NBER WORKING PAPER SERIES

UNEQUAL SOCIETIES

Roland Bénabou

Working Paper 5583

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 1996

I am grateful for helpful comments to Olivier Blanchard, Jason Cummins, John Geanakoplos, Mark Gertler, Robert Hall, Ken Judd and Julio Rotemberg, as well as seminar participants at the NBER Summer Institute, Stanford University, Princeton University and MIT. Research support from the C.V. Starr Center is also gratefully acknowledged. This paper is part of NBER's research program in Economic Fluctuations and Growth. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

© 1996 by Roland Bénabou. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

NBER Working Paper 5583
May 1996

UNEQUAL SOCIETIES

ABSTRACT

This paper aims to explain the significant variations in the social contract observed across nations. It shows how countries with similar technologies and preferences, as well as equally democratic political systems, can sustain very different average and marginal tax rates. Similarly, it provides an explanation for the striking difference between the US and European systems of education finance or health insurance. The underlying mechanism operates as follows. With imperfect credit and insurance markets some redistributive policies can have a *positive* effect on aggregate output, growth, or more generally ex-ante welfare. Examples considered here include social insurance, progressive taxation combined with investment subsidies, and public education. Aggregate efficiency gains imply very different political economy consequences from those of standard models: the extent of political support for such redistributive policies *decreases* with the degree of inequality, at least over some range. This can generate a negative correlation between inequality and growth, as found in the data, without the usual feature that transfers increase with inequality, which is not supported empirically. Moreover, capital market imperfections make future earning a function of current resources. Combined with the politics of redistribution this creates the potential for *multiple steady-states*, with mutually reinforcing high inequality and low redistribution, or vice-versa. Temporary shocks to the distribution of income or the political system can then have permanent effects.

Roland Bénabou
Department of Economics
New York University
269 Mercer Street
New York, NY 10003
and NBER

Introduction

There is considerable variation in the nature of the social contract across countries. Some have low average and flat marginal tax rates, others a steeply progressive tax system. Many nations have made the financing of education and health insurance the responsibility of the state; some, notably the United States, have left them in large part to families, local communities and employers. The extent of implicit redistribution through labor market policies or the mix of public goods provided by the government also shows persistent differences.

This paper aims to provide an explanation of these variations which is not predicated on exogenous differences in tastes, technologies, or political systems. It also seeks to reconcile two empirical findings of the recent literature on political economy and growth. Several authors such as Persson and Tabellini (1994) or Alesina and Rodrik (1994) have documented a negative relationship between initial income or wealth inequality and subsequent aggregate growth. The explanation which they have put forward is that increased inequality translates into rising numbers of poor citizens, driving the income of the median voter further below the national mean. This leads to increased pressure for redistributive policies, which hurt aggregate performance by reducing incentives for the accumulation of physical or human capital. The data, however, do not appear to support this explanation: Perotti (1994), (1996) and most of the other studies reviewed in Bénabou (1996b) find no relationship between inequality and the share of transfers or government expenditures in GDP. Moreover, the effect of transfers on growth estimated in these studies is most often significantly positive.

The point of departure of this paper is a rather different role for the state. When capital and insurance markets are imperfect, a variety of policies which redistribute wealth from richer to poorer agents can have a *positive* net effect on aggregate output, growth, or more generally welfare. Examples considered here will include social insurance, progressive taxation combined with investment subsidies, state funding of public education, and residential integration. Net efficiency gains lead to very different political economy consequences from those of standard models: the extent of support for such redistributive policies *decreases* with the degree of inequality, at least over some range. Intuitively, efficient redistributions meet with a wide consensus in a fairly homogenous society but face strong opposition in an unequal one. Moreover, if agents engage in any type of investment activity the same capital market imperfections which permit these policies make future earnings a function of current resources. Combined with the politics of redistribution, this creates the potential for *multiple steady-states:* mutually reinforcing high inequality and low redistribution, or low inequality and high redistribution. Temporary shocks to the distribution of income or the political system can then have permanent effects.

I formalize these ideas in a stochastic growth model with incomplete asset markets and dynastic agents who vote over various redistributive policies. The first ingredient in the analysis is that these enhance aggregate efficiency. Specifically, I identify policies which increase the mean and/or reduce the variance of family income. The second ingredient is a simple variant of the median voter model, capturing the idea that some agents have more influence in the political process than others. For instance, poor and less educated people have a lower propensity to vote than rich ones. They are also likely to be at a disadvantage in terms of resources available for lobbying. It should be emphasized, however, that I shall not appeal to variations in political rights or participation to explain countries' different choices: this parameter is kept fixed across steady-states.

In addition to offering potential answers to the empirical puzzles discussed above, the model has several interesting features. The first is analytical tractability. Individual transitions are linear, reflecting the absence of non-convexities. The distribution of wealth remains log-normal, thus allowing for intuitive, closed form solutions. Second, whereas all the literature has focused on proportional taxation, I show how to incorporate progressivity. In particular, the model predicts that more progressive income taxes will be accompanied by a higher rate of consumption taxation and investment subsidization. Third, I formalize political influence in an intuitive and analytically convenient manner. These results and modelling devices could be useful in other settings.

This paper ties in to three strands of the recent literature on income distribution and growth. The first one emphasizes the political economy of redistribution (Bertola (1993), Perotti (1993), Saint-Paul and Verdier (1993), Alesina and Rodrik (1994), Persson and Tabellini (1994), Benhabib and Rustichini (1996)). The second one is concerned with the financing and accumulation of human capital (Loury (1981), Glomm and Ravikumar (1992), Galor and Zeira (1993), Bénabou (1993), Durlauf (1996a), Cooper (1992), Fernandez and Rogerson (1994)). The third one stresses the wealth and incentive constraints which bear on entrepreneurial investment (Banerjee and Newman (1991), (1993), Aghion and Bolton (1996), Piketty (1996)). Most directly related are the model in Bénabou (1996a), upon which I build, and the paper by Saint-Paul (1994), which identifies another politico-economic mechanism with properties similar to those obtained here. Saint-Paul points out that increases in inequality which disproportionately affect the lower tail of the income distribution can reduce aggregate income without much effect on the median; such is the case if some agents fall from the ranks of the middle class into poverty. The middle class, which remains politically decisive, will then reduce its transfers to the poor. Saint-Paul also shows that if exit from poverty requires some kind of investment, capital market imperfections may lead to multiple steady-states: a large underclass which persists due to low redistribution, or one which is kept small by significant transfers. Piketty (1995) provides another explanation for international differences in redistribution, similar to a collective form of the bandit problem. Because individual experimentation is costly, agents never fully learn the extent to which income is affected by effort rather than predetermined by social origins. The citizens of otherwise identical countries may then end up with different distributions of beliefs concerning social mobility, which translate into different perceived tradeoffs between the insurance and incentive effects of redistribution.

The paper is organized as follows. Section 1 explains the main ideas using the simplest possible setup, which treats the aggregate impact of redistribution as exogenous. Section 2 presents the actual model, which combines imperfect asset markets and progressive taxation. The case where there are no opportunities for accumulation is solved in Section 3. In this endowment economy the extent of redistribution chosen by voters is U-shaped with respect to initial inequality. In steady-state, tax progressivity increases with the variance of the endowment process but is U-shaped with respect to its persistence. Section 4 solves the full model with endogenous income dynamics. In particular, I show how multiple steady-states arise when the political weight of the rich is neither too large nor too small. The more redistributive equilibrium has greater social mobility and lower inequality than the one with the less progressive tax system. It also has higher output, although a variant of the model can yield the reverse configuration. I also explain how the results can help account for the negative correlation between inequality and growth found in the data, without any significant correlation between transfers and inequality. Section 5 recasts the model so as to explain differences in countries' systems of education finance (public versus decentralized), which represent the most striking example of divergence in the social contract. Section 6 discusses other applications such as residential integration and the mix of public goods, and Section 7 concludes. All proofs are gathered in the appendix.

1 Stylized Models and Basic Ideas

1.1 The Standard View

Let there be a continuum of agents $i \in [0,1]$, with log-normally distributed endowments: $x^i \equiv \ln y^i \sim \mathcal{N}(m, \Delta^2)$. The log-normal is a good approximation of empirical income distributions, and will lead to analytically tractable results. It allows for an unambiguous definition of inequality, as increases in Δ^2 shift the Lorentz curve outward. This variance also measures the distance between median and per capita income: $m = \ln y - \Delta^2/2$, where $y \equiv E[y^i]$. Suppose now that agents are faced with the choice between the following two stylized policies:

$(\mathcal{P})$ *laissez-faire:* each agent consumes his endowment, $c^i = y^i$, for all i.

$(\widehat{\mathcal{P}})$ *complete redistribution:* agents pool resources, and each consumes $c^i = E[y^i]$.

I restricting sharing to be either zero or one for simplicity. Nothing changes when agents can choose from a menu of tax rates $\tau \in [0,1]$. How many are in favor of the second policy? Clearly, all those with endowment below the mean , i.e. a proportion

$$p = \Phi\left(\frac{\Delta^2/2}{\Delta}\right) = \Phi\left(\frac{\Delta}{2}\right) \tag{1}$$

where $\Phi(\cdot)$ is the c.d.f. of a standard normal. Because the income distribution is left-skewed, so that the median is below the mean, $p > 1/2$. A strict majority rule would thus predict that redistribution should always take place. In reality the poor vote with lower probability than the rich, and to some extent money buys political influence. Therefore the relevant threshold for redistribution to occur may not be $p^* = 50\% = \Phi(0)$, but $p^* = \Phi(\lambda)$, $\lambda > 0$. For instance if the π poorest agents never vote, $\Phi(\lambda) = 1/2 + \pi$.

What is important and robust in (1) is therefore not the level effect, $p > 1/2$, but the comparative statics, $\partial p / \partial \Delta > 0$: in a more unequal society there is greater political support for redistribution. For *any* degree of bias λ in the voting system, positive or negative, the likelihood that redistribution takes place increases with inequality (or more specifically, skewness). This result is only reinforced under the standard assumption that redistribution entails some deadweight loss, reducing available resources from y to ye^{-B}, $B > 0$. Given the choice between laissez-faire and sharing this reduced pie, the extent of political support for redistribution is:

$$p = \Phi\left(\frac{-B + \Delta^2/2}{\Delta}\right) = \Phi\left(-\frac{B}{\Delta} + \frac{\Delta}{2}\right). \qquad (2)$$

Note that now $p \gtrless 1/2$ but $\partial p/\partial\Delta$ is even more positive than before. As inequality increases, so does the likelihood that a policy which reduces aggregate income gets implemented. This, in essence, is the mechanism by which inequality reduces growth in models such as Alesina and Rodrik (1994), Persson and Tabellini (1994) or some cases of Bertola (1993) and Perotti (1993).[1] The idea that distributional conflict hampers economic performance is supported by the evidence: several studies have confirmed Persson and Tabellini's and Alesina and Rodrik's findings of a negative effect of inequality on growth (e.g. Clarke (1992)). This correlation, however, does not seem to arise through increased redistribution.[2] Perotti (1992), (1994) (1996), Lindert (1996) and Keefer and Knack (1995) show that there is no relationship between the income share of the middle class (measuring the relative income of the median voter) and the main government transfers as a fraction of GDP. Clarke (1992) finds no correlation between various measures of inequality and government consumption. As casual empiricism suggests, more unequal countries do not redistribute more. Moreover, the coefficients on transfers in growth regressions are most often significantly *positive:* see among others Sala-i-Martin (1992), Devarajan, Swaroop and Zou (1993), Lindert (1996) and especially the three papers by Perotti, who explicitly controls for the endogeneity of redistribution.

