“Education for the Masses?:
The Interaction between Wealth, Educational and Political Inequalities”

Francisco H.G. Ferreira¹

Keywords: Inequality, Education, Political Economy

JEL Classification Numbers: D31, D63

Abstract: This paper presents a simple model of distribution dynamics, in which the distributions of wealth, education and political power are circularly endogenous. Different levels of education translate into different income and wealth levels. Political power may (or may not) vary with wealth, and in turn affects decisions on the level of public expenditure on education. Since the market for education credit is imperfect, some people might need to rely on public schooling, the quality of which depends on those expenditure levels. As a result, educational opportunities differ along the wealth distribution. The dynamic system displays multiple equilibria, some of which are characterised by a vicious circle of interaction between educational, wealth and political inequalities. These particular equilibria, which are more unequal, are also shown to be inefficient in terms of aggregate output levels. Switching equilibria may be achieved through redistribution of political power.

¹ Department of Economics, Catholic University of Rio de Janeiro. I am grateful to Maurício Bugarin, Raquel Fernandez, Danny Quah, Martin Ravallion, an anonymous referee and seminar participants at the London School of Economics, the World Bank, the Catholic University of Rio de Janeiro, the Getúlio Vargas Foundation at Rio de Janeiro, and the University of Brasília for helpful comments. All remaining errors are mine.
1. Introduction.

The persistence of wealth or income inequality in the long-run is an empirical regularity worthy of mention alongside Nicholas Kaldor’s (1961) six stylised facts of economic growth. A considerable body of both theoretical and empirical literature has recently proposed and tested a variety of mechanisms through which inequality may not only persist among agents with identical preferences and initial skill endowments, but also lead to inefficient aggregate outcomes in steady-state.

At the risk of oversimplifying, these mechanisms can be grouped under two broad categories. The first relies on imperfect capital markets and non-convex production sets, which combine to prevent a group of agents (usually but not always the poor\(^2\)) from exercising their full economic potential.\(^3\) The second relies on political economy channels, through which greater inequality may either lead to more inefficient redistribution (Alesina and Rodrik, 1994; Persson and Tabelini, 1994) and thus lower growth rates, or to greater political instability and social conflict (e.g. Rodrik, 1997). More recently, the two strands have prom deservedly combined, as researchers seek to capture the interplay between the political determinants of (redistribution) policies, and the effect of capital market imperfections on the efficiency with which different agents use their productive resources.

In particular, Bénabou (2000) shows that risk-averse agents in a stochastic environment with incomplete insurance markets will derive positive utility from redistribution as an imperfect, publicly provided substitute for insurance. The combination of efficient (i.e. welfare-increasing) redistribution and unequally distributed political power generates the possibility of multiple political-economic equilibria, where the most equal is characterized by the highest degree of redistribution, and the most unequal has the least redistribution. This result is at odds with the predictions of some of the earlier political economy models of the dynamics of income distribution (e.g. Alesina and Rodrik, 1994 and Persson and

\(^2\) For an example where the richest are constrained by the absence of a full insurance market, see Banerjee and Newman (1991).

\(^3\) The list is now great, but seminal papers were written by Galor and Zeira (1993), Banerjee and Newman (1993), Aghion and Bolton (1997) and Piketty (1997).
Tabellini, 1994), and conforms better with the cross-country empirical evidence that inequality and redistribution (particularly when the latter is measured by public investment in education) are negatively correlated, as found by Perotti (1996) and others.

Bénabou (2000) also considers the financing of public education in an extension of the model. The market for education loans is missing, and tax revenues are redistributed as education subsidies (effectively through demand-side vouchers). The nature of the equilibria remains unchanged from his main model: although the high inequality equilibrium displays less redistribution, the two equilibria are not rankable by efficiency, either in a Pareto sense or in terms of aggregate output. This result depends on specific assumptions, such as the convexity of the education production set and the lognormality of the p.d.f.s of the random shocks.

This paper extends those results in three directions. First, it shows that the main result that redistribution can lead both to greater efficiency and to lower inequality need not rely on the presence of uncertainty. Instead, in this model redistribution leads to higher output levels by reducing the efficiency costs of indivisibilities in productive investments (as in Galor and Zeira, 1993). The model is thus also made considerably simpler. Second, abstracting from risk and uncertainty (and thus from specific functional forms for the distribution functions of shocks) allows us to derive stronger equity-efficiency relationships between the long-run equilibria of the system. In particular, a welfare comparison across different classes of steady-states, with different levels of inequality, allows us to rank them unambiguously in terms of aggregate output.

