

Oligarchy, democracy, inequality and growth

François Bourguignon^{*}, Thierry Verdier¹

*Département et Laboratoire D' Economie Théorique & Appliquée, 48, Bd. Jourdan, ENS,
Paris 75014, France*

Abstract

This paper analyses the dynamics of inequality, democratization and economic development in a political economy model of growth where education is both the engine of growth and a determinant of political participation. In a context with imperfect capital markets, we investigate the incentives for an educated oligarchy to subsidize the poor's education and to initiate a democratic transition. We characterize the equilibrium patterns of political institutions, income distribution and growth as a function of the initial income and inequalities. In particular, we identify circumstances under which the Elite promotes the endogenous emergence of a middle class for purely political economy reasons. A simple linear infinite horizon framework is then presented. In this setting, we discuss the importance of historical dependence for long-run social stratification and redistribution. © 2000 Elsevier Science B.V. All rights reserved.

JEL classification: O11; O15; D72

Keywords: Inequality; Economic development; Political economy; Redistribution; Democratization

1. Introduction

A substantial part of the recent literature on the relationship between income inequality and growth is based on a very simple political economy mechanism through which a pivotal voter in society decides on the value of some redistribu-

^{*} Corresponding author. Tel.: +33-1-4313-6300; fax: +33-1-4313-6310.

¹ Ehess and Delta, and Ceras and Delta, Paris, respectively. Delta is a joint research unit of CNRS, Ehess and ENS.

tive policy instrument, which in turn determines directly or indirectly the rate of growth of the economy. As the redistribution decision generally depends on the distance between the income or wealth of the pivotal agent and the mean income in the economy, the democratic rule explains that growth depends in effect on the primary distribution of income among agents. For instance, inegalitarian societies are shown to generate more redistribution than less inegalitarian ones, but because redistribution goes through the tax system, it may diminish incentives to invest and slow down economic growth. It is this kind of mechanism that one finds in models like those of Bertola (1993), Alesina and Rodrik (1994), Persson and Tabellini (1994) and others. However, the opposite may also hold if progressive redistribution actually permits increasing the investment rate in the economy. This is the case when this is done through an increase in the budget of public education, as in Saint-Paul and Verdier (1993) or if redistribution helps beneficiaries to overcome the effects of some capital market imperfection or liquidity constraint which was initially preventing them to invest in profitable projects or in human capital. This is the type of mechanisms put forward in the models by Galor and Zeira (1993), Perotti (1993), and potentially in all the models built along the lines of Banerjee and Newmann (1991) and in particular Benabou (1996), Piketty (1997) and Aghion and Bolton (1998).

Most political economy equilibrium models of the previous types rely on some form of the median voter theorem and often consider pure steady state situations where the distribution of (primary) incomes is constant, and independent of the rate of growth of the economy. The causality then runs in the direction of inequality determining growth through income redistribution which is itself ultimately decided by the median voter. Other models are based on less simple mechanisms where the initial distribution of income determines growth which produces itself some transformation in the distribution of income. If, at each stage, some redistribution decision is taken by the median voter who finds him/herself at a varying distance from the mean income, as for instance in Perotti (1993), then the dynamics of the economy becomes indeed much more complex and the causality relationship between the inequality of the primary distribution of income and the growth rate of the economy runs in both directions, at least as long as the economy has not reached some steady state where both the growth rate and the distribution of income are simultaneously determined by the fundamental parameters of the economy as well as the political decision mechanism behind redistribution decision.

In the present paper, we explore yet another complication of the original model, which is that the political economy decision mechanism may itself be endogenous to the growth process. Benabou (1996) offers a very nice generalization of many of the existing models in the literature by assuming that the pivotal agent in the society is not necessarily the median voter as in the standard democratic majority voting mechanism but an agent with some rank λ in the distribution of income or personal resources. He then shows that societies where λ is very different from

1/2 — the median voter or strict democratic case — may in some instances behave in terms of growth and redistribution very differently from what would imply the median voter model.² This is an important point, but then the problem becomes that of what determines the value of λ and how it might change over time. This is precisely the issue we tackle in the present paper by investigating models where λ changes endogenously during the growth process, or, more fundamentally, where political institutions change in accordance with the will of those political actors who happen to control them at some point of time. All these are done very simply by assuming that political participation depends on the educational level of economic agents, which is itself indirectly under the strict control of an oligarchy.³

A limited amount of work has been done so far to endogenize political participation in growth models. Ades and Verdier (1996) consider that entry into the political Elite is costly and determined by an exogenous fixed cost. Political participation depends then on the individual's ability to pay and characteristics of the "political technology". Gradstein and Justmann (1995) assume alternatively the existence of an exogenous (relative) income franchise. The growth of income of successive generations therefore determines the evolution of the number of voters. Acemoglu and Robinson (1996) are closer in spirit to the present paper. They study the incentives for the rich in power to enlarge the income franchise because of a threat of insurrection (and expropriation) by the poor when inequalities are too big. Here we also explicitly recognize that democratization is, to some extent, decided by the people already in power but we take the alternative route that political participation and activity depend on the educational level — or the income — of economic agents. It follows that the political Elite control, at the same time, the growth process of the economy, the evolution of the distribution of income and that of the structure of political power. It may adopt a pure predatory behavior, in the sense of Robinson (1997), and simply oppose the development of education or any other type of modernization that would represent a future threat to its political power. It may also be in its interest to strategically "promote" an educated middle class that would accelerate economic growth and minimize the danger of future expropriation. Finally, it may find it advantageous to give up political control and to progressively establish democracy.

Even though most of the arguments in the present paper rely on the simple assumption that political participation is solely determined by the educational level, or more generally by the socioeconomic status of citizen — a hypothesis which has been subject, rather successfully, to intensive testing in political

² With Benabou's notations, the median voter value for α is 0 rather than 1/2.

³ Gradstein and Justmann (1995) have recently derived from this relationship between political participation and income a justification of the well known Kuznets curve relating the extent of income inequality and the development level of a society.

sociology⁴ — and also that human capital accumulation is the sole engine of growth, we provide a framework able to describe the complex interaction between the distribution of income, the political system and economic growth. This offers interesting clues to understand some important features in the process of economic and political development. In particular, we show how, depending on initial conditions, some developing economies seem to be locked in a low-income, slow-growth and politically authoritarian regimes, whereas others find themselves on a dynamic path characterized by rapid economic growth, democratization and the emergence of a middle class. The former description conforms rather well to several countries in Africa, but there also are obvious examples in Asia or Latin America. The latter would fit better the recent experience of some fast-growing East Asian countries, although the extent of the democratization process is often difficult to appreciate, especially in countries facing external difficulties like Taiwan or South Korea.

Other illustrations of the basic mechanisms analyzed in this paper may be found in history. For instance, the resistance to developing economic and cultural relations with the Western world in pre-Revolutionary China was partly motivated by the fear that this might inoculate the democracy virus in the population. Interestingly enough, it also partly explains the outbreak of the revolution. Examples of the development of an educated middle class aimed at both fostering the economic interest of a ruling Elite and protecting its political power are also numerous. At one stage or another, such a strategy has been used by most colonial powers in the countries under their control. The examples of development of “Oxbridge” raised local Indian Elites by England or the promotion of the “Pieds noirs” community in North Africa by France are well known examples. Within industrial countries, education was also sometimes seen as part of a strategy to keep political control while accelerating economic progress. In end of 19th century France, for instance, the main objective of Jules Ferry and politicians of that time was not only to diminish the religious influence and power of the Church but also to build up a class of middle income and white collar people less inclined to revolutions and other redistributive conflicts.⁵ But, of course, this could also be interpreted as simply one intermediate step in the progressive and irreversible process of democratization observed in industrial countries since the end of the 19th century.

The model explored in the present paper is also helpful in understanding the many attempts by economists, sociologists and political scientists at identifying a

⁴ See, in particular, Verba (1990), Verba et al. (1978), Brady et al. (1995). On the role of income as a determinant of political participation, see also the debate around the early paper by Frey (1971) in *Public Choice* (Fraser (1972), Russel (1972) and Frey (1972)).

