility. This is because a neutral policy maker is not interested in "who is who" in the social parade, i.e. he is utilitarian. Thus, he maximizes the objective function imposing the maximum tax rate ($\tau = 1$) and equalizes the marginal utility of all individuals[15].

When $\omega \neq 1$ and education has a positive impact on exchange mobility $(d > 0)$, the share of GDP going into public education increases with $\omega \in (0, +\infty)$, meaning that when the poor have more political influence, education spending is larger than when the two social classes have equal influence, which in turn is larger than the case in which the rich have more political influence. Redistribution instead is zero if the rich have a very high political influence ($\omega_t < \frac{y_t^P}{y_t^R}$); it turns positive and increases with ω when the rich trade-off more redistribution with less education, in their attempt to stop social mobility; and then decreases with ω when the poor have more political influence, since they prefer to increase education.

If instead education would play no role on exchange mobility $(d = 0$, dotted lines in the figure), the effects that induce the poor to prefer more education and the rich less, would obviously disappear, and education would not depend on ω. Interestingly, in this case, a government under the political influence of the rich would reduce the level of redistribution with respect to the case when $d > 0$. This is because, with the same tax revenues, the rich prefer now to spend more for public education and less for redistribution, since the effect of public education that they dislike has disappeared. On the other hand, redistribution would be maximum for $\omega \geq 1$, because when the poor have more political influence, they have no incentive to choose less redistribution in exchange of more education and mobility.

Notice also that, for time-invariant ω_t, the equilibrium GDP share going into public education does not depend on time. This follows from the property of the Cobb-Douglas utility function and from the fact that both the rich and poor would in any case put a positive amount of resources in public education, for its positive effects on human capital. This important property of the equilibrium policy will be useful in the next section on the dynamics. Conversely, when the rich have more political influence and the policy maker spends some positive amount in redistribution (only

[15]Introducing distortionary taxation would lead to otpimal tax rate lower than 1, but wouldn't affect the main message of the paper.

when $\omega_t \geq \frac{y_t^P}{y_t^R}$), the GDP share of redistribution may change when $\frac{y_t^P}{y_t^R}$ is changing, in particular it increases when there is more pre-tax inequality, even for a constant ω_t.

4 The dynamics of the economy

This section compares the dynamics of all major endogenous variables of the system, for political regimes parametrized by the same ω_t of Proposition 1. We focus on economic growth, measured by the changes in per-capita income $\overline{y}_{t+1} = 0.5y_{t+1}^P + 0.5y_{t+1}^R$; pre-tax inequality, measured by the difference between the two gross incomes[16] $I_{t+1} = (y_{t+1}^R - y_{t+1}^P)$; and social mobility, measured by $(1 - \widetilde{p}_{t+1})$. We also look at the evolution of the mismatch in society, i.e. the dynamics of $\overline{\alpha}_{t+1}$.

We study the dynamics of the system for the three time-invariant political conditions of ω_t: $\omega^N = 1$ when rich and poor have the same political influence (neutral regime), $\omega^P > 1$, when the political bias favors the poor, and $\omega^R < 1$ when the bias favors the rich. To identify the various macroeconomic variables under the three conditions, we will use the capital index $J = N, P, R$ in the obvious way[17].

4.1 Growth

In this economy the average income, $\overline{y}_{t+1} = 0.5{e_t}^{\xi}\overline{H}_t^{\delta}(A^L + A^H)$, is equal to average human capital in society, $\overline{y}_{t+1} = \overline{H}_{t+1}$. Thus, by substitution we obtain the following dynamics equation for average income[18]:

$$\overline{y}_{t+1} = 0.5{e_t}^{\xi}\overline{y}_t^{\delta}(A^L + A^H) \tag{17}$$

Using now the results of Proposition 1, we can establish the following implications

[16] Since in our economy for all t half of the population is poor and half is rich, the only source of inequality is the difference of the two levels of incomes.

[17] This new index only applies to macroeconomic variables: for example, average income is $\overline{y}_{t+1}^N$ when $\omega = \omega^N$, $\overline{y}_{t+1}^P$ when $\omega = \omega^P$, and $\overline{y}_{t+1}^R$ when $\omega = \omega^R$. Since we do not need to identify micro variables, such as individual income, under different political regimes, we will continue to use y_{t+1}^P and y_{t+1}^R to indicate the income of the two social classes, poor and rich respectively, independently of the political regime.

[18] Notice that $\overline{y}_{t+1}$ does not depend on $\overline{\alpha}_{t+1}$, i.e. the mismatch does not affect growth. In a different specification in Section 5 we will include the cost of the mismatch.

for economic growth under the three regimes $J = N, P, R$.

Proposition 2 . *Given a fixed initial condition for the average income* $\overline{y}_0 = 0.5k_0(A^L + A^H)$ *equal for all* $J = N, P, R$*, and given time-invariant* ω^J *under regimes* $J = N, P, R$*, economic growth evolves according to:*

$$\overline{y}^J_{t+1} = B^J(\overline{y}^J_t)^{\xi+\delta} \tag{18}$$

where $B^J = 0.5(A^L + A^H)[(1 - \gamma^J)\tau^J]^\xi$*, constant under all* J*, and with* $B^P > B^N > B^R$*. Thus,* $\overline{y}^P_{t+1} > \overline{y}^N_{t+1} > \overline{y}^R_{t+1}$ *for all* $t = 0, 1, 2, ...$

Equation (18) is virtually identical to that studied by Glomm and Ravikumar (1992)[19]. Growth depends on the sum $\xi + \delta$. We can distinguish three cases under which comparing the different political regimes (see Fig. 3): *a)* if $\xi + \delta < 1$, under all political conditions there are unique, globally stable, steady states with $\overline{y}^P_s > \overline{y}^N_s > \overline{y}^R_s > 0$. Notice also that in this case $\lim_{t\to\infty} \overline{y}^J_{t+1}/\overline{y}^J_t = 1$ for all J; *b)* if $\xi + \delta = 1$, there is no steady state under the regime $J = N, P, R$ for which $B^J \neq 1$. In this case $\overline{y}^J_{t+1}/\overline{y}^J_t = B^J$; *c)* if $\xi + \delta > 1$, under all political conditions there are unique unstable steady states with $\overline{y}^R_s > \overline{y}^N_s > \overline{y}^P_s > 0$. In this case $\overline{y}^J_{t+1}/\overline{y}^J_t > 1$ and $\overline{y}^J_{t+1}/\overline{y}^J_t$ increases over time if $\overline{y}_0 > \overline{y}^J_s$.[20]

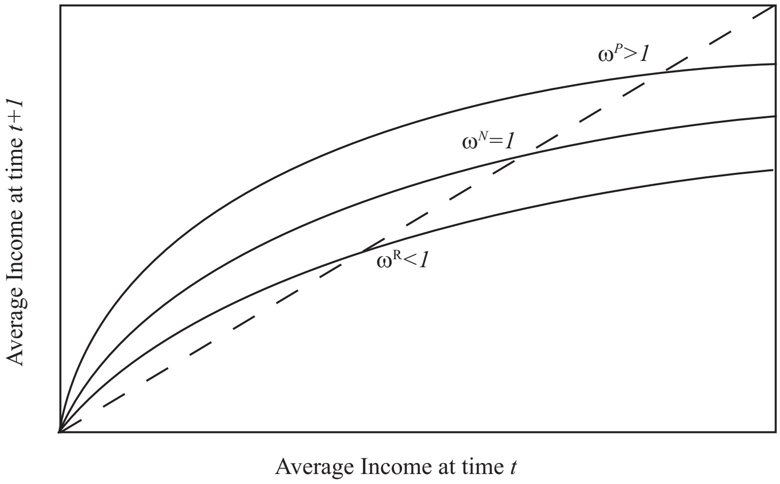
Thus, as in Glomm and Ravikumar (1992), education boosts growth through its impact on human capital, so that economic growth is higher when education is higher. However, since the effect of education on social mobility induces the poor to support education more than the rich, political regimes supported by the poor are also more effective to sustain economic growth. In fact, in case *a)* the long-run growth rates are zero under all political regimes, while in cases *b)* and *c)* the highest long-run growth rate is when the poor have more political influence, followed by the neutral case and then by the situation in which the rich have more influence.

[19] Notice, however, that Glomm and Ravikumar (1992) compare economic growth in a public versus a private education system, while we compare within a public education system the consequences on growth of the different political conditions.

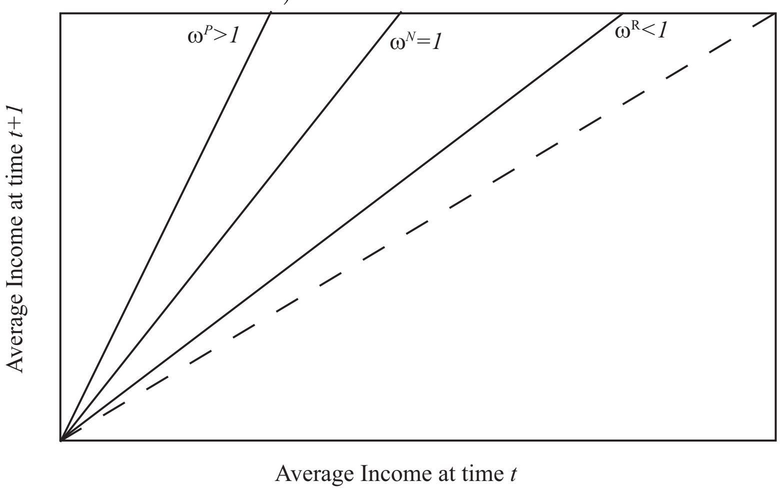
[20] In particular, this is the case in which the economy gets unbounded growth under regime J; otherwise the economy may also end up in the trivial steady-state in which income is zero. (This trivial steady-state applies to all cases *a)*, *b)* and *c)* under all regimes $J = N, P, R$; see Fig. 3).