Finally, Bénabou (2000) does not address the possible implications of changes in the nature of the link between wealth and political power. By modelling the link between the distributions of wealth and power explicitly, we hope to shed light on some of the economic consequences of shifts in political regimes. These include the possible role of democratisation and empowerment in moving economies from an unequal and inefficient equilibrium to one that may dominate it both in output and Lorenz terms.
The model combines a voting mechanism, through which policies are determined endogenously, with an imperfect credit market, to shed light on a vicious circle of interaction between educational, wealth and political inequalities, which may lead to the existence of inefficient high-inequality equilibria. If educational opportunities differ for people along the wealth distribution, and the quality of the education available to the poor depends on an endogenously determined redistribution scheme (such as funding public schools from general proportional taxation), then a distribution of power that mirrors an unequal distribution of income may lead to persistent and inefficient levels of inequality.

The focus on the nature of redistribution through public education is shared with a number of recent papers, such as Fernandez and Rogerson (1995, 1998). Unlike these, however, I am less concerned with the degree of centralization of the public financing of education, and more with its quality relative to that of its private alternatives, and the effect this has on effort levels, the distribution of human capital, aggregate productivity, and the distributions of income and political power. Unlike Fernandez and Rogerson (1995), assumptions about the inferior quality of public education are made which are more compatible with primary and secondary schooling in a country like Brazil or the United States, rather than with higher education in the same countries. Although the basic problem is similar to that studied by Glomm and Ravikumar (1992), our results are quite different, largely because we abstract from human capital bequests (as opposed to physical ones). Finally, we also share the endogeneity of the political equilibrium with Bourguignon and Verdier (2000).

2. The Model

Let there be a continuum of agents with initial wealth levels $w$ distributed over $(0, z)$, according to $G(w)$. These agents consist of households which are ex-ante identical in every respect, except their initial wealth. Their size is normalised to one, and no intra-household issues are considered. Household labour supply is inelastic, and I assume that the nature of available projects is such that labour can not be pooled across households. Generations are

---

4 Like these papers - but unlike most of the work on inequality, redistribution and growth - we consider transfers in kind (rather than in cash), through the public provision of private goods. See Besley and Coate (1991) and Ferreira (1995).
successive and do not overlap. Their finite lives have two periods, according to the linear sequence in Figure 1 below.

[Figure 1 here.]

These agents maximise a utility function based on Andreoni’s (1989) “warm glow” bequest motive:

$$U(c, b) = c^a b^{1-a}$$

(1)

where, as denoted in figure 1, both consumption and bequeathing take place at life’s end, time $t_2$. Consumption and bequest are of a single consumption good, which is chosen as the numeraire good. This good is produced by skilled and unskilled workers, whose productivity levels and remuneration rates differ, but each unit of the good emerging from both processes is identical. It is costlessly storable and does not depreciate over time.

In the first period, they allocate their time fully between two activities: studying or working as an unskilled worker, whose deterministic productivity (and remuneration rate) is $u$. At time $t_0$ they choose $\sigma$ ($0 < \sigma < 1$): the fraction of time in the first period spent studying. Studying can take place through either one of two mutually exclusive education technologies. These are distinguished by different prices and productivity levels, but produce the same good: embodied human capital $S$, an excludable and non-transferable attribute of each individual student. The price of enrolling in the education production function 1, which we will associate with private schooling, is $p_1 = p^* > 0$. The price of enrolling in the education production function 2, which we will associate with public schooling, is $p_2 = 0$. The choice of school type, which is distinct from the time allocation decision, is denoted in Figure 1 as a choice of $p$. The choice is made under the knowledge that:

$$S = r \sigma^{1/2} \quad \text{if } p = p^* \quad \text{and}$$

(2a)

$$S = r \sigma^{1/2} \tau^{1/2} \quad \text{if } p = 0.$$  

(2b)
The education productivity parameter $q > 0$ converts time spent in school into actual human capital. I assume that $\bar{q}^2 - \hat{q}^2\tau > 4\pi^{-2}u\pi$ (Assumption 1), where $\pi$ is defined immediately below. $\tau$ is total public spending on education.

In the second period, agents dedicate their full time endowment to skilled work. Each agent’s productivity (and remuneration) is assumed to be an increasing linear function of their acquired level of human capital: $\pi S$. Since agents pay tax on their inheritance at time $t_0$, final income available at $t_2$ for consumption and bequests is given by:

$$ y(w,t, p) = (1 - t)w - p(w) + \left[1 - \sigma(w)\right]u + \pi S(t, p) $$

(3)

Credit markets are assumed to be completely absent in this economy, as a result of extreme problems in enforcing contracts.

The political system functions as follows. Lump sum taxes and transfers are administratively unfeasible and the constitution mandates that all taxation be in the form of proportional wealth taxation at the beginning of each generation’s life. All public expenditures must be directed towards financing public education, through $\tau$. Budget balance at every generation is also constitutionally mandated so that $\tau = t^* \int w dG(w)$.