⁵ Another widely recognized reason for the endogenous emergence of politically powerful middle classes is urbanization and the formation of trade unions during the industrial revolution (Huber et al., 1993).

significant empirical relationship between the extent of democracy in a country, its rate of economic growth and the degree of income inequality. The simple model we analyze in what follows suggests that this relationship is indeed much more complex than implicitly assumed in simple cross-sectional regressions where essentially one of the three variables appears on the left-hand side and the two others on the right-hand side.⁶

Finally, it must be noted that the mechanisms analyzed here in the context of education and political rights are in fact more general than it appears. As suggested by the historical examples mentioned above, they are also present in other political economy contexts like, e.g. trade reforms or land reform. As a matter of fact, the arguments developed in this paper are valid for any economic reform or policy which increases the economic payoff of the incumbent Elite but also, at the same time, reduces its political power by enabling new segments of society to be politically effective and ask for downward redistribution.

The paper is organized as follows. The basic setting of the model, inspired from Perotti (1993), is presented and discussed in Section 2 which also analyzes a simple two-period version of the model. A more general dynamic formulation which emphasizes the transitory role of a middle class in leading the society towards full democracy is briefly considered in Section 3. Section 4 concludes by considering the possibility of broadening the scope of the model so as to include in it, on one hand, the possibility of costly repression by the ruling oligarchy and, on the other hand, the accumulation of physical capital.

2. A two-period model of democracy, inequality and growth

2.1. General setting of the model

Consider an economy with population scaled to unity and initially two classes of individuals. The oligarchy consists of individuals who have inherited from their educated parents: (a) higher earnings abilities; (b) the desire to influence and participate in public choices through voting. It is assumed that this oligarchy represents a minority, $1 - p$, of the population with $p > 1/2$. All members of this class are identical and their initial earnings is y^f . The rest of the population, in proportion p , are born of uneducated parents; their earnings y^p is lower than that

⁶ There is indeed a rather huge literature in this field. Useful recent surveys of the relationship between the nature of the political system, on one hand, and economic growth or inequality, on the other hand, include Sirowy and Inkeles (1990) and Przeworski and Limongi (1993). For an analysis of the role of democracy in the spirit of the recent empirical growth literature, see also Barro (1996). The complete growth–democracy–inequality relationship has been analyzed inter alia by Persson and Tabellini (1994), Perotti (1996) or Benabou (1996). A critical analysis emphasizing the complexity of the relationship between democratization and development is found in Bardhan (1997).

of the oligarchy ($y^p < y^r$) and they are assumed not to participate in the political decision process.

Political participation in the first period — i.e. period A — is thus fully exogenous. It is determined by parents' education. During the second period — i.e. period B — it will be assumed to depend on the level of education that the individuals will have acquired during the preceding period. As the decision to become educated is endogenous, so will be political participation in period B. This is the main feature of this simple model.

This assumption about the link between political participation and education — or equivalently in the present framework income level — has been the subject of a good deal of investigation by sociologists and political scientists and seems to be confirmed by available evidence⁷, although it could certainly not be taken as a law. Examples of massive political participation in poorly educated countries do exist. Even in this case, however, it may be held that, through the diffusion of their own ideas and views in the less educated population, the vote of educated people 'weighs' more than that of others. This is the stylized fact that we seek to represent here by this simple relationship between voting participation and education. It must also be noted that the following model would also be consistent with the hypothesis that political participation is determined by the level of income, rather than education.

The basic framework of the model is borrowed from Perotti (1993). All individuals during the first period must make a decision on whether they get educated or not. There is a fixed cost to education which is arbitrarily set to unity, and there are two types of returns to education. The private return amounts to R . It accrues to individuals in the second period only and it is assumed to be larger than the initial cost ($R > 1$). Education also yields a 'public' return in the sense that all individual earnings in the second period, educated or not, are augmented by an amount equal to μE , where E is the overall proportion of educated people in the population during the second period and μ is a positive coefficient. This is a simple way of accounting for an educational externality à la Lucas.

There is no capital market where individuals who do not have enough in the first period to cover the fixed cost of education may borrow. This may reflect pure moral hazard problems linked to borrowing against human capital, or other capital market imperfections. For the model to be of some interest, we assume that the poor do not invest in education because they are liquidity-constrained whereas members of the oligarchy are not. In other words:

$$y^p < 1 < y^r.$$

One possibility at this stage would be that some redistribution be decided for an amount $1 - y^p$ per capita which would permit the poor to become educated, thus

⁷ See, for instance, Brady et al. (1995).

leading to fast growth, possibly at the cost of some distortion due to redistribution. This is the route taken by Perotti who assumes that the poor can impose, in some circumstances, such a redistribution to the oligarchy through ‘universal’ majority voting. In a different framework where education is public, this is also what is done by Saint-Paul and Verdier (1993). In the present model, however, the poor are assumed to be politically inactive so that they cannot impose any redistribution, at least in the first period. Thus, if some redistribution takes place, it must be only because it is in the interest of the oligarchy. This will prove to be the driving mechanism of the whole model.

Members of the oligarchy will always get educated because they can afford it ($y^r > 1$) and because it is assumed to be privately profitable ($R > 1$) — to simplify the notation, we assume here no discount rate between the two periods A and B. They may also subsidize the education of some members of the lower class by transferring to them the amount they need, $1 - y^p$. If they decide to help a number e of poor, they need to transfer to them $T = e(1 - y^p)$. However, we shall assume that there is some efficiency cost linked to any transfer in the economy which is proportional to the square of the share of the transfer in the total income of the population. In order to transfer T to the lower class, the oligarchy needs to disburse:

$$\left[\frac{T}{\bar{y}} + a \left(\frac{T}{\bar{y}} \right)^2 \right] \bar{y}, \tag{1}$$

where \bar{y} is the mean income of the population in period A:

$$\bar{y} = py^p + (1 - p)y^r. \tag{2}$$

But, of course, the oligarchy will actually make this transfer if the return that it will get in the second period from the education externality, μe , is larger than the amount (Eq. 1) it has to pay.

There is another cost in subsidizing the education of individuals in the lower class. It comes from the fact that, by doing so, the oligarchy weakens its future political power because new voters will appear in the second period. If they are numerous enough, newly educated poor shall be able to pass a law that will redistribute income away from the oligarchy to the rest of the population. Again, this might not be a problem for the oligarchy if the educational externality is large enough so as to cover the loss they will incur in the redistribution that will follow their loss of political power by voting dilution. But it may also be a net loser in this process, in which case it will retain political power while benefiting from the educational externality by educating a number of poor which will not change the voting majority, or the median voter. By doing so, the oligarchy will favor the appearance of a middle class, while reducing the growth rate of the economy in comparison with what it could have been if all the poor had received education.

This line of modeling should provide interesting insights into the interdependence between political and economic development. The most interesting feature of the model, however, is that this interdependence is strongly affected by the initial distribution of income and the size of the oligarchy. Indeed, the more unequal the initial distribution of income, the more the oligarchy will have to lose if the lower class gets educated and gains political control. It may thus be expected that unequal oligarchic societies tend to develop and democratize more slowly than more equal societies.^{8,9}

2.2. Detail of the model

Let us make a change of variables that will put into evidence in all what follows the crucial role of the initial income inequality between the oligarchy and the poor. Let x be the difference between the initial earning levels of both classes. Together with the definition (Eq. 2) of the mean income, it then comes that:

$$x = y^r - y^p; \quad y^p = \bar{y} - (1 - p)x; \quad y^r = \bar{y} + px. \quad (3)$$

We now analyze in more detail the behavior of the various agents and derive from it the optimal behavior of the oligarchy in period A, i.e. the proportion of people, e , whose education it decides to subsidize. To do so, we start from the situation in period B. Two cases must be considered. Either e is small enough so that the oligarchy retains its political control, or the newly educated class reaches a voting majority. We consider these two cases in turn.

2.2.1. Case (i): the oligarchy retains political power: $e < 1 - p$

With a zero discount rate, the total income of oligarchy members over the two periods as a function of the number of educated poor, e , is given by:

$$\bar{Y}(e) = \left[(y^r - 1) - \frac{e(1 - y^p) + ae^2(1 - y^p)^2/\bar{y}}{1 - p} \right] + [y^r + R + \mu(1 - p + e)],$$

⁸ The result on inequality leading to slower economic development is in agreement with the recent literature which attempts to link inequality and growth through political economy mechanisms — see in particular Alesina and Rodrik (1994) and Persson and Tabellini (1994). The contribution of this paper is to make the latter endogenous.