Figure 3: Income dynamics under time-invariant political regimes

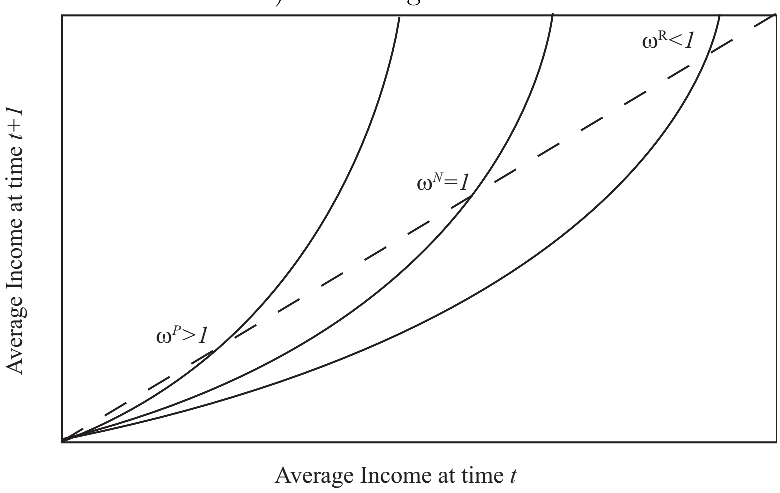
a) Decreasing returns



b) Constant returns



c) Increasing returns



4.2 Inequality

Pre-tax inequality is measured by the difference $I^J_{t+1} = \left(y^R_{t+1} - y^P_{t+1}\right)$ which we know to be positive as long as condition (10) is satisfied and $\overline{\alpha}_{t+1} \in [0.5, 1]$. Thus, under regime J, with $J = N, P, R$, inequality can be written as:

$$I^J_{t+1} = \left(y^R_{t+1} - y^P_{t+1}\right) = (\overline{H}^J_t)^\delta (e^J_t)^\xi (2\overline{\alpha}_{t+1} - 1)(A^H - A^L) \quad (19)$$

For all t, the greater is the mismatch, the lower is inequality[21]. Intuitively, more mismatch implies a lower income for the rich (reduced by the presence of low talented people in their class) and a higher income for the poor (increased by the presence of high talented people in their class)[22]. Therefore, the evolution of inequality depends on the dynamic of the mismatch under the three political regimes.

Proposition 3 . *Given time-invariant ω^J under political regimes $J = N, P, R$, the fractions of people with the "right" talent in each class converge to values of steady state given by $\overline{\alpha}^N_v$, $\overline{\alpha}^P_v$, $\overline{\alpha}^R_v$, with $1 \geq \overline{\alpha}^P_v > \overline{\alpha}^N_v > \overline{\alpha}^R_v \geq 0.5$. Further, given some initial condition $\overline{\alpha}_0 \in [0.5, 1]$ under all regimes, it is also $1 \geq \overline{\alpha}^P_{t+1} > \overline{\alpha}^N_{t+1} > \overline{\alpha}^R_{t+1} \geq 0.5$ for all $t = 0, 1, 2, ...$*

Since the rich are those who mostly (only) benefit from the mismatch, when they have more political influence, the mismatch is comparatively higher.

Substituting now in equation (19) the values of $e^J_t = (1 - \gamma^J)\tau^J \overline{y}^J_t$ found in Proposition 1 for the different regimes $J = N, P, R$, dating the same equation one

[21] In particular, $\frac{\partial I^J_{t+1}}{\partial \overline{\alpha}_{t+1}} = 2(\overline{y}^J_t)^\delta (e^J_t)^\xi (A^H - A^L) > 0$.

[22] Thus, in our model education and inequality are positively related (more education reduces the mismatch, thus increasing inequality). Notice that the relation between education and inequality is not uncontroversial. An opposite, negative relation is found by Glomm and Ravikumar (1992) and Hassler, Rodriguez Mora and Zeira (2006), due to the fact that public education increases the number of skilled workers and reduces the number of unskilled (see also Tamura 1991). However, Hassler, Rodriguez Mora and Zeira (2006) also argue that more public education requires higher taxes. Skilled parents pay higher taxes, but they may also benefit more from education than unskilled parents, because they use education better (on the different returns from education see also Connolly and Gottschalk 2006). Thus, they may prefer more public education than unskilled parents. In this case, more public education may even increase inequality. On the regressivity of optimal education policies see also De Fraja (2002).

period back, we can derive the following dynamic equation for the inequality:

$$I_{t+1}^{J} = I_{t}^{J} \left(\frac{\overline{y}_{t}^{J}}{\overline{y}_{t-1}^{J}} \right)^{\xi+\delta} \left(\frac{2\overline{\alpha}_{t+1}^{J} - 1}{2\overline{\alpha}_{t}^{J} - 1} \right) \tag{20}$$

Under all political regimes $J = N, P, R$, when $\overline{\alpha}_{t+1}^{J}$ are in steady-states $\overline{\alpha}_{v}^{J}$, inequality evolves with economic growth. Before the $\overline{\alpha}_{t+1}^{J\prime}s$ have reached their respective steady-states $\overline{\alpha}_{v}^{J}$, growth and inequality may move in opposite directions depending on whether the initial condition $\overline{\alpha}_{0}$ is greater or lower than the steady states $\overline{\alpha}_{v}^{J}$ themselves[23]. Fig. 4 illustrates three examples, under the three regimes, of the relationships between growth and pre-tax inequality for the case of decreasing returns.

Notice also that the fact that $\overline{\alpha}_{t+1}^{J\prime}s$ converge under all regimes to their steady-states, does not imply that in the long-run the impact of $\overline{\alpha}_{t+1}^{J\prime}s$ is irrelevant for inequality. On the contrary, given that a higher $\overline{\alpha}_{t+1}^{J}$ directly increases inequality for generation $t + 1$, which evolves according to equation (20), under all regimes $J = N, P, R$, a higher trajectory of $\overline{\alpha}_{t+1}^{J}$ implies a higher $\left(y_{t+1}^{R} - y_{t+1}^{P}\right)$ at all $t = 0, 1, 2, ...$ Thus, we can unambiguously compare inequality under the different political regimes.

Proposition 4 *Given initial condition $\overline{\alpha}_{0} \in [0.5, 1]$ and time-invariant ω^{J} under the three regimes $J = N, P, R$, then $I_{t+1}^{P} > I_{t+1}^{N} > I_{t+1}^{R}$ at all $t = 0, 1, 2, ...$*

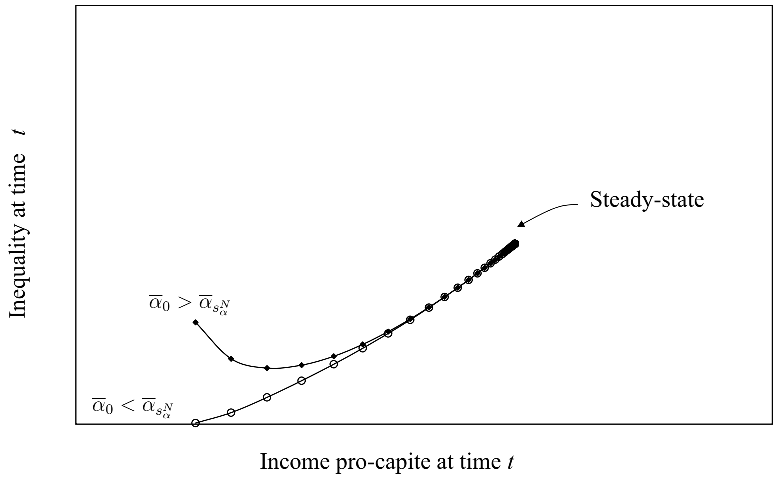
When the poor are more politically influent, the economy is characterized by a higher pre-tax inequality than under a “neutral” political regime, which in turn shows higher inequality than the regime supported by the rich.

These results deserve some comments. Technically they arise because both mismatch and economic growth are positively correlated with inequality, with both growth and inequality being higher in the first political regime (poor), followed by

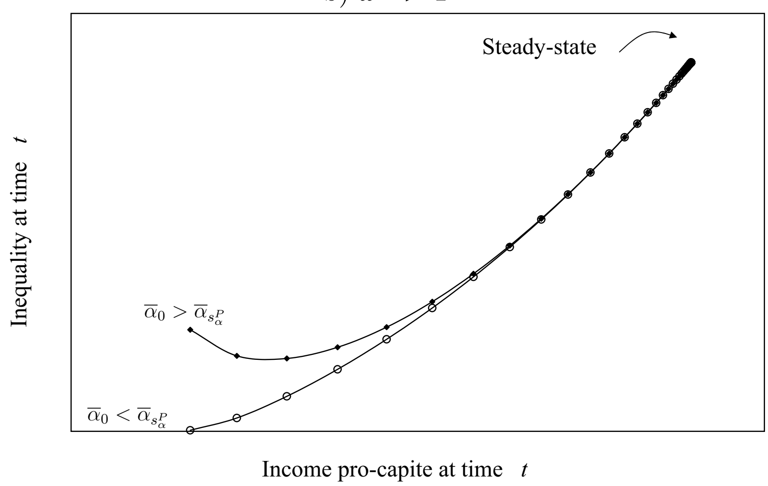
[23]Notice the different subscripts v and s, for the steady-steates of $\overline{\alpha}_{t+1}^{J}$ and $\overline{y}_{t+1}^{J}$. This is because the two variables will typically reach the steady-state at different times. (In addition, while the steady-state of $\overline{\alpha}_{t+1}^{J}$ will always be reached and under all regimes — see Proposition 3 —, the steady-state of the average income $\overline{y}_{t+1}^{J}$'s may well fail to exist or to be reached under different conditions — see Section 4.1).