Individually preferred tax rates are monotonically declining with wealth, so that the single tax rate $t^*$ is chosen by the critical voter:

$$ t^* = \text{Arg max}_{t} \left[(1 - t)w_c - p(w_c) + \left[1 - \sigma(w_c)\right]u + \pi S(t, p)\right] $$

(4)

where the wealth of critical voter is $w_c$ and the critical voter is such that:

$$ \int_{0}^{w_c} v(w) dG(w) = 0.5 $$

(5)

---

5 Human capital $S$ is the only form of capital in this economy.
6 This initial wealth taxation can be interpreted as an ex-post inheritance tax. Very little changes if the tax is collected on final incomes at $t_2$ instead, but it is then harder to reconcile positive taxation with the utility function in (1).
7 See Section 5 for a proof. This result allows us to rely on the version of the median voter theorem implicit in (4), as established by Roberts (1977). See also Gans and Smart (1996).
v(w) is the voting power function, which is assumed here to depend (weakly) positively on the individual’s initial wealth level. V(w) may take many forms, provided it satisfies:

\[ v(w) \geq 0, \quad v'(w) \geq 0, \quad \forall w, \quad \text{and} \quad \int_{0}^{z} v(w)dG(w) = 1. \]

Two examples of voting power functions which satisfy these two properties and have plausible empirical interpretations are:

(i)  \[ v(w) = 1 \rightarrow \text{“one person, one vote” or democracy} \]

(ii)  \[ v(w) = \frac{w}{\mu_w} \rightarrow \text{“money talks” or oligarchy.} \]

3. **The Static Equilibrium.**

There are three control variables in the model: c (or b), p and \( \sigma \). At time \( t_0 \) agents choose p and \( \sigma \). Through voting, the critical agent chooses \( t^* \), which can be taken as given by all other agents. The consumption/bequest allocation of final income is decided at time \( t_2 \) and is independent of the remaining choices.\(^8\)

**Lemma 1**: Given Assumption 1, \( p = 0 \) for any agent with wealth \( w < p^*(1-t^*)^{-1} \), and \( p = p^* \) for any agent with wealth \( w \geq p^*(1-t^*)^{-1} \).

**Proof**: See appendix.

This means that the quality (or productivity) differential between private and public schools is so large, that any agent capable of affording private education will choose to do so. Since credit markets have been assumed away, these are only those agents whose initial wealth level net of taxes at least equals the (exogenously given) private school fee.

Lemma 1 partitions each generation at the outset, into those which will attend private school (using education technology 2a) and those which will attend public school (using technology 2b). In conjunction with the fixed good-production productivity parameters \( u \)

---

\(^8\) Although (1) is written in terms of consumption and bequests, the Cobb-Douglas functional form implies that final income will be shared proportionately between the two uses, and every prior decision can be seen as taken to maximise end-of-period income.
and $\pi$, this knowledge of the education technology at their disposal allows each agent to determine an optimal first-period time allocation.

For agents with $w < p^*(1-t^*)^{-1}$, the problem is to $\max_{\sigma} \left\{ (1-t)w + [1-\sigma]u + \pi q^\sigma \sqrt{\tau} \right\}$, which is obtained by substituting (2b) into (3). The first order condition implies that:

$$\sigma_v^* = \left[ \frac{\pi q \sqrt{\tau}}{2u} \right]^2$$  \hspace{1cm} (6)

Similarly, for agents with $w \geq p^*(1-t^*)^{-1}$, the problem is to $\max_{\sigma} \left\{ (1-t)w - p^* + [1-\sigma]u + \pi q^\sigma \sqrt{\tau} \right\}$, which is obtained by substituting (2a) into (3). The first order condition implies that:

$$\sigma_{p^*}^* = \left[ \frac{\pi q}{2u} \right]^2$$  \hspace{1cm} (7)

Once $t^*$ is determined through the $t_v$ voting process, and agents have chosen $p$ and $\sigma$, we can complete a full characterisation of the static equilibrium by describing the final incomes accruing to each agent. Replacing the appropriate values $p(w)$ - from Lemma 1 - and $\sigma(w)$ - from (6) or (7) - into equation (3), the total income function is as follows:

$$y(w, t^*, p) = (1-t^*)w + \left[ 1 - \left( \frac{\pi q \sqrt{\tau}}{2u} \right)^2 \right] u + \left( \frac{\pi q \sqrt{\tau}}{2u} \right)^2$$  \hspace{1cm} if $w < p^*(1-t^*)^{-1}$,  \hspace{1cm} (8a)