1.2 Efficiency Gains and Redistribution: the Static Case

In a world of incomplete insurance and loan markets, a variety of policies with redistributive features can have a positive effect on aggregate welfare and even output.[3] The evidence on transfers and growth suggests that this theoretical possibility be taken seriously. Net welfare

[1] Naturally, this simple reduced form fails to capture the richness of the original models. Note that B could be a function of Δ, as long as $B(\Delta)/\Delta$ did not increase too fast.

[2] See Bénabou (1996b) for a more detailed survey of empirical studies on inequality, redistribution and growth. That paper also discusses the extent to which public transfers are appropriate measures of redistribution.

[3] For simplicity I focus on the choice between two policies, or more generally on cases where output is monotonic in the extent of redistribution. If it varies according to some Laffer-type curve the argument can be rephrased as saying that countries find themselves on the rising side of the curve rather than the declining one, as generally assumed (e.g., Alesina and Rodrik (1994)). Efficient redistributions (tax-financed public education) can also occur as the political outcome in Perotti (1993) and Saint-Paul and Verdier (1993).

gains, in turn, imply that political support for such policies varies with inequality in a radically different way from the traditional one. Indeed, suppose that by redistributing income agents achieve increased efficiency, so that each gets to consume ye^B. For now I continue to take $B > 0$ as exogenous, but later on I shall derive it from a variety of channels: insurance, altruism, or credit constraints on the accumulation of human and physical capital. The fraction of people who support an *efficient redistribution* is:

$$p = \Phi\left(\frac{B + \Delta^2/2}{\Delta}\right) = \Phi\left(\frac{B}{\Delta} + \frac{\Delta}{2}\right) \tag{3}$$

Naturally, it is always higher than under (1) and (2). But the important point is the one illustrated on **Figure 1.**

Proposition 1 *When a redistributive policy leads to efficiency gains, political support for it is U-shaped as a function of inequality.*

The intuition is simple: when dispersion is relatively small compared to the average gain, there is near unanimous support for the policy. As inequality rises, the proportion of those who stand to lose from the redistribution increases. At high enough levels of inequality, however, the standard skewness effect eventually dominates: there are so many poor that they impose redistribution, no matter what its aggregate impact.[4] There is thus *no monotonic relationship* between income inequality or the relative position of the median agent (both measured here by Δ^2) and the likelihood of redistribution.

To relate the extent of political support for redistribution to actual outcomes, one needs to specify the political mechanism through which preferences are aggregated. I shall continue to assume that redistribution occurs if p exceeds a threshold $p^* \equiv \Phi(\lambda)$, where $\lambda > 0$ reflects the degree to which financial or human wealth contributes to political influence.[5] **Figure 1** then

[4] The fact that inequality affects political support for redistribution in opposite ways, depending on the sign of its aggregate impact, is essentially *independent* of distributional assumptions. For any symmetric distribution $F(x)$ with mean m, a symmetric, mean preserving spread leads to a decline in $F(m+B)$ for all $B > 0$, and an increase for all $B < 0$. In the case of log-normally distributed wealth, a rise in Δ combines this general effect of dispersion with a fall in m due to increased skewness; hence the U-shape. These properties remain when B is a function of Δ, as long as it increases less than one for one with Δ; such will be the case when it is endogenized later on.

[5] This assumption is formalized and further discussed in Section 2, Proposition 4. More generally, the probability that policy $\widehat{\mathcal{P}}$ is adopted can be some increasing function of p.

shows that *only if* redistribution has a positive aggregate effect can it be abandoned as the result of increased inequality.

Let us now think about the dynamic implications of this result. A society which starts with enough wealth disparity to find itself in the area of **Figure 1** below p^* will not implement the kind of redistribution discussed earlier; as a result, high inequality will persist into the next period or generation. Conversely, low inequality creates wide political support for efficient policies which prevent disparities from growing. This feedback mechanism requires only that current resources affect future earnings, meaning that some form of investment is credit–constrained. Thus the *very same* market imperfections which permit efficient redistributions, hence a (partly) decreasing relationship between inequality and transfers, also provide the second ingredient required for multiple steady-states. I now turn to the core part of the paper, which will show that these simple intuitions carry over to fully specified intertemporal models of individual and collective decisions. In particular, the aggregate welfare gain will now be endogenous.

2 A Model with Incomplete Asset Markets and Progressive Taxes

2.1 The Economy

Consider an economy populated by infinitely-lived agents or dynasties $i \in [0,1] \equiv I$, with preferences given by:

$$U_0^i = E\left[\sum_{t=0}^{\infty} \rho^t \ln c_t^i\right] \tag{4}$$

These are maximized subject to the following constraints:

$$c_t^i + e_t^i = \hat{y}_t^i \tag{5}$$

$$y_t^i = k_t^i \tag{6}$$

$$k_{t+1}^i = \kappa \cdot \xi_{t+1}^i \cdot (k_t^i)^\alpha (\hat{e}_t^i)^\beta \tag{7}$$

where c_t^i, e_t^i and k_t^i denote agent i's consumption, savings and capital in period t, while $\hat{y}_t^i$ and $\hat{e}_t^i$ represent his *post tax* income and investment. The specifics of fiscal policy are described below. Capital (physical or human) depreciates geometrically at the rate $1-\alpha > \beta$, and investment is subject to i.i.d. productivity shocks ξ_t^i with $\ln \xi_t^i \sim \mathcal{N}(-s^2/2, s^2)$, hence $E[\xi_t^i] = 1$. There are no markets where individual risks could be diversified away, and no loan markets for intertemporal consumption smoothing. Naturally, all that matters is that there be some degree of incompleteness. *Pre-tax* income y_t^i is produced by the agent using his capital stock, with constant returns for simplicity. It is then subject to the following tax-and transfer scheme:

$$\hat{y}_t^i = (1-\theta_t)\,(y_t^i)^{1-\tau_t}\,(\tilde{y}_t)^{\tau_t} \tag{8}$$

where $\tilde{y}_t$ is defined by the balanced budget constraint

$$\int_0^1 (y_t^i)^{1-\tau_t}\,(\tilde{y}_t)^{\tau_t}\,di = y_t\,. \tag{9}$$

The elasticity τ_t of post-tax income measures the degree of *progressivity* (or regressivity) of the tax scheme.[6] As usual, confiscatory rates $\tau_t > 1$ must be excluded as not incentive-compatible: an agent with high wealth will simply throw some of it away if that can increase his welfare. Nothing in principle prevents a regressive tax, so I shall allow it; restricting τ_t to lie in $[0,1]$ would not change the nature of the results. The other parameter of fiscal policy is the net or *average tax rate* $\theta_t \in [0,1]$, that is, the share of total output taken by the government. These tax revenues are used to finance a proportional investment subsidy:

$$\hat{e}_t^i = (1+a_t)e_t^i \tag{10}$$

$$\theta_t\,y_t = \int_0^1 a_t\,e_t^i\,di \tag{11}$$

The presence of this subsidy is not arbitrary. While agents differ in their preferences over redistribution, all agree on the necessity to offset its negative effect on savings.

Proposition 2 *Given any anticipated sequence of progressive taxation rates* $\{\tau_t\}_{t=0}^{\infty}$,

1) Agents choose in every period a common savings rate $\nu_t = e_t^i/\hat{y}_t^i$, *with:*

$$\frac{\nu_t}{1-\nu_t} = \beta\sum_{k=1}^{\infty}\rho^k\,(1-\tau_{t+k})\prod_{j=1}^{k-1}(\alpha+\beta(1-\tau_{t+j}))$$

[6]Differences in agents' post tax incomes are less than proportional to those in pre-tax incomes, to an extent which increases with τ_t. The progressivity of the geometric scheme is also apparent in the fact that agents with the average level of income are made richer (provided of course θ_t is not too high): (9) implies $\tilde{y}_t > y_t$ for $\tau_t > 0$. Note finally that (8) can also be interpreted as *wage compression* through labor market institutions favorable to workers with relatively low skills.

2) Agents are unanimous (within and across generations) in wanting to set average tax rates and investment subsidies $\{\theta_t\}_{t=0}^{\infty}$ so as to maintain a constant investment rate:

$$\theta_t + (1-\theta_t)\,\nu_t = \frac{\rho\beta}{1-\rho\alpha} \equiv s$$

With logarithmic utility, proportional taxes have no effect on savings but progressive ones do. This intertemporal distortion can nonetheless be offset by taxing consumption to subsidize investment. The proposition shows that this in fact Pareto–optimal, so that any conceivable political system will produce the sequence $\{\theta_t\}_{t=0}^{\infty}$ which restores agents' savings rate to its socially optimal level s. The policy debate bears *only* on progressivity, that is, on $\{\tau_t\}_{t=0}^{\infty}$. The intuition for the unanimity result can also be seen from the fact that combining progressive income taxes with a consumption tax and investment subsidy is similar to taxing only consumption at a flat rate and rebating the proceeds progressively. Proposition 2 thus provides a simple incomplete–markets analogue to the standard public finance result that consumption taxes are Pareto–superior to income taxes under very general conditions (e.g., Atkinson and Stiglitz (1980), Judd (1985), Chamley (1985)).[7] Given Proposition 2, capital accumulation simplifies to:

$$\ln k_{t+1}^i = \ln \xi_{t+1}^i + (\alpha + \beta(1-\tau_t)) \ln k_t^i + \beta\tau_t \ln \tilde{y}_t + \beta \ln s + \ln \kappa, \tag{12}$$

so that wealth and income remain log-normally distributed over time. If $\ln k_t^i \sim \mathcal{N}(m_t,\, \Delta_t^2)$, then (9) leads to:

$$\ln \tilde{y}_t = \ln y_t + (1-\tau_t)\Delta_t^2/2 = m_t + (2-\tau_t)\Delta_t^2/2$$

and the economy's dynamics are governed by the simple difference equations:

[7]The next section uses a scheme of flat income taxes and progressive investment subsidies (say, to education) with the same properties. In either case, Proposition 2 shows that endogenizing *both* private decisions *and* the choice of policy instrument results in a constant savings rate. By contrast, distortions on the intratemporal margin become unavoidable when agents supply labor elastically. In Bénabou (1995) I extend the present model to allow for endogenous effort and show that its structure remains essentially unchanged. Moreover, for empirically plausible values of the supply elasticity –and in the case of education finance equalization even for infinitely elastic labor supply– there remains a significant interval where the efficiency costs of redistribution are dominated by its benefits: aggregate output increases with progressivity, and a fortiori so does ex–ante welfare. By abstracting from endogenous labor decisions, the present paper essentially focuses on the upward–sloping part of this "Laffer" curve.