Figure 4: Interaction between inequality and growth (example of decreasing returns)

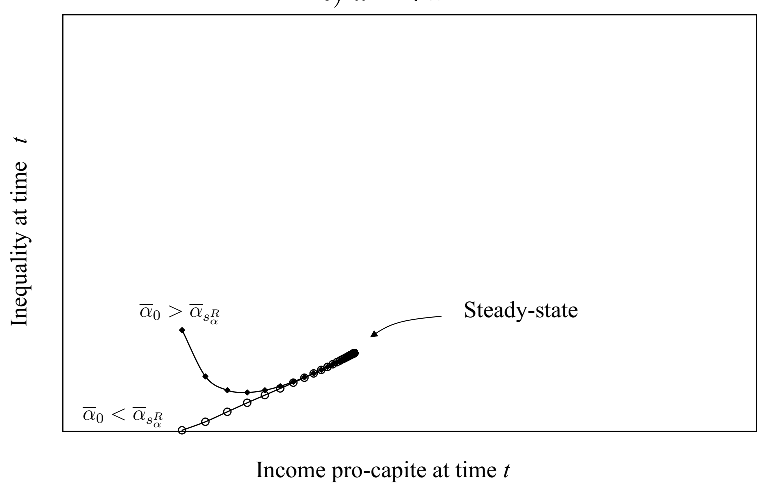
a) $\omega^N = 1$



b) $\omega^P > 1$



c) $\omega^R < 1$



the second (neutral) and then by the third (rich). That higher mismatch is associated with more pre-tax inequality is (as it has been noted) intuitive. The relationship between growth and inequality is, on the other hand, one of the most debated in the literature[24]. In this paper, the nature of this relationship is based on the idea that a higher public education on one side increases growth by increasing the level of human capital, while on the other side it better shapes differences in human capitals due to talent, thus increasing pre-tax inequality. For the reason related to the mismatch, the political regime supported by the poor is the most inclined to public education, which induces more growth and more pre-tax inequality. At the same time, together with the "neutral" regime, the regime of the "poor" is also the most favorable of redistribution; so under the regimes run by the poor, there is both maximum pre-tax inequality and minimum post-tax inequality[25].

4.3 Social mobility

In our two social classes economy, social mobility is simply given by the probability $(1 - \widetilde{p}_{t+1})$ of class transition.

Proposition 5 . *Under all regimes $j = N, P, R$, social mobility is given by:*

$$(1 - \widetilde{p}^J_{t+1}) = 1 - c + d(1 - \gamma^J_t)\tau^J_t \tag{21}$$

For a time-invariant ω^J under each regime, the corresponding $\widetilde{p}^J_{t+1}$ is time-invariant with $\widetilde{p}^R_{t+1} > \widetilde{p}^N_{t+1} > \widetilde{p}^P_{t+1}$.

As expected, social mobility is the highest when the poor are more politically influent; it reaches an intermediate value when rich and poor have the same political influence; and it is lowest when the rich are more politically influent. Social mobility

[24]See e.g. Perotti (1996). In the ninenties various theories of endogenous growth, stimulated by the renewed interest in the Kuznets' curve, have theorized a negative relationship between inequality and growth. As however recently noticed by Forbes (2000) such negative relationship is far less definitive than generally believed; see also (Barro 2000).

[25]A similar result is obtained by Saint-Paul and Verdier (1993) in a median voter model where people vote only on public education as redistributive program.

is in fact good for the poor (upward mobility), while it is bad for the rich (downward mobility).

5 The cost of the mismatch

In the previous section the mismatch had no impact on average productivity. However, when people with low talent are placed in jobs with high responsibility, the mismatch of talents may also have a direct efficiency cost in production. Though very intuitive, this effect has not been enough emphasized by the economic literature[26].

We illustrate this point in a simple model. We assume that individuals' human capitals, rather than determining directly their productivity, are perfect complement in the production process. Though this hypothesis would seem specific, it clarifies a general point. Suppose that there is a single industry employing all workers and producing all GDP using a basic Leontieff technology, which can be reproduced as many times one wishes. That is, pairing any two workers l and f of generation $t+1$, the technology produces an homogenous output $q_{t+1,lf}$ according to the production function:

$$q_{t+1,lf} = 2Min\{h_{t+1,l}; h_{t+1,f}\} \tag{22}$$

where $h_{t+1,l}$ and $h_{t+1,f}$ are the human capitals of the two workers. Clearly, since there are only two qualities of human capital in the economy, namely $h^L_{t+1} = e^{\xi}_t H^{\delta}_t A^L$ and $h^H_{t+1} = e^{\xi}_t H^{\delta}_t A^H$, it follows that any pair of workers can provide only two levels

[26]Only recently some papers have addressed similar issues, but in rather different approaches. For example, Murphy, Shleifer and Vishny (1991) study the implications for growth of the allocation of talent when individuals may choose between occupations (entrepreneurs or rent-seeking). Similarly, Galor and Tsiddon (1997) and Hassler and Rodriguez Mora (2000) show that when individuals have more returns from becoming entrepreneurs (due, for instance, to technological changes), intelligence is more important than social backgrounds for the allocation of individuals, and thus a more efficient allocation emerges, which is associated with a higher growth. In a similar line, Gennaioli and Caselli (2005) also show that failures of meritocracy, such as dynastic management, may reduce growth. Maoz and Moav (1999) focus on liquidity constraints. They show that in societies in which education is privately acquired, liquidity constraints may prevent high talented people of poor families to access higher education, with a loss of efficiency. In most of the above studies, however, the misallocation of resources is before investing in human capitals, while in our model it occurs also after the financing of public education. Its costs may therefore be relatively higher, as we explain in this section.

of output: a low output $q^L_{t+1} = 2e^\xi_t H^\delta_t A^L$, when either both or even only one of the two workers has low talent; or a high output $q^H_{t+1} = 2e^\xi_t H^\delta_t A^H$, when both workers have high talent.

Thus, if society wishes to obtain the maximum overall output from all workers, it would be necessary to pair all individuals with low talent on one side, and all individuals with high talent on the other side. If society could recognize people's talent without any mistake, social classes could be formed accordingly, with incomes of people with low talent (namely the "poor") given by $y^P_{t+1} = \frac{q^L_{t+1}}{2} = e^\xi_t H^\delta_t A^L$ and income of people with high talent (namely the "rich") given by $y^R_{t+1} = \frac{q^H_{t+1}}{2} = e^\xi_t H^\delta_t A^H$. These would also be the incomes of the poor and of the rich, respectively, in a society in which people 's productivity is given by their human capital and all individuals are put in the correct social class.

Suppose, however, that some mismatch of the form described in Section 2.3 occurs when forming the social classes. Thus, a fraction $\overline{\alpha}_{t+1}$ of workers with h^L_{t+1} and a fraction $(1-\overline{\alpha}_{t+1})$ of workers with h^H_{t+1} enter the group of people "recognized" with low talent, while symmetric proportions enter the group of people "recognized" with high talent. By applying the Hardy-Weinberg Principle of the allele frequencies we then have that[27]: i) among the people recognized with low talent , namely the poor, there are $(1-(1-\overline{\alpha}_{t+1})^2)$ pairs producing $q^L_{t+1} = 2e^\xi_t H^\delta_t A^L$ and $(1-\overline{\alpha}_{t+1})^2$ pairs producing $q^H_{t+1} = 2e^\xi_t H^\delta_t A^H$; while ii) among the people recognized with high talent, namely the "rich", there are $\overline{\alpha}^2_{t+1}$ pairs producing $q^H_{t+1} = 2e^\xi_t H^\delta_t A^H$ and $(1-\alpha^2_{t+1})$ pairs producing $q^L_{t+1} = 2e^\xi_t H^\delta_t A^L$.