$$y(w, t^*, p) = (1-t^*)w + \left[ 1 - \left( \frac{\pi q}{2u} \right)^2 \right] u + \left( \frac{\pi q}{2u} \right)^2 - p^*$$  \hspace{1cm} if $w \geq p^*(1-t^*)^{-1}$.  \hspace{1cm} (8b)
Considering \( t^* \) (and thus \( \tau \)) as given for each agent\(^9\), the second and third terms on the right hand side of (8a), as well as the second, third and fourth terms on the right hand side of (8b) consist solely of exogenous parameters. Let the sum of those terms in (8a) be denoted by \( A \) and the sum of those terms in (8b) be denoted by \( B \). These equations can then be rewritten in short form as:

\[
y(w, t^*, p) = (1 - t^*)w + A \quad \text{if } w < p^*(1-t^*)^{-1}, \\
y(w, t^*, p) = (1 - t^*)w + B \quad \text{if } w \geq p^*(1-t^*)^{-1},
\]

(9a) (9b)

Note that \( A < B \), from Lemma 1 (see proof in the appendix). This implies that incomes rise monotonically in initial wealth (with derivative less than one), but with a discontinuity at the wealth level which allows agents to attend the superior private schooling system.

Finally, it is worth noting the following comparative properties of the time allocation choices across the two different groups of agents, from equations (6) and (7). For both private and public school students, \( \frac{\partial \sigma^*}{\partial u} < 0 \) and \( \frac{\partial \sigma^*}{\partial t} > 0 \), as one would expect. The first inequality shows that effort spent in human capital accumulation declines with its opportunity cost, the remuneration for unskilled work. The second inequality shows that the effort (or time) allocation for human capital accumulation increases with its expected benefit, the rate of remuneration of skilled labour.

It is also the case that \( \frac{\partial \sigma^*}{\partial \tau} > 0 \), implying that public expenditures on schooling raise the time (or effort) spent by students in acquiring human capital. This reflects the assumption of complementarity between the two inputs, which is implicit in the education production function (2b). It implies that the impact of higher public education spending is therefore twofold: there is a direct productivity impact, which raises the output \( S \) for a given level of agent effort, through equation 2b. This impact is positive but concave, so that decreasing

---

\(^9\) In this and the next section, we solve the model conditional on the tax rate \( t^* \). We return to the endogeneity of \( t^* \) in Section 5.
returns to this kind of spending hold. There is also, however, an indirect impact through the
behavioural response of agents to higher school quality. Since a higher $\tau$ makes time spent
studying more productive, agents respond by allocating more such time, and giving up a
little more first period income.$^{10}$

Finally, for levels of $\tau$ consistent with Assumption 1 (which places an upper bound on $\tau$,
given the educational productivity parameters), $\sigma^*(p^*) > \sigma^*(0)$. This means that in
societies where the gap between the ‘quality’ of private education ($\tilde{q}$) and that of public
education systems ($\alpha \tau$) is sufficiently large with respect to the price of private education
($p^*$) – see Assumption 1 – one will observe the poorest people in society attending worse
schools and dedicating less time (or effort) to their studies. In this model, this arises from
optimising behaviour in the face of missing credit markets and exogenous ‘quality’
differences, despite identical preferences across agents. An observationally equivalent
outcome (of less dedicated students of public schools being the poorest) might be generated
from a very different model, in which there were no missing markets (i.e. opportunities did
not depend on initial wealth), but some people were lazier or less intelligent than others.
The implication is that before judging which of the two models is correct – whether the
poor study less as an optimal response to unequal opportunities or whether they are poor
because they are lazy people living in a fair system – one would have to test the two models
with respect to other predictions, or directly with respect to whether their assumptions hold
in practice.

4. Transitional Dynamics and the Multiple Steady-States

Maximisation of the Andreoni “warm glow” utility function in (1) implies that each agent
sets $c = \alpha y(w, t^*, p)$ and $b = (1 - \alpha) y(w, t^*, p)$ at time $t^*$, where the final income function
is given by equations (8) or (9). Since bequests constitute the only link between generations
in the model, the law of motion of the state variable wealth is fully characterised by
equation (10) below:

$^{10}$ The endogeneity of effort is included in the model for greater realism, and because it is of interest in itself.
It should be noted that although it reinforces the main results of the paper, it is not necessary for them.
Since all parameters in terms A and B are exogenous and constant, equations (10a,b) are a pair of linear first-order difference equations in \( w_t \), with different intercepts and domains that partition \( \mathbb{R}^+ \), to either side of \( w^* = p^*(1-t^*)^{-1} \). Since the system is entirely deterministic, the roots of these equations are attractors for the whole mass of any initial wealth distribution \( G_0(w) \) supported by \( \mathbb{R}^+ \). In other words, solving (10) for \( w_{t+1} = w_t \) yields the support of the limiting distribution of the system, which will perforce be degenerate. \(^{11}\) The wealth level which supports the limiting distribution will be a function of the exogenous parameters in (10). Because of the discontinuity at \( w = p^*(1-t^*)^{-1} \), this function will be: 

\[
\begin{align*}
  w_p &= \frac{1 - \alpha}{1 - (1 - \alpha)(1 - t^*)} \cdot A \quad \text{if the root of equation (10b) lies outside its domain;} \\
  w_R &= \frac{1 - \alpha}{1 - (1 - \alpha)(1 - t^*)} \cdot B \quad \text{if the root of equation (10a) lies outside its domain;} \\
       &\quad \text{or (} \begin{cases} w_p, & \text{if} \quad w_t < p^*(1-t^*)^{-1}, \\ w_R, & \text{if} \quad w_t \geq p^*(1-t^*)^{-1}, \end{cases} \end{align*}
\]

Since there are therefore three possible supports for the limiting distribution function \( G^*(w) \), depending on the values of the exogenous parameters of the model, there must accordingly be three different classes of steady-state distributions themselves. Formally, these limiting distributions, \( G^*(w) \), are fixed points in the space \( \mathcal{S} \) of univariate distribution functions \( F(w) \), defined over supports in the sigma-algebra of \( \mathbb{R}^+ \), or of its subset \( S = (0, w) \). In the case of this simple, deterministic system, these fixed points are degenerate distributions supported by the roots of (10) - \( w_p \) and \( w_R \), as defined above. This system is of some interest only because of the three possible cases which may arise, depending on the values of the exogenous parameters. Each case corresponds to one

\(^{11}\) If the dynamic process in (10) were stochastic, it would be a first-order Markov process. Its behaviour would be more complex and its limiting distribution would not, in general, be degenerate. There exist well-established theorems describing its convergence behaviour. See Stokey and Lucas (1989). Aghion and Bolton (1997) and Ferreira (1995) investigate the limiting distributions of wealth processes similar to (10), but stochastic in nature.
separate class of equilibrium, with different limiting distribution functions. Each class is described below:

Class I: If the indivisible cost of private education is sufficiently large, so that 

\[ p^* > \frac{(1-a)(1-t^* - t^*)}{1-(1-a)(1-t^*)} B, \]

then the root of (10b) lies outside its domain. The entire mass of any initial distribution \( G_0(w) \) converges to the root of (10a): 

\[ w_p = \frac{1-a}{1-(1-a)(1-t^*)} A. \]

The limiting distribution \( G^*(w) \) is a Dirac distribution at \( w_p \). This distribution is unique for specific exogenous parameter values satisfying the preceding inequality.

Class II: If \( p^* \) is sufficiently small, so that 

\[ p^* < \frac{(1-a)(1-t^* - t^*)}{1-(1-a)(1-t^*)} A, \]

then it is the root of (10a) which lies outside its domain. The entire mass of any initial distribution \( G_0(w) \) converges to the root of (10b): 

\[ w_R = \frac{1-a}{1-(1-a)(1-t^*)} B. \]

The limiting distribution is another Dirac distribution, at \( w_R \). Again, for specific values of the exogenous parameters satisfying the above inequality, this distribution is unique.

Class III: In intermediate cases, for values of \( p^* \) such that 

\[ \chi A \leq p^* \leq \chi B, \]

where 

\[ \chi = \frac{(1-a)(1-t^* - t^*)}{1-(1-a)(1-t^*)}, \]

the limiting distribution \( G^*(w) \) is a two point distribution \((n_p, 1-n_p)\) on the income levels \((w_p, w_R)\). In this case, the distribution depends on initial conditions \( G_0(w) \), since 

\[ n_p = G_0\left( p^*(1-t^*)^{-1} \right). \]

Existence of an equilibrium is guaranteed by the linearity of (10). The three classes just described exhaust the set of possible equilibria. And because Assumption 1 implies that \( A < B \), it is also sufficient to ensure that all three equilibrium classes are possible. Which one of them exists depends only on the relationship between the fixed price of private education, \( p^* \), and the other exogenous parameter values, in accordance with the inequalities described in the preceding paragraphs.
Figure 2 below depicts one example of the transition process (and its limiting distribution), for each of the three classes above. The three panels of Figure 2 plot the intergenerational transition function (10) for different values of the intercepts A and B, which reflect different underlying values for the unskilled and skilled productivity parameters (u and π), and/or for the educational production function parameters (q̄, ̄q). Since there is no stochastic term in the model, any initial distribution defined over (0, z) – or indeed ℜ+ - will converge to the attractor wealth levels where \( w_{t+1} = w_t \). Panel (a) of Figure 2 diagrammatically depicts a Class I equilibrium. This sees a disappearance of the privately educated class. Returns to schooling relative to its private market price are insufficient to sustain bequests capable of preventing this inexorable downward mobility. Equilibrium will be characterised by perfect wealth equality, and universal public education.