⁹ It may be stressed that the economy, which has just been described, is clearly inefficient. Indeed, efficiency would require that all poor be educated — at least if the transfer cost parameter, a , is not too high — since both rich and poor would be better off in period B. But this would be possible only if the poor could commit themselves not to redistribute away from the rich in the second period, or in other words, to leave the political control to the oligarchy even though they could overthrow it. We assume here that such a commitment not to exercise political control when feasible is not possible.

which, given Eq. 3, may also be written in terms of x as:

$$\bar{Y}(e) = \left[(\bar{y} - 1 + px) - e \frac{1 - \bar{y} + (1 - p)x}{1 - p} - ae^2 \frac{(1 - \bar{y} + (1 - p)x)^2 / \bar{y}}{1 - p} \right] + [\bar{y} + px + R + \mu(1 - p + e)]. \tag{4}$$

The two terms in square brackets correspond to the net income in the two periods. The first term is the income during period A net of own education costs and the cost per oligarchy member of the transfer necessary to educate e poor. The second term corresponds to the direct income in period B augmented by the educational externality. According to this expression, the oligarchy will actually subsidize the education of some poor if the marginal benefit of it is larger than the cost at $e = 0$, i.e. if:

$$\mu > \frac{1 - y^p}{1 - p} = \frac{1 - \bar{y} + (1 - p)x}{1 - p}. \tag{5}$$

If this condition holds, then it will choose the value e^* of e which maximizes its total income:

$$e^* = \frac{\mu(1 - p) - (1 - \bar{y} + (1 - p)x)}{2a(1 - \bar{y} + (1 - p)x)^2 / \bar{y}}, \tag{6}$$

provided of course that it retains political control in period B, i.e. $e^* \leq 1 - p$.

2.2.2. Case (ii): the oligarchy loses political control in the second period: $e \geq 1 - p$

The oligarchy will then be unable to veto a redistribution policy voted by the middle class, i.e. the e poor who have been educated during period A. We assume that redistribution policies are all linear with a flat marginal tax rate, τ , and a lump sum uniform transfer equal to:

$$c = \tau \bar{y}_B (1 - a\tau), \tag{7}$$

where, \bar{y}_B , the mean income in the population in period B, is given by:

$$\bar{y}_B = \bar{y} + (\mu + R)(1 - p + e). \tag{8}$$

As before, the bracketed term in Eq. 7 accounts for the distortionary cost of the tax and/or transfer.¹⁰ Thus, the redistribution which will be voted by the middle

¹⁰ To simplify, we assume the same distortion for the education-oriented redistribution in period A and the pure redistribution policy in period B. As the identity of transferors and transferees is not the same in both cases, this need not be so.

class is the tax rate and associate transfer (Eq. 7) which maximizes its income in period B, i.e. the solution of:

$$\begin{aligned} \max_{\tau} Z(\tau, e) &= [\bar{y} - (1 - p)x + R + \mu(1 - p + e)](1 - \tau) + \tau(1 - a\tau) \\ &\quad \times [\bar{y} + (\mu + R)(1 - p + e)], \end{aligned} \tag{9}$$

where the first term in square bracket is the income of the middle class after tax and the second term corresponds to the lump sum transfer c . The solution of this problem is given by:

$$\begin{aligned} \tau &= 0 && \text{if } x(1 - p) \leq R(p - e), \\ \tau &= \frac{x(1 - p) - R(p - e)}{2a[\bar{y} + (\mu + R)(1 - p + e)]} = \tau^*(e) && \text{if } x(1 - p) > R(p - e). \end{aligned} \tag{10}$$

The first condition in this system says that, when the decisive agent is in the middle class, he/she will vote in favor of redistribution only when the net transfer he/she may receive from it is positive. Given that the lump sum transfer in the redistribution mechanism depends on the mean income of the population, it is not difficult to see that the middle class is in favor of some redistribution if and only if its income is below the mean income of the whole population in period B, \bar{y}_B . The gap between both incomes is:

$$\Delta = [\bar{y} + (R + \mu)(1 - p + e)] - [\bar{y} - (1 - p)x + R + \mu(1 - p + e)]. \tag{11}$$

It may be seen that, after simplification, the first part of Eq. 10 means that no redistribution takes place if this gap is negative. If it is positive, the second part then says that the tax rate which is preferred by the middle class depends precisely on the size of this gap — expressed relatively to the mean income of the population, \bar{y}_B — and on the distortionary cost of the tax.

A point which will be of some importance below is whether the optimal tax rate above is increasing or decreasing with the size of the middle class, e . In the absence of the educational externality, any increase in the proportion of educated people would clearly raise the mean income of the population without changing individual incomes within a particular class. It follows that the tax rate, τ^* , chosen by the middle class would increase with its size, provided that its income is below the mean income of the whole population. However, the educational externality causes the relative gap between the income of the middle class and the mean income of the population to decrease with e . Putting these two effects together, it may be expected that the optimal tax rate chosen by the middle class

will be an increasing function of its size, e , only if the educational externality, μ , is small enough.¹¹ More rigorously, it may be shown that:

$$\frac{d\tau^*(e)}{de} > 0 \quad \text{if } R - (1 - p)x \geq 0$$

or

$$R - (1 - p)x < 0 \quad \text{and} \quad 1 + \frac{\mu}{R} < \frac{\bar{y}}{(1 - p)x - R}, \tag{12}$$

where $R - (1 - p)x$ is the difference between the income of the middle class and the mean income of the population exclusive of all educational returns.

We may now come back to the optimizing behavior of the oligarchy over the two periods. In case the middle class does not decide to impose any redistribution, i.e. in the first case of the system (Eq. 10), the total income of the oligarchy is the same as in case (i) above, i.e. Eq. 4. In the second case, it may be shown that it is equal to:

$$\begin{aligned} \tilde{Y}(e) &= \bar{Y}(e) - L(e) \quad \text{with} \\ L(e) &= \tau^*(e)[px + R(p - e)] - a\tau^*(e)^2[\bar{y} + (R + \mu)(1 - p + e)], \end{aligned} \tag{13}$$

where $L(e)$ is the loss for the oligarchy arising from the redistribution policy $\tau^*(e)$ imposed by the middle class, and given by Eq. 10.

Putting cases (i) and (ii) together, the total income of the oligarchy as a function of the number of poor whose education it decides to subsidize is thus given by:

$$Y(e) = \begin{cases} \bar{Y}(e) & \text{for } e \leq \max[1 - p, e_\tau] \\ \tilde{Y}(e) & \text{for } e > \max[1 - p, e_\tau] \end{cases} \quad \text{with } e_\tau = p\left(1 + \frac{x}{R}\right) - \frac{x}{R}. \tag{14}$$

The critical value e_τ in this expression corresponds to the size of the middle class such that its income is equal to the mean income of the population in period B, and its optimal marginal tax rate, τ^* , is zero — see Eq. 10. If the conditions in Eq. 12 hold, the tax rate becomes strictly positive and then increases with e above this threshold.¹²

It now remains to maximize the preceding function with respect to e . Fig. 1 shows the various possible shapes of $Y(e)$ according to whether the switch

¹¹ This argument and the last part of the second condition in Eq. 12 depends very much on the assumption that the education externality is uniform and independent of own human capital. That the optimal tax rate decided by the middle class be a decreasing function of e would be less likely if the externality were smaller on uneducated than educated people.

¹² It is assumed in Eq. 14 that the tax rate imposed by the middle class cannot make the educational investment by the oligarchy privately unprofitable.

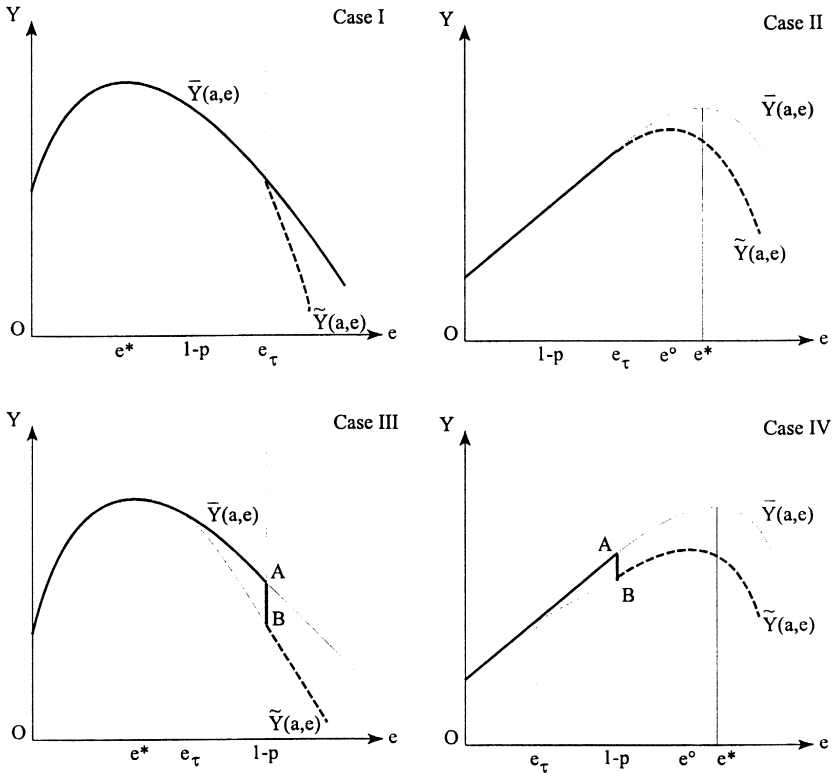


Fig. 1.