Individual incomes for people of each class, which we take to correspond to the average levels of output produced by all people of the same class, are then given for

[27]In its simplest form used here, the Hardy-Weinberg Principle of population genetics implies that randomly pairing all workers from a set containing a proportion q of workers with human capitals h^L_{t+1} and a proportion $(1-q)$ of workers with human capital h^H_{t+1}, the genotypic frequency of (h^L_{t+1}, h^L_{t+1}) is q^2, that of (h^L_{t+1}, h^H_{t+1}) is $2(1-q)q$, and that of (h^H_{t+1}, h^H_{t+1}) is $(1-q)^2$. By definining q and $(1-q)$ in terms of $\overline{\alpha}_{t+1}$ according to the proportions specified for the two social classes (and applying the Leontieff production function to the various pairs), one obtains the results given in text.

the poor and the rich by, respectively:

$$y^P_{t+1} = (1-(1-\overline{\alpha}_{t+1})^2)e^\xi_t H^\delta_t A^L + (1-\overline{\alpha}_{t+1})^2 e^\xi_t H^\delta_t A^H \quad (23)$$

$$y^R_{t+1} = (1-\overline{\alpha}^2_{t+1})e^\xi_t H^\delta_t A^L + \overline{\alpha}^2_{t+1} e^\xi_t H^\delta_t A^H \quad (24)$$

The overall average output is equal to:

$$\overline{y}_{t+1} = 0.5e^\xi_t H^\delta_t[(1+2\overline{\alpha}_{t+1} - 2\overline{\alpha}^2_{t+1})A^L + (1-2\overline{\alpha}_{t+1} + 2\overline{\alpha}^2_{t+1})A^H] \quad (25)$$

Thus, in this new setting, per-capita income depends on the extent of the mismatch $\overline{\alpha}_{t+1}$: since $\frac{\partial \overline{y}_{t+1}}{\partial \overline{\alpha}_{t+1}} > 0$ (when $\overline{\alpha}_{t+1} \in [0.5, 1]$), the greater is the mismatch, the lower is average output.

The mismatch generates a loss of output because, when workers with high talent are paired with workers with low talent, the higher productivity of the former is constrained by the lower productivity of the latter. This is here due to the Leontieff technology, but the point is clearly more general.

This example provides also a simple setting to analyze the cost of the mismatch in terms of the waste of human capital it generates. To see this, first of all notice that the new setting has not affected the way in which human capitals in society are formed, so that average human capital continues to be determined according to formula $\overline{H}_{t+1} = 0.5e^\xi_t H^\delta_t[A^L + A^H]$. Substituting in equation (25), we obtain the following relationship between current average income and current human capital (the index J indicates the political regime):

$$\overline{y}^J_{t+1} = \overline{H}^J_{t+1} \cdot F(\overline{\alpha}^J_{t+1}) \quad (26)$$

where $F(\overline{\alpha}^J_{t+1}) = \frac{[(1+2\overline{\alpha}_{t+1}-2\overline{\alpha}^2_{t+1})A^L + (1-2\overline{\alpha}_{t+1}+2\overline{\alpha}^2_{t+1})A^H]}{[A^L+A^H]}$, so that $\overline{y}^J_{t+1} = \overline{H}^J_{t+1}$ if and only if $\overline{\alpha}^J_{t+1} = 1$; whereas (since $\frac{\partial \overline{y}_{t+1}}{\partial \overline{\alpha}_{t+1}} > 0$, for $\overline{\alpha}^J_{t+1} \in [0.5, 1]$) the lower is $\overline{\alpha}^J_{t+1}$, the lower is $\overline{y}^J_{t+1}$ relative to $\overline{H}^J_{t+1}$.

Moreover, we can compare the dynamics of the average human capital and the average income to see how the waste of human capital evolves in a society with costly mismatch. After simple manipulations, the dynamics equation of the average income can be written as:

$$\overline{y}_{t+1}^J = B^J(\overline{y}_t^J)^{\xi+\delta} \cdot \frac{F(\overline{\alpha}_{t+1}^J)}{F(\overline{\alpha}_t^J)^\delta} \tag{27}$$

where $B^{J\prime}s$ are under all regimes the same as in economies without costly mismatch (see Proposition 2)[28]. Thus, when $\overline{\alpha}_{t+1}^J$ have reached their values of steady-state $\overline{\alpha}_v^J$, the conditions under the three political regimes for the economies to be growing, for existence of steady-state incomes $\overline{y}_s^J$'s and for characterizing the relationships amongst steady-states and long-run growth rates are here the same as in the economy without costly mismatch. Namely, they only depend on the sum $\delta + \xi$[29].

Moreover, since we also know that in an economy without costly mismatch it is $\overline{y}_t^J = \overline{H}_t^J$ at all t, the distance between average income and average human capital in an economy with costly mismatch increases, stays constant or decreases, depending on whether the economy is growing, is in steady-state, or it is contracting[30]. Fig. 5 illustrates the point with three examples: a) decreasing returns (in the formation of human capital) with the economy growing to the steady-state; b) decreasing returns with the economy contracting to the steady-state; and c) constant returns with the waste of human capital increasing over time.

Finally, although the above arguments hold equally true in all political regimes, they apply more strongly depending on the extent of the mismatch carried by $\overline{\alpha}_{t+1}^J$ under the three regimes. Thus, since we already know from Proposition 3 that $\overline{\alpha}_{t+1}^P > \overline{\alpha}_{t+1}^N > \overline{\alpha}_{t+1}^R$ at all t, the arguments of the costly mismatch reinforce the

[28]In particular, from equation (25), mean income can be rewritten as:

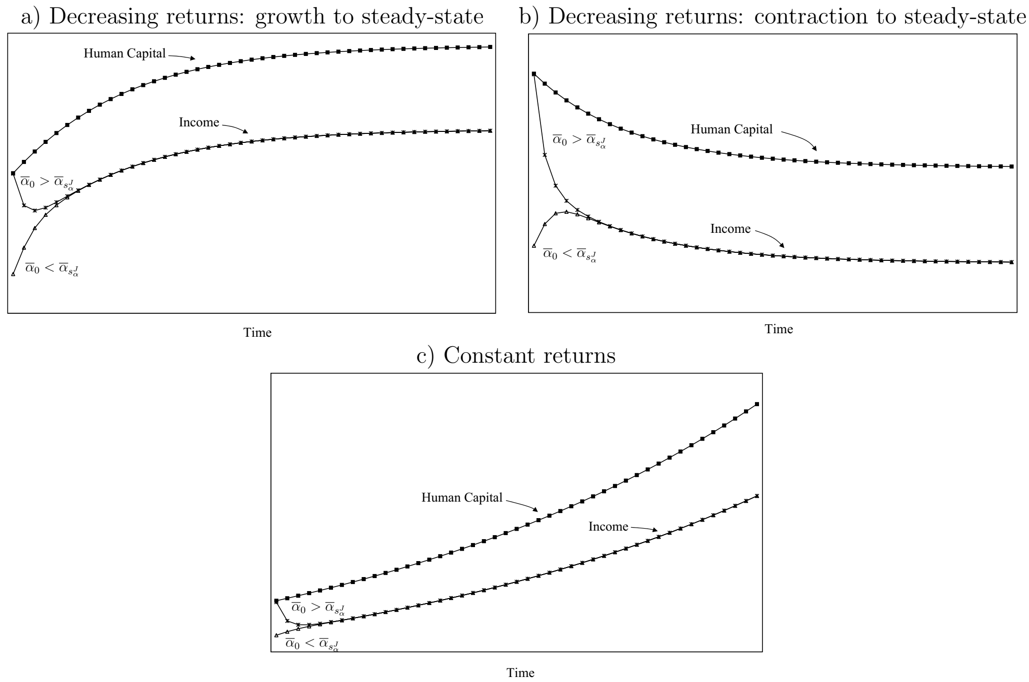
$$\begin{aligned} \overline{y}_{t+1} &= 0.5e_t^\xi H_t^\delta[F(\overline{\alpha}_{t+1}^J)(A^L + A^H)] \\ &= B^J(\overline{y}_t^J)^\xi(H_t^J)^\delta F(\overline{\alpha}_{t+1}^J) \end{aligned}$$

where $B^J = 0.5(A^L + A^H)[(1-\gamma^J)\tau^J]^\xi$ as in Proposition 2. Substituting now equation $\overline{H}_t^J = \frac{\overline{y}_t^J}{F(\overline{\alpha}_t^J)}$ dated one period backwards one obtains equation (27).

[29]This depends on the constant returns of the production function. But the point is more general and it can be easily accomodated for production functions with increasing or decreasing returns.

[30]This in particular applies when $\overline{\alpha}_{t+1}^J$ are in steady-state $\overline{\alpha}_v^J$. When $\overline{\alpha}_{t+1}^{J\prime}$'s are not yet in steady-states, growth in the economies with and without costly mismatch may for sometimes be uncoordinated (that is, one economy may be growing while the other is contracting, and viceversa) depending on whether the initial condition $\overline{\alpha}_0$ is greater or lower than the steady states $\overline{\alpha}_v^J$ of the different political regimes (see also Fig. 5).

FIGURE 5: Economic dynamics when mismatch is costly



conclusions of the previous section, that regimes supported by the poor are better for growth than neutral regimes, which in turn are better than regimes favored by the rich[31].

6 Conclusions

We presented a political economy model in which public spending in redistribution and education depends on how social mobility affects the preferences of two classes, rich and poor. In our story, public education has a positive impact on economic growth, but affects also the "mismatch of talents". In particular, it reduces the probability that individuals with low talent but coming from rich families are placed

[31] A similar discussion holds regarding inequality. In particular, inequality is now given by:

$$I^J_{t+1} = I^J_t \left(\frac{\overline{y}^J_t}{\overline{y}^J_{t-1}}\right)^{(\xi+\delta)} \left(\frac{\overline{\alpha}^J_{t-1}(A^H-A^L)+2A^L}{\overline{\alpha}^J_t(A^H-A^L)+2A^L}\right)^{\delta} \left(\frac{\overline{\alpha}^J_{t+1}}{\overline{\alpha}^J_t}\right)$$

When $\overline{\alpha}^J_t$ is in steady-state, since all factors containing $\overline{\alpha}^J_t$ are equal to 1, inequality evolves with economic growth as in the economy without costly mismatch, but with the different growth rates resulting from the effect of the mismatch. (When $\overline{\alpha}^{J\prime}_t$'s are not yet in steady-states, the dynamics may be a bit more complex due the interaction between the two factors containing $\overline{\alpha}^J_t$ at different t).

in good jobs, which should instead go to highly talented individuals (and viceversa). Since education promotes equality of opportunity, the poor prefer public education to pure redistribution in order to improve upward mobility, while the rich oppose large spending in public education to prevent downward mobility.