$$m_{t+1} = (\alpha+\beta)\, m_t + \beta\, \tau_t\, (2-\tau_t)\, \Delta_t^2/2 + \beta \ln s + \ln \kappa - s^2/2 \tag{13}$$

$$\Delta_{t+1}^2 = (\alpha + \beta(1-\tau_t))^2\, \Delta_t^2 + s^2. \tag{14}$$

In particular, aggregate income growth is:

$$\ln(y_{t+1}/y_t) = \ln\kappa + \beta \ln s - (1-\alpha-\beta)\ln y_t - [\alpha + \beta(1-\tau_t)^2 - (\alpha+\beta(1-\tau_t))^2]\Delta_t^2/2 \tag{15}$$

where the last term measures the extent to which inequality slows down growth due to the incompleteness of the credit market.[8] Finally, consumption is proportional to disposable income, $\ln c_t^i = \ln(1-\nu_t)(1-\theta_t) + (1-\tau_t)\ln k_t^i + \tau_t \ln \tilde{y}_t$. Therefore, taking expected present values in (12)–(14) leads to:

Proposition 3 *The intertemporal welfare of agent i which results from a policy profile $\{\tau_t\}_{t=0}^{\infty}$ is*

$$U_0^i = \bar{u}_0 + \left(\sum_{t=0}^{\infty} \rho^t (1-\tau_t) \prod_{k=0}^{t-1} (\alpha + \beta(1-\tau_k))\right)(\ln k_0^i - m_0) + \frac{1-\rho\alpha}{1-\rho(\alpha+\beta)}\left(\sum_{t=0}^{\infty} \rho^t\, \tau_t\, (2-\tau_t)\Delta_t^2/2\right) \tag{16}$$

where $\bar{u}_0 = [m_0 + \rho(\beta \ln s + \ln\kappa - s^2/2)/(1-\rho)]/(1-\rho(\alpha+\beta)) + \ln(1-s)/(1-\rho)$.

In a representative agent economy ($\Delta_0^2 = s^2 = 0$) welfare reduces to $\bar{u}_0$, which reflects the value of the initial capital stock and the social returns to accumulation determining the *endogenous* investment rate s. More generally, $\bar{u}_0$ corresponds to ex–ante welfare in the absence of redistribution and therefore also includes the disutility from uninsured risks s^2. The second term in (16), which disappears through aggregation, makes clear the redistributive effects of tax policy. The third one captures the associated efficiency gains, which arise from both better insurance and the redistribution of resources from low to high marginal–product investments.

[8] See Bénabou (1996a), (1995) for a discussion of similar expressions, which reflect the concavity of the investment technology (7) and tax schedule (8). Note in particular that the bracketed term in (15) is minimized for $\tau = (1-\alpha-\beta)/(1-\beta)$. Thus for a given pre–tax distribution Δ_t^2, the potential gains from redistribution reflect the extent of decreasing returns to investment.

2.2 The Political System

I now specify the mechanism by which the sequence $\{\tau_t\}_{t=0}^{\infty}$ is chosen. Ideally, one would want to solve for the subgame-perfect equilibrium where each generation chooses the current rate (before shocks are fully realized), taking into account how this will affect the distribution of income and therefore also the political outcome in subsequent periods. This problem is notoriously difficult, so that except in special cases the standard practice is to assume that voters are "myopic", in the sense of failing to recognize their influence on future events or not caring about it due to an imperfect bequest motive. I shall assume here a somewhat different form of political myopia, which greatly simplifies the problem but leads to a weaker intertemporal consistency requirement than subgame perfection.[9]

I first consider an arbitrary generation, say generation zero, and allow it to choose among all *constant sequences* $\{\tau_t = \tau\}_{t=0}^{\infty}$ the one to which it would like to commit the economy. One can think of this as *specifying a constitution*, or making some other long-lasting societal choice from within a set of simple rules. In a second step I require consistency between the initial income distribution, which determines the choice of τ, and that which results from the policy, through the law of motion (14). This occurs when the economy is in steady-state, $\Delta_0^2 = \Delta_\infty^2$; every generation t, if given the choice among constant sequences $\{\tau_{t+k} = \tau\}_{k=0}^{\infty}$, will then opt for the same "social contract" as its predecessors.[10]

Within each generation, individual preferences are aggregated through a simple variant of the median voter model which captures the idea that some agents have more influence in the political process than others. For instance, it is a fact that poorer and less educated agents have

[9] In Bénabou (1996b) I show that the more standard assumption of myopic preferences (overlapping generations with a non-dynastic bequest motive) can yield results analogous to those derived in this paper. Among the problems which make a truly dynamic solution extremely difficult is the fact that once agents internalize the effect of taxation over future political outcomes there is no presumption that their preferences remain single-peaked. Notable exceptions to the common practice of making voters myopic are Saint-Paul (1994) and Grossman and Helpman (1996). An alternative approach is to look for a dynamic equilibrium numerically, as done by Krusell, Quadrini and Ríos-Rull (1994).

[10] In the literature on dynamic taxation it is common to look for the steady-state associated to a fixed tax rate, then examine agents' preferences over potential tax sequences for current and future periods (e.g., Chamley (1985), Judd (1985)). I impose here the additional consistency requirement that no sequence within the allowed set be preferable by the politically decisive coalition to remaining at the steady-state. The "self-restraint" implicit in the restriction to constant sequences was discussed by Chamley (1985) (although he did not solve for a consistent steady-state) and can be given choice-theoretic foundations based on the work of Cohen and Michel (1991) and Caillaud, Cohen, and Julien (1994).

a relatively low propensity to vote (Edsall (1984), Conway (1991)). It also seems plausible that they would be at a disadvantage in terms of resources available for lobbying. I do not seek to explain the source of such biases, only to model them in a convenient manner.[11] Let each agent's opinion be affected by a relative weight, or probability of voting, $\omega^i / \int_0^1 \omega^j \, dj$. With log-normality, the following schemes yields particularly simple results.

Proposition 4 *Suppose each agent has single peaked preferences $V(\tau, x^i)$ over some policy variable $\tau \in \Upsilon$, and that the preferred value τ^i decreases with (log) wealth x^i.*

1) If an agent's political weight depends on his rank in the wealth distribution, $\omega^i = \omega(r^i)$, the pivotal voter is the one with rank $r = \Phi(\lambda)$ and wealth $x = m + \lambda\Delta$, where $\lambda \lesseqgtr 0$ is defined by $\left(\int_0^{\Phi(\lambda)} \omega(r)dr\right) / \left(\int_0^1 \omega(r)dr\right) = 1/2$.

2) If an agent's weight depends on the absolute level of his wealth, with $\omega(x^i) = e^{\lambda x^i} = (y^i)^\lambda$ for some $\lambda \lesseqgtr 0$, the pivotal voter has rank $r = \Phi(\lambda\Delta)$ and wealth $x = m + \lambda\Delta^2$.

Ordinal schemes ensure that each person's weight and the identity of the pivotal voter remain invariant when the distribution of wealth shifts due to growth or when it becomes more unequal.[12] Previous discussions of political rights have generally assumed that influence depends on one's absolute level of wealth. Historical examples include voting franchises restricted to citizens owning a minimum amount of property (Saint-Paul and Verdier (1993), Persson and Tabellini (1994)) or membership in a ruling elite which requires some investment expenditures (Verdier and Ades (1993)). I find it somewhat more plausible that even such cutoff levels should be relative ones, keeping up with aggregate growth and reflecting the competitive nature of bids for political influence. I shall therefore focus on ordinal schemes, but the second part of Proposition 4 shows that absolute income effects are also easy to capture; the case $\lambda = 1$, for instance, corresponds to a "one dollar, one vote" rule. Moreover, this alternative formulation would only make the paper's results stronger.[13]

---

[11] Roemer (1995) shows how the presence of a second dimension in the political game (morals, religion) can result in a similar bias, by limiting the size of the coalition which can be mustered in favor of redistribution. It should be emphasized that I shall not appeal to differences in relative voting power or in political institutions to explain why the extent of redistribution varies across countries .

[12] Note also that for any ordinal scheme $\omega(\cdot)$, the associated λ is a sufficient statistic: it is as if the bottom $\Phi(\lambda) - 1/2$ voters (or the top $1/2 - \Phi(\lambda)$ when $\lambda < 0$) systematically abstained. One can also allow λ to be a random variable and formalize the idea that redistribution occurs with a probability increasing in p.

[13] This case is somewhat similar to Verdier and Ades (1993), in that a mean-preserving rise in inequality pushes

3 Redistribution in an Endowment Economy

3.1 Social Insurance

I first highlight the welfare gains from redistribution due to pure insurance, by focusing on the case where there are no opportunities for accumulation.[14] With $\beta = 0$ there is of course no purpose for savings nor investment subsidies: $e_t^i = a_t = \theta_t = 0$. Each agent's endowment stream, $x_t^i = \ln y_t^i$, follows an exogenous AR(1) process with serial correlation α and shocks $\ln \xi_t^i$; see (12). Given a constant τ, intertemporal welfare simplifies to:

$$U_0^i = \sum_{t=0}^{\infty} \rho^t \ln y_t + (1-\tau)\frac{x_0^i - m_0}{1-\rho\alpha} - (1-\tau)^2 \sum_{t=0}^{\infty} \rho^t \Delta_t^2/2 \tag{17}$$

The first term measures utility in a representative agent economy. The second term captures the persistence of initial relative positions. The third combines the demand for insurance against idiosyncratic shocks with the effect of initial inequality, which is explained below. From the law of motion (14) it is easy to compute

$$\sum_{t=0}^{\infty} \rho^t \Delta_t^2 = \frac{\Delta_0^2 + \rho s^2/(1-\rho)}{1-\rho\alpha^2}.$$

The utility function U_0^i is quadratic in τ. Given the constraint $\tau \leq 1$, agent i's preferred tax rate is given by:

$$\frac{1}{1-\tau_0^i} = \frac{1-\rho\alpha}{1-\rho\alpha^2} \cdot \frac{\Delta_0^2 + \rho s^2/(1-\rho)}{(x_0^i - m_0)_+} \tag{18}$$

where $(\cdot)_+ \equiv \max\{\cdot, 0\}$. Voters below the median desire the maximum feasible tax rate, $\tau_0^i = 1$. Voters above the median desire a tax rate $\tau_0^i < 1$ which decreases with their initial wealth and increases with the variance of uninsurable shocks s^2. It is U-shaped with respect to the persistence parameter α, which makes the initial draw $x_0^i - m_0$ more long-lived but also makes future income more risky. As to the effect of initial inequality, it can be given either an ex-ante or an ex-post interpretation. If initial wealth x_0^i is subject to a random draw, its variance Δ_0^2 is added to the present value of risks $\rho s^2/(1-\rho)$. Alternatively, Δ_0^2 reflects the skewness in the distribution of

a greater fraction of the population into a situation of political disenfranchisement (low ω^i, due to low wealth), thereby concentrating power on a smaller "elite".