between $\bar{Y}(e)$ and $\tilde{Y}(e)$ in Eq. 14 occurs at e_τ or $(1 - p)$. It is drawn in the case where e_τ is positive so that condition (12) above holds and the tax rate $\tau^*(e)$ decided by the middle class as well as the cost inflicted to the oligarchy, $L(e)$, increases with e .¹³ The function $\bar{Y}(e)$ is a parabola with a maximum, e^* , given by Eq. 6 — or possibly at zero if condition (5) does not hold. The function $\tilde{Y}(e)$ coincides with $\bar{Y}(e)$ below e_τ and then diverges increasingly from it. It follows that the function $Y(e)$ is discontinuous at $(1 - p)$ when $e_\tau < (1 - p)$, as in cases (iii) and (iv). It can be seen in Fig. 1 that the income of the oligarchy is the highest either at the maximum e^* of $\bar{Y}(e)$ — cases (i) and (iii) — or at the maximum e^o of $\tilde{Y}(e)$ — as in case (ii). There is only one ambiguous case — case (iv) — where the discontinuity of the function $Y(e)$ makes it possible that the maximum occurs at $(1 - p)$ rather than at e^o or e^* .

¹³ The case where e_τ is negative does not lead to fundamentally different shapes for the function $Y(e)$.

Fig. 1 leads to distinguishing the following situations according to the political structure of the economy in period B and to the growth of the mean income of the population between the two periods, which essentially depends on the model on the number e of poor that the oligarchy decides to educate in period A.

2.2.2.1. (α) *Pure oligarchy and no growth.* This is the case where $\bar{Y}(e)$ is decreasing at $e = 0$ because condition (5) does not hold.

2.2.2.2. (β) *Oligarchy with a minority middle class and medium growth.* The maximum of $Y(\cdot)$ occurs on $\bar{Y}(e)$ at $e^* < 1 - p$ as in cases (i) or (iii).

2.2.2.3. (γ) *Ruling oligarchy with a middle class of (almost) equal size and medium growth.* This is the case where the maximum in case (iv) occurs at the discontinuity point A. The oligarchy would like to educate more poor people but, by giving them the political majority, it would then lose in the redistribution process that the new majority would impose.

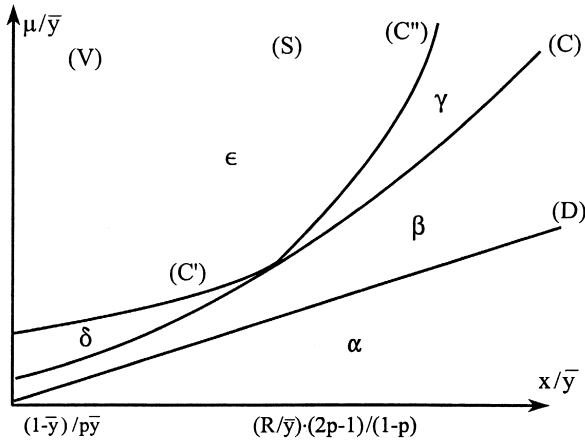
2.2.2.4. (δ) *Democracy with an accommodating ruling middle class and medium growth.* The maximum of $Y(\cdot)$ occurs on $\bar{Y}(e)$ at e^* between $1 - p$ and e_τ in case (i). The oligarchy accepts to lose political control in favor of the middle class because it knows that the latter will not exercise its redistributive power.

2.2.2.5. (ε) *Democracy, income redistribution and fast growth.* The maximum of $Y(\cdot)$ occurs on $\bar{Y}(e)$ at e^0 as in case (ii) or in case (iv). The oligarchy accepts to lose political control despite the redistribution imposed by the ruling middle class, presumably because the educational externality is large enough.

The solution of the model thus covers a rather wide range of situations, from the stagnating economy organized around an oligarchy seeking to maintain full control of the society to the developed economy controlled by an educated middle class imposing a more or less severe redistribution from the oligarchy. We labeled the latter situation — i.e. cases (δ) and (ε) — ‘democracy’ even though it may be rather far from universal voting. It may happen, in particular, that the ‘middle class’ is numerically only slightly bigger than the oligarchy, democratization corresponding then to a mere extension of the ruling Elite. These various situations maximize the intertemporal welfare of the oligarchy — in the two-period setting considered in this section — for some configurations of the parameters of the model and initial conditions of the economy. We now turn to an analysis of these configurations.

2.3. Comparative statistics

Among the determinants of which solution (α) to (ε) actually holds, we first single out in what follows the educational externality, μ , which measures the



- solution α : Pure oligarchy regime
- solution β : Minority middle class regime
- solution γ : Balance of power regime
- solution δ : Accomodating ruling middle class regime
- solution ϵ : Democracy regime

Fig. 2.

benefit the oligarchy may get from educating the poor, and, on the other hand, the initial income inequality, x , which defines the extent of redistribution which will be undertaken by the middle class if it gets to power. Fig. 2 maps the five solutions of the model in the (x, μ) space.¹⁴ The effect of the other parameters will be analyzed later as shifts of the curves appearing in Fig. 2.

The pure oligarchy solution occurs when condition (5) does not hold, i.e.:

$$\mu \leq x + \frac{1 - \bar{y}}{1 - p}.$$

The corresponding region lies below line (D) on Fig. 2.

Curve (C) above line (D) delimits the area where some limited educational transfers are made to the poor in the first period but the oligarchy retains (strict) political majority in the second period — solution (β) above. This curve is obtained from the condition that the maximum of $\bar{Y}(e)$ occurs for a value e^* smaller than $(1 - p)$. It may be shown that its parabolic shape comes from the quadratic efficiency cost of transfers.

Above (C), the educational externality is big enough so that it would be in the interest of the oligarchy to educate a proportion of the poor which would get political majority in the second period. Two cases are to be distinguished. In the

¹⁴ Note that this space must actually be restricted to values of x such that $x > (1 - \bar{y})/p$, i.e. such that the liquidity constraint for the poor's investment in education is binding.

first case, the new majority would not impose any income redistribution in the second period so that the oligarchy has nothing to fear — solution (δ) above. This case is delimited by the condition that $\bar{Y}(e)$ is maximum at e^* , above $(1 - p)$ — which corresponds to the area above curve (C) — but below the value e_τ at which redistribution becomes profitable for the middle class. Curve (C') in Fig. 2 is defined by:

$$e^* \leq e_\tau \tag{C'}$$

But it is relevant only when:

$$1 - p < e_\tau \quad \text{or} \quad x < \frac{2p - 1}{1 - p} R, \tag{S}$$

which defines the vertical line (S). It may be shown that curves (C) and (C') cross precisely on (S).

In the second case, the new majority would impose a redistribution of income which would leave the oligarchy worse off. To prevent this from happening, the latter keeps political control and educates only $(1 - p)$ of the poor. The condition for this case — solution (γ) above — to hold is that $\tilde{Y}(e^o)$ be such that:

$$\tilde{Y}(e^o) \leq \bar{Y}(1 - p) \tag{C''}.$$

This requires the opposite of condition (S) above and defines the curve (C'') in Fig. 2.

Finally, if the educational externality is large enough, or income inequality is small enough for the representative point of the economy in the (x, μ) plan to lie above curves (C') and (C''), then solution (ε) above holds and redistribution with democracy takes place.

The evolution of the economy depends also on the other parameters of the model which are taken as given in Fig. 2. We briefly consider their effects in turn.

As far as the initial level of income is concerned, it is not difficult to see that all the curves (D), (C) and (C') in Fig. 2 will shift downward as the mean income in the first period goes up, so that, other things being the same, educational expansion and its positive effects on the whole economy becomes more likely.¹⁵

A change in the private return to education has no impact on line (D) nor on curve (C). However, since the line (S) shifts toward the right, curve (C') shifts upward and curve (C'') downward. Other things being the same, it is thus making democratization and fast growth less likely for relatively egalitarian countries and more likely for relatively inegalitarian countries.