In the model, we deliberately abstract from private education in order to emphasize that public education may have an impact on mobility, besides the one related to capital markets imperfections (see, e.g. Checchi 2006, and references therein). Yet, when private education exists, rich parents can send their kids to private schools and need not to limit public education to restrict competition from the poor. Indeed, if they are able to make public education only partial (as in Fernández Rogerson 1995) or to skew it towards higher education (as in the literature stemmed from early study by Hansen and Weisbrod 1969), publicly subsidised education may even result into a redistributive flow from the poor to the rich. This argument however is not in contrast with our story, but they may complement each other in a more general picture. When private education is available, a large redistribution could allow even low income parents to send their kids to private schools and thus to reduce the mismatch of talents. As a consequence, for low levels of redistribution the rich may prefer a system with only (or mainly) private education, while for high levels of redistribution they may prefer one with only (or mainly) public education, in order to stop the poor buying privately even more education to increase exchange mobility. These more general results could for example explain some well-known differences in the mix of redistribution and overall education expenditures between most European countries and the US[32].

Another simplification of our approach is that individuals in the same social class have identical beliefs on their kids' prospect of mobility. Introducing different beliefs would not change the main message, but would deliver further interesting possibilities[33]. If, for instance, we allow a fraction of poor individuals to believe

[32]Notice in particular that, though redistribution and public education expenditures are both higher as shares of GPD in most European countries than in the US, the overall education expenditure is higher in the US than in Europe, when one adds the private components (see e.g. OECD 2006).

[33]See Piketty (1995) and Bénabou and Ok (2001), as seminal theoretical papers on the various

that their kids have high talent while the other parents believe to low talent kids, the former parents will prefer more education than others. Hence, even with public education only, a majoritarian coalition of rich and poor individuals, who believe to have low talent kids, may emerge in a political equilibrium with high redistribution and low education and taxation. However, even in this case, the latent conflict between the two social classes on mobility would not disappear. The final outcome would depend on the fractions of poor agents having different beliefs of upward mobility.

Our model provides material for empirical investigations. Governments supported by the rich are increasingly more hostile to public education spending, while softer on pure redistributive policies the greater is the level of taxation and the higher the effect of education on mobility. Moreover, as argued in section 5, the mismatch of talents, when people with low talent are allocated in jobs of higher potential productivity, and viceversa, generates a loss of resources, in particular of human capital. This may represent an important determinant in explaining the countries' economic performance which is often neglected by the empirical growth literature.

References

[1] Alesina, A. and E. La Ferrara (2005). "Preferences for Redistribution in the Land of Opportunities", Journal of Public Economics 89, pp. 897-931.

[2] Atkinson (1981) "The Measurement of Economic Mobility", in A.B. Atkinson, Essays in Honor of Jan Penn (editor), reprinted in Social Justice and Public Policy, Chapter 3, (1983). London: Wheatsheaf Books Ltd..

[3] Barro, R.J. (2000). "Inequality and Growth in a Panel of Countries", Journal of Economic Growth 5, pp. 5-32.

factors which may influence people beliefs of mobility; see Konrad and Spadaro (2006) for a more recent theoretical analysis; see Fong (2001) and Alesina and La Ferrara (2005), for empirical studies.

[4] Becker, G.S. and N. Tomes (1979). "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility", Journal of Political Economy 87, pp. 1153-1189.

[5] Becker, G.S. and N. Tomes (1986). "Human Capital and the Rise and Fall of Families", Journal of Labour Economics 4, pp. s1-39.

[6] Bénabou, R. (1996). "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance", American Economic Review 86, pp. 584-609.

[7] Bénabou, R. (2000). "Unequal Societies: Income Distribution and the Social Contract", American Economic Review 90, pp. 96-129.

[8] Bénabou, R., and E. Ok (2001) "Social Mobility and the Demand for Redistribution", Quarterly Journal of Economics 116, pp. 447-487.

[9] Bowles, S. and H. Gintis (2002). "The Inheritance of Inequality", Journal of Economic Perspectives 16, pp. 3-30.

[10] Breen, R. and J.O. Jonsson (2005). "Inequality of Opportunity in Comparative Perspective: Recent Research on Educational Attainment and Social Mobility", Annual Review of Sociology 31, pp. 223-243.

[11] Caselli, F. and Gennaioli, N. (2005) "Dynastic Management", mimeo London School of Economics.

[12] Checchi, D., A. Ichino and A. Rustichini (1999). "More Equal but Less Mobile? Education Financing and Intergenerational Mobility in taly and in the US", Journal of Public Economics 74, pp. 351-393.

[13] Checchi, D. (2006). The Economics of Education: Human Capital, Family Background and Inequality. Cambridge: Cambridge University Press.

[14] Connolly, H. and Gottschalk, P. (2006). "Differences in Wage Growth by Education Level: Do Less-Educated Workers Gain Less from Work Experience?" IZA WP No 2331.

[15] Coughlin, P.C. and S. Nitzan (1981a). "Directional and Local Electoral Equilibria with Probabilistic Voting", Journal of Economic Theory 34, pp. 1-12.

[16] Coughlin, P.C. and S. Nitzan (1982a). "Electoral Outcomes with Probabilistic Voting and Nash Social Welfare Maxima", Journal of Public Economics 15, pp. 113-121.

[17] Cremer, H. and Pestieau, P. (2004) "Intergenerational Transfer of Human Capital and Optimal Education Policy", mimeo University of Toulouse and University of Liege.

[18] Davies, J.B., J. Zhang and J. Zeng (2005). "Intergenerational Mobility under Private versus Public Education", Scandinavian Journal of Economics 107, pp. 399-417.

[19] De Fraja, G. (2002). "The Design of Optimal Education Policies", The Review of Economic Studies 69, pp 437-466.

[20] Fernández, R. and R. Rogerson (1995). "On the Political Economy of Education Subsidies", Review of Economc Studies 62, pp. 249-262.

[21] Fernández, R. and R. Rogerson (1998). "Public Education and Income Distribution: a Dynamic Quantitative Evaluation of Education-Finance Reform", American Economic Review 88, pp. 813-833.

[22] Fields, G.S. and E. Ok (1999). "The Measurement of Income Mobility: an Introduction to the Literature", in J. Silber (ed.), Handbook of Inequality Measurement. Dordrecht: Kluwer Academic Publishers.

[23] Fong, C. (2001). "Social Preferences, Self-Interest, and the Demand for Redistribution", Journal of Public Economics 82, pp. 225-246.

[24] Forbes, K.J. (2000). "A Reassessment of the Relationship between Inequality and Growth", American Economic Review 90, pp. 869-887.

[25] Galor, O. and D. Tsiddon (1997). "Technological Progress, Mobility, and Growth", American Economic Review 60, pp. 35-52.

[26] Galor, O. and J. Zeira (1993). “Income Distribution and Macroeconomics”, Review of Economic Studies 60, pp. 35-52.

[27] Glomm, G. and B. Ravikumar (1992). “Public versus Private Investment in Human Capital: Endogenous Growth and Income Inequality”, Journal of Political Economy 100, pp. 818-834.

[28] Hansen, W.H. and B.A. Weisbrod (1969) “The Distribution of Costs and Direct Benefits of Public Higher Education: the Case of California”, Journal of Human Resources 4, pp.176-91

[29] Hassler, J. and J.V. Rodriguez Mora (2000) “Intelligence, Social Mobility and Growth”, The American Economic Review 90, pp. 888-908.

[30] Hassler, J., J.V. Rodriguez Mora and J. Zeira (2006) “Inequality and Mobility”, mimeo Stockolm University, Universitat Pompeu Fabra an Hebrew University of Jerusalem.

[31] Konrad, K.K. and A. Spadaro (2006). “Education, Redistributive Taxation and Confidence”, Journal of Public Economics 90, pp. 171-188.

[32] Levy, G. (2005). “The Politics of Public Provison of Education”, Quarterly Journal of Economics 120, pp. 1507-1534.

[33] Lindbeck, A. and J.W. Weibull (1987). “Balanced-budget Redistribution as the Outcome of Political Competition”, Public Choice 52, pp. 273-297.

[34] Loury, G. (1981). “Intergenerational Transfers and the Distribution of Earnings”, Econometrica 49, pp. 843-867.

[35] Maoz, Y.O. and O. Moav (1999). “Intergenerational Mobility and the Process of Development”, Economic Journal 109, pp. 677-697.

[36] Meltzer, A. H. and Richard, S. F. (1981) “A Rational Theory of the Size of Government”, Journal of Political Economy 89, pp. 914-27.

[37] Murphy, K., Shleifer, A. and Vishny, R. (1991) "The Allocation of Talent: Implications for Growth", Quarterly Journal of Economics 106, pp.503-530.

[38] Perotti, R. (1996). "Growth, Income Distribution and Democracy: What the Data Say", Journal of Economic Growth 1, pp. 149-187.

[39] Persson, T. and G. Tabellini (2000). Political Economics: Explaining Economic Policy. Cambridge, MA: Cambridge University Press.