An example of a Class II steady-state is depicted in panel (b). It is also marked by complete wealth equality, but with education being provided solely by the private market. Returns to schooling, relative to its private market price, are high enough that everyone eventually acquires sufficient education and makes sufficiently large bequests so that no one is left with inheritances below the cut-off wealth value, \( p^*(1-t^*)^{-1} \).

The third type of equilibrium, corresponding to Class III, is depicted in panel (c), and it is the one on which I wish to focus. Here, there are two unequal wealth classes. Any lineage whose initial wealth level is below the critical value \( p^*(1-t^*)^{-1} \) will converge towards a ’poor’ attractor at \( w_p \). Those lineages fortunate enough to start off with levels of wealth above the threshold converge instead to a ‘rich’ attractor at \( w_R \). The ‘poor’ can not afford private schooling, even though this would make them more productive. The absence of credit markets prevent them from exploiting that possibility. All the ‘rich’ choose private schooling. Bequests are such that once this situation is reached, it is a stable equilibrium. Unless it is perturbed – say, by a change in the political equilibrium that determines public spending on education – such a society would remain thus economically and educationally divided forever.
We now turn to a fuller description of the political processes which underpin the model, and consider how they may account for switches across the three classes of equilibria described in this section.

5. Political Equilibria

The one variable as yet undetermined, and which has an effect on the incomes of publicly educated agents (through the intercept term $A$ in Figure 2), is the level of public spending $\tau$ or, equivalently, the tax rate $t^*$. Assumption 1, which determined the constellation of parameter values within which the model would be investigated, aimed to exclude high values of $t^*$ (relative to $p^*$), so that $y(w, t^*, p^*) > y(w, t^*, 0)$, $\forall w$. If it were relaxed, it would clearly be possible to set $\tau$ at an arbitrarily high level, so that $A = B$, and agents were indifferent between public or private education at every wealth level. Or indeed, to drive private education out of the model ex-ante (rather than in equilibrium, as in panel (a)), by having $A > B$.

Our concern, however, is not with governments that can set arbitrary values of $\tau$. Instead, as set out in Section 2, we have in mind a political economy equilibrium where a critical agent (in terms of voting power) takes a decision about $t$, based on her own selfish interests. If we now restrict our attention to the stable unequal equilibrium depicted in panel (c) of Figure 2, it is clear that if $w_c > p^*(1-t^*)^{-1}$, $t^* = 0$, since those agents do not benefit from public expenditure at all.

A non-zero value of $t^*$ will in general be obtained from

$$t^* = \arg\max_t [(1-t)w_c - p(w_c) + [1-\sigma(w_c)]\mu + \pi \delta(t, p)]$$

(4')

---

12 See the proof of Lemma 1 in Appendix 1.
for $w_c < p^*(1-t^*)^{-1}$. Setting $S(t, p) = \hat{q}\sigma^{\frac{1}{2}}\tau^{\frac{1}{2}}$, and noting that $\tau = t^* \int_0^w wdG(w)$ implies that $\tau = t^* \mu_w$, we find that $t^*(w_c) = \frac{\pi^2 \hat{q}^2 \sigma \mu_w}{4w_c^2}$. This preferred tax rate is duly monotonically declining in personal wealth for all $w_c < p^*(1-t^*)^{-1}$, as claimed earlier. It goes to zero for $w \geq p^*(1-t^*)^{-1}$. To eliminate the possibility that $t^* \geq 1$, a lower bound on the relative poverty of the pivotal voter is required: $\frac{w_c}{\mu_w^{\frac{1}{2}}} \geq \frac{\pi\hat{q}\sigma^2}{2}$ (Assumption 2). This can reasonably be interpreted as establishing an upper bound on inequality, beyond which the outcome of the political process is total expropriation. Below this level of inequality (proxied by the ratio of the wealth of the critical voter to the square root of mean wealth) the critical agent chooses an internal solution for the tax rate, since there are diminishing returns to public spending in the education production function, whereas taxes are linear.\(^{13}\)