The effect of a change in the cost of redistribution, a , is somewhat ambiguous. On one hand, it reduces the cost of educating the poor. On the other hand, it also

¹⁵ The shift in curve (C'') is not totally clear. It surely shifts downward in a neighborhood of (S), but things are quite intricate for the upper part of the curve.

increases redistribution by the middle class when it gains political control and thus worsens the situation of the oligarchy in period 2. Likewise, an increase in the initial weight of the oligarchy in the population ($1 - p$) has ambiguous effects.¹⁶

On the whole, Fig. 2 and the preceding remarks suggest then a relatively complex relationship between income distribution, democracy and growth. The complexity comes primarily from the fact that this relationship must be considered dynamically, unlike in many models which have been developed in the recent literature on growth and distribution and focus on steady state situations. Of course, it remains true that, if the analysis is restricted to initially poor oligarchic societies, the state of the economy in period B is entirely determined by its initial conditions, i.e. its level of economic development and its degree of inequality, and a few basic parameters like the importance of the educational externality.

As in other models, the more inegalitarian is the economy in a first period, the less it invests in education and the less it grows. The novelty, however, is that one must consider simultaneously not only the level of national income in the second period, but also its primary distribution among the various classes of individuals, the extent of redistribution, and the degree of democratization. The cross-sectional prediction of the model is that, starting from an oligarchic regime, the economies which are initially richer and less inegalitarian not only grow faster but also democratize and redistribute income sooner.

Checking empirically this prediction may be somewhat uneasy because of the necessity of using a sample of countries resembling what is assumed to be an oligarchy in the present model. It must be stressed, in particular, that the model does not predict that democracies necessarily exhibit a more egalitarian distribution of income than oligarchies. In the case where the oligarchy decides to educate only a fraction of the poor — i.e. $e < p$ — it can be seen that the distribution of primary incomes may become more unequal in period B than it was in period A. This is so because, even though everybody benefits from the educational externality, only the rich and the middle class do benefit from a direct increase in their income due to education. On that basis, the partial democratization of the economy does not necessarily improve the distribution of primary incomes and there is no reason to expect that, in a cross-section of countries, democracies should be less inegalitarian than oligarchies, when controlling for initial conditions. It is not even sure that such a conclusion would apply when considering the distribution of disposable incomes. In period B, the redistribution imposed by the middle class may indeed cancel the preceding unequalizing effect, but this will not always be the case since that redistribution may in fact be quite limited — and even zero in region (δ) of Fig. 2.

¹⁶ (D) rotates upward when $(1 - p)$ increases, (C) shifts upward, and line (S) moves to the left. But nothing can be said about curves (C') and (C'').

An interesting outcome of the model in matter of income redistribution is that, contrary to what is usually assumed or found in the recent literature, redistribution is not necessarily associated with a high degree of primary income inequality. In Fig. 2, redistribution occurs only on the left-hand side of curves C' and C'' , i.e. other things being the same, for relatively low levels of income inequality. The reason for this result is simply that if the degree of income inequality were initially higher, then the oligarchy would block the democratization process or would only permit a ruling middle class which would not redistribute. This suggests that there may be significant simultaneity biases in cross-sectional analyses where income redistribution is explained by initial income inequality and the nature of the political system.¹⁷

3. Dynamics of democratization, distribution and growth

The preceding conclusions may be influenced by the simple static framework used in this section. The size of the middle class in the preceding analysis depends on the initial size of the oligarchy, the mean income of the population and the degree of inequality in primary incomes. But once this middle class has appeared, the model is modified. Initial conditions are not the same. In particular, initial incomes are modified. There are more people educated so that all incomes are higher. In an intergenerational framework, one may also assume that the children of the oligarchy or those of the middle class will have inherited part of the human capital of their parents. So, the decisions of the 'enlarged' oligarchy will be different from those analyzed in Section 2. Dynamically, the middle class may keep expanding until possibly all members of the population are educated and politically active. But the process may also stop short of this full democratization in some long-run equilibrium. What is interesting is that this dynamic sequence of political decisions on how many of the poor are educated will also be shaping the full distribution of income in the economy and the aggregate economic growth path. In the long run, the distribution of human capital and the cumulated mean income growth will actually depend on the democratization path that will have been followed.

Potentially, a dynamic version of the model of Section 2 should thus permit a full integration of the dynamics of the political structure, growth and distribution. This section offers some insights on this process by simplifying and extending the initial model in several ways.

Consider now that agents live one period which actually consists of two sub-periods A and B as described in Section 2. Consider also that each individual has one offspring living also one period with two sub-periods A and B. Assume

¹⁷ See, for instance, Perotti (1996), Persson and Tabellini (1994).

perfect transmission of human capital across generations so that the offspring of an “educated” parent is “educated” and receives at the beginning of his/her own life an income identical to the one of his/her parent at the end of his/her life. Hence, someone whose parent was able to vote at the end of sub-period B is also able to vote at the beginning of his/her own life’s sub-period A. Finally, suppose that in the second generation only the “uneducated” poor are liquidity-constrained in their investment in education and consider the situation at the beginning of the second generation’s lifetime. Given these assumptions and the two-period analysis of Section 2, three situations are possible: complete democracy with two classes (rich and poor), oligarchy with two classes (rich and poor) and partial democracy with three classes (rich, middle and poor). In the last case, the existence of a middle class completely changes the model. Even if the first generation oligarchy retained political control, initial conditions for the second generation are not the same as for the first generation and decisions to be made by the second generation ‘enlarged’ Elite (rich and middle class) will be different from their elders. While a precise characterization of the equilibrium is analytically difficult, the ‘enlarged’ Elite will, for some parameter values, decide to promote again the education of a certain fraction of the poor, creating in this way another (low) middle class between them and the still “uneducated” poor. Therefore, at the end of the second generation, a situation of partial democracy with four social classes (rich, high middle class, low middle class and poor) may prevail. Another possibility is that the “enlarged” Elite vote for the education of all poor, thus leading to complete democracy with three social classes. Interestingly, associated with the process of increased social stratification, the pattern of redistribution at the end of the second generation will be different, for certain initial conditions, to the one prevailing at the end of the first generation. Adding therefore simple dynamics to the previous analysis suggests three additional features. First, one may get additional social classes and more complex steady state income distributions. Second, the pattern of redistribution evolves through time according to the initial conditions of the economy. Finally, there is path dependence of the long-run pattern of political institutions and redistribution. In particular, countries with similar political institutions (democracies) may have different long-run income distributions and consequently different redistributive conflicts.

A complete analysis of democratization, redistribution and growth would require the extension of the two-period framework of Section 2 to an infinite horizon and the characterization of the different equilibrium paths from oligarchy to democracy. Providing such an analysis in full generality is analytically untractable. The preceding informal discussion suggests, however, that essentially three types of equilibria are likely to emerge in the long run: permanent oligarchy, partial democratization and complete democratization. We present here a linear infinite horizon example which captures these features. In this extension, democratization never occurs or settles in, at most, two periods. While highly stylized, this

framework permits analyzing interesting aspects of the dynamics of income inequalities, redistribution and democratization.

3.1. A linear infinite horizon model of democratization, inequality and growth

The two sub-periods A and B are now embedded into unit periods t ($= 0, 1, \dots$), the sequence of which defines the dynamics of the model. In each of these periods, there is a stationary population (again normalized to unity) of non-overlapping generations which live exactly one unit period and transfer perfectly their skills to the next generation. In sub-period A_t , private as well as public decisions on education are made. In sub-period B_t , society votes on redistributive taxation. More precisely, in A_t , individuals start by inheriting the skills of their parents. This determines their initial income level and their ability to be “politically active”. Following is a vote between “politically active” agents on the transfers T to be made to the poor who cannot directly invest in education. Then individuals who are not liquidity-constrained invest in human capital. In the second stage of their life, B_t , “educated individuals” vote on income redistribution. Rather than assuming, as in Section 2, a quadratic distortionary cost function of taxation, we simply suppose here that the tax rate τ is restricted to be in the interval $[0, \tau^+]$ where τ^+ is the maximum admissible tax rate.¹⁸ Also let E_t be the total number of educated individuals at the beginning of sub-period B_t .¹⁹

This framework clearly implies a certain degree of myopia as individuals are concerned only about their welfare and not the one of their offspring. It is well known in the literature on the political economy of growth (Verdier, 1994; Boldrin, 1995; Krusell and Rios-Rull, 1996) that having forward-looking political agents greatly complicates the analysis without necessarily improving the intuition of the mechanisms at work. In the present case, it is true that what matters is not so much the decentralized behavior of agents as that of a cohesive and politically powerful group. It is thus better to simply interpret the myopia of the oligarchy as a high discount rate of the future in a fully intertemporal framework. We shall see that this specification has the advantage of remaining analytically tractable and yet allows us to capture the crucial feature of the model. We shall also come back briefly on the implications of a lower discount rate at the end of this section.