[40] Piketty, T. (1995). "Social Mobility and Redistributive Politics", Quarterly Journal of Economics 110, pp. 551-585.

[41] Plug, E. and W. Vijverberg (2003). "Schooling, Family Background, and Adoption: Is it Nature or Is it Nurture?", Journal of Political Economy 111, pp. 611-641.

[42] Restuccia, D. and C. Urrutia (2004). "Intergenerational Persistence of Earnings: the Role of Early and College Education", American Economic Review 94, pp. 1354-1378.

[43] Roemer, J. (1998). Equality of Opportunity. Cambride, MA: Harvard University Press.

[44] Sacerdote, B. (2002). "The Nature and Nurture of Economic Outcomes", American Economic Review, Papers and Proceedings 92, pp. 344-348.

[45] Saint-Paul, G. and T. Verdier (1993). "Education, Democracy and Growth", Journal of Development Economics 42, pp. 399-408.

[46] Shorrocks, A.F. (1978). "Measurement of Mobility", Econometrica 46, pp. 1013-1024.

[47] Solon, G. (1999). "Intergenerational Mobility in the Labour Market", in O.C. Ashenfelter and D.Card (eds.), Handboook of Labour Economics, vol. 3A, pp. 1761-1800. Amsterdam: North-Holland.

[48] Tamura, R. (1991). "Income Convergence in an Endogenous Growth Model", Journal of Political Economy 99, pp. 522-540.

Appendix

Proof of Proposition 1

The first order conditions for the maximization problem are:

$$\begin{aligned} \text{FOC}_\tau \quad : \quad 0 = & \frac{\omega\left(-y_t^P + \gamma_t \overline{y}_t\right)}{(1-\tau_t)y_t^P + \gamma_t\tau_t\overline{y}_t} + \frac{\left(-y_t^R + \gamma_t \overline{y}_t\right)}{(1-\tau_t)y_t^R + \gamma_t\tau_t\overline{y}_t} \\ & + \frac{(\omega+1)\,\xi}{\tau} - d(1-\gamma_t)(1-\omega)\left(\ln A^H - \ln A^L\right) \end{aligned} \tag{28}$$

$$\begin{aligned} \text{FOC}_\gamma \quad : \quad 0 = & \frac{\omega\tau_t y_t}{(1-\tau_t)y_t^P + \gamma_t\tau_t\overline{y}_t} + \frac{\tau_t\overline{y}_t}{(1-\tau_t)y_t^R + \gamma_t\tau_t\overline{y}_t} \\ & - \frac{(\omega+1)\,\xi}{1-\gamma_t} + d\tau_t(1-\omega)\left(\ln A^H - \ln A^L\right) \end{aligned} \tag{29}$$

All variables are indexed at time t. Thus, we henceforth suppress the index t. Notice that $\text{SOC}_\tau < 0$ and $\text{SOC}_\gamma < 0$, for all γ and τ. Notice also that the solution of the maximization problem can be of three types: i) $\tau = 1$ and $\gamma \in (0,1)$, ii) $\gamma = 0$ and $\tau \in (0,1)$, iii) $\tau \in (0,1)$ and $\gamma \in (0,1)$.

- Case $\omega \geq 1$.

It is useful to define the following functions:

$$C(\tau,\gamma,\omega) = \frac{\omega\left(-y^P + \gamma\overline{y}\right)}{(1-\tau)y^P + \gamma\tau\overline{y}} + \frac{\left(-y^R + \gamma\overline{y}\right)}{(1-\tau)y^R + \gamma\tau\overline{y}}$$

$$\xi(\tau,\gamma,\omega) = \frac{\omega+1}{\tau}\xi - d(1-\gamma)(1-\omega)\left(\ln A^H - \ln A^L\right)$$

Notice that:

$$\frac{\partial C(\tau,\gamma,\omega)}{\partial\tau} = \frac{-\omega\left(-y^P + \gamma\overline{y}\right)^2}{[(1-\tau)y^P + \gamma\tau\overline{y}]^2} + \frac{-\left(-y^R + \gamma\overline{y}\right)^2}{[(1-\tau)y^R + \gamma\tau\overline{y}]^2} < 0$$

$$\frac{\partial C(\tau,\gamma,\omega)}{\partial\gamma}=\frac{\omega y^P\overline{y}}{[(1-\tau)y^P+\gamma\tau\overline{y}]^2}+\frac{\omega y^R\overline{y}}{[(1-\tau)y^R+\gamma\tau\overline{y}]^2}>0$$

and that $\xi(\tau,\gamma,w)>0$ (since $\omega\geqslant 1$, and $A^H>A^L$).

We proceed in four steps. In step 1) we show that for $\gamma=1$ and $\omega=1$, τ solving for $C(\tau,1,1)=0$ is $\tau=1$; in step 2) we prove that for any $\omega\geqslant 1$ and any γ^* satisfying $FOC_\gamma=0$ (equation 29), then

$$C(\tau,\gamma^*,\omega)+\xi(\tau,\gamma^*,\omega)>C(\tau,1,1) \tag{30}$$

In step 3) we show that given condition (30), if an equilibrium exists when $\omega\geqslant 1$, then it is given by the pair $(\tau=1,\gamma=\gamma^*)$; in step 4), we show that such an equilibrium exists and is unique; in particular, we show that when τ is optimally chosen to be equal 1, then there is a unique γ^* internal to the interval $(0,1)$ both when $\omega=1$ and when $\omega>1$; and it is $\gamma^*=\frac{1}{1+\xi}$ when $\omega=1$, and $0<\gamma^*<\frac{1}{1+\xi}$ when $\omega>1$.

Step 1. When $\gamma=1$ and $\omega=1$,

$$\begin{aligned}C(\tau,1,1)\quad=\quad&\frac{\left(-y^P+\overline{y}\right)}{(1-\tau)y^P+\tau\overline{y}}+\frac{\left(-y^R+\overline{y}\right)}{(1-\tau)y^R+\tau\overline{y}}\\&\frac{0.5\left(y^R+y^P\right)}{(1-\tau)y^P+\tau 0.5\left(y^R+y^P\right)}-\frac{0.5\left(y^R+y^P\right)}{(1-\tau)y^R+\tau 0.5\left(y^R+y^P\right)}\end{aligned}$$

Thus, it is $C(\tau,1,1)=0$ if and only if $\tau=1$.

Step 2. To prove condition (30) notice that for γ^* solving FOC_γ (equation 29), it must be:

$$\frac{\omega\overline{y}(1-\gamma^*)}{(1-\tau)y^P+\gamma^*\tau\overline{y}}+\frac{\overline{y}(1-\gamma^*)}{(1-\tau)\overline{y}+\gamma^*\tau\overline{y}}=\xi(\tau,\gamma^*,\omega) \tag{31}$$

Also notice that $C(\tau,\gamma^*,\omega)+\xi(\tau,\gamma^*,\omega)$ is the right-hand side of $FOC\tau$ (equation 28) when γ^* solves FOC_γ. Hence, substituting from equation (31), $FOC\tau$ can be written as:

$$\begin{aligned}
C(\tau,\gamma^*,\omega)+\xi(\tau,\gamma^*,\omega) &= C(\tau,\gamma^*,\omega)+\frac{\omega\overline{y}(1-\gamma^*)}{(1-\tau)y^P+\gamma^*\tau\overline{y}}+\frac{\overline{y}(1-\gamma^*)}{(1-\tau)\overline{y}+\gamma^*\tau\overline{y}} \\
&= \frac{\omega\left(-y^P+\gamma^*\overline{y}\right)}{(1-\tau)y^P+\gamma^*\tau\overline{y}}+\frac{\left(-y^R+\gamma^*\overline{y}\right)}{(1-\tau)y^R+\gamma^*\tau\overline{y}} \\
&\quad +\frac{\omega\overline{y}(1-\gamma^*)}{(1-\tau)y^P+\gamma^*\tau\overline{y}}+\frac{\overline{y}(1-\gamma^*)}{(1-\tau)\overline{y}+\gamma^*\tau\overline{y}} \\
&= \frac{\omega\left(-y^P+\overline{y}\right)}{(1-\tau)y^P+\gamma^*\tau\overline{y}}+\frac{\left(-y^R+\overline{y}\right)}{(1-\tau)y^R+\gamma^*\tau\overline{y}}
\end{aligned}$$

With such substitution, condition (30) can then be written as:

$$\frac{\omega 0.5\left(y^R+y^P\right)}{(1-\tau)y^P+\gamma^*\tau\overline{y}}-\frac{0.5\left(y^R+y^P\right)}{(1-\tau)y^R+\gamma^*\tau\overline{y}}>\frac{0.5\left(y^R+y^P\right)}{(1-\tau)y^P+\tau\overline{y}}-\frac{0.5\left(y^R+y^P\right)}{(1-\tau)y^R+\tau\overline{y}}$$

which is always satisfied when $\omega \geqslant 1$, $\gamma^* \in (0,1)$, and $y^P < y^R$.

Step 3. Step 3 is straightforward. In particular, given: *a)* that $C(\tau,\gamma,\omega)$ is decreasing in τ and increasing in γ, *b)* that $\xi(\tau,\gamma,w) > 0$ (when $\omega \geqslant 1$), *c)* that τ solving $C(\tau,1,1) = 0$ is $\tau = 1$, *d)* that condition (30) $C(\tau,\gamma^*,\omega)+\xi(\tau,\gamma^*,\omega) > C(\tau,1,1)$ holds, and *e)* that $C(\tau,\gamma^*,\omega)+\xi(\tau,\gamma^*,\omega)$ is the right-hand side of FOC_τ when γ^* solves for $FOC_\gamma = 0$, then it follows that when the optimal γ is internal, then the optimal tax rate $\tau^* = 1$.

