Review of Economic Studies (1993) 60, 755–776 0034-6527/93/00380755$02.00

Political Equilibrium, Income Distribution, and Growth

ROBERTO PEROTTI
Columbia University

First version received February 1991; *final version accepted April* 1993 (*Eds.*)

This paper analyzes the impact of income distribution on growth when investment in human capital is the source of growth and individuals vote over the degree of redistribution in the economy. The model has three main features. First, very different patterns of income distribution are conducive to high growth at different levels of per capita income. Second, growth is associated with an externality whereby investment in human capital by one group increases the productivity of other groups, thus potentially enabling them to invest in human capital. Third, the initial pattern of income distribution and the resulting political equilibrium are crucial in determining whether the transmission of this externality is promoted, in which case growth is enhanced, or prevented, in which case growth is stopped.

Using a non-overlapping generations model with voting, I derive several empirical implications. In particular, the model implies an inverted-U relation between levels of inequality and levels of income in cross-sections, but not necessarily in time series, a result that seems consistent with a number of empirical studies.

1. INTRODUCTION

In the voluminous literature on income distribution and growth, two basic frameworks can be identified. A tradition going back at least to Kaldor (1956) emphasizes the causal effect of income distribution on capital accumulation and therefore on growth. The development economic literature that flourished in the 1960's and 1970's following the seminal work of Kuznets (1955) concentrated mainly on the opposite causal link, from growth to income distribution.

Both mechanisms are at work in the model of this paper. However, the focus here is not on capital accumulation, but on the effects of redistribution on investment in human capital. Specifically, this paper starts from the observation that income distribution is not a given, but it can be modified to some extent in an economy where the tax system redistributes income. By affecting the post-tax income of the various income groups, redistribution determines which groups will be able to invest in human capital and which groups will remain unskilled. In turn, this affects growth and how income distribution evolves over time.

When preferences are aggregated through a voting process, the initial pattern of income distribution plays a crucial role in the evolution of the economy because it determines the degree of redistribution that prevails in the political equilibrium. In a static setting, it is intuitive that when the decisive voter is poor relative to the average, she faces a relatively low tax price of redistribution; thus, as shown by Romer (1975), Roberts (1977) and Meltzer and Richard (1981), in a voting model inequality (i.e., a poor median voter relative to the average voter) tends to be positively associated with the level of taxation and redistribution. In the model of this paper, this simple intuition can generate some interesting dynamics.

The essence of the model is very simple. Individuals can belong to one of three different income groups. Growth and changes in pre-tax income distribution are the effect of investment in education. The latter benefits the investor directly, and all the other agents indirectly through a production externality. As in Galor and Zeira (1993), in the absence of perfect capital markets those individuals whose post-tax income is below the cost of acquiring education will be unable to invest in human capital, and the next period will earn the same pre-tax income. By contrast, those who can afford the expenditure needed to obtain education will have a higher income.[1]

This simple structure has a first important implication: economies with different per capita incomes have very different patterns of income distribution that are most favourable to growth. In a very poor economy *total* resources may be so scarce that at most the upper class can invest. Thus, in this case only a very unequal income distribution that concentrates resources in the upper class may be consistent with growth. Alternatively, given the share of the upper class in total pre-tax income, the median voter should not have too large an incentive to set a very progressive tax rate and expropriate the upper class. This requires that the middle class should not be too distant from the upper class.

The configuration that maximizes income growth in a rich economy is exactly the opposite (with some qualifications spelled out in the formal analysis of the model). Here, redistribution might matter only for the investment of the lower class. A precondition is that the middle and the lower class should not be too far apart. Otherwise, it will be too costly for the median voter to redistribute the resources the lower class needs to invest.

These basic ideas of the paper can be formalized in a simple two-period model. A non-overlapping generations extension of the two-period model develops more fully the implications of the strong path-dependence embedded in the framework sketched above. Specifically, it formalizes a concept of growth as a "trickle down" process by which investment by one class increases the future income of all other classes as well, thus enabling an increasing number of classes to invest in education over time. The basic message is that in the absence of a central planner the transmission of this positive externality can stop if it is too costly to the median voter to bring them about. For instance, consider an economy that has grown up to the stage where the middle class has invested in education. Now the lower class will invest in education if the median voter has an incentive to enact enough redistribution. However, if the low income class is much poorer than the middle class, the median voter does not have such an incentive and growth will stop. Thus, the political outcome generated by the initial pattern of income distribution is crucial in determining whether the "trickle down" process of growth will be stopped before the economy has reached the highest possible steady-state where all classes have invested in education.

This framework can also provide a possible explanation of the famous inverted-U relation between levels of income and measures of inequality in cross-section regressions; and of the fact that the same relation is more difficult to observe in time-series. It was seen above that a very egalitarian poor economy will not be able to start the growth process. By contrast, an economy with a very unequal income distribution is in the best position to achieve a high initial rate of growth. However, once this economy reaches a higher level of per capita income, the very same income distribution pattern that fuelled the initial

1. Besides Galor and Zeira, Bannerjee and Newman (1991) and Aghion and Bolton (1991) also develop models where capital market imperfections open up the possibility for income distribution to affect the pattern of growth of an economy.

spurt of growth will hamper further growth. Thus, a very unequal society will get stuck at an intermediate level of income, because the extreme concentration of resources in the hands of the upper class prevents the lower class and possibly even the middle class from reaching a post-tax income that allows investment in education. In a more egalitarian society all classes will eventually invest in education, so that inequality will decrease as per capita income reaches its highest level. In a cross-section, this will generate an inverted-U curve, even though only a subset of all countries will present an inverted-U pattern in time series.

The role of income distribution in endogenizing the level of taxation and growth has been the subject of some recent research. The common element to Alesina and Rodrik (1991), Bertola (1991) and Persson and Tabellini (1991) is that a higher tax rate reduces the private after-tax marginal product of capital and therefore acts as a disincentive to investment and growth. In turn, income inequality and the tax rate resulting from the voting process are positively related through a dynamic extension of the standard median voter result; the reason is that, as in my model, a relatively poor median voter faces a lower tax price of the productive public good (Alesina and Rodrik) or of the redistributive subsidy (Bertola, Persson and Tabellini). A similar mechanism operates in Saint-Paul and Verdier (1991), except that now a higher tax rate might have beneficial effects by making possible a larger expenditure on public education and therefore more accumulation of human capital. The interaction of these two opposite effects generates a hump-shaped relation between inequality and growth.[2]

The rest of the paper is organized as follows. Section 2 introduces the basic two-period model. Section 3 analyzes the existence of a non-cycling majority and proves that the median voter is the decisive voter even if preferences are not single-peaked. Section 4 characterizes the political equilibrium and studies its effects on growth depending on the initial income distribution and on the level of income. After sketching the infinite-horizon, non-overlapping generations model, Section 5 illustrates why it might be relevant in discussing the issues outlined in this introduction. Section 6 discusses the role of some crucial assumptions and draws some conclusions. Since the formal treatment of the model is rather notation-intensive and in order not to hamper the intuition behind the results, all the proofs appear in the appendices.

2. THE MODEL

There are two periods, 1 and 2 and three groups of agents, h, m and l characterized by different earning abilities, i.e. different pre-tax incomes. Let n_t^i be the earning ability of an agent belonging to pre-tax income class i in period t, with $i=h$, m or l. In every period, a proportion p^i of agents belongs to group i. In period 1, pre-tax incomes are characterized by the following inequalities: $0 \leq n_1^l \leq n_1^m \leq n_1^h$. Finally, let $\bar{n}_j$ represent the mean of the distribution of pre-tax incomes in period j. The distribution of pre-tax incomes satisfies two conditions:

(i) $p^i < 0.5$, $i=l, m, h$
(ii) $n_1^m \leq \bar{n}_1$

2. A different class of models is based on bargaining rather than a formal voting equilibrium to represent the political process. Thus, in Benhabib and Rustichini (1991) high inequality might induce individuals to ask for a lot of redistribution and therefore might generate an equilibrium where incentives to invest are small or even absent. In Chang (1992) inequality affects growth by determining how the bargaining process ends up allocating and using government revenues.

By preventing a single class from having more than half the agents of the economy, assumption (i) is a necessary condition for the existence of non-trivial majorities. In addition, assumption (i) implies that the median voter is in the middle class. Assumption (ii) ensures that the median is initially below the mean.[3]

In period 1, agents can invest a certain amount in education. As a normalization, let this amount be equal to 1. The only choice is between investing in education the amount 1, and not investing. Investment in education by an agent has a positive externality on the second period productivity of the other agents. Let μ be the proportion of agents that invested in education in period 1. Thus, μ can take the values 0, p^h, p^h+p^m, 1. Pre-tax income of agent i in period 2 is

$$n_2^i = n^i + Re + \phi(\mu)R \tag{1}$$

where e is an indicator function taking the value of 1 if the agent invested in education, and 0 otherwise, and $\phi(\mu)$ is any monotonically increasing function of μ with $\phi(0) \leq 0$.[4] Therefore, $\phi(\mu)$ represents the externality from the investment in education by a measure μ of other agents on the productivity of each agent. Note that this externality occurs even if the agent in question has not invested in education. For simplicity, from now on I will assume that $\phi(x)$ is the identity function $\phi(x) = x$.