Clearly, in steady-state, the wealth of the critical agent, $w_c$ is either $w_p$ or $w_R$. If it is $w_R (> p^*(1-t^*)^{-1})$, then $t^* = 0$. Values of $t^*$ such that $0 < t^* < 1$ are obtained when the critical voter is in the lower attractor, $w_p$. In that case, we can now incorporate the endogenous nature of $t^*$ into the system by noting that:

\[
\begin{align*}
W_p &= \frac{1-\alpha}{1-(1-\alpha)(1-t^*)} A \quad (11) \\
\tau^*(w_c) &= \frac{\pi^2 \hat{q}^2 \sigma \mu_w}{4w_c^2} \quad (12)
\end{align*}
\]

are a system of simultaneous equations in $t^*$ and $w_p$. The system solves for:

$$t^* = \frac{\pi^2 \hat{q}^2 \sigma \mu_w (1-\alpha)^{-1} - 1}{4A - 1}$$

and

\(^{13}\) I am grateful to an anonymous referee for pointing this out.
which are consistent with parameter values satisfying Assumptions 1 and 2.\textsuperscript{14}

Insofar as Assumptions 1 and 2 hold, all three classes of equilibria described in the previous section may exist, depending essentially on how large the cost of private schooling \( p^\ast \) is, in relation to the ratio of skilled to unskilled wage rates (see equation 8 for the definitions of terms A and B). Because of the simple, degenerate form of the limiting distributions of our system, it is immediate to establish the Lorenz ranking between them. Classes I and II equilibria have Lorenz curves along the diagonal, and both strictly dominate the Lorenz curves for Class III equilibria, which are characterized by inequality. Additionally, however, since education raises productivity in this model, it is possible to rank the three classes of steady-state by aggregate output, where this ranking is determined only by the mass of privately educated individuals in steady-state, \( 1 - G^\ast(p^\ast(1-t^\ast)^{-1}) \). Hence, Class II equilibria have higher aggregate output than Class III equilibria, and these in turn display higher aggregate output than Class I equilibria.

It follows that a society which found itself in an unequal, Class III steady-state such as that depicted in panel (c) of Figure 2, could simultaneously reduce its inequality and increase efficiency, as measured by aggregate output, if only it could move to a Class II equilibrium, such as that in panel (b).

But how could such a regime change be brought about, if a (c)-type equilibrium is stable? Such an economic and educational regime change may come about as a result of a change in the political power function \( v(w) \). Consider, for instance, a case in which equilibrium (c)

\[ W_p = \frac{(1-\alpha)(4A^2 - A)}{\pi^2 \dot{q}^2 \alpha \mu_w + 4A \alpha - 1} \]

\textsuperscript{14}We have thus shown that a steady-state in which the critical voter is in the lower attractor is fully consistent with a positive level of taxation derived from the endogenous political equilibrium of the model. This reconciles the solution for the steady-states of the model conditional on \( t^\ast \), as obtained in the previous section, with the endogenous nature of \( t^\ast \) political system of the model. We have not investigated, however, the path dependence of the transitional dynamics on tax-choices off steady-state. Although an analysis of the impact of initial distributions on final equilibria through the effect of fiscal choices along that transition would indeed make the description of the model richer and more complete, it lies beyond the objectives of the current study.
holds, generating a wealth Lorenz curve $L(w)$ of the general shape given by the kinked line in Figure 3 below:

![Figure 3 here]

Suppose that initially $v(w) = \frac{w}{\mu_w}$. In this case, the critical agent has rank $G(w_{co})$, where $o$ stands for “oligarchy”. Since preferences for $t^*$ decline monotonically with $w$, this outcome will yield a lower level of public expenditure $\tau_o$ than that which would arise if $v(w) = 1$. In that case, the critical agent is poorer, with wealth level $w_{cd}$ and rank $G(w_{cd})^{15}$. $t^*$ and $\tau$ rise as a result. Although the magnitude of the increase depends on the specific closed-form solution, it is clear that a constellation of parameter values $\{u, \pi, p^*, \bar{q}, \hat{q}\}$ exists such that an increase in $t^*$ of this nature will result in a change of regime from the inefficient, high-inequality Class III equilibrium to the more efficient and egalitarian Class II equilibrium.\(^{16}\)

6. Conclusions.

This paper presents a simple model of the joint determination of the distributions of education, wealth and power, in a setting where publicly provided education is of inferior quality than its privately provided substitute. Educational inequality may persist in steady state if: (i) a missing credit market prevents the poorest agents from attending the better schools, reducing their lifetime wealth; and (ii) the voting equilibrium fails to generate levels of public expenditure which are sufficient to increase the quality of the public schools attended by the poor.

If voting power is not distributed uniformly, but increases with private wealth, a self-sustaining high-inequality trap may arise, whereby educational inequality ensures the persistence of wealth inequality, which in turn ensures the persistence of political inequality, which in turn guarantees the continuation of educational inequality.

\(^{15}\) The subscript $d$ stands for “democracy”.