Step 4. With step 3 we have shown that if there is a $\gamma^* \in (0,1)$ solving for $FOC_\gamma = 0$ when $\omega \geqslant 1$, then the corresponding optimal $\tau^* = 1$. We now show that such a γ^* exists and is unique in the interval $(0,1)$. In doing this we also show that a $\gamma = 0$ with a τ solving $FOC_\tau = 0$ in the interval $(0,1)$ doesn't exist. This implies that the unique political equilibrium when $\omega \geqslant 1$ is indeed of type $(\tau^* = 1, \gamma^* \in (0,1))$. Next we compute the equilibrium γ^* when $\omega = 1$, obtaining $\gamma^* = \frac{1}{1+\xi}$; and show that $\frac{1}{1+\xi}$ is larger than the equilibrium γ^* for $\omega > 1$.

To prove existence and uniqueness of γ^* in the interval $(0,1)$, first of all notice that $\lim_{\gamma\to 1} FOC(\gamma) = -\infty$. Together with the assumption that SOC_γ holds (hence $\frac{\partial FOC_\gamma}{\partial\gamma} < 0$), it follows that for the existence and uniqueness of $\gamma^* \in (0,1)$, it is sufficient to show that $FOC_\gamma(\tau = 1, \gamma = 0) > 0$. The latter is always true since:

$$FOC_\gamma(\gamma = 0) : \frac{\omega\tau\overline{y}}{(1-\tau)y^P} + \frac{\tau\overline{y}}{(1-\tau)y^R} - (\omega+1)\,\xi + d\tau(1-\omega)\left(\ln A^H - \ln A^L\right)$$

so that $\lim_{\tau\to1} FOC_\gamma(\gamma = 0) = +\infty$. More generally, we now prove that $FOC_\gamma(\gamma = 0) > 0$ also for any $\tau \in (0,1)$ chosen so as to satisfy $FOC_\tau = 0$. Since $\frac{\partial FOC(\gamma)}{\partial\gamma} < 0$ always, this excludes the possibility of an optimum when $\gamma = 0$ and τ internal in the interval $(0,1)$. When $\gamma = 0$ and τ internal, from equation (28) we have:

$$FOC_\tau(\gamma = 0) : \frac{-\tau(\omega+1)}{(1-\tau)} = -(\omega+1)\,\xi + d\tau(1-\omega)\left(\ln A^H - \ln A^L\right)$$

Substituting $-(\omega+1)\,\xi + d\tau(1-\omega)\left(\ln A^H - \ln A^L\right)$ in FOC_γ and imposing $\gamma = 0$, one obtains the $FOC_\gamma(\gamma = 0)$ with τ chosen so as to satisfy $FOC_\tau = 0$. We define such expression as $FOC_\gamma|_{\gamma=0,\tau \text{ internal}}$ as:

$$\begin{aligned} FOC_\gamma|_{\gamma=0,\tau \text{ internal}} \quad &: \quad \frac{\tau}{1-\tau}\left[\frac{\omega\overline{y}}{y^P} + \frac{\overline{y}}{y^R} - (\omega+1)\right] = \\ &= \quad \frac{\tau}{1-\tau}\left[\frac{\omega 0.5(y^P+y^R)y^R + 0.5(y^P+y^R)y^R}{y^P y^R} - (\omega+1)\right] \\ &= \quad \frac{0.5\tau}{1-\tau}\left[\frac{\omega y^R}{y^P} + \frac{y^P}{y^R} - (\omega+1)\right] \end{aligned} \tag{32}$$

The latter expression then gives:

$$FOC_\gamma|_{\gamma=0,\tau \text{ internal}} \gtreqless 0$$

$$\text{iff } \left[\frac{\omega y^R}{y^P} + \frac{y^P}{y^R} - (\omega+1)\right] \begin{cases} > 0 \text{ if either } \frac{y^R}{y^P} < 1 \text{ or } \frac{y^R}{y^P} > \frac{1}{\omega} \\ = 0 \text{ if either } \frac{y^R}{y^P} = 1 \text{ or } \frac{y^R}{y^P} = \frac{1}{\omega} \\ < 0 \text{ iff } 1 < \frac{y^R}{y^P} < \frac{1}{\omega} \end{cases} \tag{33}$$

The case $\frac{y^R}{y^P} > \frac{1}{\omega}$ applies to the present situation (since $y^R > y^P$ and $\omega \geq 1$) which then excludes the possibility of a solution with $\gamma = 0$ and τ internal when $\omega \geq 1$. We now compute the equilibrium value of γ when $\omega = 1$ and $\tau^* = 1$, and then show that it is greater than the equilibrium value of γ when $\omega > 1$. Define the function $G(\omega, \gamma)$ as:

$$G(\omega, \gamma) = \frac{\omega + 1}{\gamma} - \frac{(\omega + 1)\xi}{1 - \gamma}$$

Notice now that for $\gamma = \frac{1}{1+\xi}$, then $G(\omega, \gamma) = 0$ all ω. Hence, since $G(\omega = 1, \gamma) = FOC_\gamma(\omega = 1, \tau = 1, \gamma)$, it follows that for $\omega = 1$ the equilibrium is given by the pair $(\tau = 1, \gamma = \frac{1}{1+\xi})$. For $\omega > 1$ and $\tau = 1$, $FOC_\gamma(\omega > 1, \tau = 1, \gamma)$ is:

$$FOC_\gamma(\omega > 1, \tau = 1, \gamma) = G(\omega, \gamma) + (1 - \omega)S \tag{34}$$

where $S = d\left(\ln A^H - \ln A^L\right)$. Since $\frac{\partial FOC_\gamma}{\partial \gamma}$ is always negative and $(1 - \omega)S < 0$, it follows that γ solving for $FOC_\gamma(\omega > 1, \tau = 1) = 0$ is strictly lower than $\gamma = \frac{1}{1+\xi}$ (which solves for $FOC_\gamma(\omega = 1, \tau = 1) = 0$). The actual value of γ (< 1) solving equation 34 is given by:

$$\begin{aligned} \gamma \quad = \quad & \frac{1}{-2S(1-\omega)} \cdot [(\omega + 1)(1 + \xi) - S(1 - \omega) \\ & -\sqrt{((\omega + 1)(1 + \xi) - S(1 - \omega))^2 + 4S(1 - \omega)(\omega + 1)}] \end{aligned} \tag{35}$$

(Equation (35) has also been used to draw the graphs of $\tau\gamma$ and $\tau(1 - \gamma)$ in Fig. 2 when $\omega > 1$, since in this case $\tau = 1$).

- Case $\omega < 1$.

Firstly we show that there isn't a solution with $\tau = 1$ when $\omega < 1$. This follows by noting that when $\tau = 1$, then a solution can only exist with γ internal. This requires finding a $\widehat{\gamma}$ which solves $FOC_\gamma(\tau = 1)$. Then, it is possible to show that for such a $\widehat{\gamma}$ (and $\omega < 1$), it is always $FOC_\tau(\tau = 1, \gamma = \widehat{\gamma}) < 0$. Hence, since $\lim_{\tau \to 0} FOC_\tau = +\infty$ for any $\gamma \in (0, 1)$ and since SOC_τ holds, it follows that a solution with $\widehat{\gamma}$ internal and $\tau = 1$ cannot hold when $\omega < 1$.

Hence, a solution when $\omega < 1$ always requires FOC_τ satisfied with equality and τ internal. Reconsider then FOC_γ and also recall the expressions in equation (33) giving the conditions for $FOC_\gamma|_{\gamma=0,\tau \text{ internal}} \gtreqless 0$. Since $\frac{\partial FOC_\gamma}{\partial \gamma} < 0$ and $\lim_{\gamma \to 1} FOC_\gamma = -\infty$, the inequalities in (33) point out the conditions when the optimal γ for τ internal can be of two types: either *i)* $\gamma = 0$, in particular applying when $FOC_\gamma|_{\gamma=0,\tau \text{ internal}} < 0$, hence $\omega < \frac{y^P}{y^R}$ (see inequalities in 33); or *ii)* $\gamma \in (0,1)$, applying when $FOC_\gamma|_{\gamma=0,\tau \text{ internal}} > 0$, then $\frac{y^P}{y^R} < \omega < 1$. We now separate the analysis in sub-case *i)* and sub-case *ii).*

Sub-case *i).* When $\omega < \frac{y^P}{y^R}$, the political equilibrium is characterized by τ internal and $\gamma = 0$. The actual value of τ solving the problem can be found simply imposing $\gamma = 0$ in the FOC_τ and solving for $FOC_\tau = 0$. This gives:

$$FOC_\tau \colon \frac{-(\omega+1)}{(1-\tau)} + \frac{(\omega+1)\,\xi}{\tau} - d\left(\ln A^H - \ln A^L\right)(1-\omega) = 0$$

When $d = 0$, $\tau = \frac{\xi}{1+\xi}$; on the other hand, given that $\frac{\partial FOC_\tau}{\partial \tau} < 0$, when $d > 0$ the optimal τ is lower than $\frac{\xi}{1+\xi}$. More specifically, it is given by:

$$\begin{aligned}\tau \;=\; & \frac{1}{2S(1-\omega)} \cdot [(\omega+1)(1+\xi) + S(1-\omega) \\ & -\sqrt{((\omega+1)(1+\xi) + S(1-\omega))^2 - 4S(1-\omega)(\omega+1)\xi}]\end{aligned} \tag{36}$$

where $S = d\left(\ln A^H - \ln A^L\right)$. (Since in this sub-case $\gamma = 0$, equation 36 has also been used to draw the graph of $\tau(1-\gamma)$ in Fig. 2 when $\omega < \frac{y^P}{y^R}$).