This externality drives the "trickle-down" property of the model, by which investment in education by one group may enable other groups to invest in education. In its absence, all the dynamics of the model would consist in a once-and-for-all investment by those groups whose pre-tax income exceeds 1, the cost of investing in education. However, this is not the only type of externality that delivers the "trickle-down" feature of the model. For example, if voting on the tax rate occurred in both periods, investment by one group could be beneficial to other groups by increasing the resources available for redistribution in the future.[5]

There is no capital market, no uncertainty, no discounting.

In period 1 the agents of this economy vote over the level of income taxes. Taxes are proportional to pre-tax income. The revenues collected in this way are redistributed as a per capita subsidy, constant across individuals. The government budget is always balanced. However, there are convex costs in collecting taxes: thus, if t is the tax rate, $t\bar{n}$ is collected but only $(t-t^2)\bar{n}$ can be redistributed to each individual.[6] Thus, no agent will ever vote for $t > \frac{1}{2}$ because the subsidy per capita is decreasing in t for $t > \frac{1}{2}$. Note that given these assumptions a higher *proportional* tax rate (in the range $[0, \frac{1}{2}]$) implies a more *progressive* tax-subsidy system.

Utility is linear in consumption. Let c_1^i and c_2^i represent consumption in period 1 and 2 of an agent belonging to class i, respectively, and let $\bar{n}_2$ represent the per capita income in period 2. Overall utility for an agent belonging to group i is:

$$c_1^i + c_2^i = n^i(1-t) + (t-t^2)\bar{n} - e + (n^i + Re + R\mu). \tag{2}$$

3. From now on, whenever a first period variable is considered, the subscript indicating the time period will be omitted if no ambiguity can result. Thus, $\bar{n}$ stands for $\bar{n}_1$, n^m for n_1^m and so on.

4. This specification of the effects of education on earning ability is not an orthodox one in the human capital literature. A multiplicative rather than an additive effect is usually assumed. An example of a paper using the additive effect specification adopted here is Chiswick (1971).

5. A model based on this "redistribution externality" was developed in an earlier version of this paper, Perotti (1990).

6. Without convex costs of collecting taxes, it is a standard result that, when labour is supplied inelastically, all voters below the mean prefer $t=1$ while all voters above the mean prefer $t=0$. Introducing a convex cost of collecting taxes allows one to avoid these corner solutions.

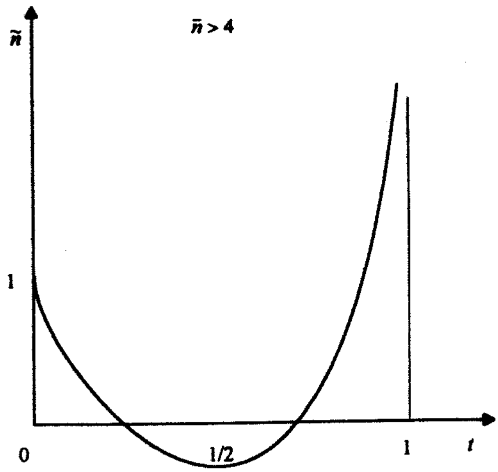

FIGURE 1(a)
$\tilde{n}>4$

It is clear that an agent who takes as given the actions of the other agents would like to invest in education as long as $R \geqq 1$. In what follows, I will assume that this inequality is satisfied. This effectively ensures that all agents would like to invest in education, independently of how many agents invest.[7]

However, given the absence of capital markets, agent i cannot invest in education if $n^i(1-t)+(t-t^2)\bar{n}<1$ (henceforth, the expression "agent i" will indicate an agent belonging to group i). Let $\tilde{n}$ denote the pre-tax income of an agent whose after-tax income is exactly 1 at the tax rate t. Then $\tilde{n}$ is defined implicitly by:

$$\tilde{n}(1-t)+(t-t^2)\bar{n}-1=0. \tag{3}$$

Thus, all agents with pre-tax income $n^i<\tilde{n}$ cannot invest in education at the tax rate t. $\tilde{n}$ as a function of t is depicted in Figures 1(a), (b) and (c), which show that the function has very different qualitative behaviours depending on whether $\bar{n}>4$ (a "rich" economy), $1<\bar{n}<4$ (an "intermediate income" economy), or $\bar{n}<1$ (a "poor economy").

Since the behaviour of the function $\tilde{n}(t)$ is crucial for the results of the model, it is important to obtain some intuition of its shape. Consider first a rich economy. At each tax rate, a large amount of resources are redistributed. Thus, however poor an agent is, there will always be a tax rate t, $t \leqq \frac{1}{2}$, such that her post-tax income exceeds 1. When *per capita* income is at an intermediate level, there might be a situation where an agent's pre-tax income is so small ($n<\tilde{n}_{min}$ in Figure 1(b)) that no tax rate will raise her post-tax income to 1 before the convexity of the cost of collecting taxes takes over. Finally, consider a very poor economy. If an agent starts with a pre-tax income below 1, no tax rate will ever enable her to invest in education: even in the absence of costs of collecting taxes she could reach at most a post-tax income equal to $\bar{n}$, which is less than 1. Moreover, by reducing the post-tax income of all agents with a pre-tax income above $\bar{n}$, income redistribution hurts all agents with pre-tax income above 1, and the more so the higher is the tax rate.

7. Note that, in the two-period model, it is irrelevant whether a storage technology exists, because of the linearity of preferences.

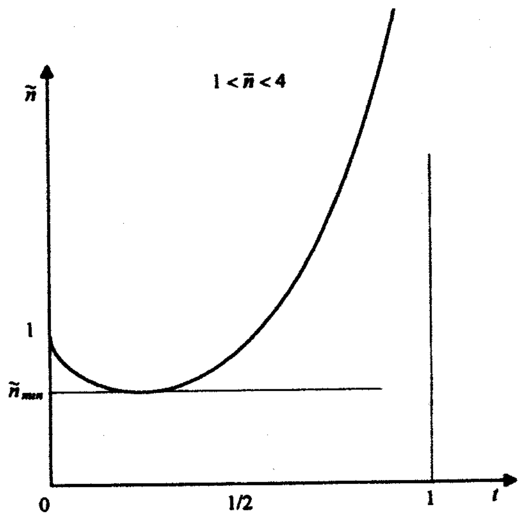

FIGURE 1(b)
$1<\bar{n}<4$

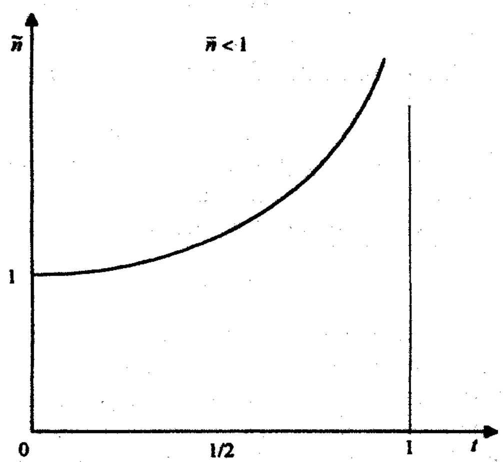

FIGURE 1(c)
$\bar{n}<1$

3. EXISTENCE OF A STABLE MAJORITY

In this section, I will prove that the median voter is the decisive voter when agents vote over the level of the tax rate in the first period.[8] The reason why a whole section is needed to establish this result is that, due to the abundance of discontinuities in the model, preferences may not be single-peaked and therefore the usual sufficient conditions for the existence of a stable majority cannot be applied directly. However, since the succeeding sections develop all the important conceptual issues, the reader uninterested in technical details can skip this section without missing any important intuition of the model.

8. In what follows, the expression "median voter" is synonymous with "an individual with first-period income n^m".

Consider the problem solved by agent i in period 1. Her proposal will be:

$$t_i = \max\{0, \operatorname{argmax}\{c_1^i + c_2^i\}\}. \tag{4}$$

Note that c_2^i depend on t for two reasons: first, the tax rate in period 1 determines whether agent i can invest in education; second, it determines how many agents invest in education, which affects the pre-tax income of agent i in the second period through the externality effect.

Now consider the term argmax $\{\cdot\}$ in equation (4). Wherever $\bar{n}_2$ and n_2^i are differentiable with respect to t, this term is found by solving:[9]

$$\frac{d[c_1^i + c_2^i]}{dt} = 0 \tag{5}$$

i.e.

$$-n^i + (1-2t)\bar{n} - \frac{dn_2^i}{dt} = 0. \tag{6}$$

Clearly, $dn_2^i/dt = 0$ whenever this derivative exists. Therefore, over all the points where n_2^i is differentiable with respect to t, the tax rate proposed by agent i will be

$$t_i^* = \max\left\{0, \frac{1}{2}\left(1 - \frac{n^i}{\bar{n}}\right)\right\} \tag{7}$$

i.e. in all points where n_2^i is differentiable, the optimal tax rate in period 1 for agent i is the tax rate that maximizes her post-tax income in the same period.

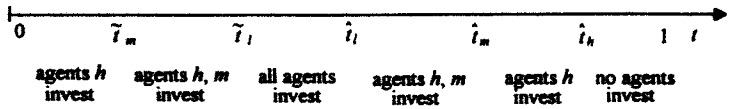

FIGURE 2

However, it is clear that there are several points of discontinuity of n_2^i as a function of t. The reason is that, whenever the tax rate reaches the level at which the post-tax income of agent i is equal to 1, all agents in group i invest in education, thereby increasing discretely their own pre-tax income in period 2 and the income of all other agents via the externality effect. The exact number of points of discontinuity depends on the values of n^l, n^m, n^h and $\bar{n}$. Figure 2 illustrates the case of $\bar{n} > 4$, $n^m < 1$, $n^l < 1$, i.e. the case with the largest number of discontinuities. $\tilde{t}_l$ and $\hat{t}_l$ are the smaller and larger root of $n^l(1-t) + (t-t^2)\bar{n} - 1 = 0$, while $\tilde{t}_m$ and $\hat{t}_m$ are defined similarly with n^m replacing n^l in the previous equation. In other words, as long as the tax rate is between $\tilde{t}_l$ and $\hat{t}_l$, agent l has enough pre-tax income to invest in education, and similarly for agent m. Since necessarily $n^h > 1$, when $\bar{n} > 1$ there is only one value of t, $\hat{t}_l$, such that $n^l(1-t) + (t-t^2)\bar{n} - 1 = 0$. If the tax rate exceeds $\hat{t}_h$, agent h will not be able to invest in education. From Figure 1(a) $\hat{t}_l$, $\hat{t}_m$ and $\hat{t}_h$ are all larger than $\frac{1}{2}$.

When making her proposal in period 1, each voter must compare the value of her overall utility when $t_i = t_i^*$ to its value when t_i is such that n_2^i changes discretely. For example, when t_m^* is between $\tilde{t}_m$ and $\tilde{t}_l$ in Figure 2, agent l cannot invest at the tax rate

9. It is easy to verify that the second-order conditions for a maximum are satisfied.

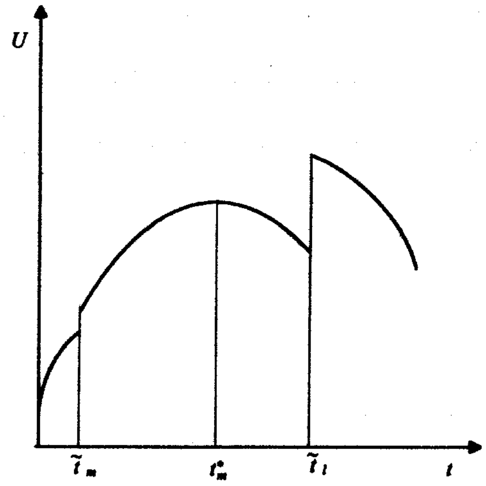

FIGURE 3(a)

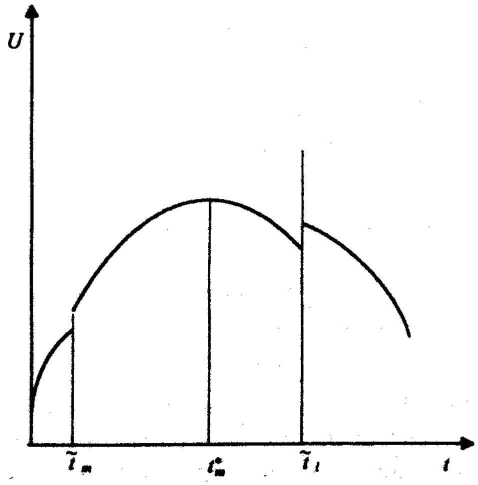

FIGURE 3(b)

that maximizes agent m's post-tax income in period 1. In this case, agent m's *overall* utility may be higher when $t=\tilde{t}_l$ than when $t=t_m^*$, because in the former case agents l can invest in education and therefore next period's pre-tax income of agent m will be higher as well through the human capital externality.[10]

It is now clear that in this model indirect utility functions are not always single-peaked as a function of t. Figure 3 illustrates the two possible qualitative behaviours of agent m's indirect utility in the case considered in Figure 2 and $t_m^* < \tilde{t}_l$ (the indirect utility is plotted only for $0 < t < \frac{1}{2}$ which will turn out to be the relevant range in equilibrium). The standard sufficient conditions for the existence of a stable majority with the median voter as the decisive voter fail to apply. However, it is still possible to show that the median voter is the decisive voter in this problem:

10. A more systematic discussion of this point is left for Appendix A and Sections 4 and 5.

Result 1. *The proposal by agent m beats all other proposals in pairwise comparison.*

Proof. See Appendix A. ‖

It can be shown that preferences in this model satisfy the condition of Order Restrictedness (see Rothstein (1989)). In fact, with three proposals it can be shown that a necessary and sufficient condition for Result 1 to hold is that preferences be Order Restricted. It is here that the importance of assuming a finite number of classes can be appreciated. In fact, the proof of Order Restrictedness requires a finite number of alternatives.

4. INCOME DISTRIBUTION, REDISTRIBUTION AND GROWTH

In this section I will investigate how the initial distribution of income affects the degree of redistribution and, through this, the potentials for growth of an economy. The analysis of the previous section established that the median voter is the decisive voter in all possible states: therefore, in what follows it is sufficient to analyze the optimal policies of the median voter in order to determine the equilibrium outcomes.

The next two sub-sections consider the two cases of a rich and a poor economy respectively. It will be shown that they have very different patterns of income distribution that are most favourable to growth. The dynamic implications of this simple fact will be more fully developed in Sections 5 and 6.

4.1. *The cases of a rich and an intermediate income economy*

Consider first an economy with a high per capita income, $\bar{n} \geqq 4$ (see Figure 1(a)). By Result A.2, $t^*_m < \hat{t}_h$ when $\bar{n} \geqq 1$; this means that agent h can always invest in education at the tax rate that maximizes the median voter post-tax income in period 1. Thus, the only situation in which the median voter might want to propose a tax rate different from t^*_m is when $t^*_m < \tilde{t}_l$, in which case agent l cannot invest in education at t^*_m because her post-tax income would be below 1 at t^*_m. When $t^*_m < \tilde{t}_l$ the median voter faces an intertemporal trade-off. If she sets $t = \tilde{t}_l$ she loses something in the first period relative to $t = t^*_m$, but clearly will gain something in period 2, since n^m_2 increases if agent l invested in period 1. How the trade-off is resolved by the median voter has important implications for growth: if $t = \tilde{t}_l$, high growth will result. Otherwise, growth will be low. Thus, in order to study the effects of income distribution on growth one must analyze two questions: (a) under what configurations of the relative shares of the low-income and middle-income groups will the median voter face an intertemporal trade-off? (b) if there is indeed a trade-off, what configurations of income will induce the median voter to set a high tax rate, so that the low-income group will invest and high growth will obtain?

Let $x(n^m, n^l)$ and $y(n^m, n^l)$ be the first-period loss and second-period gain to the median voter respectively from setting the tax rate at $\tilde{t}_l$ instead of t^*_m. Let $z(m^m, n^l) = y(n^m, n^l) - x(n^m, n^l)$ be the overall gain (if positive) or loss (if negative). Then, question (a) above is equivalent to finding the shape of the $x = 0$ locus in the (n^l, n^m) space (see Figure 4).[11] Above this locus n^l is sufficiently close to n^m that $t^*_m \geqq \tilde{t}_l$ and the median voter does not face a conflict between the short run and the long run. Below this locus there is indeed a conflict because $t^*_m < \tilde{t}_l$. Question (b) therefore corresponds to finding the locus $z = 0$ in

11. This and the following figures assume $p^l = p^m = 0{\cdot}4$ and $p^h = 0{\cdot}2$.

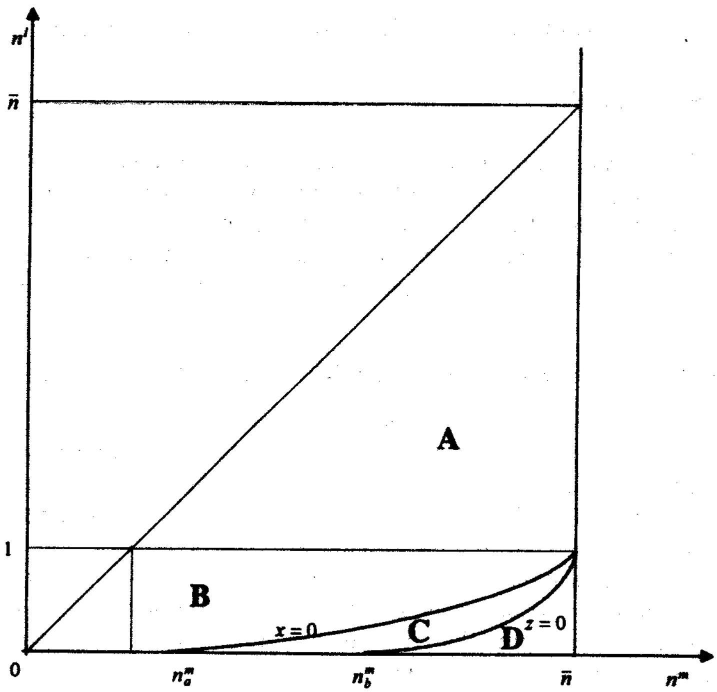

FIGURE 4
$\bar{n}=5$

the region below the $x=0$ locus. Above the $z=0$ locus n^l is sufficiently close to n^m that the extra progressivity of the tax system required to enable the low-income group to invest is small compared to the second-period gain; thus, the median voter will set the tax rate at $\tilde{t}_l$ and high growth will follow. Below the $z=0$ locus $z<0$ and the median voter will set the tax rate at t_m^* so that low growth will obtain. The following result formalizes this argument:

Result 2. *For an economy with $\bar{n} \geqq 4$:*

(a) *the $x=0$ locus is upward sloping and convex and is defined for $n^l \in [0, 1]$ and $n^m \in [n_a^m, \bar{n}]$, where n_a^m is a function of $\bar{n}$;*
(b) *the $z=0$ locus is upward sloping and convex and is everywhere below the $x=0$ locus;*
(c) *$z<0$ in the region comprised between the locus $z=0$ and $n^l=0$.*

Proof. See Appendix B ‖

Result 2 implies that the $x=0$ locus and the $z=0$ locus have the shapes depicted in Figure 4. The intuition behind it is straightforward. There are two relevant regions, $n^l>1$ and $n^l<1$. When $n^l>1$ (region A), the median voter does not face any intertemporal trade-off, since the low income group can afford investment in education even when there is no redistribution. When $n^l<1$, whether the low-income group invests depends on its position relative to the middle group. Consider fixing a value for n^l on the vertical axis; if n^m is not too distant from n^l (regions B and C) the low-income group will be able to invest in education; in region B, because the distance between n^m and n^l is so low that $t_m^* \geq \tilde{t}_l$; in region C, because it is not too costly for the median voter to deviate from t_m^* and implement

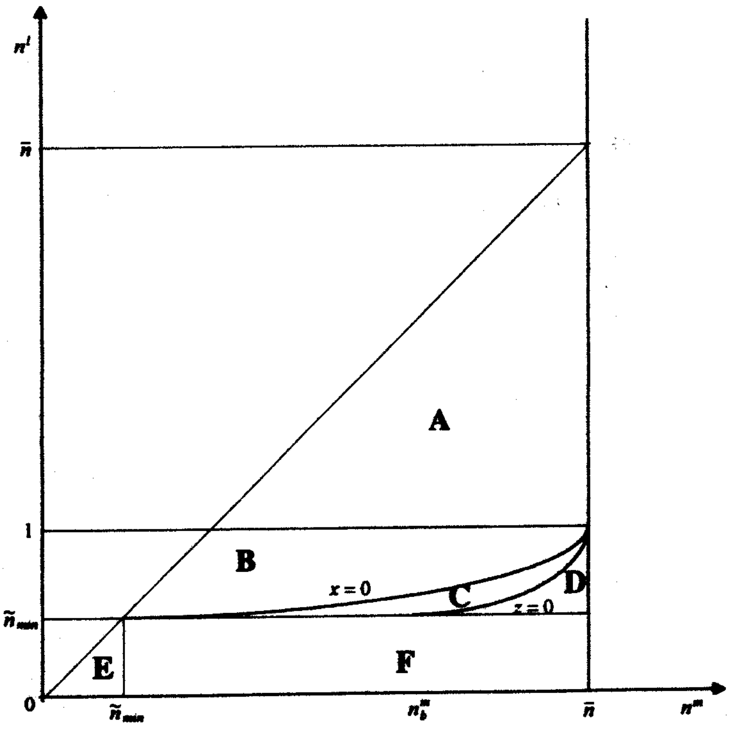

FIGURE 5

$\bar{n}=3$

$\tilde{t}_l$. However, if n^m is large relative to n^l (region D) then $z(n^m, n^l)<0$, so that the short-run cost to the median voter from high redistribution outweighs the long-run gain.

A result similar to Result 2 holds for an intermediate income economy too, with $1<\bar{n}<4$ (Figure 5). However, now the economy is poorer than the economy sketched in Figure 4. In particular, if $n^l<\tilde{n}_{min}$ no level of redistribution enables the low-income group to invest,[12] and the same is true for n^m if $n^m<\tilde{n}_{min}$. Apart from these differences, the logical structure of the problem remains the same as in the case of a rich economy. In particular, the shapes of the $x=0$ and $z=0$ loci can be explained by exactly the same considerations made above.

4.2. *The case of a poor economy*

Consider now the case of a poor economy, with $\bar{n}<1$. Only agents h can now possibly invest in education: any agent starting with a pre-tax income below $\bar{n}$ and therefore below 1 will never be able to reach a post-tax income of at least 1 (see Figure 1(c)). It is then clear that when $n^h<1$ no agent can invest in education, and therefore no growth can take place. Therefore, assume from now on that $n^h \geq 1$.

Now the only potential investors in the economy, agents h, are hurt by high tax rates. It is then intuitive that the median voter will face two relevant situations. If n^m is large given n^h (so that t_m^* is small) or n^h is large given n^m (so that $\hat{t}_h$ is large) the economy will

12. Note that this feature of the model depends crucially on the existence of convex costs of collecting taxes.

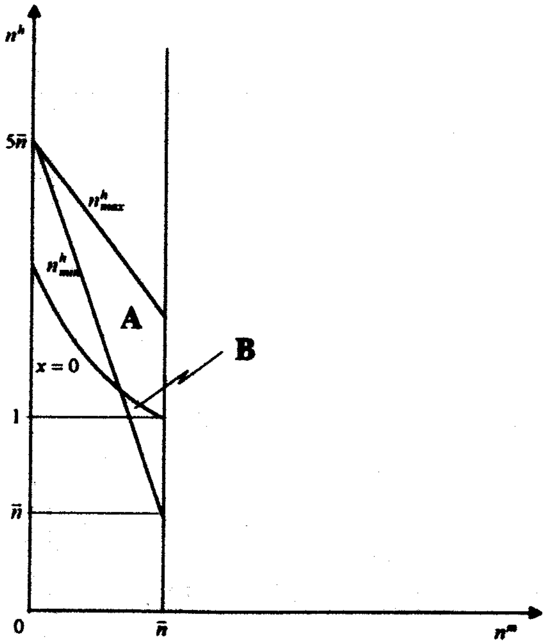

FIGURE 6

$\bar{n}=0{\cdot}5$

be in region A in Figure 6: here, $t_m^* \leq \tilde{t}_h$ and agents h will be able to invest in education at the tax rate preferred by the median voter.[13] In contrast, if n^m is low given n^h or n^h is low given n^m, the economy will be in region B, where $t_m^* > \tilde{t}_h$. Now agents h will not be able to invest in education at t_m^* and the median voter faces the familiar intertemporal trade-off. The only difference is that now she must trade *less* redistribution in period 1 for a higher per capita income in period 2.

One can therefore define two loci $x(n^m, n^h)=0$ and $z(n^m, n^h)=0$ in the (n^m, n^h) space in exact analogy to the case of a rich economy analyzed above. Thus, below the $x=0$ locus agents h cannot invest at t_m^*, while below the $z=0$ locus the short-period loss to the median voter from deviating from the optimal tax rate outweighs the long-run gain deriving from a higher second-period *per capita* income. It is therefore relatively easy to prove the following.

Result 3. *For an economy with $\bar{n}<1$:*

(a) *the $x=0$ locus is downward sloping and convex;*
(b) *the $z=0$ is downward sloping, convex and everywhere below the $x=0$ locus;*
(c) *for $\phi(p^h)$ sufficiently small, there exists an admissible region above $n^h=1$ where $z<0$ and $x>0$;*
(d) *for $\phi(p^h)$ and/or R sufficiently large, $z>0$ everywhere for $n^h \geq 1$.*

Proof. See Appendix B. ‖

13. In Figure 6 note that, for a given n^m, n^h can only take a value comprised between $n^h_{max}(n^m)$ and $n^h_{min}(n^m)$, where n^h_{max} is the value of n^h when $n^l=0$ (so that $p^h n^h_{max}=\bar{n}-p^m n^m$) and n^h_{min} is the value of n^h when $n^l=n^m$ (so that $p^h n^h_{min}=\bar{n}-(p^m+p^l)n^m$).

4.3. *Income shares, levels of income and growth*

The model developed so far delivers a clear message: economies with different per-capita incomes have very different patterns of income distribution (i.e. relative shares) that are most favourable to growth. In particular, income distribution affects growth through two channels.

First, in a very poor economy growth can occur only if the distribution of income is sufficiently unequal, so that $n^h > 1$. Similarly, in an intermediate income economy income distribution determines whether there is a tax rate at which the low-income group and the middle-income group can invest in education.

Second, income distribution affects the pre-tax income of the median voter *relative* to that of the group whose investment in education depends on the tax rate (n^h in a poor economy, n^l in a rich economy). This relative share in turn determines whether an intertemporal trade-off exists and, when it exists, whether the median voter has an incentive to set a tax system that promotes growth.

In a rich and an intermediate income economy the best preconditions for high growth (in the sense that both groups l and m invest) are a low share of group h and/or very similar shares of groups l and m (the region along the 45° line in Figures 4 and 5, assuming $p^m = p^l$). When the share of the high-income group is relatively low, the two remaining groups will start with a relatively high pre-tax income. When n^l is close to n^m, either $t_m^* \geq \tilde{t}_l$ or the median voter has relatively high incentives to let the low-income class invest through high redistribution.

Exactly the opposite configuration of income distribution favours high growth (*i.e. investment by group h*) *in a poor economy with* $\bar{n} < 1$. In Figure 6, if the share of the high-income group is very low, the economy will be below the $n^h = 1$ line, so that no tax rate will allow agents h to invest. Also, if n^l is very close to n^m (along and close to the n^h_{min} line[14]) then the economy will be more likely to be in the region where $z < 0$, if it exists. The intuition for this result is obvious: in a poor economy, not even a very progressive tax system will allow low-income agents to invest, so that only h agents can potentially invest. Thus, any pattern of income distribution that endangers the investment ability of the high-income agents can harm growth.

5. A NON-OVERLAPPING GENERATIONS MODEL WITH VOTING

In this section, I will develop an infinite-horizon, non-overlapping generations extension of the two-period model of Sections 3 and 4. This is necessary to study the effects of the mechanism analyzed so far on the dynamic path of the economy. Indeed, by considering an explicitly dynamic economy, this section has three main objectives: (i) to formalize a process of growth, implicit in the two-period model, whereby a group investing in education increases the future income of all groups and therefore may enable the groups further down the ladder to invest in education as well; (ii) to analyse the degree of persistence in the evolution of the economy stemming from the initial pattern of income distribution; (iii) finally, to relate the result of the analysis in (ii) to the empirical evidence on the relation between income distribution and growth.

The mechanisms at work in the infinite-horizon version are essentially a straightforward extension of those operating in the two-period model. In order to concentrate on the conceptual issues, in this section I will only set up the model, outline the method of

14. Recall that the n^h_{min} line in the (n^m, n^h) space represents the same points as the 45° line in the (n^l, n^m) space.

solution and then discuss its implications for the three issues listed above. A full formal treatment of the model is left for Appendix C.

Agents live for two periods. Just before dying, each individual gives birth to an individual belonging to the next generation. Thus, in each period only one generation is alive. In the first period of her life, an agent can invest in education exactly as in the two-period model, and the effects of such an investment are as in the two-period model. Every individual inherits from her parent the "innate" ability to earn income n^i but not that acquired through education. Before an individual dies, she leaves her successor a bequest. Let a generation be indexed by the time-period when it is young. Therefore, generation s is young in period s. The world starts in period 1 when the old of generation 0 are alive. Preferences are additively time-separable, linear in first-period consumption and Cobb–Douglas in second-period consumption and bequests. Lifetime utility for agent i in generation s is therefore:

$$b_s^i+n^i(1-t)+(t-t^2)\bar{n}_s-e+(n^i+Re+\mu_s-b_{s+2}^i)^\gamma(b_{s+2}^i)^{1-\gamma} \tag{8}$$

where $0<\gamma<1$, μ_s is the measure of agents that invested in education in period s, b_s^i is the bequest received by agent i in generation s and b_{s+2}^i is the bequest left by the same agent.

This specification has the very useful implication that the indirect lifetime utility is again linear in e. Indeed, it is easy to show that b_{s+2}^i is proportional to the income of agent i in generation s when old:

$$b_{s+2}^i=(1-\gamma)(n^i+Re+\mu_s R). \tag{9}$$

Therefore, the indirect utility of agent i as a function of t is:[15]

$$n^i(1-t)+(t-t^2)\bar{n}_s-e+\beta(n^i+Re+\mu_s R) \tag{10}$$

where $\beta=\gamma^\gamma(1-\gamma)^{1-\gamma}$. Since $\beta\leq\frac{1}{2}$, now $R\geq 2$ is a necessary condition for an agent to be willing to invest in education whenever her post-tax income exceeds 1. Therefore, from now on $R\geq 2$ will be assumed. Given these hypotheses, the infinite-horizon model is essentially a sequence of two-period models, with some minor modifications due to the presence of bequests (see Appendix C for details). However, by following the economy over more than two periods, it is now possible to analyze all the implications of the strong path-dependence built into the two-period model. In fact, Section 4 showed that high growth is associated with different patterns of income distribution in economies with different *per capita* incomes. In an infinite-horizon version, this property of the model has rather important implications.

Essentially, a given pattern of income distribution can be extremely appropriate for growth at a certain level of income; once the economy has reached a higher level of income, however, that same pattern of income distribution might hamper or, in extreme cases, prevent growth. This is so because pre-tax income distribution is essentially a state variable, and highly dependent on initial conditions. Also, the feasibility of changing the post-tax income distribution depends on the characteristics of the political equilibrium resulting from the pattern of pre-tax income distribution.

Thus, an important characteristic of the model is that the steady-state reached by the economy is highly sensitive to the initial distribution of income. In particular, an economy that starts out at a very low level of income ($\bar{n}(2-\gamma)<1$) with very high values of n^h

15. Notice that now the model must assume the existence of a storage technology, since each individual inherits something from her parent and the two generations do not overlap. However, no agent will want to transfer resources from the first to the second period of her life through this storage technology, because $\beta<1$ for $0<\gamma<1$.

certainly satisfies the preconditions for growth at that level of income, as described in Section 4, and therefore can move to the next level of income ($1<\bar{n}_1(2-\gamma)<4$). Once there, however, high growth can occur only with a very different configuration of income distribution. As a consequence, after an initial spurt of high growth a very unequal society might get stuck at a relatively low level of income with an even worse income distribution than the initial one.

This mechanism may be potentially relevant in connection with the long-standing debate on the existence of an inverted-U relation between inequality and per capita income. Empirically, this relation seems to be quite robust in cross-section studies, and has been consistently obtained for more than three decades.[16] However, time-series studies tend to cast doubts on the shape of the relation.[17] Essentially, the growth process seems to be consistent with a wide variety of behaviours of income distribution measures over time. The path-dependence displayed by the overlapping-generations model of this section might explain these empirical regularities. The point is best made by way of an example (Appendix C generalizes the result to a certain extent). It is assumed here that $p^l=p^m=p^h=\frac{1}{3}$, $\gamma=\frac{4}{5}$ and $R=2$. Also, $n^l=n^m$ only for simplicity. Consider now three economies, A, B and C with the same initial per capita income in period 1 $\bar{n}_1=0{\cdot}82$, so that $\bar{n}_2=\bar{n}_1(2-\gamma)=0{\cdot}98$, but with very different patterns of income distribution, as specified in the tables below. Letting j indicate the time period, the evolution of average income and income distribution in each economy is as follows:

Country A

$j=1$	$\bar{n}_1=0{\cdot}82$	$n^l=0{\cdot}82$	$n^m=0{\cdot}82$	$n^h=0{\cdot}82$
$j=2$	$\bar{n}_1(2-\gamma)=0{\cdot}98$	$n^l(2-\gamma)=0{\cdot}98$	$n^m(2-\gamma)=0{\cdot}98$	$n^h(2-\gamma)=0{\cdot}98$
$j=3$	$\bar{n}_3=0{\cdot}82$	$n^l=0{\cdot}82$	$n^m=0{\cdot}82$	$n^h=0{\cdot}82$
$j=5$	$\bar{n}_5=0{\cdot}82$	$n^l=0{\cdot}82$	$n^m=0{\cdot}82$	$n^h=0{\cdot}82$

Country B

$j=1$	$\bar{n}_1=0{\cdot}82$	$n^l=0{\cdot}81$	$n^m=0{\cdot}81$	$n^h=0{\cdot}84$
$j=2$	$\bar{n}_1(2-\gamma)=0{\cdot}98$	$n^l(2-\gamma)=0{\cdot}97$	$n^m(2-\gamma)=0{\cdot}97$	$n^h(2-\gamma)=1{\cdot}01$
$j=3$	$\bar{n}_3=2{\cdot}15$	$n^l=0{\cdot}81$	$n^m=0{\cdot}81$	$n^h=3{\cdot}67$
$j=5$	$\bar{n}_5=4{\cdot}82$	$n^l=4{\cdot}81$	$n^m=4{\cdot}81$	$n^h=4{\cdot}84$

Country C

$j=1$	$\bar{n}_1=0{\cdot}82$	$n^l=0{\cdot}0$	$n^m=0{\cdot}0$	$n^h=2{\cdot}46$
$j=2$	$\bar{n}_1(2-\gamma)=0{\cdot}98$	$n^l(2-\gamma)=0{\cdot}0$	$n^m(2-\gamma)=0{\cdot}0$	$n^h(2-\gamma)=2{\cdot}95$
$j=3$	$\bar{n}_3=2{\cdot}15$	$n^l=0{\cdot}0$	$n^m=0{\cdot}0$	$n^h=6{\cdot}46$
$j=5$	$\bar{n}_5=2{\cdot}15$	$n^l=0{\cdot}0$	$n^m=0{\cdot}0$	$n^h=6{\cdot}46$

Note that necessarily all these countries reach a steady-state at most in period 5, where a steady state is defined as a situation in which the economy repeats itself every two periods. Now let S_j^h be the share in total income of group h in period j. This is a measure of inequality frequently used in applied work, in particular in several studies of the inverted U-curve. Economy A starts with a completely egalitarian income distribution, so that

16. The story of the inverted-U curve dates back to Kuznets (1955). For a recent contribution supporting Kuznets' hypothesis see Campano-Salvatore (1988). For a less sympathetic view on the existence of an inverted-U relation in cross-sections, see Ram (1988).

17. See especially Fields-Jakubson (1990).

$S_1^h = \frac{1}{3}$. Economy B is only slightly more inegalitarian: $S_1^h = 0{\cdot}35$. Finally, economy C is characterized by a very unequal income distribution: $S_1^h = 1$. As shown above, economy A cannot grow; between B and C, B is clearly worse equipped for initial growth: in fact, it barely succeeds in starting the development process. However, once the growth process has started, B is in a better position to continue growth than C. Indeed, in the second period C reaches a steady-state where only the high income class is investing in education, and therefore income distribution is even more unequal than initially. By contrast, economy B reaches the steady state in the third period, with all classes investing. Thus, income distribution has improved after the first increase in inequality *and* steady-state income is higher than C's steady-state income.

Now suppose an econometrician observes these economies after they have reached their steady-states[18] and tries to fit the best curve in the $(S^h, \bar{n})$ space: *such a curve will be an inverted-U* (see Figure 7). The reason is simple: the economies that in steady-state have a higher income level are those whose initial income distribution enabled them to deal best with the different phases of economic development. In very egalitarian economies, like A, no investment in human capital could ever take place. In economies with a very unequal income distribution, like C, the middle and/or the lower class are so poor that not even the maximum feasible level of redistribution will enable them to invest in education. Only economies that started out sufficiently equal, but not excessively so, have the ability both to start growth and to keep growing once an intermediate level of income is reached.

Note however that the time-series behaviour of S^h presents an inverted-U pattern only in the case of economy B, while in country C it only increases and in country A it never moves. This seems to be consistent with the available empirical evidence in two respects. First, as mentioned above, the time-series behaviour of inequality measures is known to follow a variety of patterns. Second, the presence of an inverted-U pattern in

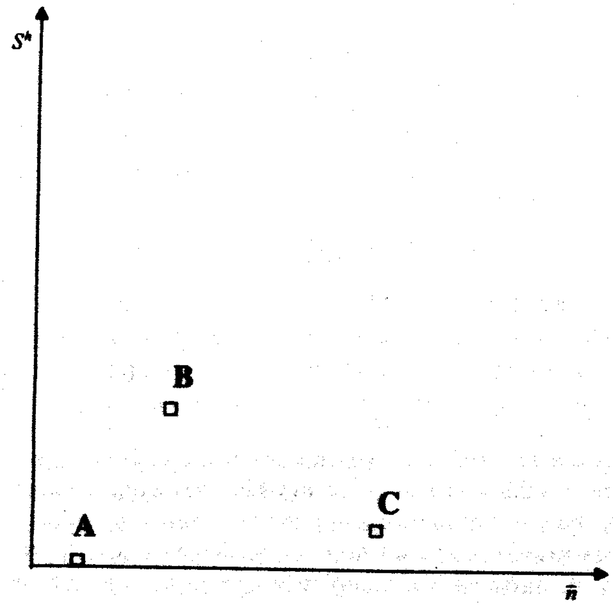

FIGURE 7

18. For simplicity, I am assuming here that the economy is observed every two periods when the old are alive.

time series has been documented quite convincingly for several currently industrialized countries, including U.S., Great Britain, Germany, Norway, Denmark, the Netherlands,[19] while high and increasing levels of inequality are more common among intermediate income economies.

The non-overlapping generations model also sheds further light on the "trickle down" process of growth that is only implicit in the two-period model. By increasing the productivity of all income groups, investment by the upper class in the first period might allow the other classes to invest in the following period. A comparison of economies B and C reveals what is a precondition for this "trickle-down" mechanism to operate: the pre-tax income of the groups that rely on this mechanism should be above a certain threshold level, below which there is no level of redistribution that allow investment in education. Similarly, under some circumstances investment in education by the middle class will enable the low-income group to invest in the next period.

This also illustrates the crucial role played by political factors in the growth process. Essentially, the political outcome resulting from a given income distribution determines whether the intertemporal transmission of the externality outlined above goes on until all classes have invested or it stops before this occurs.

6. DISCUSSION AND CONCLUSION

An important assumption of the model is that taxes are linear and revenues are rebated to individuals in a lump-sum fashion. The reason for this assumption is exclusively one of analytical tractability: it is well known that with non-linear taxes the existence of a non-cycling majority may fail. Notice, however, that if one introduced two ideologically committed parties it might be possible to have a stable majority with non-linear taxes. In such a situation one could explore the interesting possibility that, in the political equilibrium, the high-income class is made to pay almost all the costs of investment in education by the other classes.

More generally, the tax-subsidy scheme assumed in the model implies that it is not possible to subsidize the middle class without subsidizing the poor as well. Although this is a long-standing and unresolved issue, many researchers have argued that the middle class captures a disproportionate part of the benefits of government expenditure. Again, this important aspect cannot be captured in this model. However, the essence of this paper is that when growth is associated with redistribution, the benefits spill over to *some* extent to the poor; when the spillover is substantial, this enables the poor to qualitatively change their pattern of education. In this sense, the assumption on the distribution of benefits in this paper plays an important role in proving the existence of a political equilibrium, but is not strictly necessary for the *economics* of the model.

The model also assumes that the benefits of investment in education by the poor are captured by all classes. One might argue that the rich benefit the most from a better educated work force, since in general it is the rich who hire labour. In this case there would be two effects, working in opposite directions, on the degree of progressivity preferred by the high-income class. If the effect just described prevails, one would observe both the rich and the poor vote for high redistribution. Again, it is not clear whether a non-cycling majority would exist.

19. See Lindert-Williamson (1985) for a brief review of the time-series experience of currently industrialized countries.

Another situation in which both the rich and the poor might vote for similar amounts of redistribution arises when one considers the possibility of publicly-provided education. This is a case in which possibly the rich and the poor will vote for a *low* amount of redistribution, the former because their tax price is higher, the latter because, with enough curvature in the utility function, the opportunity cost of going to school may be extremely high at low levels of income. Fernandez and Rogerson (1991) develop a model of voting on public expenditure where this result obtains under certain configurations of the distribution of income.

Finally, it is clear that it would be useful to have some empirical evidence on the mechanisms of growth presented in this model and on some possible alternatives, some of which have been sketched out above. Alesina-Rodrik (1991), Perotti (1993) and Persson-Tabellini (1991) take a first step in this direction.

APPENDIX A

This Appendix proves Result 1 in Section 3. To this end, I will first prove some preliminary results.

Result A.1. *Consider the case* $1<\bar{n}<4$. *Let* $t_{\min}$ *be defined by* $\bar{n}_{\min}(1-t_{\min})+(t_{\min}-t^2_{\min})\bar{n}-1=0$, *i.e.* $t_{\min}$ *is the tax rate at which* $d\bar{n}/dt=0$ *in Figure* 1(*b*). *Then,* $t_i^*>t_{\min}$ *if and only if* $n^i<\bar{n}_{\min}$.

Proof. The proof is immediate upon manipulation of the expressions for t_i^* and $t_{\min}$. ‖

Essentially, Result A.1 says that t_i^* is on the upward-sloping part of the $\bar{n}(t)$ curve if and only if $n^i<\bar{n}_{\min}$.

Result A.2. $t_m^*\leq \hat{t}_h$, *for* $\bar{n}>1$.

Proof. Result A.1 ensures that, when $\bar{n}>1$, agents *h* might be unable to invest at t_m^* only when $n^m<\bar{n}_{\min}$ (which implies necessarily $1<\bar{n}<4$). To show that this will never occur, consider the smallest possible value of n^h corresponding to each value of n^m. Clearly, given n^m, n^h will be smallest when $n^l=n^m$, so that $n^h_{\min}$ is defined by:

$$p^h n^h_{\min}=\bar{n}-(1-p^h)n^m. \tag{11}$$

Clearly:

$$n^h_{\min}(1-t_m^*)+(t_m^*-t_m^{*2})\bar{n}-1\geqq n^h_{\min}(1-t_m^*)-1=n^h_{\min}\frac{1}{2}\left(1+\frac{n^m}{\bar{n}}\right)-1. \tag{12}$$

Define $H(n^m)=n^h_{\min}\tfrac{1}{2}(1+n^m/\bar{n})-1$, where $n^h_{\min}$ is a function of $\bar{n}$ from (11). Since H is quadratic in n^m it is sufficient to show that:

—$dH(n^m)/dn^m>0$ when evaluated at $n^m=0$;

—$H(0)>0$, $H(\bar{n}_{\min})>0$.

Now:

$$\frac{dH}{dn^m}=\frac{1}{2}-\frac{n^m}{\bar{n}}\left(\frac{1-p^h}{p^h}\right). \tag{13}$$

Thus, at $n^m=0$, $dH(n^m)/dn^m=\tfrac{1}{2}$. It is also easy to show that $H(0)\equiv\bar{n}/2p^h-1\geqq 0$ for $\bar{n}\geqq 1$. By computing $H(\bar{n}_{\min})$, one finds after some manipulations that $H(\bar{n}_{\min})$ achieves a minimum for $\bar{n}=1$, where $H(\bar{n}_{\min})=1$, so that certainly $n^h_{\min}(1-t_m^*)+(t_m^*-t_m^{*2})-1$ is strictly positive for all $\bar{n}\geqq 1$. ‖

Result A.2 ensures that, when $\bar{n}>1$, the tax rate that maximizes the post tax income of the median voter can never be so high as to prevent agents *h* from investing in education.

It is now relatively easy to prove.

Result 1. *The proposal by agent m beats all other proposals in pairwise comparison.*

Proof. It is useful to prove first the following results:

(i) agent h never proposes a tax rate higher than that proposed by agent m;
(ii) agent l never proposes a tax rate lower than that proposed by agent m.

To prove (i), consider first an economy with $\bar{n} \geq 1$. Define:

$$W(n^i, n^r) = \begin{cases} -K & \text{if } n^r < \bar{n}_{min}; \\ -K & \text{if } t_i^* \geq \tilde{t}_r; \\ z(n^i, n^r) & \text{if } t_i^* < \tilde{t}_r \end{cases} \tag{14}$$

where $K>0$, $i=h, m$, $r=m, l$ and $z(n^i, n^r)$ is defined as the difference between the overall utility of agent i when $t=\tilde{t}_r$ and her overall utility when $t=t_i^*$. Note that m and l represent the two groups whose ability to invest might depend on the tax rate. Note also that this formulation includes the case where agent m cannot invest at $t_h^*=0$, so that $W(n^i, n^r)=W(n^m, n^m)=-K$. Clearly, whenever $W(n^i, n^r)<0$, agent i prefers t_i^* to $\tilde{t}_r$. When $W(n^i, n^r) \geq 0$, n^i prefers $\tilde{t}_r$. Part (i) follows immediately from the fact that $W(n^h, n^r) \leq W(n^m, n^r)\ \forall n^r$, so that it cannot happen that agent h prefers $\tilde{t}_r$ when agent m prefers $t_m^* < \tilde{t}_r$, and from the fact that $t_h^* = 0$. In the case of an economy with $\bar{n}<1$, one can define

$$W(n^i, n^h) = \begin{cases} -K & \text{if } t_i^* \leq \hat{t}_h; \\ z(n^i, n^r) & \text{if } t_i^* > \hat{t}_h \end{cases} \tag{15}$$

and proceed as in the case $\bar{n} \geq 1$. Part (ii) can be proved following a similar procedure.

Now let t_i indicate the proposal by agent i. To prove that agent h always prefers t_m to t_l, note that this statement is obvious if agent l can invest at t_m or $n^l < n_{min}$ or $\bar{n}<1$. If this is not the case, necessarily agent m proposes t_m^* and agent l proposes $t_l > t_m^*$. Then one can easily show that $U^h(t_l) \leq U^h(t_m^*)$. Using a similar procedure, one can show that agent l always prefers agent m's proposal to agent h's proposal. ∎

In an earlier version of this paper, Result 1 was proved by showing that preferences in the model satisfy the condition of Order Restrictedness (see Rothstein (1989)). In fact, it is easy to show that the proof of Result 1 above and the one in the earlier version are essentially the same.

APPENDIX B

This Appendix proves Results 2 and 3 in Section 4.

Result 2. *For an economy with* $\bar{n} \geq 4$:

(a) *the* $x=0$ *locus is upward sloping and convex and is defined for* $n^l \in [0, 1]$ *and* $n^m \in [n_a^m, \bar{n}]$, *where* n_a^m *is a function of* $\bar{n}$;
(b) *the* $z=0$ *locus is upward sloping and convex and is everywhere below the* $x=0$ *locus*;
(c) $z<0$ *in the region comprised between the locus* $z=0$ *and* $n^l=0$.

Proof. The proof consists of several steps:

(a) The locus of points n^l, n^m such that $t_m^*=\tilde{t}_l$ is also the locus of points such that $x=0$. In fact, since $t_m^*=\text{argmax}\,\{c_i\}$ and t_m^* is unique, $c(t_m^*)>c(\tilde{t}_l)\ \forall t_m^* \neq \tilde{t}_l$. Thus, $x(t_m^*, \tilde{t}_l)>0\ \forall t_m^* \neq \tilde{t}_l$. This means that only for $\tilde{t}_l=t_m^*$ is $x=0$.[20] Using the implicit function theorem, it is easy to verify that along the $x=0$ curve $dn^l/dn^m>0$. It is also obvious that $x(n^m=\bar{n}, n^l=1)=0$, while $x(n_a^m, 0)=0$ (in the case $\bar{n} \geq 4$) and $x(\bar{n}_{min}, \bar{n}_{min})=0$ (in the case $1<\bar{n}<4$).
(b) Along the $z=0$ locus, $dn^l/dn^m>0$. This follows immediately from the implicit function theorem and the envelope theorem, since when $z=0$:

$$\frac{dn^l}{dn^m} = -\frac{t_m^* - \hat{t}_l}{[-n^m+(1-2\hat{t}_l)\bar{n}](d\hat{t}_l/dn_l)} > 0. \tag{16}$$

To show that the $z=0$ locus lies everywhere below the $x=0$ locus, notice first that $y=p^hR$ is independent of n^l and $y \geq 0\ \forall n^l, n^m$. Consider a point (n^m, n^l) such that $x(n^m, n^l)=0$. Then $z(n^m, n^l)>0\ \forall n^m \neq \bar{n}$. To obtain

20. $\tilde{t}_l=(1/2\bar{n})(\bar{n}-n_l-\sqrt{(\bar{n}+n^l)^2-4\bar{n}})$. Note that $\sqrt{(\bar{n}+n^l)^2-4\bar{n}} \geq 0$ for $n^l \geq \bar{n}_{min}$, since $\bar{n}_{min}=2\sqrt{\bar{n}}-\bar{n}$.

$z=0$ one must decrease n^l, since $x(n^m, n^l)$ is decreasing in n^l. Finally, convexity follows from direct verification.

(c) this part follows directly from the proof of part (b). ‖

I will now prove

Result 3. *For an economy with $\bar{n}<1$:*

(a) *the $x=0$ locus is downward sloping and convex;*
(b) *the $z=0$ is downward sloping, convex and everywhere below the $x=0$ locus;*
(c) *for $\phi(p^h)$ sufficiently small, there exists an admissible region above $n^h=1$ where $z<0$ and $x>0$;*
(d) *for $\phi(p^h)$ and/or R sufficiently large, $z>0$ everywhere for $n^h\geq 1$.*

Proof.

(a) By the implicit function theorem it is easy to show that the $x(n^m, n^h)=0$ locus is downward sloping; also, $x(0, 2-\frac{1}{2}\bar{n})=0$ and $x(\bar{n}, 1)=0$;
(b) Now consider the $z(n^m, n^h)=0$ locus. By the implicit function theorem and the envelope theorem, along the $z=0$ locus:

$$\frac{dn^h}{dn^m}=-\frac{t_m^*-\hat{t}_h}{[-n^m+\bar{n}(1-2\hat{t}_h)](d\hat{t}_h/dn^h)}. \tag{17}$$

It is also clear that $z=0$ below the $x=0$ locus, since $z=-x+\phi(p^h)R$. Convexity of both loci is proved by direct verification, as in Result 2.

(c) This part follows from the fact that

$$z(n^m, n^h)=-x(n^m, n^h)+\phi(p^h)R \tag{18}$$

using a simple continuity argument.

(d) $z>0$ everywhere if $z(0, 1)>0$, since the $z=0$ locus is downward sloping. The statement of part (d) is proven by showing that $z(0, 1)=-(\bar{n}/4)+\phi(p^h)R\geq-\frac{1}{4}+\phi(p^h)R$. ‖

APPENDIX C

Consider an infinite-horizon economy where in the first period, period 1, the old of generation 0 are alive and the average income is $\bar{n}$. In period 2 the average income of the young of generation 2 is $\bar{n}(2-\gamma)$. Consequently, a "poor" economy is now defined as one in which $\bar{n}(2-\gamma)<1$.