The initial income distribution in period 0 at stage A_0 is given by $1 - p$ “educated” rich with income $y_0^r = \bar{y}_0 + px$ and p “uneducated” poor with income $y_0^p = \bar{y}_0 - (1 - p)x$ and $y_0^r > 1 > y_0^p$. Initial incomes are supposed to include the education externality term $\mu(1 - p)$ due to the “educated” rich. We

¹⁸ The implicit assumption is no distortionary cost for tax rates between 0 and τ^+ and infinite costs for tax rates larger than τ^+ .

¹⁹ In the sequel, it will be more convenient to deal with E_t the total number of educated individuals than the number of educated poor as in the static model of Section 2.

assume also $R(1 - \tau^+) > 1$ to make sure that the only reason why some individuals do not invest in education is that they are liquidity-constrained. Note also that because of our assumption of perfect transferability of skills across generations and because $R > 1$, individuals whose parents have invested in education will also be able to do so without transfers. Hence in sub-periods A_t , dynasties of educated people have an income which increases by R at each generation.²⁰ Because individuals are only concerned with what happens in their lifetime, the model can be solved by usual backward induction. We start with the political equilibrium of taxation in each sub-period B_t . We then consider investment decisions in education and the vote about the education of the poor in sub-periods A_t .

Denote y_t^r , y_t^p and \bar{y}_t the rich, poor and average income at the beginning of sub-period A_t . As will be clear in the sequel, it is useful to start the analysis by considering the case where the oligarchy is the only educated and thus the only politically active class until period t , (i.e. $E_{t-1} = 1 - p$). The total net income of an oligarchy member in period t is given by:

$$\begin{aligned} \tilde{Y}_t(E_t) = & [y_t^r - 1] - \frac{(1 - y_t^p)(E_t - (1 - p))}{1 - p} \\ & + [y_t^r + R + \mu(E_t - (1 - p))](1 - \tau_t) \\ & + \tau_t[\bar{y}_t + RE_t + \mu(E_t - (1 - p))]. \end{aligned}$$

The first two terms represent the net income in sub-period A_t , the last two terms represent the after tax income of sub-period B_t with τ_t being the tax rate voted in B_t . It is clear from the absence of quadratic distortionary costs that the tax rate τ_t is a bang–bang solution taking the values 0 or τ^+ . $\tilde{Y}(E_t)$ is thus piecewise linear in E_t .

3.2. Dynamics of democratization, growth and inequality

It is then easy to derive Fig. 3²¹ which represents the equilibrium number of educated people in period 0 in the space (x, μ) of the initial dispersion and the externality parameter. Fig. 3 is similar to Fig. 2.²² The absence of quadratic costs of distortions now implies that all curves are linear. It also implies that the oligarchy will always choose corner solutions in terms of the number of poor it promotes. Therefore, there is no regime (β) with a “minority middle class and medium growth” and curve (C) disappeared. Below (D), one still finds the pure

²⁰ Note also that, because of the assumption of perfect transferability of human capital across generations, the number of educated and therefore “politically active” individuals at the beginning of sub-period A_t is E_{t-1} .

²¹ See Appendix A.

²² Note again that x has to be big enough so that $y_0^r > 1 > y_0^p$. In the limit where $\bar{y}_0 = 1$, x has to be simply positive.

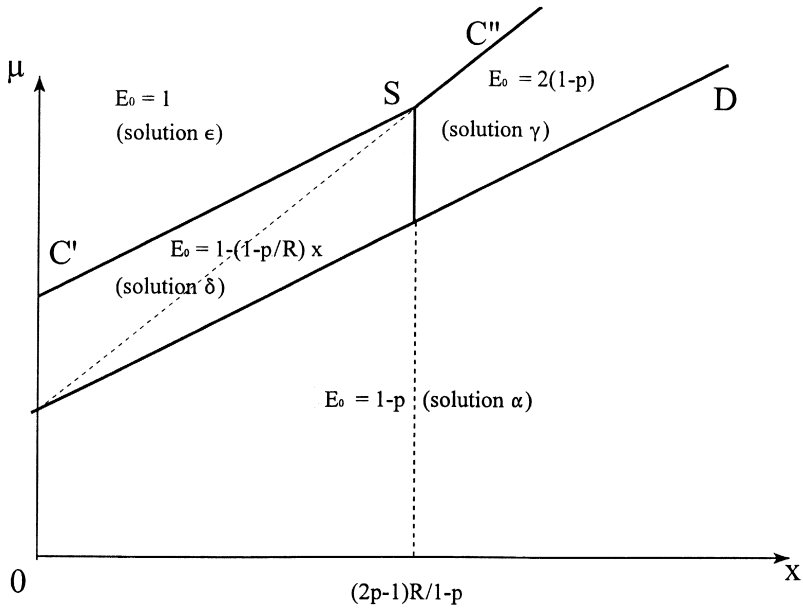


Fig. 3.

oligarchy regime (α). Above (C') and (C''), there is full education and democracy (regime ϵ). Finally, there is the balance of power regime (γ) (respectively accommodating ruling middle class regime δ) on the right (respectively left) of $(2p - 1)R/(1 - p)$.

The major difference with the two-period analysis hinges on the dynamic impact of the human capital externality on the political participation of the poor. The education of a first generation of individuals increases the income of the remaining unskilled poor through the education externality. When this effect is strong enough, this in turn allows these generations of poor to overcome their liquidity constraints, pay themselves the fixed cost of education and acquire skills without subsidies from the Elite. Doing so also allows them to become politically active. Democratization is then triggered as an irreversible process. Whether the educational externality effect succeeds in inducing such a process depends on the number of educated individuals previously subsidized by the Elite, the level of the fixed cost of education and the initial income of the unskilled poor. In the present linear specification, it can be shown (see Appendix A) that, whenever the Elite subsidizes the education of some poor, it does so to such an extent that the human capital externality is strong enough to allow democratization through widespread self-financed education of the poor. Hence, full education and democracy never occur or occur in one or two periods — i.e. generations — so that the middle class regimes analyzed in the static model appear as essentially transitory.

Conditionally, on the structural parameters of the economy μ , R , p , \bar{y}_0 and x , the dynamics of education, democratization and growth are then easily described.

(i) Below (D), democratization never occurs and only the rich are educated. There is a permanent oligarchy with slow growth and an increasing gap between the poor and the rich. The poor never become educated and there are two classes. In each period, the rich invest in skills and increase their income by R while the poor stagnate at y_0^p .

(ii) Between (D) and (C')–(C''), there is full education and democracy occurs in period 1. This is a regime with slow democratization and the emergence of a middle class with moderate growth. Full democracy is reached in two periods. This region itself is divided into two sub-regions: for small initial inequality ($x < (2p - 1)R/(1 - p)$), the middle class gets into power without taxing the rich. On the contrary, when initial inequalities are larger ($x > (2p - 1)R/(1 - p)$), the middle class takes some time to get into power — a one period lag — but then taxes the rich and redistribute income. Once full education is reached, all individuals see their income increasing by R in each period and the income gap between the three classes remains constant.

(iii) Above (C')–(C''), full education and democracy occur in period 0. There is immediate democratization, full education and fast growth. Both classes invest in education in each period and the income gap remains constant over time.

This model also provides a prediction of what the long-run social structure will be as a function of the initial parameters of the distribution of income at time 0. When full democracy is reached with some delay, the long-run social structure of society is given by a three-class income distribution (rich, middle and poor). On the contrary, with immediate democracy or permanent oligarchy, social stratification involves two classes only (rich and poor). Although highly stylized, this example shows how the social structure prevailing in the long run depends not only on technology and institutional parameters (summarized by μ) but also on the characteristics of the initial distribution of income (summarized by p , x and \bar{y}_0). Also, it implies that countries with similar political institutions in the long run (democracies) may have different steady state redistribution patterns. Given that they can have different long-run social structures (in the present case, two classes vs. three classes), redistributive conflicts and the extent of redistribution may differ in the long run and depend on the particular historical path followed by each society.