[14] The critical assumption here is the absence of insurance markets. That of missing loan markets could in principle be relaxed, but at the cost of greater analytical complexity.

initial draws, which makes the median voter desire redistribution even in the absence of shocks.[15] Denoting by λ the degree of bias from Proposition 4, we have:

Proposition 5 *The progressivity of the tax system chosen at time zero is*

$$\frac{1}{1-\tau_0} = \frac{1-\rho\alpha}{1-\rho\alpha^2} \cdot \frac{\Delta_0^2 + \rho s^2/(1-\rho)}{\lambda_+ \ \Delta_0}$$

It is increasing in s^2, but U-shaped in Δ_0.

In what follows I shall focus on the case where λ is positive, ensuring that $\tau_0 < 1$. Not only is this the empirically relevant case, but a value $\lambda_+ > 0$ can actually be derived from the incentive compatibility constraint $\tau \leq 1$. Indeed, suppose there is on average neither pro-rich nor pro-poor bias in the political system: let λ be any random variable *symmetrically* distributed around zero. It is easy to see that Proposition 5 still holds, with τ_0 now denoting the *expected* tax rate, and $\lambda_+ \equiv E[\lambda \mid \lambda \geq 0] > 0$. The economic constraint $\tau \leq 1$ effectively limits the political power of the poor, or more precisely the extent into which it translates into greater redistribution.

For any initial condition Δ_0, τ_0 gives the rate of fiscal progressivity chosen by voters who can commit to a once and for all choice of τ, or at least believe that they can. Barring unexpected aggregate shocks, however, the economy is in steady-state, and inequality remains unchanged over time: $\Delta_0^2 = \Delta_\infty^2$. Therefore in every period, agents consistently choose the policy $\tau = \tau_\infty$ which is obtained by replacing Δ_0^2 by $s^2/(1-\alpha^2)$ in the expression for τ_0:

$$\frac{1}{1-\tau_\infty} = \frac{s}{\lambda_+} \cdot \frac{1-\rho\alpha}{(1-\rho)\sqrt{1-\alpha^2}}. \tag{19}$$

Proposition 6 *The steady-state rate of progressive taxation τ_∞ increases with the variance of individual shocks s^2, but is U-shaped with respect to their persistence α. It reaches a minimum at $\alpha = \rho$ and asymptotes to 1 as α tends to 1.*

These results indicate that the relationship between inequality and redistribution is not likely to be monotonic. What matters is not just the *amount* of inequality, but also *its source*. To the

[15] The combination of logarithmic utility and a log–normal distribution has the convenient property that expected utility equals the utility of the median agent. The risk premium is therefore equal to the difference between mean and median income. Note also from (17) that when voters' choice is restricted to $\tau \in \{0, 1\}$, the fraction supporting $\tau = 1$ is $p = \Phi[(\Delta_0^2/2 + B)/\Delta_0)(1-\rho\alpha)/(1-\rho\alpha^2)]$, with $B \equiv \rho s^2/2(1-\rho)$. This is similar to (3), up to a scaling factor.

extent that high income disparities in some countries reflect larger uninsurable shocks or more imperfect insurance markets, higher tax rates should be observed. But if higher inequality is due to more persistent shocks, transfers may well be lower: empirical values of serial correlation and the discount factor are consistent with $\alpha < \rho$, and in that range, $\partial\tau_\infty/\partial\alpha < 0$. Greater persistence reduces risk-sharing, even though it implies more volatile income. The reason is that it also increases the number of agents for whom the value of insurance is more than offset by their vested interest in the status quo.[16]

3.2 Concern for Equity

Apart from social insurance, one of the main reasons for income redistribution is simply that most people dislike living in a society which is too unequal. This may be due to pure altruism or to the fact that inequality generates social tensions, crime, and similar problems which have direct costs. To capture these ideas one can simply replace (4) with:

$$U_0^i = E\left[\sum_{t=0}^{\infty}\rho^t\left(\ln c_t^i - (\mathcal{A}/2)\,var_t(\ln c)\right)\right] \tag{4'}$$

The coefficient $\mathcal{A}$ captures everyone's aversion to disparities in consumption or felicity, measured by the cross-sectional variance $var_t(\ln c) \equiv var[\ln c_t^j,\ j \in I]$. In this model, inequality-aversion and risk-aversion are clearly equivalent. Therefore all the results in the previous section remain unchanged, except that the present value $\sum_{t=0}^{\infty}\rho^t\,\Delta_t^2 = (\Delta_0^2 + \rho s^2/(1-\rho))/(1-\rho\alpha^2)$ is now multiplied by $(1+\mathcal{A})$ everywhere.[17] The same scaling factor applies to the expression (19) for the steady-state tax rate τ_∞. Progressivity increases with the degree of inequality-aversion $\mathcal{A}$ and decreases with political the power of wealth, λ. If one observed two countries, the first with low pre-tax inequality yet extensive redistribution, the other with high inequality yet limited redistribution, one would indeed be tempted to conclude that the citizens of the first country are more altruistic, or their poor more powerful. In fact it could be that preferences are identical and

[16] These results are independent of the level of per capita income: since $E[\ln\xi_t^i]$ has no effect on τ_0 or τ_∞, one can always keep steady-state income $\ln y_\infty = (\ln\kappa + E[\ln\xi])/(1-\alpha) + (s^2/2)(1-\alpha^2)$ constant while varying s^2 or α.

[17] Nothing would change if we assumed that people dislike inequities in intertemporal utility, $var[U_t^j,\ j \in I]$. Note also that (4') combines both risk- and inequality-aversion. To focus only on the latter one can reinterpret x_t^i as the level of income and let preferences be linear in own consumption: $U_0^i = E[\sum_{t=0}^{\infty}\rho^t\,(c_t^i - (\mathcal{A}/2)\,var_t(c))]$. The results are the same except that $(1+\mathcal{A})$ is replaced by $\mathcal{A}$.

political systems equally democratic, but that the second country's more unequal distribution reflects a more a more persistent income process (a higher α). The next section will show that this degree of social mobility can even be endogenized.

4 Redistribution and Accumulation

4.1 Wealth Dynamics and Social Contracts

I now turn to the paper's central theme, sketched in the discussion of dynamics at the end of Section 1: if greater inequality leads to less redistribution and pre-tax income depends on past transfers due to imperfections in loan markets, multiple equilibria can arise. To demonstrate this point I solve the full model with capital accumulation subject to wealth constraints ($\beta > 0$). With constant τ, Proposition 2 yields $\nu = \rho\beta(1-\tau)/(1-\rho\alpha)$ and $\theta = (s-\nu)/(1-\nu)$. The capital accumulation equation (12) becomes

$$\ln k_{t+1}^i = \ln \xi_{t+1}^i + (\alpha + \beta(1-\tau)) \ln k_t^i + \beta(m_t + \tau(2-\tau)\Delta_t^2) + \beta \ln s + \ln \kappa,$$

which makes clear that choosing a value of τ is equivalent to *choosing the degree of persistence* of the wealth process, $\alpha + \beta(1-\tau)$. The expression for intertemporal welfare, (16), now simplifies considerably:

$$U_0^i = \bar{u}_0 + \frac{(1-\tau)(\ln k_0^i - m_0)}{1-\rho(\alpha+\beta(1-\tau))} + \left(\frac{1-\rho\alpha}{1-\rho(\alpha+\beta)}\right)\left(\frac{\tau(2-\tau)}{2}\right)\left(\frac{\Delta_0^2 + \rho s^2/(1-\rho)}{1-\rho(\alpha+\beta(1-\tau))^2}\right). \quad (20)$$

The representative agent term $\bar{u}_0$ was explained earlier. In the second term, the numerator reflects the tax system's redistribution of consumption and the denominator its effect on social mobility. The third term captures the gains in aggregate or ex-ante welfare from income redistribution. These arise because missing markets prevent individuals who are risk-averse and face decreasing returns to investment from equating marginal rates of substitution across dates and states. When $\alpha = 0$ this makes complete pooling optimal: the third term, proportional to $\tau(2-\tau)\sum_{t=0}^{\infty} \rho^t \Delta_t^2$, is maximized at $\tau = 1$. When $\alpha > 0$, however, the differential productivity of investment expenditures for agents with different capital stocks leads to an interior maximum.[18]

[18]Note that even if the rich are powerful enough to impose regressive taxation, a rate such that $\rho(\alpha+\beta(1-\tau))^2 \geq 1$

To derive the tax scheme chosen in equilibrium one needs to examine whether U_0^i is single–peaked in τ. This is far from obvious given that it sums up a term whose variation in τ depends on the sign of $\ln k_0^i - m_0$ with one where an increasing function of τ is multiplied by a decreasing one. To keep the algebra tractable I shall make one of two simplifications. If voters choose from a menu of only two policies, $\tau \in \{\underline{\tau}, \bar{\tau}\}$, quasiconcavity holds trivially; this is the route taken in the next section, where I examine whether education and similar goods will be provided privately or publicly. Alternatively, I shall assume here that capital fully depreciates (or nearly so) and set $\alpha = 0$. This leads to:

$$\frac{\partial U_0^i}{\partial \tau} = -\frac{\ln k_0^i - m_0}{(1-\rho\beta(1-\tau))^2} + (1-\tau)\left(\frac{1-\rho\beta^2}{1-\rho\beta}\right)\left(\frac{\Delta_0^2 + \rho s^2/(1-\rho)}{(1-\rho\beta^2(1-\tau)^2)^2}\right). \tag{21}$$

While this is still a complicated function of τ, note how it embodies the same intuition as the stylized model of Section 1: for any $\tau \leq 1$, the proportion of agents who would like further redistribution ($\partial U_0^i/\partial\tau > 0$) is U-shaped in Δ_0. Moreover, one can show the following results.

Proposition 7 *Let $\alpha = 0$. Each agent's expected utility U_0^i is strictly quasi-concave in the tax rate τ, and maximized at $\tau_0^i = \max\{\tau \leq 1 \,|\partial U_0^i/\partial\tau \geq 0\}$. This τ_0^i is increasing in s^2 and Δ_0^2.*

Recall that rates greater than 1 are excluded by incentive compatibility. Proposition 7 allows us to characterize the outcome of the political process by simply replacing $\ln k_0^i - m_0$ with $\lambda\Delta_0$ in the first-order condition $\partial U_0^i/\partial\tau = 0$. This implicit equation leads to:

Proposition 8 *The rate of progressive taxation τ_0 chosen at time zero is increasing in s^2 but U-shaped in Δ_0^2. It declines from 1 at $\Delta_0^2 = 0$ to a minimum at $\Delta_0^2 = \rho s^2/(1-\rho)$, then rises back towards one as Δ_0^2 tends to infinity.*

These results provide an *endogenously derived* analogue to Figure 1, with the proportion of people supporting redistribution in a zero-one decision now replaced by the continuous policy outcome τ. Note that the non-monotonicity of the relationship between inequality and redistribution applies equally in an endowment economy ($\beta = 0$) and in the presence of accumulation. In the first case the underlying efficiency gains arise purely from insurance. In the second they also

will never be chosen: it would make the discounted sum of variances in future income Δ_t^2 diverge, yielding $U_0^i = -\infty$ for all i.

reflect the reallocation of investment funds to agents whose marginal product is higher, due to tighter liquidity constraints.[19] But the really critical difference is that the distribution of wealth is now itself a function of the tax rate, as shown by (14). Thus in the long run,

$$\Delta_0^2 = \Delta_\infty^2 = \frac{s^2}{1-\beta^2(1-\tau)^2}. \tag{22}$$

A steady-state equilibrium is an intersection of this downward-sloping locus, $\Delta = \Delta_\infty(\tau)$, with the U-shaped curve $\tau = \tau_0(\Delta)$ described in Proposition 8. This corresponds to a rate of tax progressivity solving the equation:

$$f(\tau) \equiv \left(\frac{1-\tau}{\sqrt{1-\beta^2(1-\tau)^2}}\right)\left(\frac{(1-\rho\beta(1-\tau))^2}{1-\rho\beta^2(1-\tau)^2}\right)\left(\frac{1-\rho\beta^2}{1-\rho\beta}\right)\left(\frac{s}{1-\rho}\right) = \lambda \tag{23}$$

It can be shown that there always exists at least one stable steady state.[20] Moreover,

Theorem 1 *For ρ large enough, there exists $\underline{\lambda}$ and $\bar{\lambda}$ with $0 < \underline{\lambda} < \bar{\lambda}$, such that:*

1) For any λ in $[\underline{\lambda}, \bar{\lambda}]$ there are exactly two stable steady-states. The steady-state with the more progressive tax system has greater mobility, lower inequality, and higher total output, .

2) For $0 \leq \lambda < \underline{\lambda}$ or $\lambda > \bar{\lambda}$ there is a unique, stable steady-state.

Where multiple steady-states occur, history matters: temporary shocks to the distribution of wealth or the political system (slavery, immigration, shifts in demand or technology) can permanently move the economy from one equilibrium to the other. This contrasts with traditional politico-economic models, where countries can deviate at most temporarily from a common steady-state level of inequality and redistribution (keeping technological, political and preference parameters constant). In particular, fiscal policy operates there as a stabilizing force on the distribution of wealth: more inequality today means more redistribution, hence less inequality

[19] Perotti (1994) provides direct evidence that credit market imperfections reduce investment, especially where the income share of the bottom 40% is low. More indirect evidence on wealth constraints and decreasing returns in education and farming is discussed in Bénabou (1996b). While potentially important, this source of output gains is not essential to the main result of multiplicity; see the discussion in Section 4.2 below.

[20] Note also from (21) and (22) that neither $\Delta = 0$ nor $\Delta = +\infty$ can correspond to a steady-state. In other words, the feedback mechanism from τ onto itself leads neither to complete equality nor explosive inequality, but to (an) interior, stable steady-state(s). The simple formula giving long-run output as a function of τ is provided in the appendix, as part of the proof of Theorem 1. It can also be read off equation (15).

tomorrow.[21] Multiple steady-states due to a negative impact of inequality on redistribution is the distinguishing feature which the present model shares with Saint-Paul (1994). Consistent with my general argument that this entails redistributions which increase the size of the pie, Saint Paul's model also has the property that transfers raise aggregate income.

4.2 Implications for Output and Growth

While the two steady–states are clearly not Pareto–rankable, unconditional welfare is higher in the more redistributive one because of both better insurance and higher mean income. Recall that greater progressivity τ implies a lower private savings rate $\nu = \rho\beta(1-\tau)/(1-\rho\alpha)$, but also a higher rate of consumption taxation and investment subsidization, $\theta = (s-\nu)/(1-\nu)$. The fact that both equilibria have (endogenously) the same investment rate makes clear that the higher output under the more redistributive social contract arises from a *more efficient allocation* of investment expenditures. Proposition 1 can therefore easily be translated into a ranking of *long-term growth rates*, by preserving economy–wide constant returns through spillovers or public goods complementing private investment.[22] This positive correlation between redistribution and growth is particularly interesting in light of the empirical fact mentioned earlier: regression estimates of the effects of social and educational transfers on growth are most often positive and significant.

Since that evidence is likely to be assessed very differently depending on readers' priors, it is worth pointing out that the paper's main result is also consistent with the reverse, more conventional view. European countries, it is often argued, have chosen a higher degree of social insurance and compression of inequalities at the cost of higher unemployment and slower growth

[21] In Persson and Tabellini (1994) the degree of inequality, hence also the policies implemented, depends only on the underlying distribution of talent. In Alesina and Rodrik (1994) and Bertola (1993) the deterministic nature of the models allows any distribution of initial endowments to persist indefinitely. Incorporating uninsurable idiosyncratic shocks would normally lead to a unique steady-state distribution. Uniqueness also obtains in Perotti (1993) and Saint-Paul and Verdier (1993). In Saint-Paul and Verdier (1992) greater inequality again results in higher taxes and spending on public education, but this stabilizing effect is more than compensated by a divergence in the incentives of the rich and the poor to invest privately in additional human capital. Hence two possible steady-states: high (low) inequality and public education expenditures, with private accumulation by the rich only (by both classes). This generates the same correlations as movement along the convergence path in Saint-Paul and Verdier (1993): greater inequality is associated with increased redistribution and, up to a point, with higher growth. Both correlations are somewhat problematic in view of the empirical evidence.

[22] For instance, let (7) be simply replaced by $k_{t+1}^i = \kappa \cdot \xi_{t+1}^i \cdot (k_t^i)^\alpha \, (\hat{e}_t^i)^\beta (k_t)^{1-\alpha-\beta}$, where k_t is the total capital stock. Because it aggregates individual contributions linearly this spillover is heterogeneity–neutral, i.e. does not introduce any new effects of income distribution on growth. It just makes permanent those due to imperfect credit markets, by eliminating the "convergence" term $-(1-\alpha-\beta)\ln y_t$ from (15).

(e.g., Freeman (1995)). Whether this is viewed as enlightened policy or dismal "Eurosclerosis", it begs the question of why countries with similar economic and political fundamentals would choose such different points on the equity-efficiency tradeoff. To see how a simple variant of the model can generate this scenario, observe that gains in per capita income are not essential for multiplicity and path-dependence. What must be true, as explained in Section 1, is that the more redistributive steady-state has higher average welfare. This is consistent with lower output and even growth, arising for instance from reductions in labor supply. As long as the resulting welfare loss remains dominated by the gains from insurance or inequality-reduction (which is always true over some positive interval), the two steady-states can persist.

There are of course redistributive policies which are detrimental to both accumulation and welfare ($B < 0$ in the notation of **Figure 1).** Putting them together with the beneficial ones emphasized here can help explain the simultaneous findings of a negative correlation between inequality and subsequent growth, and a zero correlation between inequality and subsequent transfers. In the standard case, (exogenous) increases in inequality lead to redistributions which discourage investment. In the present case (comparing the two steady-states), increased inequality is associated with less redistribution, but also with slower growth.

5 The Financing of Education

Education finance provides perhaps the most compelling case of a redistributive policy with positive efficiency implications. Loan market imperfections are more likely to affect investment in human that in physical capital, which can serve as collateral. The same is true for decreasing returns. The financing of primary and secondary schools also constitutes a striking example of persistent international differences in the extent of redistribution. Japan and most European countries have state-funded public education, which essentially equalizes expenditures across pupils. The United States, in contrast, relies in large part on local financing; because communities are heavily income-segregated, expenditures reflect parental resources to a large extent, making education a quasi-private good

Bénabou (1996a) demonstrates how a move from community to state funding of education could raise the economy's long-run output level, and even its long-term growth rate. Fernandez and Rogerson (1994) provide a quantitative estimate of the potential gains: calibrating a model with local funding to US data, they find that a move to state finance could raise steady-state GNP by about 3%. Whether or not one subscribes to this view, the persistent international differences in educations systems need to be explained, unless one is willing to appeal to intrinsic differences in tastes, technologies or political rights.[23]

Our model provides a potential answer, in the form of multiple equilibria and history-dependence with respect to the financing of education, or any other type of liquidity-constrained investment. Let k_t^i now represent human wealth. The term $(k_t^i)^\alpha$ in (7) captures the transmission of human capital within the family, which empirical studies have shown to be important. The shocks ξ_t^i represent unpredictable innate ability, as in Loury (1981). Whereas the previous section cast policy in terms of progressive taxes and a flat investment subsidy, let us now consider a *flat tax* rate and *progressive subsidies* to investment . Replace (8), (10) and (11) by:

$$\hat{y}_t^i = (1-\theta_t)\, y_t^i \qquad (8')$$

$$\hat{e}_t^i = e_t^i\,(1+a_t)\,(\tilde{y}_t/y_t^i)^{\tau_t} \qquad (10')$$

$$\theta_t\, y_t = \int_0^1 (\hat{e}_t^i - e_t^i)\; di \qquad (11')$$

where $\tilde{y}_t$ is still determined by (9) and a_t by the government's budget constraint (11'). Note that τ_t does not directly affect consumption any more. This elasticity of the subsidy with respect to decreases in wealth measures the extent to which investment is *publicly and equally* provided, rather than *privately*. Compare for instance the cases $\tau_t \equiv 0$ and $\tau_t \equiv 1$, keeping in mind that agents choose a common savings rate $\nu_t = e_t^i/\hat{y}_t^i$.

This new policy is almost equivalent to the previous one. Very similar derivations show that consumers and voters choose $\{\nu_t^i = \nu_t\}_{t=0}^{\infty}$ and $\{\theta_t\}_{t=0}^{\infty}$ so as to maintain the same, time-invariant investment rate s. Moreover, the dynamics of wealth are identical to (12), and with constant τ

[23]In Glomm and Ravikumar (1992), private finance of education leads to a higher long-run growth rate than public funding, because it gives individuals better incentives to accumulate human wealth. Bénabou (1996a) shows that taking into account the randomness in children's ability tends to reverse this ranking, as does the presence of economy-wide spillovers. Gradstein and Justman (1993) derive similar results in a model with labor supply, then examine voters' choice among different funding regimes. They obtain a unique equilibrium. Saint-Paul and Verdier's (1992) model yields multiplicity with respect to the level of public education funding, as explained in the previous footnote.

the expression for intertemporal utility differs only slightly from (20):

$$U_0^i = \bar{u}_0 + \frac{\ln k_0^i - m_0}{1-\rho(\alpha+\beta(1-\tau))} + \frac{\rho\beta\tau(2-\tau)/2}{1-\rho(\alpha+\beta)}\left(\frac{\Delta_0^2 + \rho s^2/(1-\rho)}{1-\rho(\alpha+\beta(1-\tau))^2}\right) \tag{20'}$$

Theorem 1 applies unchanged, but I shall instead focus on another simple case which allows $\alpha > 0$. I restrict policy to two options. Under laissez-faire or *decentralized funding,* $\tau = 0$, education expenditures are determined by family or community resources; the two are essentially equivalent when communities are stratified by income. Equal, *public funding* of education corresponds to $\tau = 1$, or more generally to $\tau = \bar{\tau}$, where $0 < \bar{\tau} \leq 1$. Given initial inequality Δ_0^2 it is adopted if $p = \Pr[i \,|\, U_0^i(\bar{\tau}) > U_0^i(0)] > p^* = \Phi(\lambda)$, or setting $\ln k_0^i - m_0 = \lambda\Delta$,

$$\lambda < \left(\frac{2-\bar{\tau}}{2}\right)\left(\frac{1-\rho(\alpha+\beta(1-\bar{\tau}))}{1-\rho(\alpha+\beta(1-\bar{\tau}))^2}\right)\left(\frac{\Delta_0^2+\rho s^2/(1-\rho)}{\Delta_0}\right). \tag{24}$$

Note the usual U-shape in Δ_0 of the right-hand side. Substituting in the asymptotic variance under public provision (partial or complete), namely $\hat{\Delta}_\infty^2 = s^2/(1-(\alpha+\beta(1-\bar{\tau}))^2)$, this regime is a steady-state if

$$\frac{\lambda}{\Omega s} < \left(\frac{1-\rho(\alpha+\beta(1-\bar{\tau}))^2}{\sqrt{1-(\alpha+\beta(1-\bar{\tau}))^2}}\right), \tag{25}$$

where $\Omega(1-\rho)$ denotes the product of the first two terms in (24). Conversely, laissez-faire or community provision of education is a steady-state if:

$$\frac{\lambda}{\Omega s} > \left(\frac{1-\rho(\alpha+\beta)^2}{\sqrt{1-(\alpha+\beta)^2}}\right). \tag{26}$$

Comparing the right-hand sides of (26) and (25) leads to the following result.

Theorem 2 *If the discount factor ρ is large enough, there exists $\underline{\lambda}$ and $\bar{\lambda}$ with $0 < \underline{\lambda} < \bar{\lambda}$ and:*

1) For any λ in $[\underline{\lambda}, \bar{\lambda}]$ both public and decentralized funding of education are stable steady-states. The steady-state with public finance ($\tau = \bar{\tau}$) has greater social mobility, lower inequality and higher total output than the one with local or private finance ($\tau = 0$).

2) For $\lambda < \underline{\lambda}$, public funding is the only steady-state. For $\lambda > \bar{\lambda}$, decentralized funding is the only steady-state.[24]

[24]Theorem 2 characterizes the set of steady-states when ρ is above some level $\bar{\rho}$; the expressions for $\bar{\rho}$ and for

These results can easily be extended to allow for parental decisions over how to allocate their time between work and child-rearing, and for a production structure where families are linked through complementarities between different types of workers or other economy-wide spillovers. I have intentionally abstracted from any form of externality throughout the paper, but Proposition 2 carries over to the model in Bénabou (1996a), which incorporates these additional features.

6 Other Applications

6.1 The Mix of Public Goods

Some public services such as the legal system, protection of property, prisons, etc., benefit citizens differentially according to their levels of wealth or investment. Others, such as public infrastructure or education, have more uniformly or even progressively distributed benefits. Deininger and Squire (1995) find in cross-country regressions that public investment affects the growth of income equally for all quintiles, while public schooling expenditures benefit the bottom 40% most, the middle class to a lesser extent, and the rich not at all. Let us therefore think of the government as choosing, at the margin, from among a menu of public goods which are complementary to private capital accumulation: if g_t is spent on a public good with characteristics $(\kappa, \alpha, \beta, \gamma)$, the private sector faces the following technology:

$$k_{t+1}^i = \kappa \cdot \xi_{t+1}^i \cdot (k_t^i)^\alpha (e_t^i)^\beta (g_t)^\gamma \tag{27}$$

The first type of good corresponds to a high value of $\alpha + \beta$ and a low value of γ; the second type has low $\alpha + \beta$ and high γ. In the simple case where the policy decision consists of choosing a single public good (broadly defined) from such a menu, the problem is analogous to the earlier ones. As a result, countries can sustain different choices without any underlying differences in tastes. In reality many public goods are provided simultaneously and the debate is over the appropriate mix.[25] Nonetheless the same intuitions should carry over to this more complicated problem.

output are given in the appendix. For $\rho < \bar{\rho}$ there is either a unique steady-state or no steady-state. Theorem 1, which also applies to education funding (for $\alpha = 0$ or small enough), indicates that non-existence is an artefact of restricting policy choice to discrete values. Conversely, Theorem 2 applies equally to progressive taxation (for $\tau \in \{0, \bar{\tau}\}$).

[25] Formally, $k_{t+1}^i = \int \kappa \cdot \xi_t^i \cdot (k_t^i)^\alpha (e_t^i)^\beta (g_t(\kappa, \alpha, \beta\gamma))^\gamma \, d\kappa \, d\alpha \, d\beta \, d\gamma$, with $\int g_t(\kappa, \alpha, \beta, \gamma) \, d\kappa \, d\alpha \, d\beta \, d\gamma$ equal to govern-

6.2 The Socioeconomic Structure of Cities

The presence in human capital accumulation of peer effects, role models and other non-fiscal neighborhood interactions implies that residential stratification increases the persistence of income disparities across families (e.g., Durlauf (1996a)). Urban ghettos are but the most extreme example of this phenomenon, which Borjas (1995) and Cooper, Durlauf and Johnson (1994) document empirically. Bénabou (1993), (1996a) shows that the long-run impact of this segregation on aggregate surplus is likely to be negative; put in another way, equilibrium segregation tends to be inefficiently high. The two conditions for multiple steady-states identified in this paper are therefore satisfied: there are redistributive policies (broadly speaking, subsidizing integration) which can improve total output and welfare. Yet a high degree of segregation can be self-perpetuating, precisely due to the increased inequality which it induces. Theories which ascribe international differences to a multiplicity of equilibria always face the challenge of identifying the original source of these diverging paths. In the case of segregation, the historical circumstances which set the United States and a few other countries on a permanently different course do not seem hard to pinpoint.[26]

7 Conclusion

This paper explained why countries with similar technologies and preferences, as well as equally democratic political systems, can nonetheless make very different choices with respect to social insurance, fiscal progressivity, education finance, or labor market institutions. The answer involves two mechanisms which arise naturally in the absence of complete insurance and credit markets. First, redistributions which could increase average welfare command less political support in an unequal society than in a more homogenous one. A lower rate of redistribution, in turn,

ment revenue. Benoît and Osborne (1994) study a related issue in the context of crime, examining what determines a country's chosen mix between the severity of punishment and preventive social expenditures.

[26] In addition to the obvious legacy of racial segregation, the post-war period saw a combination of sudden technological change (the spread of the automobile) and federal government policies (subsidies to infrastructure and home ownership) which made extensive suburbanization viable. Formally, Theorem 2 can be applied to show that segregation ($\tau = 0$) and integration ($\tau = 1$) can coexist as stable steady-states in Bénabou's (1996a) model of stratification and growth. Another mechanism through which this may occur is the one explored by Durlauf (1996b), where it is only if income disparities are not too large that rich families are willing to share with poorer ones the fixed costs of running a community and its schools.

increases inequality in future incomes due to wealth constraints on human or physical capital investment. This leads to two stable steady–states, with low inequality yet high redistribution in some countries, and the reverse configuration in others. While they are not Pareto rankable, ex–ante welfare is always higher under the more redistributive social contract. Because of the value of insurance (and possibly inequality–aversion) this is consistent with mean income being either higher or lower. Which equilibrium has higher output or faster growth will reflect the balance between tax distortions to effort and the productivity gains from reallocating investment resources to more severely wealth–constrained agents.

These ideas were formalized in a stochastic growth model with missing markets, progressive taxes or education policy, and a flexible representation of the political system. The resulting distributional dynamics remain simple enough that a variety of extensions can be considered. In Bénabou (1995), for instance, I incorporate labor supply and quantify the effects of redistribution on output. Another interesting issue arises from differences in the organization of production across developed countries; while not as striking as those in the social contract, they are nonetheless significant. The interactions between technological choice, the education system and the political sphere will be explored in future research.

Appendix

Proof of Propositions 2 and 3 Let $\nu_t^i \equiv e_t^i/\tilde{y}_t^i$ and take expected present values (conditional on $\ln k_0^i$), denoted by $EPV[\cdot]$, in (5):

$$U_0^i \equiv EPV[\ln c_t^i] = EPV[\ln(1-\nu_t^i)(1-\theta_t)] + EPV[(1-\tau_t)\ln k_t^i] + EPV[\tau_t \ln \tilde{y}_t]. \tag{A.1}$$

so the first-order condition for the optimal ν_0^i is:

$$\frac{1}{1-\nu_0^i} = EPV\left[(1-\tau_t)\frac{\partial \ln k_t^i}{\partial \nu_0^i}\right] \tag{A.2}$$

Next, substituting (8) and (10) into the accumulation equation (12) yields:

$$\ln k_{t+1}^i = \ln \kappa + \ln \xi_{t+1}^i + \beta \ln[\nu_t^i(1-\theta_t)(1+a_t)] + (\alpha + \beta(1-\tau_t))\ln k_t^i + \beta\tau_t \ln \tilde{y}_t \;\Rightarrow \tag{A.3}$$

$$\frac{\partial E[\ln k_t^i]}{\partial \nu_0^i} = \frac{\beta}{\nu_0^i}\prod_{k=1}^{t-1}(\alpha + \beta(1-\tau_k)). \tag{A.4}$$

Substituting into (A.2) yields the first result of the Proposition, as the choice of date $t = 0$ was clearly arbitrary. The symmetry of agents's savings decisions, $\nu_t^i = \nu_t$, implies by (A.3) that if wealth is initially log-normally distributed across agents, this remains true in every period. In turn, $\ln k_t^i \sim \mathcal{N}(m_t, \Delta_t^2)$ implies that the level of transfers $\tilde{y}_t$ which satisfies the first budget constraint (9) is $\ln \tilde{y}_t = m_t + (2-\tau_t)\Delta_t^2/2$. As to the government's second budget constraint (11), with $\nu_t^i = \nu_t$ it becomes:

$$\theta_t\, y_t = a_t\, \nu_t \int_0^1 \hat{y}_t^i\, di = a_t\, \nu_t\, (1-\theta_t)\, y_t \;\Rightarrow\; \nu_t(1-\theta_t)(1+a_t) = \theta_t + \nu_t(1-\theta_t) \tag{A.5}$$

The accumulation equation can therefore be rewritten:

$$\ln k_{t+1}^i = \ln \kappa + \ln \xi_{t+1}^i + (\alpha+\beta(1-\tau_t))\ln k_t^i + \beta \ln(\theta_t + \nu_t(1-\theta_t)) + \beta\tau_t(m_t + (2-\tau_t)\Delta_t^2/2)\,. \tag{A.6}$$

We can now evaluate intertemporal welfare, under any sequence $\{\tau_t, \theta_t\}_{t=0}^{\infty}$. First, note that:

$$\ln c_t^i = \ln(1-\nu_t)(1-\theta_t) + (1-\tau_t)(\ln k_t^i - m_t) + m_t + \tau_t(2-\tau_t)\Delta_t^2/2 \tag{A.7}$$

To compute the expected present value of the terms on the right hand-side, note that:

$$\begin{aligned}
\ln k_{t+1}^i - m_{t+1} &= (\alpha+\beta(1-\tau_t))(\ln k_t^i - m_t) + \ln \xi_{t+1}^i \Rightarrow \\
EPV[(1-\tau_t))(\ln k_t^i - m_t)] &= EPV\left[(1-\tau_t)\prod_{k=0}^{t-1}(\alpha+\beta(1-\tau_k))\right](\ln k_0^i - m_0); \\
m_{t+1} &= (\alpha+\beta)m_t + \ln\tilde{\kappa} + \beta\ln(\theta_t+\nu_t(1-\theta_t)) + \beta\tau_t(2-\tau_t)\Delta_t^2/2 \Rightarrow \\
EPV[m_t](1-\rho(\alpha+\beta)) &= m_0 + \rho EPV[\ln\tilde{\kappa} + \beta\ln(\theta_t+\nu_t(1-\theta_t)) + \beta\tau_t(2-\tau_t)\Delta_t^2/2]
\end{aligned}$$

where $\tilde{\kappa} = \kappa - s^2/2$. Therefore:

$$\begin{aligned}
U_0^i &= EPV[\ln(1-\nu_t)(1-\theta_t)] + (1-\rho(\alpha+\beta))^{-1}(m_0 + EPV[\rho\beta\ln(\theta_t+\nu_t(1-\theta_t)) + \ln\tilde{\kappa}]) \\
&\quad +(1-\rho\alpha)(1-\rho(\alpha+\beta))^{-1}EPV[\tau_t(2-\tau_t)\Delta_t^2/2] \\
&\quad +(\ln k_0^i - m_0)\,EPV\left[(1-\tau_t)\prod_{k=0}^{t-1}(\alpha+\beta(1-\tau_k))\right] \qquad \text{(A.8)}
\end{aligned}$$

Recall that $\{\theta_t\}_{t=0}^{\infty}$ does not affect the evolution of the distribution of income $\{\Delta_t^2\}_{t=0}^{\infty}$, hence has no possible feedback on the choice of $\{\tau_t\}_{t=0}^{\infty}$. Therefore (A.8) shows that given $\{\tau_t\}_{t=0}^{\infty}$, all agents in generation zero have the same preferences over sequences $\{\theta_t\}_{t=0}^{\infty}$. If any of them is free to choose the whole sequence, he will set for each t :

$$\frac{1}{1-\theta_t} = \frac{\rho\beta}{1-\rho(\alpha+\beta)} \cdot \frac{(1-\nu_t)}{\theta_t+\nu_t(1-\theta_t)} \Leftrightarrow \frac{1-(1-\theta_t)(1-\nu_t)}{(1-\theta_t)(1-\nu_t)} = \frac{\rho\beta}{1-\rho(\alpha+\beta)} \qquad \text{(A.9)}$$

Since this is a constant, all members of all generations agree that given $\{\tau_t\}_{t=0}^{\infty}$, the optimal sequence $\{\theta_t\}_{t=0}^{\infty}$ is the one which equalizes every period's investment rate $\theta_t + \nu_t(1-\theta_t)$ to $s = \rho\beta/(1-\rho\alpha)$. This concludes the proof of Propositions 2 and 3.

Proof of Proposition 4 For any weighting scheme with $\omega^i = \tilde{y}(x^i)$, the proportion of (expected) votes cast by agents i with $x^i \leq x$ (their total political weight) is $G(x)/G(\infty)$ where:

$$G(x) \equiv \int_{-\infty}^{x} \tilde{y}(z)\,dF(z)$$

and F denotes the c.d.f. of x. For an ordinal scheme, $\tilde{y}(z) = \omega(F(z))$, so $G(x) = \int_0^{F(r)} \omega(r)\,dr$. Given the properties satisfied by preferences, the agent with rank $r \equiv \Phi(\lambda)$ such that $G(r)/G(\infty) = 1/2$ is pivotal. Moreover, in the log-normal case $F(z) = \Phi((z-m)/\Delta)$, so $G(x)/G(\infty) = F(x - \lambda\Delta)$; the whole distribution is simply shifted by $\lambda\Delta$. For the cardinal scheme with $\tilde{y}(z) = e^{\lambda z}$, simple derivations show that:

$$G(x) = \int_{-\infty}^{x} e^{\lambda z}\, dF(z) = e^{\lambda(m+\lambda\Delta^2/2)} \cdot F(x - \lambda\Delta^2), \tag{A.10}$$

hence $G(x)/G(\infty) = F(x - \lambda\Delta^2)$. The distribution is now shifted by $\lambda\Delta^2$, and so is the solution to $G(x^*)/G(\infty) = 1/2$.

Proof of Proposition 5 The right-hand side of (19) is proportional to $\varphi(\alpha) \equiv (1 - \rho\alpha)/\sqrt{1-\alpha^2}$. The result follows from $\varphi'(\alpha)/\varphi(\alpha) = \alpha/(1-\alpha^2) - \rho/(1-\rho\alpha) = (\alpha-\rho)/(1-\alpha^2)(1-\rho\alpha)$.

Proof of Proposition 7 I will show quasi-concavity of U_0^i in τ by proving that, for all τ: if $\partial U_0^i(\tau)/\partial\tau = 0$, then $\partial^2 U_0^i(\tau)/\partial\tau^2 < 0$. Let $C \equiv ((1-\rho\beta^2)/\beta(1-\rho\beta))(\Delta_0^2 + \rho s^2/(1-\rho))$, and rewrite the expression in Proposition 7 as $\partial U_0^i/\partial\tau = \Psi(\beta(1-\tau); m_0 - x_0^i, C)$, where

$$\Psi(u; v, \Delta^2, s^2) \equiv \frac{v}{(1-\rho u)^2} + \frac{Cu}{(1-\rho u^2)^2} \tag{A.11}$$

Consider now Ψ as a function of u, for fixed values of the other arguments. Then

$$\Psi'(u) = \frac{2\rho v}{(1-\rho u)^3} + \frac{C(1+3\rho u^2)}{(1-\rho u^2)^3}$$

and if $\Psi(u) = 0$ this becomes:

$$\Psi'(u) = \frac{C[(1+3\rho u^2)(1-\rho u) - 2\rho u(1-\rho u^2)]}{(1-\rho u)(1-\rho u^2)^3} = \frac{C(-\rho^2 u^3 - 3\rho u(1-u) + 1)}{(1-\rho u)(1-\rho u^2)^3}. \tag{A.12}$$

Let $P(u)$ denote the third-degree polynomial multiplying C in the numerator. Its derivative, $P'(u) = -3\rho(\rho u^2 - 2u + 1)$, reaches its maximum at $1/\rho > 1/\sqrt{\rho}$. Therefore, $P'(u) > 0$ for all $u \in [0, 1/\sqrt{\rho}]$; since $P(0) = 1$, this implies that $P(u) > 0$ on that same interval. Thus $\Psi'(u) > 0$ (i.e., $\partial^2 U_0^i(\tau)/\partial\tau^2 < 0$) whenever $\Psi(u) = 0$ (i.e., $\partial U_0^i(\tau)/\partial\tau = 0$), which proves the result. Finally, note that $\Psi(1/\sqrt{\rho}) = +\infty$, or $\partial U_0^i/\partial\tau = +\infty$ at the minimal value $\underline{\tau} = 1 - 1/\beta\sqrt{\rho}$; conversely at the maximal value $\tau = 1$, $\Psi(0) = 2\rho v$. Therefore U_0^i is maximized at $\tau = 1$ for an agent with $v = m_0 - x_0^i > 0$, and at an interior point where $\partial U_0^i/\partial\tau = 0$ for one with $v < 0$. In that latter case, comparative statics follow immediately from the implicit function theorem.

Proof of Proposition 8 The tax rate τ_0 is obtained by setting $\ln k_0^i - m_0 = \lambda\Delta_0$ in the first-order condition $\partial U_0^i/\partial\tau = 0$ or, with the notation defined above:

$$\Psi(\beta(1-\tau_0); -\lambda\Delta_0, \Delta_0^2 + \rho s^2/(1-\rho)) = 0 \tag{A.13}$$

Therefore, with all derivatives evaluated at that same point where $\Psi = 0$, $\partial\tau_0/\partial\Delta_0 = (-\lambda\Psi_2 + 2\Delta_0\Psi_3)/(-\Psi_4)$. Strict quasiconcavity of U_0^i implies that the denominator is positive, while the numerator equals:

$$-\lambda\Psi_2 + 2\Delta_0\Psi_3 = \frac{-\lambda}{(1-\rho u)^2} + \frac{2\Delta_0 u}{(1-\rho u^2)^2} = \frac{u(\Delta_0^2 - \rho s^2/(1-\rho))}{\Delta_0(1-\rho u^2)^2}$$

where I have used the fact that $\Psi = 0$. Therefore, τ_0 is U-shaped in Δ_0^2, with a minimum at $\Delta_0^2 = \rho s^2/(1-\rho)$. The monotonicity in s^2 follows immediately from $\Psi_3 > 0$.

Proof of Theorem 1 Defining $u = \beta(1-\tau)$, the function $f(\tau)$ in (23) is proportional to

$$\varphi(u) \equiv \frac{u}{\sqrt{1-u^2}} \frac{(1-\rho u)^2}{1-\rho u^2}; \tag{A.14}$$

this function is such that $\varphi(0) = 0$ and $\varphi(1) = +\infty$ (for $\rho < 1$). I will show that for ρ high enough, $\varphi(u)$ has exactly one local maximum u_1 and one local minimum $u_2 > u_1$ in $(0,1)$; in other words, it is N-shaped. This will prove the claims concerning the number of equilibria. Differentiating shows that $\varphi'(u)$ has the sign of

$$P(u;\rho) \equiv (1-\rho u)(1-\rho u^2) - 2\rho u(1-u)(1-u^2) \tag{A.15}$$

which is a fourth-order polynomial in u. Therefore φ can have at most four local extrema in $[0,1]$. If it had four, the second derivative $\partial^2 P(u;\rho)/\partial u^2 = -12u^2 + 3(2+\rho)u + 1$ would have two zeros in that interval. However, inspection of this polynomial in u shows that it has two real roots of opposite sign; since it takes values $+1$ at $u = 0$ and $3\rho - 5 < 0$ at $u = 1$, only one of these belongs to $[0,1]$. Therefore $P(u;\rho)$ cannot have four roots in that interval. Neither can it have three, u_1, u_2, u_3: since $\partial P(0;\rho)/\partial u > 0$, u_1 and u_3 would correspond to local maxima, and u_2 to a local minimum. But since $\partial P(u;\rho)/\partial u > 0$ as $u \to 1$, there would have to be another local minimum in $(u_3, 1)$, which is not possible. Similarly, $P(u;\rho)$ cannot have a unique zero in $(0,1)$. Having proved that it has either none or two, I now show that the first case applies when ρ is low enough, and the second ρ is high enough. For $\rho \leq 1/2$, $P(u;\rho) \geq (1-\rho u)(1-\rho u^2) - u(1-u)(1-u^2) \geq 0$; φ is then increasing on $[0,1]$. To consider the case where ρ is close to 1, rewrite:

$$P(u;\rho) \equiv (1-u)(1-u^2)(1-2u) + (1-\rho)u(2u^3 - (3+\rho)u^2 - u + 3). \qquad \text{(A.16)}$$

For $\rho = 1$, this expression equals zero for $u = 1/2$. Let now $\varepsilon \equiv 1-\rho << 1$ and $\eta = u - 1/2 << 1$. A Taylor expansion of $P(u;\rho)$ yields:

$$P(u;\rho) \approx \frac{3\eta}{4} - \frac{7\varepsilon}{8} + o(\eta) + o(\varepsilon). \qquad \text{(A.17)}$$

Therefore if we set η/ε to be a number smaller than 7/6, then $P(u;\rho) < 0$. This implies that $P(u;\rho)$ has one root in $(0,1)$, hence in fact two, due to our earlier result. Thus the function φ has exactly on local maximum and one local minimum on that interval, which completes the proof of the first part of Theorem 1. It remains to show that steady-state output is higher in the equilibrium with a higher value of τ. Taking limits in equation (13) and (14) yields:

$$m_\infty = \bar{m} + \frac{\beta\tau(2-\tau)}{1-\alpha-\beta} \cdot \frac{\Delta_\infty^2}{2} \qquad \text{(A.18)}$$

$$y_\infty = \bar{m} + \frac{1-\alpha-\beta(1-\tau)^2}{1-\alpha-\beta} \frac{s^2/2}{1-(\alpha+\beta(1-\tau))^2} \qquad \text{(A.19)}$$

where $\bar{m} \equiv (\ln\kappa - s^2/2 + \beta \ln s$. Recall now that we are considering the case where $\alpha = 0$; thus y_∞ varies with $1-\tau$ as $(1-\beta u^2)/(1-\beta^2 u^2)$ varies with u, which is negatively since $\beta < 1$.

Proof of Theorem 2 The right-hand side of (25) is larger than that of (26) if and only if $\rho \geq \underline{\rho}$, where $\underline{\rho} \in (0,1)$ is defined by:

$$\underline{\rho} \equiv \frac{\sqrt{1-\rho(\alpha+\beta(1-\bar{\tau}))^2} - \sqrt{1-\rho(\alpha+\beta)^2}}{(\alpha+\beta)^2\sqrt{1-\rho(\alpha+\beta(1-\bar{\tau}))^2} - (\alpha+\beta(1-\bar{\tau}))^2\sqrt{1-\rho(\alpha+\beta)^2}} \qquad \text{(A.20)}$$

Output in the two steady-states is given by (A.19) with the values $\tau = \bar{\tau}$ and $\tau = 0$ respectively. Denoting $u = 1-\bar{\tau}$, the difference between the two has the sign of $(1+\alpha+\beta)(1-\alpha-\beta u^2) - 1 + (\alpha+\beta u)^2 = \beta(1-u)(1-\alpha+(1+\alpha)u) > 0$. Hence the result. When $\rho < \underline{\rho}$ the right-hand side of (25) is smaller than that of (26); a (unique) steady-state exists if and only if λ is outside this interval.

References

Aghion, P. and Bolton, P. (1996) "A Trickle-Down Theory of Growth and Development with Debt Overhang," *Review of Economics Studies*, forthcoming.

Alesina, A. and Rodrik, D. (1994) "Distributive Politics and Economic Growth," *Quarterly Journal of Economics, 109*, 465-490.

Atkinson, A. and Stiglitz, J. (1980) "Lectures on Public Economics," Mc–Graw Hill, New York.

Banerjee, A. and Newman, A. (1991) "Risk-Bearing and the Theory of Income Distribution," *Review of Economic Studies*, 58, 211-236.

Banerjee, A. and Newman, A. (1993) "Occupational Choice and the Process of Development," *Journal of Political Economy*, 101, 274-298.

Bénabou, R. (1993) "Workings of a City: Location, Education and Production," *Quarterly Journal of Economics*, 108, 619-652.

Bénabou, R. (1995) "Meritocracy, Redistribution and Efficiency," New York University mimeo, November.

Bénabou, R. (1996a) "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance," *American Economic Review, forthcoming.*

Bénabou, R. (1996b) "Inequality and Growth," *NBER Macro Annual*, forthcoming.

Benoît, J.P., and Osborne, M. (1994) "Crime, Punishment and Social Expenditures," New York University mimeo, October.

Bertola, G. (1993) "Factor Shares and Savings in Endogenous Growth," *American Economic Review*, 83, 1184-1198.

Borjas, G. (1995) "Ethnicity, Neighborhoods, and Human-Capital Externalities," *American Economic Review*, 85, 365–390.

Caillaud, B., Cohen, D. and Julien, B. (1994) "Towards a Theory of Self-Restraint," CEPREMAP Discussion Paper 9421, June.

Chamley, C. (1985) "Efficient Tax Reform in a Dynamic Model of General Equilibrium," *Quarterly Journal of Economics*, May, 335–356.

Cohen, D. and Michel, P. (1991) "Which Rules Rather Than Discretion: An Axiomatic Approach," CEPR Discussion Paper 537, April.

Clarke (1992) "More Evidence on Income Distribution and Growth," World Bank Working Paper 1064, Policy, Research and External Affairs, December.

Conway, M. (1991) "Political Participation in the United States," US Congressional Quarterly Press, Washington D.C.

Cooper, S. (1992) "A Positive Theory of Income Redistribution," Stanford University mimeo, July.

Cooper, S., Durlauf, S. and Johnson, P. (1994) "On the Evolution of Economic Status Across Generations," *American Statistical Association, Business and Economics Section, Papers and Proceedings,* 50-58.

Deininger, K. and Squire, L. (1995) "Inequality and Growth: Results from a New Data Set," World Blank Mimeo, December.

Devarajan, S. , Swaroop, V. and Zou, H. (1993) "What Do Governments Buy?" World Bank Working Paper 1082, Policy, Research and External Affairs, February.

Durlauf, S. (1996a) "A Theory of Persistent Income Inequality," *Journal of Economic Growth,* forthcoming.

Durlauf, S. (1996b) "Neighborhood Feedbacks, Endogenous Stratification, and Income Inequality," forthcoming in *Proceedings of the Sixth International Symposium on Economic Theory and Econometrics,"* W. Barnett, G. Gandolfo and C. Hillinger, eds., Cambridge University Press.

Easterly, W. and Rebello, S. (1993) "Fiscal Policy and Economic Growth," *Journal of Monetary Economics,* 32, 417-458.

Edsall, T. (1984) *"The New Politics of Inequality,"* W.W. Norton & Company, New York.

Fernandez, R. and Rogerson, R. (1994) "Public Education and the Dynamics of Income Distribution: A Quantitative Evaluation of Education Finance Reform," NBER Working Paper 4883, October.

Freeman, R. (1995) "Labour Market Institutions and Earnings Inequality," Harvard University mimeo, November.

Galor, O. and Zeira, J. (1993) "Income Distribution and Macroeconomics," *Review of Economic Studies,* 60, 35-52

Glomm, G. and Ravikumar, B. (1992) "Public vs. Private Investment in Human Capital: Endogenous Growth and Income Inequality," *Journal of Political Economy,* 100, 818-834.

Gradstein M. and Justman, M. (1993) "Education, Inequality and Growth: A Public Choice Perspective," Ben Gurion University mimeo, October.

Grossman, G. and Helpman, E. (1996) "Intergenerational Redistribution with Short-Lived Governments," NBER Working paper 5447, January.

Judd, K. (1985) "Redistributive Taxation in a Perfect Foresight Model," *Journal of Public Economics,* 28, 59–83.

Keefer, P. and Knack, S. (1995) "Polarization, Property Rights and the Links Between Inequality and Growth," World Bank mimeo, October.

Krusell, P., Quadrini, V. and Ríos-Rull, J.-V. (1994) "Politico -Economic Equilibrium and Economic Growth," CARESS working paper 94-11, University of Pennsylvania, March.

Lindert, P. (1996) "What Limits Social Spending?" *Explorations in Economic History*, 33, 1-34.

Loury, G. (1981) "Intergenerational Transfers and the Distribution of Earnings," *Econometrica*, 49, 843-867.

Perotti, R. (1992) "Fiscal Policy, Income Distribution, and Growth," Columbia University Working Paper 636, November.

Perotti, R. (1993) "Political Equilibrium, Income Distribution, and Growth," *Review of Economic Studies*, 60, 755-776.

Perotti, R. (1994) "Income Distribution and Investment," *European Economic Review*, 38, 827-835.

Perotti, R. (1996) "Growth, Income Distribution and Democracy: What the Data Say," *Journal of Economic Growth*, forthcoming.

Persson, T. and Tabellini, G. (1991) "Is Inequality Harmful for Growth? Theory and Evidence," *American Economic Review*, 48, 600-621.

Piketty, T. (1995) "Social Mobility and Redistributive Politics," Quarterly Journal of Economics, 110, 551-442.

Piketty, T. (1996) "The Dynamics of the Wealth Distribution and Interest Rate with Credit-Rationing," *Review of Economic Studies*, forthcoming.

Roemer, J. (1995) "Why the Poor Do Not Expropriate the Rich in Democracies: A New Argument," mimeo, University of California - Davis, May.

Saint-Paul, G. (1994) "The Dynamics of Exclusion and Fiscal Conservatism," CEPR Discussion Paper 998, July.

Saint-Paul, G. and Verdier, T. (1993) "Education, Democracy and Growth," *Journal of Development Economics*, 42, 2, 399-407.

Saint-Paul, G. and Verdier, T. (1992) "Historical Accidents and the Persistence of Distributional Conflicts," *Journal of the Japanese and International Economies*, 6, 406-422.

Sala-i-Martin, X. (1992) "Transfers," NBER Working Paper 4186, October.

Verdier, T. and Ades, A. (1993) "The Rise and Fall of Elites: Economic Development and Social Polarization in Rent-Seeking Societies," Harvard University mimeo, February.

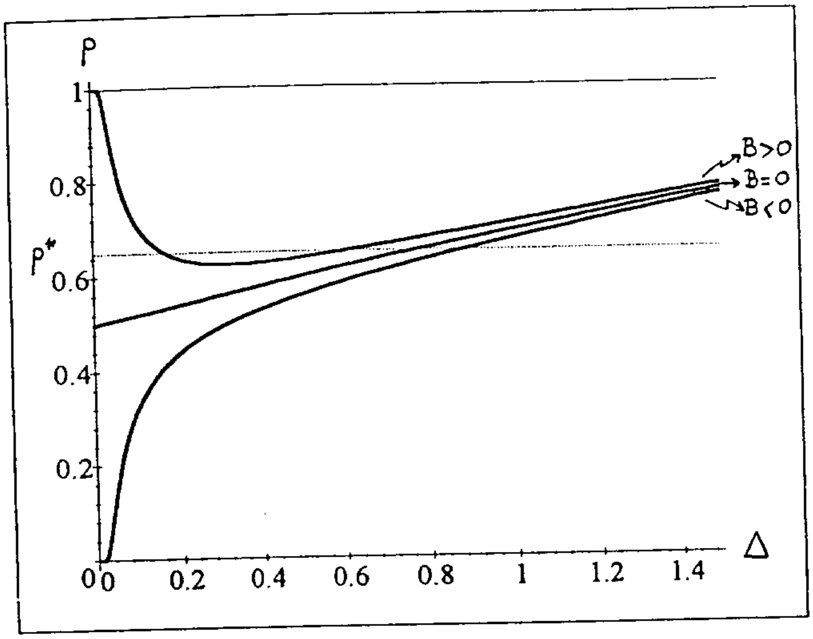

Figure 1: Inequality and political support for redistribution (B=5%, 0, -5%)