Sub-case *ii).* When $\frac{y^P}{y^R} < \omega < 1$, the equilibrium is with both γ and τ internal. Hence, both FOC_τ and FOC_γ must equal 0, namely:

$$\begin{aligned} FOC_\tau \;:\; & \frac{\tau\omega\left(-y^P + \gamma\overline{y}\right)}{(1-\tau)y^P + \gamma\tau\overline{y}} + \frac{\tau(-y^R + \gamma\overline{y})}{(1-\tau)y^R + \gamma\tau\overline{y}} = -(\omega+1)\,\xi \\ & + d(1-\gamma)\tau(1-\omega)\left(\ln A^H - \ln A^L\right) \end{aligned} \tag{37}$$

$$\begin{aligned} FOC_\gamma \;:\; & -\frac{(1-\gamma)\omega\tau\overline{y}}{(1-\tau)y^P + \gamma\tau\overline{y}} - \frac{(1-\gamma)\tau\overline{y}}{(1-\tau)y^P + \gamma\tau\overline{y}} = -(\omega+1)\,\xi \\ & + d(1-\gamma)\tau(1-\omega)\left(\ln A^H - \ln A^L\right) \end{aligned} \tag{38}$$

Substituting gives:

$$\frac{\tau\omega\left(-y^P + \gamma\overline{y}\right)}{(1-\tau)y^P + \gamma\tau\overline{y}} + \frac{\tau(-y^R + \gamma\overline{y})}{(1-\tau)y^R + \gamma\tau\overline{y}} = -\frac{(1-\gamma)\omega\tau\overline{y}}{(1-\tau)y^P + \gamma\tau\overline{y}} - \frac{(1-\gamma)\tau\overline{y}}{(1-\tau)y^P + \gamma\tau\overline{y}}$$

and after few manipulations:

$$\frac{\omega}{(1-\tau)y^P+\gamma\tau\overline{y}}=\frac{1}{(1-\tau)y^R+\gamma\tau\overline{y}}$$

Substituting in FOC_τ and FOC_γ, one obtains:

$$\tau\gamma = \frac{(1-\tau)\cdot(\omega y^R-y^P)}{\overline{y}(1-\omega)} \tag{39}$$

$$\tau(1-\gamma) = \frac{\xi 2\overline{y}}{2\overline{y}(1+\xi)+S(1-\tau)(y^R-y^P)} \tag{40}$$

(where $S=d\left(\ln A^H-\ln A^L\right)$) .

Expressions (39) and (40) can be studied as functions of the optimal τ and of the other parameters. In particular, after few more manipulations one obtains a quadratic equation for τ :

$$\begin{aligned} h(\tau) &= -\tau^2 S(y^R-y^P)^2 0.5(\omega+1) \\ &\quad +\tau[(y^R-y^P)[(1+\xi)\overline{y}(1+\omega)+S((\overline{y}(1-\omega)+2(\omega y^R-y^P))]] \\ &\quad -\xi\overline{y}(\omega+1)(y^R-y^P)-(\omega y^R-y^P)(2\overline{y}+S(y^R-y^P)) \end{aligned}$$

Regarding this function we note the following:

1. $h(0)<0$ and $h(1)>0$: hence the function has a single root (the optimal τ^*) in the interval $\tau\in(0,1)$; in addition, at τ^*, $\partial h(\tau^*)/\partial\tau>0$.

2. Further, it is possible to show that at τ^*: $\partial h(\tau^*)/\partial\omega<0$, $\partial h(\tau^*)/\partial d>0$, and $\partial h(\tau^*)/\partial\xi<0$. Since $h(\cdot)$ is increasing at τ^*, the signs of the above derivatives imply, in the order, that: $\partial\tau^*/\partial\omega>0$, $\partial\tau^*/\partial d<0$, and $\partial\tau^*/\partial\xi>0$.

3. From equations (39) and (40) one can then obtain the signs for the derivatives:

$$\frac{\partial\tau\gamma}{\partial\omega}=\frac{\partial\tau}{\partial\omega}[1-\frac{S(y^R-y^P)}{[2\overline{y}(1+\xi)+S(1-\tau)(y^R-y^P)]^2}]>0 \qquad \frac{\partial\tau(1-\gamma)}{\partial\omega}=\frac{S(y^R-y^P)}{[2\overline{y}(1+\xi)+S(1-\tau)(y^R-y^P)]^2}>0$$

$$\frac{\partial\tau\gamma}{\partial d}=-\frac{(\omega y^R-y^P)}{\overline{y}(1-\omega)}\frac{\partial\tau}{\partial d}>0 \qquad \frac{\partial\tau(1-\gamma)}{\partial d}=\frac{\partial\tau}{\partial d}[1-\frac{(\omega y^R-y^P)}{\overline{y}(1-\omega)}]<0$$

$$\frac{\partial\tau\gamma}{\partial\xi}=-\frac{(\omega y^R-y^P)}{\overline{y}(1-\omega)}\frac{\partial\tau}{\partial\xi}<0 \qquad \frac{\partial\tau(1-\gamma)}{\partial\xi}=\frac{\partial\tau}{\partial\xi}[1-\frac{(\omega y^R-y^P)}{\overline{y}(1-\omega)}]>0$$

which are used to draw the solid lines in Fig. 2 when $\frac{y^P}{y^R} < \omega < 1$ (and $d > 0$).

Alternatively, the same predictions can be obtained studying the explicit forms for $\tau\gamma$ and $\tau(1-\gamma)$, which can also be obtained from equations (39) and (40). In particular, after few manipulations, one can obtain the following quadratic equations for $\tau\gamma$ and $\tau(1-\gamma)$, respectively:

$$\begin{aligned} 0 &= -(\tau\gamma)^2 S(1-\omega)(y^R - y^P)[(\omega y^R - y^P) + (1-\omega)\overline{y}] \\ &\quad -\tau\gamma(\omega y^R - y^P)(y^R - y^P)[(1+\xi)(\omega+1) - S(1-\omega)] \\ &\quad +2(\omega y^R - y^P)^2 \qquad (41) \\ 0 &= (\tau(1-\gamma))^2 S(1-\omega) - \tau(1-\gamma)\left[(1+\xi)(1+\omega) + S(1-\omega)\right] \\ &\quad +\xi(1+\omega) \qquad (42) \end{aligned}$$

Regarding the latter equation for $\tau(1-\gamma)$ also notice that since it does not depend on the pair (y^R, y^P), but only on the parameters, it is time-invariant. (Indeed, equation 42 delivers for $\tau(1-\gamma)$ the same root as both equation 35 — holding when the optimal solution is with $\tau = 1$—, and equation 36 — holding when in the optimal solution $\gamma = 0$). Finally, also notice that when $d = 0$ (hence $S = 0$), equations (41) and (42) give:

$$\tau\gamma = \frac{2(\omega y^R - y^P)}{2(\omega y^R - y^P) + (\omega+1)\xi(y^R - y^P) + (y^R + y^P)(1-\omega)}$$
$$\tau(1-\gamma) = \frac{\xi}{1+\xi}$$

which are used to draw the dotted lines in Fig. 2.

Proof of Proposition 2.

Substitute from Proposition 1 the political equilibrium values of e_t^J in equation (17) for the different time-invariant ω^J. For example, for $\omega^N = 1$, $e_t^N = \frac{\xi}{1+\xi}\overline{y}_t^N$ so that $\overline{y}_{t+1}^N = 0.5(A^L + A^H)(\frac{\xi}{1+\xi})^{\xi}(\overline{y}_t^N)^{\xi+\delta}$, and $B^N = 0.5(A^L + A^H)(\frac{\xi}{1+\xi})^{\xi}$ in equation (18). Similar substitutions when $\omega^P > 1$ and $\omega^R < 1$ imply $B^P > B^N > B^R$. The rest of the Proposition follows from basic properties of difference equations. (See also Fig. 3).

Proof of Proposition 3

Substitute for the generic time-invariant $\frac{e_t}{y_t}$ of equation (13), the specific time-invariant $\frac{e_t^J}{y_t^J}$=$(1-\gamma^J)\tau^J$ derived from Proposition 1 under regimes $J = N, P, R$.

Proof of Proposition 4

Directly from: i) Proposition 2 showing that, under all possible combinations of $\xi+\delta$ (whether greater, lower or equal to 1), then $\frac{\overline{y}_t^P}{\overline{y}_{t-1}^P} > \frac{\overline{y}_t^N}{\overline{y}_{t-1}^N} > \frac{\overline{y}_t^R}{\overline{y}_{t-1}^R}$ for at least some t (otherwise they may be equal); ii) Proposition 3 indicating that $\overline{\alpha}_{t+1}^P > \overline{\alpha}_{t+1}^N > \overline{\alpha}_{t+1}^R$ at all $t = 1, 2, ...$; iii) equation (20) for the evolution of I_{t+1}^J.

Proof of Proposition 5

From Equation (12) and Proposition (1).