3.3. Comparative dynamics

As in Section 2, simple comparative dynamics can also be performed on Fig. 3. Clearly, the further we move upward in that figure — i.e. the larger the externality parameter μ — the faster the pace of democratization and the better the growth performances of the economy. Also, the further we move to the right, the larger the initial level of inequality in the economy and the slower the pace of

democratization and education. It is also easy to see that an increase in the initial level of income \bar{y}_0 induces a downward shift of all the curves. Hence, the richer initially the economy, the faster the rate of democratization and education in the population. Finally, an increase in the taxation power, τ^+ , of the poor or equivalently a lesser cost of redistribution in sub-period B does shift the curves (D) and (C')–(C'') upward and slows down, or even stops the process of democratic transition.²³ This discussion allows us therefore to say something not only on the relationship between democracy, income levels and inequality but also on how the rate of change in democratization is connected to initial conditions.

Would these results be different if the oligarchy had not been assumed to be myopic? The answer is unambiguously yes. Indeed, with a time discount rate and an educational externality, μ , sufficiently low, the oligarchy will find that the achieved cost of redistribution after full democracy is larger than the benefit from a better-educated population. It will then block the democratization process by educating less than $(1 - p)$ people. Even if the educational externality makes it profitable to accept democracy in the long run, there is no reason to believe that getting there after one generation will be optimal. The discounted gain of the oligarchy may be higher with a slower democratization process. If redistribution after it has lost political control were endogenous, instead of being arbitrarily set at t^+ , it could also be in its interest to monitor the progress of education and democracy so as to generate a distribution of income making redistribution less disadvantageous. These issues are left for future work.

4. Conclusion, empirical implications and extensions

We have explored in the present paper the consequences of allowing political institutions to be endogenous in a political economy model of income redistribution, educational investments and growth. This was done by making political participation dependent on the educational level of citizen in an economy where fixed costs of education and liquidity constraints prevent poor persons to become educated in the absence of transfers from the upper income and politically active class. Being tractable enough, the model yielded interesting empirical implications which may be summarized as follows.

(i) Initial income per capita levels (initial income inequality) affect positively (negatively) the likelihood for a country to be a democracy and its average rate of growth at any given time horizon.

²³ Note that unlike in Section 2, education subsidies in sub-period A do not involve retribution cost in the present specification. They would clearly produce effects opposite to those linked to an increase in τ^+ .

(ii) Initial income per capita levels (initial income inequality) affect positively (negatively) the speed of (full) democratization for countries which are experiencing a democratic transition.

From an empirical point of view, these conclusions may be taken as an indication of the way political institutions, income distribution and growth could be integrated within truly structural models rather than in the loose and often ambiguous reduced forms currently found in the literature. From a policy point of view, they point, like other recent models, to the social benefit to be expected from any exogenous redistribution of wealth in slow-growing and authoritarian societies. However, as it would clearly be opposed by the ruling class, such a redistribution is more wishful thinking than a true policy option. More interestingly, and assuming that aid may be strictly monitored and targeted by international agencies, the question which comes to mind in the present framework is whether this aid must take the form of an income subsidy to the poor or that of providing direct education to them. In the first case, the material cost for the oligarchy of educating the poor is reduced but the threat of redistributive taxation remains the same so that it may eventually keep impeding the full democratization of society. In this case, aid would not necessarily increase the investment in education and would provide only temporary poverty relief. Things would clearly be different if, on the contrary, aid were under the form of more education being given to the poor or, equivalently, if it were enough to allow them to escape the liquidity constraint. In this case, not only would human capital and output be increased but the economy could also shift onto a new path of self-sustained growth and democratization.

Several possible extensions of the preceding model might be envisaged. First, it must be noted that 'democratization' in the preceding model is taken in a rather weak sense. The issue is not so much that the oligarchy refuses to give voting rights to the mass of citizen as the fact that the majority of citizen do not use their voting rights. Although this assumption fits well many contemporary developing economies, it may be rather far from the true history of the democratization process. This is in contrast with the recent paper by Acemoglu and Robinson (1996) who explicitly bring into their model the costs and benefits for the oligarchy and the mass of politically inactive citizen of a revolutionary conflict which would eventually impose democracy. Viewed in such a perspective, our own model should thus be interpreted as a representation of the threat to the oligarchy that mass education may make a democratic revolution more likely. So, members of the oligarchy would like more poor to be educated to benefit from the educational externality, but at the same time they know that, by doing so, they increase the probability of a revolutionary switch to democracy.

If one adopts such a conflictive interpretation of the results obtained in this paper, there is no reason to rule out the possibility for the oligarchy to resist the democratization process pushed forth by the progress of education either by fighting successfully the revolution, or, in our original framework, by repressing

the voting liberty of the newly educated class. In such a case, the oligarchy would benefit from the educational externality but would avoid the cost of redistributive taxation, which, as a matter of fact, would be partly substituted by the cost of repression. It would be easy to introduce such a mechanism in the preceding framework, adding to the various regimes considered above one of ‘repressed democratization’, the occurrence of which would be determined by the (exogenous) cost of repression. It may be noted that such regimes may actually correspond to what is presently observed in several countries where educational progresses and economic growth have been very rapid over the last two or three decades but where supposedly democratic governments remains somewhat ‘authoritarian’.

Second, in the present model, redistribution was obtained through flat taxation and lump sum transfers. This is quite legitimate if one reinterprets our redistributive scheme in terms of uniform instruments like consumption taxes, VAT or trade taxes. If, however, taxes are literally interpreted as income taxes, then a useful extension is to allow for nonlinear redistributive schemes and progressivity in income tax rates. The main well known difficulty with such an extension concerns the determination of the political equilibrium as the median theorem may not apply. However, it is not difficult to design simple examples of nonlinear income tax systems where education and democratization patterns remain qualitatively similar to the ones developed in the main text with linear taxation.

Another limitation of the model analyzed in this paper is the fact that we totally ignore the accumulation of physical capital. By doing so, we may be overestimating the benefit that the educated poor may derive from redistributive taxation when they have political control. This is because the income tax may reduce the pace of physical capital accumulation and the productivity gains that it entails, as in several recent models of growth and income distribution. From this point of view, the present model tends to overestimate the resistance of the oligarchy to the education of the mass and democratization. This bias would be still stronger in the (likely) case of a complementarity between physical and human capital in the economy since the education of the poor would then increase the return to physical capital and thus the income of the oligarchy, if it is reasonably assumed that it owns most of the material wealth in the economy. As a matter of fact, it should be possible in such a framework to ignore the educational externality on which relies the whole argument in the present paper since the benefit for the oligarchy of more educated people in the economy would simply be the higher profit obtained with the existing physical capital stock.

Acknowledgements

We thank Philippe Aghion, Abhijit Banerjee, Roland Bénabou, Allan Drazen, Sylvie Lambert, Pierre André Chiappori, Roger Guesnerie, Jim Robinson, Gilles

Saint Paul, Carlos Winograd and participants at seminars in the World Bank, Yale University, Delta, NBER, and the Seminaire d’Economie du Developpement in Paris for useful comments. We also thank two anonymous referees for helpful critiques and remarks. All errors remain ours.

Appendix A. Appendix on the linear infinite horizon model

One can first derive the following two results:

Result 1. When $\mu < (1 - y_i^p)/(1 - p)$ and $E_{t-1} = 1 - p$, then $E_t = 1 - p$.

Result 2. When $\mu > (1 - y_i^p)/(1 - p)$ and $E_{t-1} = 1 - p$, then $E_t \geq 2(1 - p)$.

The gain of educating a poor for a rich is equal to the externality μ , minus the per capita education subsidy $(1 - y_i^p)/(1 - p)$ and minus the cost of eventual income redistribution which will take place in period B_t , $\tau_t[y_i^r + R - \bar{y}_t - RE_t]$, which is strictly positive for all E_t . Thus, if the externality is less than the educational subsidy, the rich can only lose in subsidizing the education of the poor. This is Result 1. In the opposite case, the rich gain as long as they keep political control or, possibly, if they lose it but no taxation is voted in sub-period B_t . Therefore, the rich have an incentive to subsidize the education of at least $(1 - p)$ unskilled poor. Hence, $E_t \geq 2(1 - p)$ and Result 2 follows. QED

Proposition 1. *Democratization occurs in period 0, or 1 or the oligarchy remains indefinitely the only one with education and political power. (i.e. $E_0 = 1$, or $E_t = 1$, or $E_t = 1 - p$ for all t).*

Proof: Suppose that Result 2 holds and the rich subsidize the education of some poor. From the condition for this to be optimal:

$$\mu > \frac{1 - y_i^p}{1 - p}, \tag{X}$$

it comes that:

$$y_i^p + \mu(1 - p) > 1. \tag{Y}$$

The income of the poor in period A_{t+1} is given by $y_i^p + \mu[E_t - (1 - p)]$. But from Result 2, $E_t \geq 2(1 - p)$, so that Eq. Y implies that the poor are not liquidity-constrained anymore in period $t + 1$. It follows that $E_{t+1} = 1$. However, it is also possible that $E_t = 1$ be considered optimal by the oligarchy. Finally if Eq. X does not hold, then Result 1 implies that $E_t = 1 - p$ forever. QED.

Note that two basic features are essential for these simple democratization dynamics. The first one is the additive rather than multiplicative accumulation of human capital. The second crucial point is that when the oligarchy promotes the education of the poor after period 0, it is ready to educate at least $2(1 - p)$ poor (the condition $E_1 > 2(1 - p)$ is crucial to show that the remaining uneducated are not liquidity-constrained in the next period). This result hinges on the absence of distortionary costs of taxation and decreasing returns in the externality of education. Hence, the marginal condition that it pays to the rich to educate a fraction of poor also holds for $E_t \leq 2(1 - p)$. If one had convex costs of distortion or decreasing returns in the education’s externality, the same type of dynamics would hold but under the additional condition that in period 1, $E_1 \geq 2(1 - p)$.

To characterize precisely the dynamic path of education, democracy and taxation, one has to solve explicitly for the equilibrium income tax τ_t and the optimal E_t in each period. Given Proposition 1, it boils down to consider only what happens in period 0.

In sub-period B_0 , as long as rich people represent a majority of “politically active” agents (i.e. $E_0 \leq 2(1 - p)$), the political equilibrium tax rate is $\tau_0(E_0) = 0$. On the other hand, when $E_0 > 2(1 - p)$, the median voter is the middle educated class and the political equilibrium tax rate is given by the maximization of $Z(\tau, 0, E_0 - (1 - p))$ — see Eq. 9 in the text — with respect to τ under the additional constraint that τ has to be in the interval $[0, \tau^+]$. This maximization yields the bang–bang solutions 0 and τ^+ represented by the equivalent of Eq. 10:

$$\begin{aligned} \tau_0(E_0) &= 0 \quad \text{if } E_0 \leq \max\left[2(1 - p), \quad 1 - \frac{(1 - p)x}{R}\right] \\ \tau_0(E_0) &= \tau^* \quad \text{if } E_0 > \max\left[2(1 - p), \quad 1 - \frac{(1 - p)x}{R}\right] \end{aligned} \tag{15}$$

In stage A_0 in period 0, the total net income of a rich over his lifetime is given by $\tilde{Y}_0(E_0)$ and the equilibrium number of educated individuals in period 0 is the solution of the maximization of this function under the constraint that E_0 belongs to $[1 - p, 1]$. Depending on the value of the externality μ , different shapes for $\tilde{Y}_0(E_0)$ are possible. Clearly, from Result 1, as long as $\mu < (1 - y_0^p)/(1 - p)$, the optimal number E_0 is $E_0^* = 1 - p$. Noting $\alpha = \max\{2(1 - p), 1 - (1 - p)x/R\}$, in the case $(1 - y_0^p)/(1 - p) < \mu$, it can be seen that $\tilde{Y}_0(E_0)$ is increasing in E_0 except in $E_0 = \alpha$ where it has a downward jump due to the fact that $\tau(E_0)$ jumps from 0 to τ^+ . In that case, the equilibrium number of educated individuals is determined by the comparison of the lifetime income levels $\tilde{Y}_0(1)$ and $\tilde{Y}_0(\alpha)$ with:

$$\tilde{Y}_0(\alpha) = y_0^r - 1 - \frac{(1 - y_0^p)(\alpha - (1 - p))}{1 - p} + y_0^r + R + \mu\alpha, \tag{17}$$

and there is full (respectively partial) education whenever $\tilde{Y}_0(1) >$ (respectively \leq) $\tilde{Y}_0(\alpha)$. This last condition can be rewritten as:

$$\mu < (>) \tau * \frac{y_0^r - \bar{y}_0}{1 - \alpha} + \frac{(1 - y_0^p)}{1 - p}. \quad (18)$$

From the condition $(1 - y_0^p)/(1 - p) < (>) \mu$ and Eq. 18, Fig. 3 in the text can be easily constructed. QED.

References

- Acemoglu, D., Robinson, J., 1996. Why did the West extend the franchise? Democracy, inequality and growth in historical perspective. Mimeo, MIT.
- Ades, A., 1996. The rise and fall of elites: economic development and social polarization in rent-seeking societies. CEPR Discussion Paper No. 1495.
- Aghion, A., Bolton, D., 1998. A trickle-down theory of growth and development with debt overhang, forthcoming. *Review of Economic Studies*.
- Alesina, R., Rodrik, D., 1994. Distributive politics and economic growth. *Quarterly Journal of Economics* 109, 465–490.
- Bardhan, P., 1997. Democracy and development: a complex relationship. In: Shapiro, I. (Ed.), *Rethinking Democracy for a New Century*. Cambridge Univ. Press, Cambridge, MA.
- Banerjee, A., Newmann, A., 1991. Risk-bearing and the theory of income distribution. *Review of Economic Studies* 58 (194), 211–236.
- Barro, R., 1996. Democracy and economic growth. *Journal of Economic Growth* 1 (1), 1–27.
- Benabou, R., 1996. Inequality and growth. CEPR Discussion Paper No. 1450.
- Bertola, G., 1993. Factor shares and savings in endogenous growth. *American Economic Review* 83 (5), 1184–1198.
- Boldrin, M., 1995. Public education and capital accumulation. Mimeo.
- Brady, H., Verba, S., Schlozman, K.L., 1995. Beyond SES: a resource model of political participation. *American Political Science Review* 89 (2), 271–294.
- Fraser, J., 1972. Political participation and income level: an exchange. *Public Choice* 13, 115–118, Fall.
- Frey, B., 1971. Why do high-income people participate more in politics? *Public Choice* 11, 101–106.
- Frey, B., 1972. Political participation and income level: an exchange: reply. *Public Choice* 13, 119–122.
- Galor, O., Zeira, J., 1993. Income distribution and macroeconomics. *Review of Economic Studies* 60 (202), 35–52.
- Gradstein M., Justmann, M., 1995. A political interpretation of the Kuznets curve. Mimeo.
- Huber, E.D.R., Stephens, J., 1993. The impact of economic development on democracy. *Journal of Economic Perspectives* 7, 71–86.
- Krusell, P., Rios-Rull, J., 1996. Vested interests in a positive theory of stagnation and growth. *Review of Economic Studies* 63 (2), 301–329.
- Perotti, R., 1996. Growth, income distribution and democracy: what the data say. *Journal of Economic Growth* 1 (2), 149–187.
- Perotti, R., 1993. Political equilibrium, income distribution, and growth. *Review of Economic Studies* 60 (205), 755–776.
- Persson, T., Tabellini, G., 1994. Is inequality harmful for growth? *American Economic Review* 84 (3), 600–621.

- Piketty, T., 1997. The dynamics of wealth distribution and the interest rate with credit rationing. *Review of Economic Studies* 64 (2), 173–189.
- Przeworski, A., Limongi, F., 1993. Political regimes and economic growth. *Journal of Economic Perspective* 7, 51–70.
- Robinson, J., 1997. When is a state predatory? Mimeo. University of Southern California.
- Russel, K., 1972. Political participation and income level: an exchange. *Public Choice* 13, 113–114, Fall.
- Saint-Paul, G., Verdier, T., 1993. Education, democracy and growth. *Journal of Development Economics* 42 (2), 399–407.
- Sirowy, L., Inkeles, A., 1990. The effects of democracy on economic growth and inequality: a review. *Studies in Comparative International Development* 25, 126–157.
- Verba, S., Nie, N., Kim, J., 1978. *Participation and Political Equality*. The University of Chicago Press, Chicago.
- Verba, S., 1990. Politics, economics and equality. In: Dye, T. (Ed.), *The Political Legitimacy of Markets and Governments*. Political Economy and Public Policy Series vol. 7 JAI Press, Greenwich, CT.
- Verdier, T., 1994. Models of political economy of growth: a short survey. *European Economic Review* 38, 757–763.