To analyze the implications of the overlapping-generations model, I will follow through time the evolution of an economy that starts in period 2 with average income $\bar{n}_2=\bar{n}\ (2-\gamma)<1$. Since the main purpose of the analysis is to determine under what patterns of income distribution the various classes invest in education, I will consider economies that have a chance to reach the highest level of income where all classes have invested. This amounts to imposing the two following conditions:

$$\bar{n}_4=\bar{n}(2-\gamma)+2p^hR(1-\gamma)>1 \tag{19}$$

$$\bar{n}(2-\gamma)+p^hR(1-\gamma)>\bar{n}_{min}(\bar{n}_4) \tag{20}$$

where now $\bar{n}_{min}(\bar{n}_4)=2\sqrt{\bar{n}(2-\gamma)+2p^hR(1-\gamma)}-\bar{n}(2-\gamma)-2p^hR(1-\gamma)$. Condition (19) says that, after group h has invested, the average income of the young of the next generation is larger than 1. Condition (20) says that there is *some* level of redistribution that enables group m to invest in education once group h has invested when group m has the highest possible pre-tax income (represented by the L.H.S. of inequality (19)). Since it can be shown that condition (19) holds whenever (20) holds, from now on only the latter will be considered.

Now consider generation 2 that starts in period 2 with $\bar{n}(2-\gamma)<1$. If $n^h(2-\gamma)\leq 1$, nobody can invest in education and from now on the economy repeats itself every two periods. If agents h invest, in period 4 the pre-tax incomes of the different classes are $n^i+(n^i+(1+p^h)R)(1-\gamma)$, $i=h$, and $n^j+(n^j+p^hR)(1-\gamma)$, $j=m$, l, respectively. If $n^m+(n^m+p^hR)(1-\gamma)<\bar{n}_{min}(\bar{n}_4)$, agents m will not be able to invest in education and from now on the economy will repeat itself every two periods. If $n^m+(n^m+p^hR)(1-\gamma)>\bar{n}_{min}(\bar{n}_4)$, certainly agents m will invest. In this case, pre-tax incomes in period 6 will be $n^i+(n^i+(1+p^h+p^m)R)(1-\gamma)$, $i=h$, m, and $n^j+(n^j+(p^h+p^m)R)(1-\gamma)$, $j=l$, respectively. The problem solved by the median voter is exactly the one analyzed in Section 4. Agents l will invest in period 6 if this is possible at the tax rate preferred by agents m or, when agents m face an intertemporal trade-off, if the two pre-tax incomes are not too different. Note also that agents l might be able to invest directly in period 4, together with n^m, if $n^l+(n^l+p^h R)(1-\gamma)>\bar{n}_{min}(\bar{n}_4)$ and one of the two conditions above is realized. The economy reaches a steady-state whenever no new class invests in education. Once the steady-state is reached, the economy repeats itself every two periods.

Consider now four economies that started in period 2 with the same initial per-capita income $\bar{n}(2-\gamma)<1$ and have reached a steady state, with 0, 1, 2, 3 classes having invested in education respectively. Assume for simplicity that these economies are observed in steady-state in odd periods, i.e. when an old generation is alive. Let $S_j^h(K)$ be the share of the high income class in period j when k classes have invested in education in the previous period. Since there are too many parameters in the model to be able to give a general result for all their possible values, for illustrative purposes I will assume that $p^i=\frac{1}{3}$, $i=h, m, l$ and $\gamma=\frac{4}{3}$.[21] For this set of parameters, it is possible to prove the following.

Result C.1. *For $\bar{n}$ sufficiently high*: (i) $S^h_{max}(0)<S^h_{min}(1)$; (ii) $S^h_{min}(2)>S^h_{max}(3)$.

Proof. (i) Since agents h invest in period 2 whenever $n^h(2-\gamma)\geqq 1$, $S^h_{min}(1)>S^h_{max}(0)$ if and only if:

$$\frac{1/(2-\gamma)+(R/3)+R}{\bar{n}+(2R/3)}>\frac{1/(2-\gamma)}{\bar{n}}. \tag{21}$$

Given that $\gamma=\frac{4}{3}$, after some algebraic manipulation inequality (21) becomes $\bar{n}-\frac{5}{12}>0$.[22] (ii) $S^h_{min}(2)>S^h_{max}(3)$ if and only if

$$\frac{1/(2-\gamma)+R(p^h+p^m)+R}{\bar{n}+2R(p^h+p^m)}>\frac{\bar{n}-p^m\bar{n}_{min}+p^h 2R}{p^h(\bar{n}+2R)} \tag{22}$$

After some manipulation inequality (22) can be written as

$$F=\delta\bar{n}+2\delta R-\tfrac{13}{3}\bar{n}R-3\bar{n}^2+\tfrac{2}{3}R^2+\bar{n}_{min}\bar{n}+\tfrac{4}{3}\bar{n}_{min}R>0. \tag{23}$$

where $\delta=1/(2-\gamma)$. It is easy to show that $\partial F/\partial\bar{n}<0\ \forall R$ and $\partial F/\partial R>0\ \forall\bar{n}$ so that $F(\bar{n}, R)$ reaches a minimum when $\bar{n}$ is at a maximum and when R is at a minimum, i.e. when $\bar{n}=\delta$ and $R=2$. For these values, it can be shown that F is positive. ‖

Acknowledgements. I am grateful to Alberto Alesina, Nathaniel Beck, Olivier Blanchard, Sugato Dasgupta, Peter Diamond, Rudi Dornbusch, Alessandro Missale and Federico Sturzenegger for their insightful comments and helpful discussions, and to seminar participants at Boston University, Brown, Caltech, Columbia, Harvard, London School of Economics, MIT, NBER and Princeton. I also thank two anonymous referees for their comments that led to substantial improvements in the paper.

REFERENCES

ADELMAN, I. and ROBINSON, S. (1988), "Income Distribution and Development", in H. Chenery and T. N. Srinivasan, *Handbook of Development Economics* (Amsterdam: North Holland).

AGHION, P. and BOLTON P. (1991), "A Trickle-Down Theory of Growth and Development with Debt Overhang" (Working paper, DELTA, Paris).

ALESINA, A. and RODRIK, D. (1991), "Distributive Politics and Economic Growth" (NBER working paper 3668).

BANERJEE, A. and NEWMAN, A. (1991), "Risk Bearing and the Theory of Income Distribution", *Review of Economic Studies*, **58**, 211–235.

BENHABIB, J. and RUSTICHINI, A. (1991), "Social Conflict, Growth and Income Distribution" (mimeo, New York University).

BERTOLA, G. (1991), "Market Structure and Income Distribution in Endogenous Growth Models" (NBER working paper 3851).

---

21. Note that now $z>0$ always whenever $\bar{n}(2-\gamma)<1$ and $t_m^*>\hat{t}_h$, so that m agents will always prefer to let h agents invest in education whenever the latter cannot at t_m^*. This result follows from the fact that $z(0,1)=-[\bar{n}(2-\gamma)/4]+R/3>0$ since necessarily $R\geqq 2$.

22. Note that, for the economy to be able to grow at all, it must be the case that the maximum possible value of $n^h(2-\gamma)$ is at least equal to 1. This implies $\bar{n}(2-\gamma)/p^h>1$, which in turn implies $\bar{n}>\frac{5}{18}$. Therefore, this last condition is not binding for the proof of part (i) of Result C.1.

CAMPANO, F. and SALVATORE, D. (1988), "Economic Development, Income Distribution and the Kuznets Hypothesis", *Journal of Policy Modelling*, **10**, 265–280.

CHANG, R. (1992), "Political Party Negotiations, Income Distribution, and Endogenous Growth" (New York University discussion paper).

CHISWICK, B. R. (1971), "Earnings Inequality and Economic Development", *Quarterly Journal of Economics*, **85**, 21–39.

FERNANDEZ, R. and ROGERSON, R. (1991), "On the Political Economy of Educational Subsidies" (mimeo, Hoover Institution).

FIELDS, G. S. and JAKUBSON, G. H. (1990), "The Inequality-Development Relationship in Developing Countries" (mimeo, Cornell University).

GALOR, O. and ZEIRA, J. (1993), "Income Distribution and Macroeconomics", *Review of Economic Studies*, **60**, 35–52.

KALDOR, N. (1956), "Alternative Theories of Distribution", *Review of Economic Studies*, **23**, 83–100.

KUZNETS, S. (1955), "Economic Growth and Income Inequality", *American Economic Review*, **45**, 1–28.

MELTZER, A. H. and RICHARD, S. F. (1981), "A Rational Theory of the Size of Government", *Journal of Political Economy*, **89**, 914–927.

PEROTTI, R. (1990), "Political Equilibrium, Income Distribution, and Growth" (mimeo, MIT).

PEROTTI, R. (1993), "Fiscal Policy, Income Distribution, and Growth" (mimeo, Columbia University).

PERSSON, T. and TABELLINI, G. (1991), "Is Inequality Harmful for Growth? Theory and Evidence" (mimeo, Berkeley).

RAM, R. (1988), "Economic Development and Income Inequality: Further Evidence on the U-curve Hypothesis", *World Development*, **16**, 1371–1375.

ROBERTS, K. W. S. (1977), "Voting Over Income Tax Schedules", *Journal of Public Economics*, **8**, 329–40.

ROMER, T. (1975), "Individual Welfare, Majority Voting, and the Properties of a Linear Income Tax", *Journal of Public Economics*, **14**, 163–185.

ROTHSTEIN, P. (1989), "Order Restricted Preferences and Majority Rule" (mimeo, Washington University in St. Louis)

SAINT-PAUL, G. and VERDIER, T. (1991), "Education, Democracy, and Growth" (mimeo, DELTA Institute).