Such an equilibrium is inefficient in the sense that an alternative equilibrium with higher aggregate output (or total wealth) is attainable, through a temporary increase in taxes and public expenditures, so as to lift the poor out of the low education – low productivity trap. During the transition, this change is not Pareto improving, since higher taxes make the richer agents worse off. The new steady-state however, would have higher output, first-order stochastically dominate and Lorenz dominate the former. If tax-and-spend decisions are fully endogenous, the only way to escape the initial, inefficient equilibrium is through a political regime change, which transfers political power from richer to poorer agents. This is a case when democratisation (or empowerment of the disenfranchised) has economic returns along two economic dimensions – equity and efficiency.

This potentially important role for political power in determining the outcome of distribution dynamics suggests that there may be high returns to further research on the nature of the relationship between economic variables (such as wealth or income) and political influence. In this paper, we treated this relationship as a simple function v(w), the functional form and properties of which we considered exogenously given. It is clear, however, that this function subsumes and obfuscates various different modalities of political interaction. While we modelled the political process as consisting exclusively of a single voting process, real political decisions are affected by economic agents through other activities, such as lobbying, corruption, or demonstrating. The relationship between wealth and political power differs in each of these political activities, and the overall relationship will clearly depend on their relative importance. As the classical political economists of two centuries ago well knew, it seems that in order to properly understand the nature of the economic outcomes we observe, we may need to comprehend the manner in which political power is distributed and wielded.

---

16 That such a constellation exists can be seen intuitively from the fact that these parameter values can be chosen to make A and B arbitrarily close.

17 See Dixit, Grossman and Helpman (1997) and the references therein for examples of recent treatments of non-voting political behaviour and its economic consequences. Note, however, that in that paper, the authors are not concerned with the distributional implications of their analysis. I am grateful to Filipe Campante for many helpful conversations on this topic.
Appendix: Proof of Lemma 1.

1. \( p \in \{0, p^*\} \). An agent with wealth level \( w \) will choose \( p^* \) over 0 iff \( y(w, t^*, p^*) > y(w, t^*, 0) \). That is, iff:

\[
(1-t^*)w + \left[ 1 - \left( \frac{\pi q}{2u} \right)^2 \right] u + \frac{(\pi q)^2}{2u} - p^* > (1-t^*)w + \left[ 1 - \left( \frac{\pi \hat{q} \hat{\tau}}{2u} \right)^2 \right] u + \frac{(\pi \hat{q} \hat{\tau})^2}{2u} \]

\[
u \left[ \left( \frac{\pi \hat{q} \hat{\tau}}{2u} \right)^2 - \left( \frac{\pi \hat{q}}{2u} \right)^2 \right] + \frac{1}{2u} \left[ (\pi \hat{q})^2 - (\pi \hat{q} \hat{\tau})^2 \right] > p^* \]

\[
\frac{1}{4u} (\pi^2 \hat{q}^2 \tau - \pi^2 \bar{q}^2) + \frac{1}{2u} (\pi^2 \bar{q}^2 - \pi^2 \hat{q}^2 \tau) > p^* \]

\[
\frac{1}{2} \pi^2 (\bar{q}^2 - \hat{q}^2 \tau) > 2 up^* \]

\[
\bar{q}^2 - \hat{q}^2 \tau > 4u \pi^2 p^* \]

which is Assumption 1. This assumption thus sets an upper bound on \( \tau \), given the values of \( p^*, \bar{q}, \hat{q}, u \) and \( \pi \), such that Lemma 1 holds for any \( w \).

2. Under this assumption, \( p = p^* \) is preferred to \( p = 0 \), for any \( w \). But since \( p^* > 0 \) and there are no credit markets, \( p \) must be paid entirely out of initial after-tax wealth \((1-t^*)w\).

It follows that only agents with \( w \geq p^* (1-t^*)^{-1} \) can afford to effectively exercise their choice of \( p = p^* \). Agents with \( w < p^* (1-t^*)^{-1} \) are constrained to choose \( p = 0 \).

\[\blacksquare\]
References.


Figure 1:

<table>
<thead>
<tr>
<th>$t_0$ (study / work)</th>
<th>$t_1$ (work)</th>
<th>$t_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>receive $b_{t-1} = w_t$</td>
<td>finish school</td>
<td>receive $\pi S$</td>
</tr>
<tr>
<td>choose $p \in {0, p^*}$</td>
<td>receive $(1 - \sigma)u$</td>
<td>consume</td>
</tr>
<tr>
<td>choose $\sigma$</td>
<td>reproduce</td>
<td>bequeath</td>
</tr>
<tr>
<td>vote on $t$ and pay taxes</td>
<td></td>
<td>die</td>
</tr>
</tbody>
</table>
Figure 2:

(a) 

(b) 

(c)
Figure 3: