

# Luxembourg Income Study Working Paper Series

Working Paper No. 142

Inequality and Growth

Roland Benabou

June 1996

*(scanned copy)*



# Inequality and Growth

Roland Bénabou<sup>\*</sup>

New York University

March 1996

Revised June 1996

<sup>\*</sup> This paper is forthcoming in the *NBER Macroeconomics Annual*. I am indebted to Jess Benhabib, Jordi Gali and Julio Rotemberg for helpful conversations at early stages in this project. I am also grateful for their detailed comments to my discussants, Roberto Perotti and José-Víctor Ríos-Rull, as well as to Alberto Alesina, Abhijit Banerjee, Andrew Bernard, Giuseppe Bertola, Patrick Bolton, Raquel Fernandez, Herschel Grossman, Andrew Newman, Torsten Persson and Thomas Piketty. I also wish to thank Klaus Deininger and Lynn Squire for providing me with their data set, and Tim Smeeding for providing me with specific series from the Luxembourg Income Study. Research Assistance from the C.V. Starr Center is gratefully acknowledged.

## Abstract

Using two unifying models and an empirical exercise, this paper presents and extends the main theories linking income distribution and growth, as well the relevant empirical evidence. The first model integrates the political economy and imperfect capital markets theories. It allows for departures from perfect democracy and embodies the tradeoff between the growth costs and benefits of redistribution through taxes, land reform or public schooling: such policies simultaneously depress savings incentives and ameliorate the wealth constraints which impede investment by the poor. The second model is a growth version of the prisoner's dilemma which captures the essence of theories where sociopolitical conflict reduces the security of property rights, thereby discouraging accumulation. The economy's growth rate is shown to fall with interest groups' rent-seeking abilities, as well as with the gap between rich and poor. It is not income inequality per se that matters, however, but inequality in the relative distribution of earning and political power. For each of the three channels of political economy, capital markets and social conflict, the empirical evidence is surveyed and discussed in conjunction with the theoretical analysis. Finally, the possibility of multiple steady-states leads me to take up a new empirical issue: are cross-country differences in inequality permanent, or gradually narrowing? Equivalently, is there convergence not just in the first moment of individual income (GDP per capita), but convergence in distribution?

**Keywords:** Growth, Inequality, Income Distribution, Political Economy, Redistribution, Credit Constraints, Political Instability, Property Rights.

**JEL classification:** O11, D31, E62, D72.

# 1 Introduction

## 1.1 Korea and the Philippines, circa 1960: Not So Similar After All

To introduce the theme of this paper I shall revisit the puzzle raised by Lucas (1993) at the beginning of his article. In the early 1960's, South Korea and the Philippines were similar with respect to all major economic aggregates: GDP per capita, population, urbanization, primary and secondary school enrollment. The Philippines had a somewhat higher share of GDP in manufacturing, but in exports both countries had the same proportions of primary commodities and manufactures. Given nearly identical starting points, how can it be that over the next quarter century Korea experienced "miraculous" growth averaging about 6% per annum, while the Philippines stagnated at about 2%?

If one looks beyond first moments, however, initial conditions were in fact quite different. As shown by Table 1, the distribution of income was considerably more unequal in the Philippines:

|             | Gini (%) | Q1   | Q2    | Q3    | Q4    | Q5    | Q3+Q4 | Q5/Q1 | Q5/(Q1+Q2) |
|-------------|----------|------|-------|-------|-------|-------|-------|-------|------------|
| 1965        |          |      |       |       |       |       |       |       |            |
| Korea       | 34.34    | 5.80 | 13.54 | 15.53 | 23.32 | 41.81 | 38.85 | 7.21  | 2.16       |
| Philippines | 51.32    | 3.50 | 12.50 | 8.00  | 20.00 | 56.00 | 20.50 | 16.00 | 3.50       |
| 1988        |          |      |       |       |       |       |       |       |            |
| Korea       | 33.64    | 7.39 | 12.29 | 16.27 | 21.81 | 42.24 | 38.08 | 5.72  | 2.15       |
| Philippines | 45.73    | 5.20 | 9.10  | 13.30 | 19.90 | 52.50 | 33.20 | 10.10 | 3.67       |

Table 1: Korea and the Philippines

Source: Deininger and Squire (1995a) data set;  $Q_i$  denotes the share of the  $i$ -th quintile.

The Philippines' Lorentz curve lay everywhere below that of Korea.<sup>1</sup> The Gini coefficient was seventeen percentage points higher, about 1.8 standard deviations in the world distribution of Ginis, or 2.5 among East-Asian countries. Most strikingly, the ratio of the income share of the top 20% to the bottom 20%, or even to the bottom 40% was about *twice as large* in the Philippines. Similar disparities characterized land ownership: the Gini coefficient for farmland was 38.7 for Korea and 53.4 for the Philippines in 1961 and 1960 respectively.<sup>2</sup> This greater concentration of income and wealth, which persists to this day, cannot be blamed on the "kleptocratic" nature of the Marcos regime. Ferdinand Marcos was first elected president in 1965, and declared a state

<sup>1</sup>The year 1965 is the earliest one for which data for both countries are available. Moreover, these figures were based on the same survey method (gross household income), making them comparable. The 1988 Philippine number is for gross personal income, but very close to the 1985 figure which still used household income.

<sup>2</sup>The source for land Ginis is Taylor and Hudson (1972). The twelve-point gap corresponds to about one standard deviation. While it is common knowledge that Korea implemented a land reform following World War II, it may be worth mentioning that the Philippines also had its own around the end of World War I, during the American occupation period. Following the disestablishment of the Catholic Church, a large part of its land holdings were purchased and redistributed. The relatively high land Gini for 1960 suggests that either land reform was not as egalitarian as in Korea, where individual holdings were limited to 3 hectares, or that land ownership became re-concentrated during the following forty years

of emergency only in 1972. Preceding him in power at the end of the US-assisted reconstruction period were two presidents whose main policies are described in the US State Department's background notes as seeking to "expand Philippine ties to its Asian neighbors, implement domestic reform programs, and develop and diversify the economy".

Of course Table 1 does not constitute proof that greater initial equality was the reason – certainly not the sole reason – why the "miracle" occurred in one country and not the other. But the facts which it documents do suggest that the answer to the puzzle may lie outside the representative agent framework.

## 1.2 Empirical Regularities

South Korea and the other East-Asian "dragons" are usually contrasted not to the nearby Philippines, but to Latin American countries. It has long been part of development economists' conventional wisdom that the very equal distribution of income and land in the first group played a significant role in their take-off, whereas the high levels of wealth concentration in the later were a serious impediment to growth. In recent years the literature has moved from anecdotal to formal empirical evidence, impelled by the findings of Alesina and Rodrik (1994) and Persson and Tabellini (1994) of a negative effect of inequality in cross-country growth regressions and by Perotti's (1992), (1994), (1996) methodical testing of the main theories.

Table 2 summarizes the main results from twenty-three recent studies of the links from inequality to growth or investment. I shall refer to it throughout the paper. Column (1) corresponds to the basic, reduced form regression, where the average growth rate of per capita GDP over some long period (twenty years or so) is regressed on initial inequality and several controls. These typically include: (a) initial income, whose coefficient is always negative and significant; (b) the initial stock of human capital, or proxies in the form of initial school enrollment ratios; as shown by column (8) this effect is systematically positive;<sup>3</sup> (c) regional dummies. These regressions, run over a variety of data sets and periods with many different measures of income distribution, deliver a consistent message: initial inequality is detrimental to long-run growth. The magnitude of this effect is consistent across most studies: a one standard deviation decrease in inequality raises the annual growth rate of GDP per capita by .5 to .8 percentage points. Whether this is large or small may be in the eye of the beholder. On one hand it amounts to between 30% and 45% of the standard deviation of growth rates found in most samples. It also implies an income gap of about 25% after 30 years, which is far from inconsequential. On the other hand, this does not come close to accounting for the growth differential observed between the Philippines and Korea, or East Asia and Latin America. However, several models in the literature predict a non-linear effect, possibly leading to a multiplicity of equilibria. In that case small differences in

---

<sup>3</sup>With the exception of the puzzling result that when male and female stock are distinguished, the latter appears to contribute negatively (e.g. Barro (1996), who also discusses possible explanations).

initial inequality can have marginal effects well in excess of the average slope estimated by linear regressions, and significantly influence long-run outcomes.

The inclusion of controls (b) and especially (c) typically reduces the coefficient on inequality. Sometimes it becomes insignificant, as in Benhabib and Spiegel (1993) when the stock of human capital is included, or in Deininger and Squire (1995b) when dummies for Latin America and sub-Saharan Africa are added. In most other studies inequality remains significant even in the presence of these controls. In any case, enrollments in and stocks of secondary education have a substantial negative correlation with inequality, and in some of the theories discussed below the link between income distribution and growth arises precisely through human capital investment. Note finally that a large part of the cross-country variation in inequality comes from the high levels observed in Africa and, particularly, Latin America. These persistent inter-regional differences need to be explained, and in fact point in the direction of the models with multiple long-run distributions mentioned above.

### 1.3 Overview

This paper presents and extends the main theories linking income distribution and growth, as well the evidence on their relevance. This is done through two unifying models, a survey of the empirical literature, and an econometric exercise. The first model integrates the political economy and imperfect capital markets theories. The second one deals with social conflict and the security of property rights. The empirical section asks whether countries are converging to the same level of inequality. The paper focuses primarily on how the distribution of income (whether exogenous or endogenous) can affect output growth, rather than on reverse effects from the level of development to inequality. It thus barely touches on the Kuznets (1955) hypothesis.<sup>4</sup>

I start in Section 2 with theories where asset markets are complete and distributional effects arise solely through the balance of power in the political system (Bertola (1993), Alesina and Rodrik (1994), Persson and Tabellini (1994)). The idea is that by lowering the income of the median voter or pivotal middle class relative to the national average, greater inequality increases the pressure for redistribution. This, in turn, discourages investment. The model developed here, which combines features of Persson and Tabellini (1994) and Bénabou (1995a), displays this mechanism but also formalizes departures from the one person, one vote, ideal. Whether or not there is a role for productive public investment, growth in this class of models always increases with the degree of pro-wealth bias in the political system. I also provide new results, together with a general caveat, on the issue of whether inequality affects democracies and non-democracies differentially. Compared to a perfect democracy, income disparities are shown to have a lesser impact on redistribution and growth only in right-wing or wealth-biased regimes, and a greater impact in left-wing or populist ones.

---

<sup>4</sup>See Fields and Jakubson (1994) for a recent review of this issue as well as some new evidence, which does not support the Kuznets curve hypothesis.

In Section 3 I turn to another set of theories, where the distribution of wealth has macroeconomic implications due to imperfections in asset markets. I use the very same model as before but simply “turn off” the loan market. Indeed, the main idea in this literature which starts with Loury (1981) is that credit constraints prevent the poor from undertaking the efficient amount of investment. With decreasing returns their marginal product is higher, so that redistributions can increase total output or growth. My model embodies the tradeoff between these growth benefits of land reform, public schooling or other progressive transfers, and their traditional costs due to depressed incentives for savings or labor supply. The other central issue in the literature is intergenerational mobility, for which uninsurable shocks are of the essence. Some papers focus on the decentralized equilibrium (Galor and Zeira (1993), Banerjee and Newman (1993), Bénabou (1996a), Aghion and Bolton (1996), Piketty (1996)), while others combine market incompleteness with the politics of redistribution (Perotti (1993), Saint-Paul and Verdier (1993), Verdier and Ales (1993), Saint-Paul (1994), Bénabou (1995a)). In a related class of models the macroeconomic effects of distribution arise from, or are magnified by, the endogenous sorting of agents into homogeneous communities or other “clubs” (Bénabou (1993), (1996b), Durlauf (1996a), (1996b), Fernandez and Rogerson (1996), (1994), Maskin and Kremer (1994)). The results obtained in this paper are representative of this entire set, and include a number of new propositions as well. I explain for instance why maximizing growth requires some (interior) degree of democracy, in contrast to the complete markets case. I also demonstrate how the combination of credit constraints and a less than perfectly democratic political system can lead to multiple long-run equilibria, differing both in growth rates and levels of inequality. I then use this example to explain the common principle by which multiple steady-state distributions arise, through a variety of general equilibrium feedbacks, in several of the models mentioned above.

Section 4 turns to theories based on the idea that sociopolitical conflict reduces the security of property rights, thereby discouraging accumulation. In particular, when the gap between rich and poor widens the latter presumably have a greater temptation to engage in rent-seeking or predatory activities at the expense of the former. The security of property rights is the main focus of the models of Grossman (1991), (1994), Acemoglu (1995), Tornell and Velasco (1992), Tornell (1994), Grossman and Kim (1996) and Benhabib and Rustichini (1996). Only the last two papers, however, explicitly link inequality and growth. I propose here a simple growth version of the prisoner’s dilemma which captures the essence of this class of models. The economy’s maximum sustainable growth rate is shown to be negatively related to interest groups’ rent-seeking abilities, as well as to income disparities between them. It may then be profitable for the rich to collectively transfer wealth to the poor through land reform, education subsidies, or trade protection. More generally, the analysis reveals that what really matters is not income inequality per se, but inequality in the relative distribution of earning and political power. I also point out that actual instability or uncertainty is not part of the story told by most models in this class, contrary to common interpretation. As in the case of taxation, growth is reduced through

a decline in the expected return on investment, due to a higher threat of expropriation.

Finally, the possibility of multiple long-run distributions and history-dependence discussed earlier leads me to raise in Section 5 an empirical question which has not been taken up previously: are countries converging to the same level of inequality, or are there permanent differences? Equivalently, is there convergence not only in the first moments of their income distributions – GDP per capita – but also in the second one (and higher), as would be predicted by most versions of the neoclassical model once it is enriched with uninsurable idiosyncratic shocks? Using a new data set put together by Deininger and Squire (1995a) I make a first pass at this issue of convergence in distribution. Although it is insufficient to resolve the issue this exercise uncovers some interesting puzzles, including some evidence of mean-reversion in Ginis.<sup>5</sup>

The general strategy adopted in this paper is one of simplification and unification. The models are thus stripped of many elements which contribute to the richness and realism of the literature, but are not essential to conveying the main ideas. These are briefly discussed in the concluding section, before turning to directions for further research. All proofs are gathered in the appendix.

## 2 Political Economy with Complete Markets

The model to be developed here and in the next section will bring together the political economy and imperfect capital markets theories. There is a continuum of non-altruistic overlapping generations families, indexed by  $i \in [0, 1]$ . The utility of a member  $i$  in generation  $t$  is

$$U_t^i = \ln c_t^i + \rho \ln d_t^i \quad (1)$$

where  $c_t^i$  and  $d_t^i$  denote consumption when young and old respectively. This person is born endowed with resources  $w_t^i$ , distributed independently across agents with mean  $w_t \equiv E[w_t^i]$ . Agents can invest in human or physical capital according to the technology

$$y_t^i = r (k_t^i)^\beta (w_t^i)^{1-\beta}, \quad (2)$$

where  $r$  is constant and  $0 < \beta \leq 1$ ;  $k_t^i$  is the amount invested and  $y_t^i$  is second-period pre-tax income. Note that individuals face decreasing returns but the aggregate technology is linear. One can interpret  $w_t^i$  as agent  $i$ 's basic level of skill or human capital, which can be supplemented through investment. The higher the general level of basic skills, the easier it is to develop or acquire new knowledge. I shall focus for a while on what happens within a given generation, in which case time subscripts can be omitted.

---

<sup>5</sup>The paucity and sometimes poor quality of international data on income distribution remain binding constraints here, as in all empirical work on these issues; see Deininger and Squire (1995a) and Perotti (1996) for discussions of data quality. Another general problem is the lack of almost any data on the distribution of wealth (land is sometimes used as an imperfect proxy), even though in most theories it is this distribution rather than that of income which is determinant.

## 2.1 Savings and Intragenerational Growth

There is a frictionless credit market where agents in each generation borrow from and lend to each other, at some endogenous interest rate  $\tilde{r}$ . The amount borrowed by  $i$  is denoted  $b^i \geq 0$ . Finally, there is a government who redistributes second-period income. I depart here from the standard case of linear taxes and focus instead on a simple log-linear scheme which will yield explicit solutions with both perfect and missing capital markets.<sup>6</sup> Let post-tax and transfer income be

$$\tilde{y}^i = (y^i)^{1-\tau}(\tilde{y})^\tau, \quad (3)$$

where the break-even income level  $\tilde{y}$  is defined by the balanced budget constraint

$$\int_0^1 (y^i)^{1-\tau}(\tilde{y})^\tau di = \int_0^1 y^i di \equiv y. \quad (4)$$

Note that the scheme is *progressive* (e.g.,  $\tilde{y} > y$ ) when  $0 < \tau \leq 1$ , and *regressive* when  $\tau < 0$ ; the maximum rate compatible with free disposal is  $\tau = 1$ . I allow  $\tau < 0$ , which will be relevant when studying imperfectly democratic regimes, but for technical reasons I restrict it to  $\tau \geq -\sqrt{1 + 1/\rho\beta} \equiv \underline{\tau}$ . Given an expected redistribution rate  $\tau$ , agent  $i$ 's maximization problem is

$$\max_{b^i, k^i} \{ \ln(w^i + b^i - k^i) + \rho \ln[ (r(k^i)^\beta w^{1-\beta})^{1-\tau}(\tilde{y})^\tau - \tilde{r}b^i ] \}, \quad (5)$$

and the first-order conditions are:  $d^i/c^i = \rho\tilde{r} = \rho\beta(1-\tau)(\tilde{y}^i/k^i)$ . The second one implies that everyone invests the same amount,  $k^i = k$ , hence  $\tilde{y}^i = y^i = rk^\beta w^{1-\beta} = y = \tilde{y}$ , and  $\tilde{r} = r\beta(1-\tau)(w/k)^{1-\beta}$ . The first Euler equation then becomes

$$rk^\beta w^{1-\beta} - \tilde{r}b^i = \rho r\beta(1-\tau)(w/k)^{1-\beta}(w^i + b^i - k).$$

Summing up over agents and using the loan market-clearing condition  $\int_0^1 b^i di = 0$  yields:

$$k = \frac{\rho\beta(1-\tau)w}{1 + \rho\beta(1-\tau)} \equiv s(\tau)w \quad (6)$$

Each individual thus invests the same fraction  $s = s(\tau)$  of *aggregate* resources. Consequently everyone has the same second-period income, so no actual transfers take place in equilibrium. The critical feature of redistribution under complete markets is the *threat* of expropriation of part of the return to investment, as opposed to actual transfers of wealth.<sup>7</sup> In particular, the growth

<sup>6</sup>This scheme was introduced in Bénabou (1995a) in a context of dynastically altruistic families facing missing credit and insurance markets. Applied here to complete markets, it yields the same results as would be obtained from the standard scheme of linear taxes and lump-sum transfers.

<sup>7</sup>One could easily introduce some multiplicative form of ability into the investment function (2), perhaps correlated with  $w^i$ . Optimal investment levels would then differ across individuals and equilibrium transfers would be positive, but none of the essential results would change.

rate of aggregate income within each generation,

$$g(\tau) \equiv \ln(y/w) = \ln r + \beta \ln s(\tau), \quad (7)$$

declines with the tax rate  $\tau$ .

## 2.2 Intergenerational Linkages

To extend this result to long-run growth rates it remains to specify the dynamic linkages between generations. The simplest kind is an aggregate spillover (as in Persson and Tabellini (1994)) through which the productivity achieved by generation  $t$  becomes embodied into the basic human capital endowment of generation  $t + 1$ :

$$w_{t+1}^i = \epsilon_{t+1}^i y_t, \quad (8)$$

where  $\epsilon_t^i$  is an i.i.d. shock with mean normalized to 1, representing for instance innate ability. Equivalently, let altruistic parents devote some of their second-period resources to children's education (say, primary schooling), which is provided as a public good. Specifically, let parents care about old age consumption, now denoted  $c_t^i$ , and about the human capital endowment of their child:

$$U_t^i = \ln c_t^i + \rho \gamma \ln c_t^i + \rho(1 - \gamma) E_t[\ln w_{t+1}^i]. \quad (1')$$

The latter is a combination of ability and public expenditures on schooling,

$$w_{t+1}^i = \kappa \epsilon_{t+1}^i e_t, \quad (8')$$

with the latter financed through a proportional tax on second-period income. With logarithmic preferences the unanimously preferred tax rate is  $1 - \gamma$ , so that  $c_t^i = \gamma d_t^i$  and  $e_t = (1 - \gamma) d_t = (1 - \gamma) y_t$ . Nothing else in parents' behavior or maximized utility is affected.<sup>8</sup> Thus  $w_{t+1}^i = \kappa(1 - \gamma) \epsilon_{t+1}^i y_t$ , and normalizing the constant to one yields the same result as (8), namely that the intra- and intergenerational growth rates are equal:

$$g_t = \ln(y_t/y_{t-1}) = \ln(w_{t+1}/w_t) = \ln(y_t/w_t) = \ln r + \beta \ln s(\tau_t) \quad (9)$$

where  $\tau_t$  is the tax rate chosen in generation  $t$ . Moreover, the distribution of relative incomes  $w_{t+1}^i/w_t = \epsilon_{t+1}^i$  is stationary, which will result in a time-invariant equilibrium tax rate. In what follows I shall therefore omit time subscripts once again.

---

<sup>8</sup>See Glomm and Ravikumar (1992) and Bénabou (1996a) for models of private, local and public education with a similar specification. The former uses overlapping generations with a "warm-glow" bequest, as I do here; the latter has dynastic preferences.

### 2.3 Inequality and Redistribution

I now examine the determination of taxes within each generation. Agent  $i$ 's optimal investment entails a net borrowing of

$$b^i = \frac{\rho}{1+\rho}(w - w^i) \quad (10)$$

at the interest rate

$$\tilde{r} = r\beta(1-\tau)^\beta(1+\rho\beta(1-\tau))^{1-\beta}.$$

Redistribution through what is essentially a tax on capital income lowers this equilibrium return, thereby favoring borrowers and hurting lenders. It leaves unaffected the "representative" agent with resources equal to the per capita average  $w$ , as he does not make use of the loan market. Given optimal borrowing and investment decisions, the intertemporal utility of individual  $i$  is equal to:

$$U^i(\tau) = V(\tau) + (1+\rho) \ln [1 + (w^i/w - 1)(1 + \rho\beta(1-\tau))/(1+\rho)] \quad (11)$$

where

$$V(\tau) \equiv (1+\rho) \ln w + \ln(1 - s(\tau)) + \rho \ln(rs(\tau)^\beta) = \ln c + \rho \ln d, \quad (12)$$

with  $c$  and  $d$  representing aggregate consumption in the first and second periods. For an agent with the average endowment  $w$  the second term in (11) vanishes. His preferences thus coincide with those of a social planner concerned only with *intertemporal efficiency*, in the sense of being indifferent to the distribution of consumption across individuals. Since

$$V'(\tau) = -\frac{\rho\beta\tau}{(1-\tau)(1+\rho\beta(1-\tau))}, \quad (13)$$

the preferred tax rate of this representative individual or social planner is zero.<sup>9</sup> Individuals poorer than average would like (progressive) taxes on capital: it is easily seen that for all  $w^i < w$ ,  $U^i(\tau)$  is single-peaked and reaches its maximum at the solution  $\tau^i$  to  $U'^i(\tau) = 0$ , or:

$$\varphi(\tau) \equiv \frac{\tau(1+\rho)}{(1+\rho(1-\tau))(1+\rho\beta(1-\tau))} = 1 - \frac{w^i}{w}. \quad (14)$$

This implies that agent  $i$ 's preferred tax rate  $\tau^i$  is positive and declines as his relative income falls, as in Meltzer and Richards (1981). The problem is more complicated for agents who are richer than average, as  $\varphi$  is not monotonic over all negative values of  $\tau$ . One can nonetheless show that  $U^i(\tau)$  remains strictly concave on its domain  $[\underline{\tau}, 1]$ , where  $\underline{\tau} < -1$  was defined earlier. Thus for  $w^i/w < 1 - \varphi(\underline{\tau})$  agent  $i$ 's preferred tax rate is still given by the first-order condition (14); for

<sup>9</sup> Note that  $V(\tau)$  is different from ex-ante efficiency  $E_{w^i}[U^i(\tau)]$ . Because agents are risk-averse, maximizing this "behind-the veil" criterion would always call for  $\tau > 0$ . By assuming that people vote only once they know their type I am intentionally abstracting from the insurance value of redistribution, in order to highlight its effects on the path of output. See Loury (1981) or Bénabou (1996a), (1995) for dynamic models incorporating this insurance motive.

still richer individuals it is the corner solution,  $\tau^i = \underline{\tau}$ .

To abstract from time consistency issues, I follow Persson and Tabellini (1994) in assuming that  $\tau$  is chosen irrevocably during the first-sub-period. If the political process can be represented by majority voting over the single issue of taxation the median voter is decisive, and his preferred policy is implemented. Hence the central implication of this class of models:<sup>10</sup>

### Proposition 1

*Let inequality be measured by, or correlated with, a lower ratio of median to mean pre-tax wealth. Then more inequality leads to more redistribution, slower growth and reduced efficiency:  $\tau$  rises, causing  $g$  and  $V$  to fall.*

These results on the personal distribution of wealth in a one-factor model can be extended to the functional distribution of income, say between capital and labor. Bertola (1993) and Alesina and Rodrik (1994) show more generally that a lower median-to-mean ratio in agents' relative endowments of accumulable and non-accumulable factors reduces the growth rate. The mechanism is the same, namely increased taxation of capital, and so are the efficiency implications. The fact that these are models with infinitely-lived agents also makes clear that the results are independent of the assumption of myopic preferences.<sup>11</sup> Note finally that what is key in Proposition 1 is how inequality affects the desire for redistribution of whoever is the decisive agent in the political system. Because the latter is likely to depart from the one person, one vote abstraction, I will later on consider variations in political institutions.

## 2.4 Empirical Evidence

I now turn to the evidence on the general idea that inequality increases redistribution, which in turn reduces growth. Column (4) of Table 2 reports the results of empirical studies which have examined the links between income inequality and a variety of measures of redistribution: share of transfers in GDP (either as a whole or decomposed into different categories such as welfare, unemployment, health and social security), average and marginal tax rates, or education expenditures. Column (5) then reports the effects which these policies were found to have on the economy's growth or investment rate. Some studies examine only one of these two relationships; others, most notably Perotti (1992), (1996) estimate a simultaneous system.

In column (4) the measure of inequality is generally taken to be the share of the third quintile, or the share of the third plus fourth quintiles. This reflects the emphasis placed by the models on the median voter and middle class. The results are rather disappointing: the effect of income

---

<sup>10</sup>The main difference between the present model and that of Persson and Tabellini (1994), apart from the fact that I allow  $\beta < 1$  and make explicit the functioning of the loan market ( $b^i$ 's and  $\bar{r}$ ), is the use of a geometric rather than linear redistribution scheme. While this modification will prove most convenient when considering missing credit markets, it is here inessential. All formulas derived above also obtain in the traditional, linear, case.

<sup>11</sup>The quantitative magnitude of the effect, on the other hand, does depend on how far into the future agents look when making savings decisions and casting their ballots; see Krusell, Quadrini and Ríos-Rull (1996).

distribution on transfers and taxes is rarely significant, and its sign varies from one study or even one specification to the other. The same is true for education expenditures. Both Alesina and Rodrik (1994) and Persson and Tabellini (1994) raise an objection to interpreting such regressions as a test of the theory. They point out that redistribution can take many forms other than straight transfers: progressive income taxation, minimum wage laws, trade and capital restrictions, the composition of government expenditure, patent legislation, regulation, and the protection of property rights. This is certainly true but not entirely persuasive. First, the studies reported in Table 2 already include some of these variables such as tax rates, public employment or education expenditures. Second, it is hard to see why the increased pressure for redistribution would manifest itself *only* through indirect channels and not through direct ones. Presumably the pivotal voter or class is equalizing the marginal costs and benefits or pressure along the different dimensions of redistribution. Finally, one would want to see at least some evidence that minimum wage laws, trade and capital restrictions or regulation are higher in more unequal countries. Such is not the case within the OECD: these indirect forms of redistribution are lowest in the United States, which has the greatest pre-tax inequality, and much more extensive in continental Europe (welfare state and labor legislation) or Japan (regulation and trade protection).

Column (5) provides even more surprising news. The coefficient in growth regressions of most of the transfers described above is positive, and frequently statistically significant. This is true not only for education expenditures, but also for social insurance and even pensions. These results stand in sharp contrast to those found for pure government consumption. Most studies which include this variable find it to be negatively related to growth, as in Barro (1990). Government consumption, however, is conceptually different from redistribution, at least on the expenditure side; I shall come back to this point later on. On the revenue side both are eventually financed through taxes, but even tax rates fail to have the predicted growth effect. Easterly and Rebello use twelve different average and marginal tax rates; the coefficients are always negative, but only one is significantly different from zero. Perotti (1996) finds that both average and marginal tax rates have a significantly positive effect on growth. One might be tempted to dismiss the whole set of results in column (5) as due to reverse causality, based on the plausible idea that the welfare state is a luxury good – a form of “Wagner’s law”. The fact that Perotti estimates simultaneous equation systems which allow for the level and growth rate of GDP to affect transfers and tax rates precludes such as an easy way out. Together, the results of columns (4) and (5) indicate that the median voter, complete markets model is missing some important factors which bear on both the causes and consequences of redistribution.

## 2.5 Democracies and Other Regimes

Another dimension of the political economy model which has generated substantial empirical effort concerns its differential implications across political regimes. Perhaps because some of these have not been formally modelled but simply argued on the basis of intuition, the literature manifests

some disagreement as to what exactly these implications are and how they can be tested. I shall therefore explicitly formalize departures from the “one person, one vote” ideal, in a manner which is quite specific but allows me to demonstrate some more general points.

Instead of the voter at the 50<sup>th</sup> percentile of the wealth distribution being decisive, let it be the agent or pressure group located at the  $p^{\text{th}}$  percentile. The case  $p > 1/2$  corresponds to a system biased against the poor due to a wealth-restricted voting franchise, unequal lobbying power, vote-buying or simply the fact that poor and less educated individuals have lower participation rates in elections, even in industrialized countries (e.g., Edsall (1984) or Conway (1991) for the United States). Conversely,  $p < 1/2$  corresponds to a populist bias reflecting either the ideology of a non-democratic leftist regime (“dictatorship of the proletariat”) or the bargaining power of powerful unions. To get simple expressions, assume moreover that the distribution of initial endowments in each generation is log-normal:

$$\ln w^i \sim \mathcal{N}(m, \Delta^2). \quad (15)$$

The pivotal voter or pressure group then corresponds to the wealth level  $w^*$  defined by  $\Phi((w^* - m)/\Delta) = p$ , where  $\Phi$  is the c.d.f. of a standard normal. Equivalently,  $w^* = m + \lambda\Delta$ , where  $\lambda \equiv \Phi^{-1}(p)$ . Regimes with positive values of  $\lambda$  could be called “elitist”, those with negative values “populist”. But since even Western democracies exhibit departures from  $\lambda = 0$ , I will use a milder terminology and simply speak of positive or negative “wealth bias” in the political institutions.<sup>12</sup> Substituting  $\ln(w^*/w) = \lambda\Delta - \Delta^2/2$  into (14) yields the following results:

## Proposition 2

- (1) *Growth is higher, the more biased against the poor is the political system:  $\partial g/\partial \lambda > 0$ .*
- (2) *Intertemporal efficiency is maximized when the individual setting the tax rate is the one with the average endowment  $w$ .*
- (3) *Achieving intertemporal efficiency requires more pro-wealth bias in a more unequal country:  $w^* = w$  for  $\lambda = \Delta/2$ .*

The very stark nature of these results may lead one to think that they reflect special assumptions rather than general insights. They do not. These are in fact robust features of most political economy capital taxation models with complete asset markets. This includes not only models where there is no useful role for the government (Persson and Tabellini (1994)) but also those where taxes or subsidies are required to correct an externality (Bertola (1993)) or to finance some productive public investment (Alesina and Rodrik (1994)). The underlying intuitions are indeed straightforward. With complete markets wealth heterogeneity has no efficiency consequences; if preferences are homothetic it does not even affect aggregates. To equalize the social

---

<sup>12</sup>See Verdier and Ales (1993) for a dynamic model of political elites, and Bénabou (1995a) for a discussion of alternative representations of the political system. I show there in particular that if political influence reflects absolute rather than relative wealth, with an elasticity  $\lambda$ , the pivotal agent has rank  $\lambda\Delta$  and log-wealth  $m + \lambda\Delta^2$ .

marginal costs and benefits of taxation one need therefore only look at averages, which means adopting the point of view of the individual with the “representative” endowment.<sup>13</sup> Complete markets also imply that aggregate consumption growth rises with the interest rate; in steady-state, the same applies to output growth. Maximizing growth thus means maximizing the rate of return on capital, thereby adopting the perspective of the individual with highest wealth relative to other sources of income. This includes his valuation of any complementarities, public inputs or regulations which might affect the private return.

The empirical literature seems to have somewhat shied away from the most direct testable implication, namely (1), focusing instead on a more indirect, second-order effect: the differential effects of inequality on redistribution and growth across political regimes. The claim, first made by Persson and Tabellini (1994) and taken up by Perotti (1992), (1996) and others, is that the negative impact of inequality discussed in Proposition 1 should be stronger in “democracies” rather than “non-democracies”. While it does relate to some valid intuitions, there are two problems with this argument.

(a) I know of no formal proof or model. Upon reflection, this is probably no accident. Such claims concern the *cross-partial second derivative* of the growth rate, with respect to income distribution and regime. Not only is it rare when cross-partial derivatives can be signed unambiguously, but in this instance the growth rate (or equivalently, the tax rate) is determined by the first-order condition of the decisive agent –whether a dictator, landowning elite or powerful union. Its cross-partial thus inevitably involves the *third* derivative of that agent or group’s objective function, and therefore cannot be signed without restrictive functional assumptions.<sup>14</sup>

(b) Whatever its sign, any differential impact of income distribution across political regimes must surely depend on the direction in which one moves away from pure democracy. A similar criticism was formulated by Alesina and Rodrik (1994) who pointed out that even dictators are subject to political pressure, whether through riots or bribes. What should matter is how inequality affects the desire for redistribution of their constituency.

Elaborating on these remarks, I shall now formally examine the cross-effects of political institutions and inequality,  $\lambda$  and  $\Delta$ , on redistribution and growth. When voters or constituents located at the  $p^{th}$  wealth percentile are decisive, the equilibrium tax rate is given by their first-order condition:  $\ln(1 - \varphi(\tau)) = \lambda\Delta - \Delta^2/2$ , where  $\varphi$  was defined in (14). This implies

$$\frac{\partial \tau}{\partial \Delta} = (\Delta - \lambda) \left( \frac{1 - \varphi(\tau)}{\varphi'(\tau)} \right) \quad \text{and} \quad \frac{\partial \tau}{\partial \lambda} = -\Delta \left( \frac{1 - \varphi(\tau)}{\varphi'(\tau)} \right), \quad (16)$$

<sup>13</sup>This reasoning implicitly assumes linear taxation, as in all the literature under discussion. My model yields the same result even with progressive taxation because in equilibrium everyone invests the same amount (all that matters is the threat of taxation).

<sup>14</sup>Saint-Paul and Verdier (1993) and Persson and Tabellini in the working paper version of their 1994 article show how restricting the political rights of different classes affects growth (an effect similar to  $\partial g/\partial \lambda$  here). But neither paper provides results on how the political regime alters the impact of inequality on redistribution and growth ( $\partial^2 \tau/\partial \lambda \partial \Delta$  and  $\partial^2 g/\partial \lambda \partial \Delta$  in my model), which is the issue of contention here.

hence

$$\frac{\partial^2 \tau}{\partial \lambda \partial \Delta} = \left( \frac{1 - \varphi(\tau)}{\varphi'(\tau)} \right) \left( -1 + (\Delta - \lambda) \Delta \left( \frac{\varphi'(\tau)^2 + \varphi''(\tau)(1 - \varphi(\tau))}{\varphi'(\tau)^2} \right) \right). \quad (17)$$

The first term in the second bracket is negative, but the second one depends on the curvature of  $\ln(1 - \varphi)$ .<sup>15</sup> This is a complicated expression, even for  $\lambda$  close to zero. One could of course evaluate it numerically. Alternatively, observe that when  $\Delta$  is small, the second term is dominated by the first. Together with similar derivations for the growth rate, this allows me to prove an amended version of the “democracy and inequality” conjecture:

### Proposition 3

*As long as inequality is not too large, its positive effect on redistribution and negative effect on growth are weaker, the less favorable to the poor is the political system:  $\partial^2 \tau / \partial \lambda \partial \Delta < 0$  and  $\partial^2 g / \partial \lambda \partial \Delta > 0$ .*

Compared to democracies, inequality should have *more* impact on redistribution and growth in left-wing, populist regimes and less impact *only* in right-wing or wealth-biased systems.<sup>16</sup> Note also from (16) that in this latter case taxes first decline, then rise with inequality:  $\partial \tau / \partial \Delta \leq 0$  as  $\Delta \leq \lambda$ . Distributional conflict initially shifts policy to inefficiently low levels of taxation, but the growing skewness of the income distribution eventually tilts the balance of power back towards the poor. The regressive nature of taxes or capital subsidies in this range ( $\tau < 0$  for  $\Delta \leq 2\lambda$ ) need not be taken literally, because one could incorporate some insurance value or public good which ensured that  $\tau$  remained positive. What is important, given the lack of any empirical link between inequality and redistribution evident in Table 2, is the non-monotonicity of  $\tau$  with respect to  $\Delta$ . The possibility of redistribution decreasing with inequality will play a still larger role once combined with imperfect capital markets in Section 3, and will therefore be discussed in more detail at that point.

Propositions 2 and 3 make clear that what matters for growth (both directly and through the impact of income distribution) is not the just the degree to which the political system departs from perfect democracy, which corresponds to  $|\lambda|$ , but also whose political rights or influence are being curtailed. The empirical literature has nonetheless followed Persson and Tabellini (1994) in testing for differential effects of absolute democracy. Typically, countries are classified into “democracies” and “non-democracies” on the basis of some index of political rights, freedom of expression and association, etc., such as those of Gastil (1982). The growth regression is then run separately for each subsample, or for the whole sample with a “democracy” dummy variable included –both by

<sup>15</sup>The expression multiplying  $(\Delta - \lambda) \Delta$  measures the concavity of the optimal tax rate with respect to the log-income of the tax-setter; it reflects the third derivative of preferences, as argued earlier. Note that by using the first-order condition to characterize the tax rate preferred by agents with  $w^i/w = e^{\lambda\Delta - \Delta^2/2}$ , I am implicitly assuming that it is always interior. A simple sufficient condition is  $e^{\lambda^2/2} < 1 - \varphi(\tau)$ .

<sup>16</sup>The earlier caveat about the inevitable reliance of *any* result on such cross-derivatives upon specific functional and distributional assumptions naturally applies here as well.

itself and interacted with income distribution. The results of these tests, summarized in columns (2) and (3) of Table 2, have generally been negative. First, there is no consistent direct effect of democracy on growth (recall that this class of models predicts  $\partial g/\partial \lambda > 0$ ). The latest study, Barro (1996), finds a non-linear effect which I shall revisit in the next section. Turning to column (3), most other studies have not reproduced Persson and Tabellini's (1992), (1994) finding of a statistically significant differential impact of inequality across democracies and non-democracies. Moreover, the sign of the estimates varies from one study to another. But since what is really measured in such regressions is the effect of  $|\lambda|$  and  $|\lambda|\Delta$  rather than  $\lambda$  and  $\lambda\Delta$ , I see these results as less damaging for the theory than those of the direct tests discussed in Section 2.4. It should be feasible in future work to provide more discriminating tests by distinguishing between left- and right-wing regimes, at least for some subset of countries.<sup>17</sup>

## 2.6 Consumption Taxes and Labor Supply

One potential explanation for the fact that redistribution does not appear to depress investment or growth is the presence of credit constraints, as I will show later on. But even with complete markets this need not happen: nothing in principle prevents progressive taxation, transfers, minimum wages and the like from being combined with a consumption tax or investment subsidy to maintain accumulation at the appropriate level. Indeed, value-added and sales taxes account for a much higher share of government revenue in "welfare state" European countries than, say, in the US or Japan.

Going back to the model, suppose that a proportional consumption tax is introduced and its revenue used to subsidize investment. First period consumption is now  $c^i = (1 - t)(w^i + b^i - k^i)$ , investment  $(1 + a)k^i$  and second period pre-tax income  $y^i = [r((1 + a)k^i)^\beta (w)^{1-\beta}]^{1-\tau} (\tilde{y})^\tau$ . Balanced budgets require  $t(w - k) = ak$ , where letters without superscripts still denote aggregates. It is easily verified that, for any rate of redistribution  $\tau$ :

1. Agents still save the same fraction  $s(\tau)$  of their income, so that the investment rate is  $s(\tau)(1 + a)$ .
2. The consumption tax rate *unanimously* preferred by agents is  $t^*(\tau) = \rho\beta\tau/(1 + \rho\beta\tau)$ . Thus the more progressive the income tax, the greater the recourse to consumption taxes.
3. The resulting investment rate never departs from its *first best* level:  $s(\tau)(1 + a) = \rho\beta/(1 + \rho\beta) = s(0)$ .

---

<sup>17</sup>One case which does not suffer from this problem is the historical sample of Persson and Tabellini (1994), where political institutions are represented by the fraction of the population which did not have access to an income-based or wealth-based voting franchise. This corresponds closely to  $\lambda$  itself, and is therefore a more appropriate regressor than any post-war democracy index. The resulting estimate of  $\partial g/\partial \lambda$  is nonetheless negative, although not significant. The cross-effect is not estimated in this set of regressions.

4. Introducing elastic labor supply, by replacing first-period income  $w^i$  with  $w^i l^i$  and preferences (1) with  $U = \ln c - \delta l + \rho \ln d$ , leaves these results unchanged. As to the growth rate, it becomes  $g(\tau) = \ln \tau + \beta \ln s(0) + \ln l(\tau)$ , where  $l(\tau) = (1 + \rho\beta(1 - \tau))/\delta$  is agents' common labor supply.<sup>18</sup>

This simple extension of the model makes clear that it is only when redistributive policy is restricted to capital taxation that increases in redistribution automatically translate into intertemporal distortions, thereby generating results like Proposition 1. A related conclusion emerges from Bertola's (1993) two factor model, where labor supply is fixed. He shows that when redistribution takes the form of investment subsidies rather than factor taxation, wealthy agents want less of it than poor ones, as it depresses the value of the existing capital stock. Therefore the poorer the median voter relative to the mean, the *higher* the growth rate. Of course when labor supply is elastic, redistribution depresses effort and this decline in  $l(\tau)$  still lowers the growth rate. Note, however, that the magnitude of this effect is smaller than that of  $s(\tau)$ —even with linear disutility of labor, which puts an upper bound on distortions to labor supply.

## 2.7 What of Land Reform and Public Education?

Formalizing the broad idea of a link between distributive conflict and growth is an important contribution of the pure political economy models discussed above (e.g., Bertola (1993), Alesina and Rodrik (1994), Persson and Tabellini (1994)). The specific mechanism on which they focus, however, appears somewhat at odds with some of the motivating issues and evidence. For instance when interpreting their empirical results, Alesina and Rodrik (1994) refer to the idea that land reform was an important factor in the growth performance of Japan, Korea and other Asian countries, especially when compared to Latin America. Similar claims are often made concerning the role of public investment in primary and secondary education, which mostly benefits the poor and lower middle class. This means that if Latin American or African countries did not grow as fast as Asian ones it is—among other things—because they failed to redistribute, at least initially (land) and perhaps also along the way (education).<sup>19</sup> Yet the *only* transfers allowed in this class of models are those detrimental to growth. Public schooling and any kind of redistribution affecting people's *ability to invest* are just not feasible under complete asset markets. Land reform and other reallocations of property rights which could reduce future distributional conflict are only envisioned as off-the-equilibrium path events, in the form of one-time, unanticipated capital levies. But if one-time redistributions are so beneficial they should occur systematically at high levels of inequality (Brasil and South Africa come to mind). To avoid this time-inconsistency one needs to explain what is impeding such transfers, and perhaps why this constraint may be relaxed

---

<sup>18</sup>Similar results obtain when labor is less than infinitely elastic, but the formula for aggregate labor supply is now more complicated because agents choose different levels of effort  $l^i(\tau)$ .

<sup>19</sup>In 1945 the adult literacy rate in South Korea was only 22%; thirty years later, it was over 90%.

following a war or major crisis. More generally, countries face a tradeoff between the benefits of redistribution and its incentive or reputation costs; both aspects should be explicitly incorporated into the politico-economic game. This is the task to which I now turn.

For redistribution to potentially have any growth benefit, in addition to costs, it must of course take place before investment is completed. But altering the timing of transfers is not sufficient: the reader will easily verify that none of the earlier results change if the government redistributes investment expenditures (or capital itself but prior to production) according to the scheme:

$$\hat{k}^i = (k^i)^{1-\tau}(\tilde{k})^\tau \quad (3')$$

where  $\tilde{k}$  is defined by the budget constraint

$$\int_0^1 (k^i)^{1-\tau}(\tilde{k})^\tau di = \int_0^1 k^i di \quad (4')$$

and second-period income is left untaxed:  $\hat{y}^i = y^i = r(k^i)^\beta(w)^{1-\beta}$ . This policy corresponds for instance to a land reform (redistributing plots of size  $k^i$ ), but one which is anticipated by agents. Equivalently, one can think of redistributing education budgets across rich and poor school districts, as done implicitly in countries which have a national system of education finance, and explicitly in an increasing number of U.S. States in response to constitutional court mandates. Finally, if one thinks of physical capital, redistributing investment expenditures is equivalent to taxing large investments and subsidizing small ones: agents investing  $k^i$  face a price  $p(k^i) = (k^i/\tilde{k})^{\tau/(1-\tau)}$ .<sup>20</sup> But whatever the context may be, as long as agents are not liquidity-constrained such redistributions –or the threat thereof– can only distort first-best investment decisions.

### 3 Growth and Political Economy with Imperfect Asset Markets

#### 3.1 Direct Effects of Inequality and Redistribution

Let me now consider the very same model as above, but with missing credit markets: agents have no possibility of borrowing from one another, for instance due to a severe moral hazard problem (there is no way to enforce repayment). Later on I shall relate the model to others which make credit frictions explicit. I naturally focus on the case where redistribution occurs in time to affect the resources available for investment, namely (3')–(4'). With no access to a loan market, agent

---

<sup>20</sup>It might seem more natural to simply tax first-period wealth  $w^i$ . But since it is here exogenously determined for each generation this would be equivalent to an unanticipated capital levy. Formalizing redistribution as affecting investment expenditures enables me to capture its disincentive effects while simultaneously allowing it to occur in time to affect potential liquidity constraints. Straightforward income taxation with dynastic or infinitely-lived agents leads to very similar expressions (Bénabou (1995a)).

$i$ 's optimization problem is the same as (5) before, except that  $b^i$  is now constrained to zero:

$$\max_{k^i} \{ \ln(w^i - k^i) + \rho \ln(\tau [(k^i)^{1-\tau} (\tilde{k})^\tau]^\beta w^{1-\beta}) \}. \quad (5')$$

The optimal investment is therefore:

$$k^i = \frac{\rho\beta(1-\tau)w^i}{1 + \rho\beta(1-\tau)} = s(\tau)w^i. \quad (18)$$

Whereas everyone previously invested a fraction  $s$  of *aggregate* income, they now invest that same fraction of their *own* income. The government's balanced budget condition (4') then defines  $\tilde{k}$  to be  $\tilde{k} = s\tilde{w}$ , with

$$(\tilde{w}/w)^\tau \equiv w^{1-\tau} / E[(w^i)^{1-\tau}]. \quad (19)$$

This yields second-period income for agent  $i$ :

$$y^i = \tau s^\beta (w^i)^{\beta(1-\tau)} \tilde{w}^{\beta\tau} w^{1-\beta}. \quad (20)$$

Summing over all agents, the growth rate of aggregate income is given by:

$$g(\tau) \equiv \ln(y/w) = \ln \tau + \beta \ln s(\tau) - \ln \left( E[(w^i)^{1-\tau}]^\beta / E[(w^i)^{\beta(1-\tau)}] \right), \quad (21)$$

where  $\tau$  is the tax rate chosen by the current generation. With dynamic linkages still specified as in Section 2.2, this is also the intergenerational growth rate:  $g(\tau_t) = \ln(w_{t+1}/w_t) = \ln(y_t/y_{t-1})$ . The first two terms represent the growth rate with frictionless loan markets, derived in (9). The last term is negative for  $\beta < 1$  due to Jensen's inequality, and vanishes for  $\beta = 1$ . A more unequal distribution of resources, in the usual sense of a mean-preserving spread in  $w^i$ , tends to increase this loss and thereby slow down growth. A particularly simple expression obtains when endowments are log-normally distributed,  $\ln w^i \sim \mathcal{N}(m, \Delta^2)$ , which I will assume from now on:

$$g(\tau) = \ln \tau + \beta \ln s(\tau) - \beta(1-\beta)(1-\tau)^2 \Delta^2/2. \quad (22)$$

The intuition is straightforward. With  $\beta < 1$  individuals face decreasing returns to investment. Reallocating funds from rich to poor relaxes the latter's credit constraints, allowing them to earn a higher return. Because the government is in no better position than potential financial intermediaries to monitor borrowers and ensure that they repay, this more efficient allocation of investment can only be achieved together with a net transfer to the poor. This in turn depresses savings incentives, as in the case of perfect capital markets. The same two effects are present when evaluating intertemporal efficiency: the objective function of a planner with no distributional concern now equals

$$W(\tau) \equiv \ln c + \rho \ln d = V(\tau) - \rho\beta(1-\beta)(1-\tau)^2 \Delta^2/2, \quad (23)$$

where  $V(\tau)$  corresponds to the complete markets case.

#### Proposition 4

*(1) Under any given policy  $\tau$ , inequality reduces growth and intertemporal efficiency. This loss decreases with the extent of pre-investment redistribution.*

*(2) Growth is hill-shaped with respect to redistribution, and the growth-maximizing tax rate increases with inequality.*

These results differ somewhat from those of Galor and Zeira (1993), Banerjee and Newman (1993), Perotti (1993) or Aghion and Bolton (1996). In these models the effects of inequality and redistribution on growth depend critically on the initial distribution of wealth. This is because investment involves a minimum project size, generating a threshold level of wealth below which agents do not invest, or do not leave enough to their offspring for them to invest. Increasing growth means maximizing the number of people above this threshold, and in the early stages of development this may require concentrating resources on a lucky few or on some emerging middle class. This non-convexity was also originally seen as a key ingredient in models with explicit credit rationing, especially in generating multiple long-run equilibria (poverty traps).<sup>21</sup> To demonstrate that multiplicity can arise solely through the feedback from distribution to factor prices (see Section 3.6), Piketty (1996) does away with indivisibility by simply augmenting a standard Solow model with moral hazard. His neoclassical, concave technology again implies that inequality is always harmful and redistribution towards the poor beneficial —abstracting of course from incentive effects, which he does. Thus Piketty's model, incorporating the micro-foundations underlying the credit rationing, validates the similar result obtained by simply closing down the loans market, as I have done here following Loury (1981), Tamura (1991) and Bénabou (1996a).

### 3.2 Empirical Evidence

Because asset market incompleteness is difficult to measure, there has been little work testing its growth implications directly.<sup>22</sup> Columns (7) and (8) in Table 2 show the results from Perotti (1994), who uses the loan-to-value ratio for mortgages as an indicator of credit availability and finds it to affect the investment rate positively and significantly. Moreover this effect is stronger where the share of the bottom two quintiles is lower (that coefficient is significant only at the 10% level). Similarly, the negative effect of inequality on investment rises with credit frictions, as predicted by the model.

---

<sup>21</sup> Another potential appeal is its ability to generate Kuznets-type dynamics, as the range of increasing returns initially causes some measures of inequality to increase. Eventually, inequality decreases as aggregate growth pulls everyone out of the poverty trap through a decline in the scarcity of capital (Aghion and Bolton (1996)) or through some positive spillovers from the better educated segments of society (Galor and Tsiddon (1993)).

<sup>22</sup> There is of course an empirical literature linking financial development and growth, but it does not directly connect to the personal or functional distribution of income and the consequences of redistributive policies, which represent our main concerns. More closely related is the literature devoted to credit constraints on the investment of small firms, e.g., Bernanke and Gertler (1989).

Evidence of a more indirect nature can also be considered. For instance, if liquidity constraints are impeding investment by the poor or lower middle-class, *any* form of progressive transfer will contribute to relaxing them. By contrast, government consumption will not, except perhaps for civil servants. If anything, the taxes needed to finance it will aggravate wealth constraints. This differential effect is remarkably consistent with the empirical evidence discussed earlier: the contribution of transfers to growth is generally positive, that of government consumption generally negative. The effects of public education expenditures also deserve attention. Column (5) of Table 2 shows that these generally promote growth. In the absence of credit constraints, increases in public spending on education would be offset by decreases in private spending (unless the two are complements), and therefore would not contribute to growth. To the extent that they require higher levels of distortionary taxation they would even have a negative effect, as in Becker and Tomes (1979).<sup>23</sup>

The development literature also provides some valuable evidence, even if it does not directly relate to aggregate growth. First, it appears to be a robust finding that credit rationing and the lack of adequate insurance significantly constrain the investments of poor farmers away from profit-maximizing levels and compositions. See for instance Rosenzweig and Wolpin (1993), or the comprehensive surveys by Binswanger and Deininger (1995) and Binswanger and Rosenzweig (1995). Informal risk-sharing and consumption-smoothing arrangements play an important role but are only limited substitutes for well-functioning markets (Townsend (1995)). This literature also provides evidence of the decreasing returns which are central to the resulting inefficiency: whether in developing or developed countries, the family farm is usually the most efficient unit of production. Inadequate access by peasants to credit for buying land, agricultural inputs or equipment is one of the main causes of concentration in land ownership; others are the favorable tax treatment of agricultural income and various subsidies which benefit wealthy farmers most. Diminishing returns also occur in education, according to most empirical estimates of the private and social returns to primary, secondary and tertiary schooling (e.g., Psacharopoulos (1993)).

### 3.3 Political Economy with Missing Asset Markets

Let us now examine how inequality affects the efficient and equilibrium amounts of redistribution and the corresponding growth rates, by studying the interactions of political economy and imperfect asset markets. I continue to focus on the case where first period endowments are log-normal

---

<sup>23</sup> An interesting issue is the degree to which government transfers are actually progressive. Alesina (1995) shows that a significant fraction of health and especially education expenditures is often captured by the middle 40%, with even the top 20% appropriating a large share in some Latin American countries. The size of transfers is thus not as good an indicator of social equity as might appear. With respect to growth, however, these observations are not inconsistent with the idea that redistribution tends to relax wealth constraints on investment. First, even Alesina's data suggests that a significant amount of resources generally reaches the bottom 40%. Moreover, the key issue is what segment of the population has the highest marginal return on investment; in a poor country this could easily include parts of the middle class, as observed earlier when discussing the effects of fixed costs. Naturally, institutional reforms which could minimize transfers to the truly well-off would always be beneficial.

and to omit time subscripts. Using (20), individual  $i$ 's welfare is equal to:

$$U^i(\tau) = V(\tau) - (1 + \rho\beta(1 - \tau)^2)\Delta^2/2 + (1 + \rho\beta(1 - \tau))(\ln w^i - m). \quad (24)$$

It can be shown that  $U^i$  is strictly concave over all  $\tau \geq \underline{\tau}$ , where  $\underline{\tau} = -(1 + 1/\rho\beta)^{1/2} < -1$  was taken to be the most regressive tax feasible. Except for the richest individuals who prefer this corner solution ( $\tau^i = \underline{\tau}$  for  $\ln w^i - m > (\rho\beta)^{-1}V'(\underline{\tau}) + (1 - \underline{\tau})\Delta^2 > 2\Delta^2$ ), each agent's ideal tax rate  $\tau^i$  is then given by the first-order condition  $U^i(\tau) = 0$ ; it decreases with own income  $\ln w^i$  and increases with inequality  $\Delta$ . Agents richer than  $\ln w^i = m + \Delta^2 = \ln w + \Delta^2/2$  favor regressive taxes while all others, including the individual with the average endowment  $w$ , are for progressivity. To compare these private optima to the social one, let us next rewrite:

$$U^i(\tau) = W(\tau) - \rho\beta^2(1 - \tau)^2\Delta^2/2 + (1 + \rho\beta(1 - \tau))(\ln w^i - m). \quad (25)$$

This leads to the following results.

#### Proposition 5

(1) A social planner concerned only with intertemporal efficiency  $W$  would set a positive tax rate  $\tau^P(\Delta)$ , increasing in  $\Delta$ .

(2) The preferred tax rate of the agent with median wealth is above this efficient level.

(3) The preferred tax rate of the agent with average wealth is below the efficient level if

$(1 - \beta)\Delta^2/4\beta > (2\beta - 1)/(2 + \rho)$ , and above in the reverse case.

The first result is intuitive, given that redistribution now has not only efficiency costs through  $s(\tau)$  but also efficiency benefits, proportional to  $\Delta^2$ . The second one has a familiar ring, although it should be noted that the median voter's preferred policy now maximizes *ex-ante* or aggregate welfare,  $E_{w^i}[U^i(\tau)]$ . Most striking is the third result. In the previous class of models, the agent with average wealth always chose the efficient policy. Now there are two opposite forces which make him "un-representative". The first is the *progressivity* of the tax scheme, whereby for  $\tau > 0$  all those below  $\tilde{w} > w$  gain at the expense of those above. Previously, redistribution was a threat which discouraged savings but did not materialize in equilibrium, as everyone invested the same amount. With liquidity constraints investments reflect wealth levels, allowing the average individual to join in "soaking" the very rich (differential productivities would yield a similar result). The opposite effect, which tends to make him choose too little redistribution, is his *failure to internalize* the social losses from liquidity constraints, which are proportional to  $(1 - \beta)\Delta^2$ . As seen in (10), the missing credit market does not affect him at all because he does not want to borrow or lend. In this sense he is the *least representative* of all agents, as both richer and especially poorer ones are affected. When  $(1 - \beta)\Delta^2$  is large enough or simply when  $\beta < 1/2$  this effect dominates the progressivity bias and the average agent chooses an inefficiently low tax rate,

possibly even *less* efficient than that of the median voter.<sup>24</sup> Democracy has its virtues, after all.

To explore in more detail the effects of political institutions and inequality on redistribution, let the pivotal agent or group correspond once again to the  $p^{th}$  percentile, with  $p = \Phi(\lambda)$ . Substituting  $\ln w^i = m + \lambda\Delta$  into (11) and setting  $U^i(\tau) = 0$  leads to the following results:

### Proposition 6

(1) *Growth and intertemporal efficiency are hill-shaped with respect to the degree of wealth bias in the political system,  $\lambda$ . They are maximized at  $\lambda_G > \lambda_W > 0$ .*

(2) *If the pivotal agent is at or below the median ( $\lambda \leq 0$ ), taxes increase with inequality; such increases are always inefficient and reduce growth.*

(3) *If the pivotal agent is of higher rank than the median ( $\lambda > 0$ ), taxes are U-shaped with respect to inequality. Tax cuts induced by rising inequality are always inefficient and can also reduce growth if  $\beta$  is not too close to 1. Tax increases induced by rising inequality have the opposite effects up to some point, beyond which their contribution to intertemporal efficiency and growth becomes negative.*

The first result provides a potential explanation for Barro's (1996) finding that growth is maximized at some intermediate level of democracy, and more generally for the absence of any monotonic relationship between the two. See Table 2, column (2), with the usual caveat about measuring  $|\lambda|$  rather than  $\lambda$ . Note that asset market incompleteness and tax distortions are both critical for this non-linear effect of the political system on the growth rate. In models such as Alesina and Rodrik (1994) or Bertola (1994) the presence of public investment or spillovers does not change the fact that restricting the political rights of the poor *always* increases growth, as explained following Proposition 2. A polar opposite is Saint-Paul and Verdier's (1993) model where non-distortionary taxes are used to finance public education, supplementing wealth-constrained private expenditures. As a result, growth increases monotonically with redistribution, and therefore with  $-\lambda$ . The model presented here can be seen as combining both forces, resulting in an interior solution for the growth-maximizing political system.

As far as equilibrium taxes and *welfare* are concerned, results (2) and (3) also apply in the complete markets case.<sup>25</sup> The difference is only one of degree: because the optimal level of  $\tau$  now increases with  $\Delta$ , inequality-induced tax increases are now less inefficient and tax decreases more inefficient. With respect to *growth*, however, the losses from credit constraints lead to a new and potentially important explanation for its negative correlation with inequality: Proposition 6 shows that more inequality can lead to *less* redistribution, which in turn leads to *less* growth (see also Bénabou (1995a)).

<sup>24</sup>If  $\beta < 1/4$  this is always the case for  $\Delta$  small enough; see the proof of Proposition 5.

<sup>25</sup>This can be shown directly using (14) and (13) or one can simply set  $\beta = 1$  in (25), in which case credit constraints entail no efficiency loss. Both the complete and missing markets cases accord with the claim in Bénabou (1995a) that it is when redistribution is ex-ante efficient that the standard effect  $\partial\tau/\partial\Delta > 0$  can be reversed. In that model such a reversal can occur with progressive taxes  $\tau > 0$ .

### 3.4 Direct and Indirect Effects of Inequality

We saw earlier that for a given policy  $\tau$ , inequality reduces growth and intertemporal efficiency. Does this remain true when policy responds endogenously? A straightforward application of the envelope theorem to (23) shows that such is the case when redistribution is determined by an efficient planner, i.e.  $\tau = \tau^P(\Delta) \equiv \arg \max \{W\}$ . The more interesting case is when it arises through the political process. Let  $\tau^*(\Delta, \lambda)$  denote this solution to the first-order condition  $U''(\tau) = 0$  of the pivotal group:

$$(\rho\beta)^{-1}V'(\tau) + (1 - \tau)\Delta^2 - \lambda\Delta = 0. \quad (26)$$

We can write

$$\frac{dg}{d\Delta} = \frac{\partial g}{\partial \Delta} + \left( \frac{\partial g}{\partial \tau} \right)_{\tau=\tau^*} \times \left( \frac{\partial \tau^*}{\partial \Delta} \right),$$

and similarly for  $W$ . The first term, which holds policy fixed, is always negative, while the last two were analyzed in Proposition 6. In a pure democracy or a populist system ( $\lambda \leq 0$ ) taxes increases with inequality; moreover, the equilibrium rate is always above  $\tau^P$ , hence in the range where  $W$  and a fortiori  $g$  decline with redistribution. With positive wealth bias, on the other hand, taxes are  $U$ -shaped with respect to inequality. Lower taxes have a positive effect on savings incentives but a detrimental one on credit constraints. Whether growth rises or falls thus depends on the extent to which agents face decreasing returns and have unequal resources. The effects on  $W$  are similar, except that since it puts positive weight on first-period consumption it does not “overvalue” low taxes as  $g$  does. In particular, Proposition 6 established that where rising inequality leads to tax cuts, these are never efficient.

#### Proposition 7

- (1) *If the pivotal agent is at or below the median ( $\lambda \leq 0$ ), inequality reduces both growth and intertemporal efficiency.*
- (2) *If the pivotal agent is of higher rank than the median ( $\lambda > 0$ ), inequality still lowers  $W$  as long as  $\beta$  is not too large. For  $\beta$  close to 1 inequality may increase intertemporal efficiency, but this can only occur through an increase in redistribution.*
- (3) *For  $\lambda > 0$  inequality reduces growth if  $\Delta$  is not too small and  $\beta$  not too close to 1. In the opposite case, inequality increases growth through a decrease in tax rates.*

While the direct effect of inequality on growth (through wealth constraints) often dominates, its interaction with the indirect one (policy response) can generate a fairly complex relationship between these variables. Sorting out the two channels through a structural model should nonetheless be possible, and even appears desirable. Particularly interesting is the third result, which calls attention to the fact that if political power is sufficiently correlated with financial or human wealth and if capital market imperfections are not too severe, greater inequality may actually

increase growth through a decline in redistribution. A similar insight will arise from the model of social conflict discussed in Section 4 : what matters for sociopolitical stability (which promotes growth) is not income inequality per se, but the inequality of income *relative to* the inequality of political power. Where institutions or the political technology are such that increases in the former translate into increases in the latter, instability (or here, taxation) need not rise, and may actually fall.

### 3.5 Endogenous Income Distributions and Social Mobility

Up to this point I have maintained a feature inherited from the complete markets models discussed in Section 2: even though it plays a critical role, the pre-tax distribution of endowments is completely exogenous. In Persson and Tabellini (1994) as here it simply reflects, period after period, the fixed distribution of agents' abilities. In Bertola (1993) and Alesina and Rodrik (1994) the distribution of wealth is indeterminate: the constancy of factor prices combines with the absence of idiosyncratic uncertainty to ensure that any initial distribution will reproduce itself from one period to the next. By contrast, the endogenous evolution of the wealth distribution is at the core of incomplete markets models. This includes those which focus on the decentralized market outcome (Banerjee and Newman (1993), Galor and Zeira (1993), Aghion and Bolton (1996), Piketty (1996)), those which analyze potential redistributive policies (Loury (1981), Glomm and Ravikumar (1992), Bénabou (1996a), Fernandez and Rogerson (1994)) and those which endogenize redistribution as a political outcome (Perotti (1993), Saint-Paul and Verdier (1993), Verdier and Ales (1993), Saint-Paul (1994), Bénabou (1995a)). Making these distributional dynamics explicit is also necessary to address the issue of *social mobility*. The intergenerational persistence of income and wealth disparities is at least as relevant a social concern as the level of inequality per se. The question of how policy affects persistence, both in theory and in practice, is therefore of considerable interest.

The model developed in this paper is easily amended to deal with these issues. Until now it was assumed that intergenerational linkages operated only at the aggregate level. Whether through an economy-wide spillover or through universal public education, they resulted in a transmission mechanism of the form:  $w_{t+1}^i = \epsilon_{t+1}^i y_t$ . Let me now recognize that the transmission of human capital operates largely within the family or at the level of small communities which are fairly homogenous in socioeconomic status: school districts, neighborhoods, social networks, etc. Consider first the pure human capital spillover interpretation, under which  $w_{t+1}^i = \epsilon_{t+1}^i y_t$  is simply replaced by:

$$w_{t+1}^i = \epsilon_{t+1}^i y_t^i. \quad (8'')$$

Under perfect capital markets this changes nothing, because  $y_t^i = r(s_t)^\beta w_t$  for all agents. Family income remains i.i.d., with variance  $\Delta^2 = \text{Var}[\ln \epsilon] \equiv s^2$ . With imperfect capital markets, on the other hand, equation (20) now implies that family income follows the law of motion:

$$\ln(w_{t+1}^i) = \ln \epsilon_{t+1}^i + \ln r + \beta \ln s(\tau_t) + \beta(1 - \tau_t) \ln w_t^i + \beta \tau_t \ln \tilde{w}_t + (1 - \beta) \ln w_t \quad (27)$$

where  $\tau_t$  is the tax rate chosen in period  $t$  and  $\tilde{w}_t$  is given by (19). Income remains log-normally distributed, and its intergenerational persistence is given by  $\beta(1 - \tau_t)$ . As a result, inequality follows the autoregressive process:

$$\Delta_{t+1}^2 = s^2 + \beta^2(1 - \tau_t)^2 \Delta_t^2. \quad (28)$$

The alternative mechanism of intergenerational transmission is educational investment, now privately or locally funded. With the preferences (1'), agents all want to invest the same fraction  $1 - \gamma$  of second period resources in their child's human capital. Therefore  $w_{t+1}^i = \epsilon_{t+1}^i d_t^i$ , where the constant has again been normalized to one. In the absence of credit markets,  $d_t^i = y_t^i$  so (27) and (28) remain unchanged. Complete markets lead to somewhat different individual dynamics, but these retain the property that redistribution reduces the persistence of  $w_t^i$ . From here on I shall focus on the first case, since it has implications for aggregate growth whereas the second one does not. Recall in particular that

$$g(\tau_t) = \ln r + \beta \ln s(\tau_t) - \beta(1 - \beta)(1 - \tau_t)^2 \Delta_t^2/2.$$

Together with (28) this makes clear that once the dynamics of income distribution are endogenized, an additional, *long-run benefit* of redistribution appears. A marginal rise in  $\tau_t$  causes the same distortion  $s'(\tau_t)$  as before, but now the growth losses due to the combination of wealth inequality and borrowing constraints are reduced not only in the current period ( $(1 - \tau_t)^2 \Delta_t^2/2$  is lower), but also in *all future periods*:  $\Delta_{t+k}^2$  is lower, for all  $k$ . This homogenizing effect means that long-run growth is maximized at a higher level of redistribution than previously.<sup>26</sup> This, in turn, requires a lower degree of wealth bias  $\lambda$  in the political system, which must therefore be closer to pure democracy.

This brings me to the question of the long-run behavior of the economy and its income distribution, which I will discuss both theoretically and empirically. As before, let taxes be determined in each generation by the preferred policy of the pivotal group:  $\tau_t = \tau^*(\Delta_t, \lambda)$ , defined by (26). Together,  $\tau_t$  and  $\Delta_t$  then determine next period's distribution, according to (28). Recall now from Proposition 7 that when  $\lambda \leq 0$ ,  $\tau^*(\Delta_t, \lambda)$  is increasing in  $\Delta_t$ , so that endogenous redistribution acts as a stabilizing feedback on inequality.

---

<sup>26</sup> Agents' myopic preferences (1) or (1') imply that they do not internalize this variance effect when setting the current tax rate, just as they do not internalize the mean effect of  $y_t$  on  $w_{t+1}$ . With forward-looking altruism or infinite lives both are internalized to an extent which reflects the discount factor; see Bénabou (1996a).

### Proposition 8

If the median voter or some agent located at a lower percentile chooses the tax rate in every period ( $\lambda \leq 0$ ), the economy converges to a unique steady-state growth path  $(\tau_\infty, \Delta_\infty, g_\infty)$  where  $\tau_\infty = \tau^*(\lambda, \Delta_\infty)$ ,  $\Delta_\infty^2 = s^2/(1 - \beta^2(1 - \tau_\infty)^2)$  and the growth rate is

$$g_\infty = \ln r + \beta \ln s(\tau_\infty) - \frac{\beta(1 - \beta)(1 - \tau_\infty)^2 s^2/2}{1 - \beta^2(1 - \tau_\infty)^2}. \quad (29)$$

As one would expect, a stronger populist bias shifts this steady-state towards greater redistribution and less inequality.

### 3.6 Multiple Long Run Distributions

With any positive influence of wealth on the political process, on the other hand, taxes initially decrease with  $\Delta_t$ , before increasing again. By (28) this tends to increase  $\Delta_{t+1}$ , hence also  $\tau_{t+1}$ , and so on. The feed-back from policy to intergenerational dynamics is now potentially destabilizing.

### Proposition 9

If the pivotal agent is of higher rank than the median ( $\lambda > 0$ ) there can be multiple steady-states. In that case, intertemporal efficiency  $W$  increases as one moves from a less redistributive, more inequalitarian steady-state to one where  $\tau_\infty$  is higher and  $\Delta_\infty$  lower. This is a fortiori true for ex-ante welfare  $E_{w^i}[U^i]$ , and may also be true for the growth rate  $g_\infty$ .

Proposition 9 and equation (27) on which it is based illustrate more generally the *common principle* through which multiple long-run distributions arise in models with imperfect capital markets: a negative general equilibrium feedback of inequality on mobility, via a worsening of credit constraints.<sup>27</sup> In Banerjee and Newman (1993) this feedback operates through the real wage. A more unequal distribution of wealth means more people who do not have the collateral required to become either self-employed entrepreneurs or employers of labor. Instead they become workers, collectively depressing the wage and therefore the bequest which they leave to their children. Conversely, wealthy dynasties benefit from paying lower wages. This makes for low social mobility and persistently high inequality. Building on Aghion and Bolton (1996), Piketty (1996) generates a related form of “mobility trap” through the interest rate. When wealth is distributed very unequally there are few lenders and a lot of would-be borrowers. This implies high interest rates, especially if the economy’s total wealth is relatively low. High interest rates, in turn, aggravate poor agents’ borrowing constraints and make mobility more difficult; thus both low total wealth and high dispersion persist into the next period. Across steady-states, aggregate income and wealth vary negatively with the interest rate. In the present model, as in Saint-Paul

<sup>27</sup>Multiple equilibria can also arise from a non-convexity at the level of individual families, as in Galor and Zeira (1993). With enough idiosyncratic uncertainty, however, a poor families escapes the poverty (no investment) threshold with positive probability; the wealth process is then ergodic and the long-run distribution unique.

(1994) and Bénabou (1995a), the feedback is through policy; taxes and transfers or public funding of education serve as imperfect substitutes for the missing credit market. When greater inequality leads to lower redistribution this again reduces mobility and makes disparities more persistent, by worsening the credit constraints of poor agents. This is why transfers and growth can be positively correlated across steady-states in spite of savings distortions, as stated above. A related form of multiplicity occurs through segregation in models of endogenous community composition. In Durlauf (1996b) for instance, sufficient disparities in wealth and/or education cause well-off families to form small homogeneous communities rather than share the fixed costs with a larger, more heterogeneous population. Cut off from the tax base and positive local spillovers which richer, more educated neighbors would have provided, poor families again experience no or little upward mobility.<sup>28</sup>

Are countries all headed towards the same long-run distribution of income, as in Proposition 8, or are there permanently different trajectories in terms of both level and inequality, as in Proposition 9? I shall argue in Section 5 that this issue of convergence in distribution is of even broader interest and merits to be investigated empirically, then take the first few steps in that direction. But first I turn to the last of the three main theories linking distribution and growth.

## 4 Social Conflict and Property Rights

In the median voter model (at least when taken literally), expropriation occurs through an orderly process: votes are counted, then everyone submits to the majority rule. At the other extreme are models where agents or interest groups can simply grab part of someone else's wealth. It is generally presumed that this becomes more likely when the gap between rich and poor widens, and that the resulting decline in the security of property rights discourages investment, thereby impeding growth.

A number of related ideas have been formalized in the literature. Grossman (1991), (1994), Acemoglu (1995), and Grossman and Kim (1996) focus on the allocation of resources between productive, predatory and defensive activities in the context of one-shot interactions between individuals or economic classes. Using "dynamic commons problem" games with either Markovian or trigger-type strategies, Tornell and Velasco (1992), Tornell (1994) and Benhabib and Rustichini (1996) study how growth is affected by distributional conflict between long-lasting interest groups or coalitions. The first two papers emphasize the role and sources of interest groups' ability to opportunistically extract rents from others; they do not directly deal with inequality. Benhabib and Rustichini (1996) show how the extent to which social conflict constrains growth may depend on the economy's level of development, as well as on a form of inequality which relates to equilib-

---

<sup>28</sup>See also Bénabou (1993), (1996b), Durlauf (1996a) and Fernandez and Rogerson (1996), (1994) for related sorting effects, as well as Maskin and Kremer (1994) on segregation by skills in the labor market. In the latter case, when segregation occurs it is always efficient. But if one combined this production model with credit-constrained accumulation of human capital, multiplicity could occur once again.

rium selection rather than initial conditions: among the continuum of subgame-perfect equilibria of their model, the one with fastest growth involves equal utility for all players.

I shall seek here to convey the essence of this literature and make some new points by means of a very simple model, which relates the underlying structure of the problem to a familiar object: the prisoner's dilemma. The model has close links to that of Benhabib and Rustichini (1996) but also incorporates ingredients from several of the others.

#### 4.1 A Prisoner's Dilemma with Capital Accumulation

Consider an economy constituted by two homogeneous groups of agents, labelled 1 and 2; allowing for  $n$  groups would be straightforward. In period  $t$  each can choose to either moderate its claims to the economic pie ("cooperate") or try and extract a disproportionate amount, at the expense of the other group ("deviate"). Formally, let  $k_t$  be the economy's capital stock at the start of the period. The consumptions of groups 1 and 2 are given by the following shares of  $k_t$  :

| $1 \backslash 2$ | $C$                                                               | $D$                                                               |
|------------------|-------------------------------------------------------------------|-------------------------------------------------------------------|
| $C$              | $\alpha_1(1-s), \alpha_2(1-s)$                                    | $\alpha_1(1-s) - (\beta_2 - \delta_2)s, \alpha_2(1-s) + \beta_2s$ |
| $D$              | $\alpha_1(1-s) + \beta_1s, \alpha_2(1-s) - (\beta_1 - \delta_1)s$ | $\gamma_1, \gamma_2$                                              |

When both sides cooperate, a total of  $c_t = (1-s)k_t$  is consumed, while  $sk_t$  is reinvested according to the linear technology:

$$k_{t+1} = r(k_t - c_t). \quad (30)$$

The parameters  $\alpha_1 \leq \alpha_2 = 1 - \alpha_1$  capture the distribution of income in the non-conflictual outcome, which can be thought as the market solution. They could for instance reflect the competitive shares of capital and labor. In the single-period game, however,  $(C, C)$  is never a Nash equilibrium: by deviating, group  $i$  can appropriate an extra share of the pie, equal to a fraction  $0 < \beta_i \leq 1$  of the resources  $sk_t$  which would otherwise have been reinvested. Of this opportunistic gain, a fraction  $0 < \delta_i \leq 1$  comes from raiding the capital stock and  $\beta_i - \delta_i \geq 0$  is at the direct expense of the other group's consumption. Following most of the literature, it is assumed that transfers or rents unilaterally extracted from the rest of society can only be consumed.<sup>29</sup> The underlying motivation is that otherwise they could be seized back by the opposite side. That consumption is less vulnerable to expropriation than investment does not

<sup>29</sup>Or alternatively, reinvested in some asset with a lower return, such as deposits abroad; see Tornell and Velasco (1992). One could also introduce into the model direct or opportunity costs of engaging in expropriation, by making non-cooperative payoffs sum to less than  $(1-s)k_t + \delta_i sk_t$ . Thus deviator  $i$ 's private gain (net of resources spent on expropriation)  $\beta_i sk_t$  could be strictly less than the sum of the other side's (net) consumption loss, say  $\varphi_i sk_t$ , and the reduction in investment,  $\delta_i sk_t$ . I shall not do so for simplicity, thus imposing  $\beta_i = \varphi_i + \delta_i$ . All costs and payoffs in the stage game are of course specified exogenously; to endogenize them would require modelling agents' resource allocation problems between productive, offensive and defensive activities, as in Grossman (1991), (1994) and Grossman and Kim (1996). Using these papers' terminology, the  $\beta_i$ 's reflect the private efficacy of the predation technology and the  $\delta_i$ 's (more generally, the  $\delta_i + \varphi_i - \beta_i$ 's) its social or aggregate "destructiveness".

mean that it is entirely safe, however. The possibility of indirect taxes and subsidies, import levies or outright theft implies that not all of the deviator's gain need come from investment, as is often assumed in the literature;  $\beta_i$  and  $\delta_i$  are conceptually distinct, and will be seen below to have different implications.

On net, a unilateral deviation by some group  $i$  reduces the capital stock by a factor  $1 - \delta_i$ . In a situation of open conflict where both groups try to appropriate resources from each other they will be lead to overconsume even more: for simplicity I assume that under  $(D, D)$  all capital is consumed,  $\gamma_1 + \gamma_2 = 1$ , but this is not essential for the nature of the results. The shares  $\gamma_i$  represent each side's strength in the political struggle, that is, *its power*. They could be related to the  $\alpha_i$ 's and  $\beta_i$ 's, as discussed later on, but in general need not be.<sup>30</sup> Let me assume from here on that  $\alpha_i(1 - s) - (\beta_{-i} - \delta_{-i})s < \gamma_i$  for all  $i$ , which implies that the unique equilibrium of the one-shot game is  $(D, D)$ . Consider, however, agents who are infinitely-lived, with preferences

$$U_t^i = \sum_{s=0}^{\infty} \rho^s (c_{t+s}^i)^{1-a},$$

where  $1/a > 1$  is the intertemporal elasticity of substitution and  $\rho r^{1-a} < 1$ . In each period they play the game described by the above payoffs. The capital stock  $k_t$  constitutes a state variable which makes this a dynamic game, rather than simply a repeated one. I begin by focusing on the "first-best" case where there is only one group, or equivalently a central planner who can carry out lump-sum transfers. Starting with  $k_0 = k$ , playing  $(C, C)$  in every period yields the intertemporal utility  $U = u(s) k^{1-a}/(1-a)$ , where

$$u(s) \equiv \frac{(1-s)^{1-a}}{1 - \rho(r s)^{1-a}}. \quad (31)$$

The *growth rate* is constant and equal to  $rs$ . This strategy is preferable to  $(D, D)$ , where the entire capital stock is consumed immediately, if  $u(s) > u(0) = 1$ . As depicted on Figure 1 for the case  $r = 3$ ,  $\rho = .4$ ,  $a = .25$ , the function  $u(\cdot)$  is strictly concave and maximized at  $s^* \equiv \rho^{1/a} r^{1/a-1}$ . This is the value which the planner would choose if he could select the savings rate continuously. One could let  $s = s^*$ , but more generally I will simply require that  $s$  belong to the interval  $(\bar{s}, 1)$  over which  $u(s) > 1$ , so that  $rs$  is indeed the best of the two growth rates achievable in our simple economy. The main question is whether it can be sustained as a subgame perfect equilibrium when groups have the ability to behave opportunistically and extract more than their cooperative share from the common pool of resources.<sup>31</sup>

<sup>30</sup>For instance, Benhabib and Rustichini (1996) impose a symmetry restriction similar to  $\gamma_i = 1/2$ .

<sup>31</sup>Were agents also allowed to consume less than  $\alpha_i(1 - s)$ , such deviations could be shown to be unprofitable as long as  $s$  is not too far below the optimal savings rate  $s^*$  (specifically,  $\alpha_i f(s) \leq 1$  where  $f(s) \equiv (1-s)(s/s^*)^a / (1-s(s/s^*)^a)$  is a decreasing function with  $f(s^*) = 1$ ).

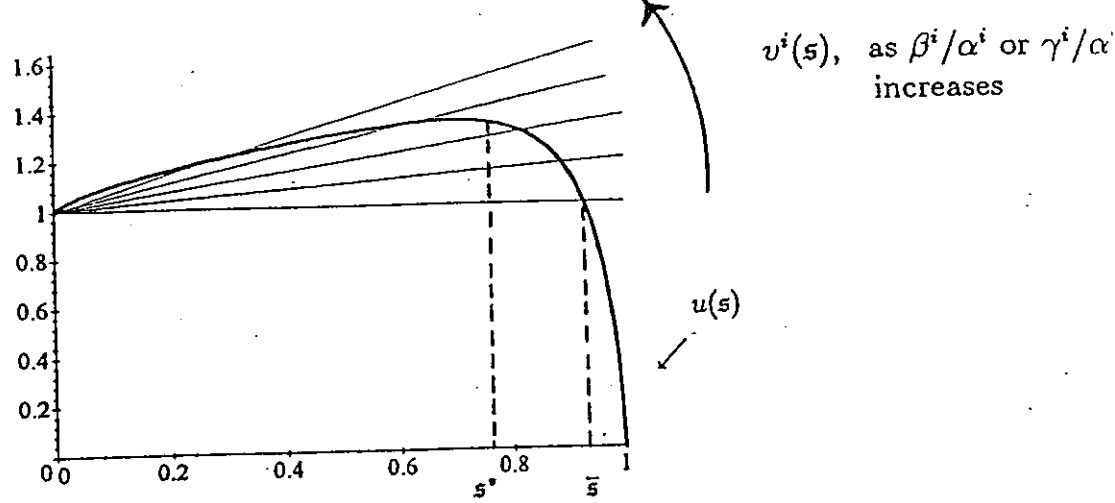


Figure 1: incentive constraints on growth

## 4.2 Income Inequality, Power Inequality and Growth

Playing  $(D, D)$  in every period is always an equilibrium. This high-conflict, low-growth outcome will naturally serve as the punishment to which players revert whenever someone has deviated from cooperative behavior. These trigger strategies will sustain the first best if, for all  $i$ ,

$$(\alpha_i)^{1-\alpha} u(s) \geq (\alpha_i(1-s) + \beta_i s)^{1-\alpha} + \rho(r s (1-\delta_i) \gamma_i)^{1-\alpha}.$$

The first term on the right hand side is group  $i$ 's current consumption when it deviates and grabs extra resources equal to a fraction  $\beta_i s$  of the capital stock. The second term measures its payoff in the following period where both groups play  $D$ , liquidating the capital stock. Rewriting this condition leads to:

### Proposition 10

*If the incentive compatibility constraint*

$$u(s) \geq (1-s + (\beta_i/\alpha_i)s)^{1-\alpha} + \rho(r s (1-\delta_i)(\gamma_i/\alpha_i))^{1-\alpha} \equiv v_i(s) \quad (32)$$

*is satisfied for both groups, continued cooperation and growth at the rate  $r s$  is sustainable as an equilibrium. If it is violated, the only equilibrium is the non-cooperative one with no growth.*

Several interesting results can now be obtained by examining condition (32), which is illustrated on Figure 1.

(1) Incentive compatibility typically holds when  $s$  is low enough, and is violated at high enough values. This implies that there is a *maximum sustainable growth rate* which is constrained by conflict over the distribution of income, as in Benhabib and Rustichini (1996). This maximum growth rate  $rs_{\max}$  falls as  $\max_i\{\beta_i/\alpha_i\}$  or  $\max_i\{\gamma_i/\alpha_i\}$  rise.<sup>32</sup>

(2) Growth becomes harder to sustain when some (or both) group's ability to expropriate the other through unilateral, opportunistic deviations is *disproportionate* to its "normal" share of the economic pie: for any given  $s$ , (32) ceases to hold as  $\beta_i/\alpha_i$  becomes large enough. Consequently,  $s_{\max}$  declines.

(3) *Income inequality* also limits the economy's growth rate, for a given allocation of power and expropriating ability. For instance if  $\beta_1 = \beta_2$  and  $\gamma = 1/2$ , the incentive compatibility constraint is binding for the group with the smaller income share,  $\alpha_1 < \alpha_2$ . Deviations come from the poor, and are more likely the poorer they are. In this case it may be in the interest of the rich to voluntarily transfer wealth to the poor through land reform, education subsidies, a minimum wage, or trade protection: group 2 is better off if its own share  $\alpha_2$  can be permanently reduced to increase  $\alpha_1$ . This scenario, formalized by Grossman (1994), (1995) in a static context, arises here from the desirability of growth.

(4) However, it could be the poor who are relatively more vulnerable to expropriation or exploitation; one can think of small peasants versus large landowners. If  $\beta_1/\alpha_1 = \beta_2/\alpha_2$  but  $\gamma_1/\alpha_1 < \gamma_2/\alpha_2$ , for instance, it is the rich's threat to deviate which reduces accumulation. The model thus makes clear that what matters is not inequality in the (equilibrium) distribution of income per se, as often claimed in the literature, but inequality in the relative *distribution of earning and political power*: since  $\alpha_1 + \alpha_2 = \gamma_1 + \gamma_2 = 1$ , the minimal value of  $\max_i\{\gamma_i/\alpha_i\}$ , which maximizes growth, is 1.<sup>33</sup> Note also that the important role played by off-the equilibrium-path payoffs is a potential problem for the empirical implementation of such game-theoretic models. I shall return to it below.

(5) The more of the expropriation comes from capital, as opposed to other people's consumption, the easier it is to sustain cooperation:  $v_i$  is decreasing in  $\delta_i$ .

<sup>32</sup>For instance as  $s$  tends to zero,  $u(s) \approx 1 + \rho(rs)^{1-a} > 1 + \rho(rs(1 - \delta_i)(\gamma_i/\alpha_i))^{1-a} \approx v_i(s)$  as long as  $(1 - \delta_i)(\gamma_i/\alpha_i) < 1$ . Conversely,  $u(1) = 0 < v_i(1)$  always and  $u(\bar{s}) = 1 < v_i(\bar{s})$  whenever  $\beta_i/\alpha_i > 1$ . In the game described here there are only two sustainable growth rates,  $rs$  and zero. But suppose now that there is in fact a continuum of such games, indexed by  $s$ , which can be played; all other parameters are invariant. Among values of  $s$  for which cooperation can be sustained, both groups will prefer the one with the constrained-optimal growth rate,  $s^{**} \equiv \min\{s^*, s_{\max}\}$ . They will also be unanimous in preferring the  $s^{**}$  game to any other with a single, uncooperative equilibrium if  $u(s^{**}) > \max_i\{(\gamma_i/\alpha_i)^{1-a}\}$ . If this condition is not satisfied the Pareto frontier includes both the  $s^{**}$  and the zero-growth outcomes.

<sup>33</sup>Deviations by the rich also occur in Barbosa, Jovanovic and Spiegel (1996), but the context is different. Agents with stochastic endowments enter into an ex-ante efficient income sharing agreement. Ex-post, they may renege and fight over the appropriation of total resources. Their decisions are based on current payoffs only.

### 4.3 Empirical Evidence

The findings from most studies which have examined the links between inequality, conflict over property rights, and growth are summarized in columns (9) and (10) of Table 2. Two broad measures of “instability” or “insecurity” have been considered. The first consists of indices of sociopolitical instability constructed from various combinations of protests, strikes, government turnover, political violence, coups, revolutions, and the like (e.g., Venieris and Gupta (1986), Londregan and Poole (1990), Alesina and Perotti (1996), Alesina, Ölzner and Roubini (1996), Perotti (1996)). Because of the potentially serious simultaneity problem most studies estimate a joint model of growth and instability. The other form of “instability” is more directly linked to the issue of property rights. It is measured by various indicators of “country risk” sold by specialized firms to international investors, such as sovereign default risk, expropriation or nationalization risk, rule of law, enforceability of contracts, quality of the bureaucracy, corruption, and so on; see Keefer and Knack (1995), Svensson (1993) or Barro (1996). Column (9) shows that both types of security deteriorate as a result of increased income inequality. In turn, column (10) shows that both greater political turmoil and lower protection of property rights decrease investment and growth. The general idea that inequality exacerbates social conflict, which in turn makes property rights less secure and reduces growth thus appears well supported by the evidence; indeed it is eminently plausible. The specific channels through which this occurs, however, remain to be clearly identified. Svensson (1993) finds that when both political instability and property rights variables are included in investment regressions, only the latter are significant; on the other hand, greater political stability implies more secure property rights, as does a more equal income distribution. Keefer and Knack (1995) report that when one controls for initial GDP in the investment equation, political instability becomes insignificant.

I also wrote support for “the general idea” rather than “the models” because the link between the latter and intuitive or empirical notions of social or political instability is still somewhat tenuous. This can be seen from the representative model presented above. First, note that contrary to common interpretation, uncertainty or *unpredictability is really not part* of the story. The mechanism formalized here and in most of the literature is not one where greater inequality leads, through more acute political conflict, to a mean-preserving spread in the return to investment (and even this is known to have an ambiguous effect). It has no up-side and thus leads, *very much as in the voting models*, to a lower mean return due increased expropriation, or the threat thereof. Second, with perfect information about actions and payoffs, equilibrium is achieved without any “friction” or resource dissipation. Just as there is no reason for delay to occur in a bargaining model with symmetric information, there is no role in these models for strikes, riots, coups, assassinations, etc. Incorporating incomplete information, as for instance in Alesina and Drazen (1991), will generate delays and resource-dissipating struggles along the equilibrium path; but the extent to which these occur will reflect not so much the distribution of income (and power) as the extent of agents’ uncertainty over what this distribution is. Why the two should be related

is not obvious. Given imperfect information, resources would be dissipated as well in standard political economy models: the voter who turns out to be pivotal (say, median) will eventually have his way, but only after costly political campaigns, lobbying and legislative maneuvers. This will make the two types of model a little more similar yet, as would the incorporation of strategic, dynamic concerns into the voting model: reputation of the pivotal class for fiscal restraint, trigger strategies on the part of the private sector, and so on.

## 5 Convergence: On to Second Moments (and Higher)

Are countries converging to the same level of inequality? This question may seem unusual, and indeed it has not been taken up by the empirical literature. Yet it is important, on several counts. First, ascertaining the facts is in itself of interest. It is common knowledge that Latin American countries tend to be more unequal than European ones, themselves less equal than East-Asian ones. Similarly, the specter of U.S.-style inequality is often used in Europe to justify high levels of redistribution and minimum wages. But are all these gaps really permanent, or inexorably narrowing?<sup>34</sup>

Second, it can shed light on the relevance of models with multiple steady-states and history-dependence of the income distribution. In particular, it provides an indirect test of the joint mechanisms of credit market incompleteness and negative feedback from inequality to social mobility which underlie these theories, as explained in Section 3.6.

Third, there is now a vast empirical literature examining international convergence in per capita incomes. This variable is after all only the *first moment* of each country's income distribution. Once augmented with idiosyncratic shocks, most versions of the neoclassical growth model imply *convergence in distribution*: countries with the same fundamentals should tend towards the same invariant distribution of wealth and pre-tax income.<sup>35</sup> Barring unexplained differences in "tastes" for equity or in the distribution of innate abilities, persistent differences in the degree of inequality would conversely indicate the presence of some form of increasing return or complementarity in the economic or politico-economic structure.

<sup>34</sup>Looking at OECD countries in the 1980's, Gottschalk and Smeeding (1995) find no relation between initial levels of inequality and changes in inequality during the 80's. If anything, there is some evidence of polarization at the extremes, of the type suggested by Proposition 9. The largest increases occurred in the UK and the US, which had the highest initial levels, suggesting that these countries may be converging to a high inequality steady-state. The smallest increases occurred in Italy, France, Portugal and Denmark; thus Denmark could be converging to a low inequality steady-state. While this exercise provides valuable information, it suffers from being based on a small number of countries, several of which are observed over a very short interval of time.

<sup>35</sup>See for instance Banerjee and Newman (1991), Ayagari (1994), Bertola (1995) or the case without credit constraints in Piketty (1996) for examples of this ergodicity property. Some of these models even include credit rationing (agents can borrow less than what they could repay with probability one) but the feedback from the wealth distribution to credit constraints discussed in Section 3.6 is either absent because the latter are specified exogenously, or as the economy grows it becomes too weak to generate multiplicity (Aghion and Bolton (1996)). Departing from homothetic preferences by introducing a minimum consumption level as in Chatterjee (1994) could probably lead to history-dependence, but the stochastic version of that model still remains to be studied.

Having raised the question of convergence in higher moments, I shall make here a first pass at trying to answer it. Ideally, one would apply to an international panel of inequality measures the same tests which are now standard in the literature on the convergence of first moments: regressing rates of change on initial values, examining whether cross-sectional dispersion falls over time (Barro (1991), Barro and Sala-i-Martin (1992)), or the more complex tests of Quah (1993) and Durlauf and Johnson (1995). The binding constraint, however, is data: no such panel exists over a long enough period. The closest substitute is a new data base recently put together by Deininger and Squire (1995a). It significantly expands the coverage of previous data sets, both over time and over countries, to the extent that one can think of it as a panel with (still many) missing cells. I shall also use data for a sample of OECD countries from the Luxembourg Income Study (LIS). It contains far fewer observations but considerable effort has been devoted to making the numbers comparable across countries.

I start with a bird's eye view of the issue, by comparing the seven main regions of the world for which Deininger and Squire (1995a) provide summary statistics on inequality: Latin America, Sub-Saharan Africa, Middle East and North Africa, East Asia/Pacific, South Asia, OECD and High Income, and Eastern Europe. For each region they constructed decadal averages of various inequality measures, by averaging all available observations from the 1960's to the 1990's. From these numbers I computed both the world average and the cross-regional standard deviation of each inequality measure. These are reported in Table 3.

|                    | Gini (%) | Adj. Gini (%) | Bottom 20% | 3 <sup>d</sup> +4 <sup>th</sup> Quintiles | Top 20% | Top 20 / Bottom20 |
|--------------------|----------|---------------|------------|-------------------------------------------|---------|-------------------|
| Mean               |          |               |            |                                           |         |                   |
| 1960's             | 39.99    | 41.52         | 6.18       | 35.65                                     | 47.84   | 9.03              |
| 1970's             | 39.34    | 40.92         | 6.03       | 36.49                                     | 47.24   | 8.89              |
| 1980's             | 37.47    | 39.92         | 6.52       | 37.56                                     | 44.73   | 8.31              |
| 1990's             | 38.28    | 41.63         | 6.43       | 37.24                                     | 45.47   | 8.46              |
| Standard Deviation |          |               |            |                                           |         |                   |
| 1960's             | 10.03    | 10.77         | 2.17       | 4.29                                      | 8.91    | 4.52              |
| 1970's             | 10.54    | 11.23         | 2.77       | 3.74                                      | 8.76    | 4.29              |
| 1980's             | 8.33     | 9.43          | 1.50       | 2.97                                      | 5.97    | 3.73              |
| 1990's             | 7.79     | 9.46          | 1.58       | 3.52                                      | 6.66    | 3.82              |

Table 3: inequality across seven main regions

Regions: Latin America, Sub-Saharan Africa, Middle East and North Africa , East Asia / Pacific, South Asia, OECD and High Income, Eastern Europe.

Source: computed from Deininger and Squire (1995a), Tables 5 to 7.

While these are somewhat crude statistics, they point to some interesting facts.<sup>36</sup> For the mean level of inequality in the world the picture is one of relative stability; some small and

<sup>36</sup> One would want to weigh regions by population or income, for instance. The "adjusted Gini" in the second column of Table 3 refers to an adjustment made by Deininger and Squire (1995a) to observations based on expenditure rather than income surveys. The former typically bias the Gini coefficient downward. The adjustment consists

gradual improvement is visible until the 1980's, after which the situation starts to deteriorate. Cross-regional variations in inequality, on the other hand, show a clear decline. The coefficient of variation falls by 19% for the Gini coefficients, 30% for the shares of the bottom quintile, 22% for the shares of the third and fourth quintiles and 37% for the share the top quintile. Note also that almost all of this "convergence" occurred between the 1970's and 1980's.<sup>37</sup>

I next went to the full sample and ran simple convergence tests for the Gini coefficients across individual countries, to the extent permitted by the data. I chose the Gini rather than particular shares or share ratios because it was available for more observations, and because it relates more directly to the transition equation (28) in the formal model.<sup>38</sup> In the first test I computed for each country the average yearly rate of change in the Gini coefficient between the first and last available observations. I then ran a regression of this variable on a constant and the initial value of the Gini. The results are reported in Table 4, for four samples.

|              | Sample 1 (N=69)  | Sample 2 (N=25)  | Sample 3 (N=19)  | Sample 4 (N=16)  |
|--------------|------------------|------------------|------------------|------------------|
| c            | .641<br>(3.07)   | .755<br>(2.65)   | 1.22<br>(4.49)   | 2.56<br>(1.48)   |
| initial Gini | -.015<br>(-3.04) | -.017<br>(-2.54) | -.036<br>(-4.75) | -.060<br>(-1.34) |
| $\bar{R}^2$  | .061             | .262             | .480             | .050             |

Table 4: convergence tests using using Ginis' average rate of change  
(heteroskedasticity-consistent t-statistics in parenthesis)

Sample 1: Deininger and Squire data set (1995a); Sample 2: subset of countries using gross household income surveys; Sample 3: subset of OECD countries; Sample 4: LIS data set.

The first sample consists of the 69 countries for which I was able to compute the rate of change.<sup>39</sup> The second one is perhaps the most reliable: it restricts attention to the 25 countries for which the data are most comparable, having been obtained from similar types of surveys (gross household income). The third one consists of the 19 OECD countries in the full sample. It is complemented by the LIS data set, where I was able to perform the same regression on 16 countries; the typical time interval is much shorter, however. The results are consistent across the four columns: the negative and generally significant signs indicate mean-reversion in inequality. To reduce the chance that these results were due to measurement error on the first observation

of adding to the Gini the average size of the bias, which Deininger and Squire compute to be about 6.6 % in their data set.

<sup>37</sup>Given that there are only seven observations in each case, the differences in standard deviations or coefficients of variation documented in Table 3 are, unsurprisingly, not statistically significant. In my view this does not make them any less intriguing, especially in the light of the other evidence discussed below.

<sup>38</sup>For a log-normal, the Gini is a monotonic function of the variance of the logs  $\Delta_i^2$ . I also ran many of the regressions with the logarithms of the Ginis instead of the levels. This led to similar results.

<sup>39</sup>Restricting the sample to those countries where at least ten years separate the first and last observations leads to throwing out nine data points, with essentially unchanged results.

(which biases the coefficient towards negative values), I ran the regression with the exact same right-hand-side variables but replacing the left-hand side by the average rate of change between the second and last available observations. For the first three subsamples the coefficient remained virtually unchanged and statistically significant (t-statistics ranged from 1.9 to 3.0). For the LIS data set, it became positive and insignificant; however, that regression contained only 12 data points (with the trend often computed over dates separated by just a few years) as opposed to between 18 and 62 for the Deininger and Squire data set.

One problem with this test is that inequality and its rate of change are computed over different periods for different countries. The incompleteness of the data set does not permit one to run the standard cross-sectional regression over a fixed, ten or twenty year period. To generate a usable panel it is necessary to interpolate some missing data. This is what I did, using neighboring observations when required to obtain Gini coefficients for 1970, 1980 and 1990. I was then able to run inequality convergence regressions for the periods 1970–1980 (Table 5a, 33 countries), 1980–1990 (Table 5b, 38 countries) and 1970–1990 (Table 5c, 25 countries). The first column presents the basic regression of the change in the Gini on the initial value and a constant. The coefficient is negative and significant at the 5% level for 1970–1980, similar in size but significant only at the 10% level for 1980–1990, and small and insignificant for 1970–1990. One source of concern is that the data from different countries may not be directly comparable. Some numbers come from expenditure surveys, others from income surveys; some pertain to household income, others to personal income; finally, most concern gross income, but some are for net (after tax) figures.<sup>40</sup> I deal with this issue in three different ways, all of which lead to similar results. First, I include in the regression dummy variables for observations generated from expenditure data, personal income data and net income data (expdum, perdum and netdum); the results are reported in the second column of Table 5. Alternatively, I first regress both the initial and final Gini coefficients on these dummies (evaluated at the relevant years), then run the regression on the residuals (“dginires”, third column of each table). Finally, I run the regression on the subset of countries which used gross family income for both dates. These results, given in the “Sample 3” column of each table, are very similar to the previous ones. Once again there is significant evidence of inequality convergence from 1970 to 1980, much less so from 1980 to 1990 and none at all for the whole period 1970 to 1990.

One problem which might explain the puzzling discrepancy between the first two results and the last one is that the data for each subperiod are for different subsets of countries. I therefore re-ran the regressions, with and without dummies, on yet another subsample consisting of the 24 countries for which 1970, 1980 and 1990 observations were available (either from the original data or by interpolation). The results are presented under the heading “Sample 2”. Once again the estimates are all negative, with a drop in significance but not in magnitude from the first decade to the second, and a drop in both when moving to the full twenty-years period. A possible

---

<sup>40</sup>Deininger and Squire (1995a) discuss these problems, of which the first one is the most serious.

explanation would for some countries to go from above to below the mean (say) during the first decade, then move back above during the second one. In the fourth, much smaller subsample of OECD countries, by contrast, the evidence of mean-reversion seems stronger in terms of both magnitude and stability across periods. Overall, the regressions reported in Tables 5a to 5c consistently point to some mean-reversion between 1970 and 1980, but the picture for the other periods is much less clear. It is interesting to recall that in Table 3, the dispersion of inequality across major regions of the world experienced most of its decline between the 1970's and the 1980's. Looking across individual countries, however, one sees no evidence of any narrowing in the range of Ginis. The standard deviation (sdev), reported in the last two lines of Table 5, shows no virtually no change during any of the two decades under consideration.

The general picture which emerges at the end of this empirical exercise is thus a mixed one: general stability in the world distribution of Ginis, within which countries and regions seem to be experiencing some non-negligible amount of relative mobility. The question of whether there is actual convergence or whether countries oscillate around distinct long-run levels of inequality still awaits a definite answer. This is perhaps not surprising given the rather preliminary nature of the analysis and, above all, the limitations of the data.<sup>41</sup> My main purpose, however, was to put forward the issue of *convergence in distribution* as an important and essentially unexplored topic for empirical research. Hopefully, future studies with more sophisticated econometrics and better data (looking across states or regions as in Barro and Sala-i-Martin (1992)) will help resolve the issue.

Two related empirical questions also deserve serious investigation. One, motivated in particular by Proposition 9, is whether there is so to speak "convergence in re-distribution". Are countries, or even industrialized countries, converging to the same relative size of the welfare state, or are there permanent differences in the social contract? Evidence on OECD countries provided by Lindert (1996) suggests the latter, but more formal tests on larger samples are needed.<sup>42</sup> Another broad issue is whether educational, fiscal, or other forms of progressive redistribution actually increase mobility, as formalized in equation (27) and suggested by most models with imperfect capital markets. Recent work by Cooper (1996) studying the effect of state financing of education on the intergenerational mobility of families in poor communities suggests a positive answer. Mulligan (1995), on the other hand, finds only mixed evidence that borrowing constraints affect the transmission of inequality.

---

<sup>41</sup>In particular, I have only looked at unconditional convergence. The potential effects of development on inequality (e.g., the Kuznets hypothesis) and the reverse links from inequality to growth studied in this paper suggest that the dynamics of mean income and income inequality should be examined jointly.

<sup>42</sup>See Piketty (1995) and Bénabou (1995a) for models which seek to explain such persistent differences.

## 6 Conclusion

Non representative-agent growth theory has developed so rapidly in recent years that this tour is necessarily incomplete. First, I focused only on the effects of income or wealth inequality on growth, rather than the reverse. The first model's income distribution evolved endogenously in response to credit constraints and policy outcomes, but due to the homotheticity of preferences, technology and policy it was not affected by the economy's level of development. I thus entirely abstracted from the Kuznets (1955) hypothesis, except when discussing the role of fixed costs. This choice reflects both space constraints and a waning interest in the Kuznets curve, arising from its lack of empirical support in most recent studies. Similar homogeneity properties in the second model precluded me from examining whether sociopolitical conflict and extensive rent-seeking are diseases more likely to afflict poor countries or rich ones (see Benhabib and Rustichini (1996)). Another level-dependent effect of income distribution arises in the "big push" theory formalized by Murphy, Shleifer and Vishny (1989). If industrialization requires a sufficiently large domestic market to make increasing returns technologies profitable, excessive concentration of wealth may represent an obstacle to growth. The relevant constraint, however, is the absolute economic weight of the middle class, which in a large country like India is considerable even with very high inequality. Keefer and Knack (1995) test the "big push" mechanism by including in growth regressions measures of market size such as population, total GDP, and openness to trade, both by themselves and interacted with income distribution. They find no support for the theory in their data. An empirically more successful direction of research concerns the links between income distribution, fertility, and development. Perotti (1996) finds that a greater income share of the middle class has a strong negative effect on fertility, and this in turn has a significant positive impact on growth. I did not discuss this mechanism due to space constraints and because there is yet no well worked out theory of distribution, fertility and growth which can robustly generate these correlations. Among the models which have begun to formalize some of these links, moreover, the kind of wealth constraints on educational investment which were extensively considered in Section 3 often play a critical role (e.g., Galor and Zang (1993)).

Even for the three main theories which were examined in some depth, the analysis necessarily involved a number of simplifications. For concision I abstracted from inherited ability, dynastic altruism, occupational choice and effort decisions of workers, entrepreneurs or rent-seekers, as well as from spillovers and non-convexities of any kind. All these elements have been incorporated in the literature, and their role mentioned at various points in the paper.

Where do we go from here? The political economy and imperfect capital market models are now well understood and have reached the quantitative stage.<sup>43</sup> While further developments

---

<sup>43</sup>See for instance Krusell, Quadrini and Ríos-Rull (1996), (1995) in the first case, Fernandez and Rogerson (1994) and Bénabou (1995b) in the second. There also many quantitative models where liquidity constraints impede consumption-smoothing but not investment (e.g., Ayagari (1994), Bertola (1995)), but their properties are rather different.

(perhaps in relation to fertility or technology) will surely prove valuable, the gap between the scope of the theoretical literature and the scarcity of direct evidence makes further empirical work a high priority: signs of the adverse effects of redistribution on growth remain elusive (to put it mildly), and inferences about the role of credit constraints far too indirect. With respect to sociopolitical conflict, there remains room for theoretical and empirical work to move closer to one another. Existing models provide many valuable insights but few robust results about the effects of income inequality on conflict, whether conflict will be open or latent (off the equilibrium path), fully predictable or uncertain. On the empirical side, the *variances* of policy and property rights variables should be included in the regressions, so as to determine whether it is instability or just the average level of extra-legal redistribution that matters. Examining the interaction between economic inequality and inequality of political power in determining the occurrence of open conflict also seems desirable.

Coming full circle, it is tempting to conclude by asking: which of the theories linking distribution and growth considered in this paper best shed light on the experiences of South Korea and the Philippines? While this is not a substitute for detailed case studies (see for example Rodrik (1995) on Korea and Taiwan), one would expect the most relevant mechanisms to clearly show up in basic indicators such as those compiled by Barro and Wolf (1989). Once again, the pure political economy hypothesis of excessive redistributive pressure emanating from the poor, whether through the ballot box or the street, does not fare well. Public transfers in Korea were low but still two-and-a-half times higher than in the Philippines: 3.71% versus 1.46% of GDP on average between 1970 and 1980. Education expenditures show the same pattern, with shares of 5.13% and 1.97% respectively.<sup>44</sup> While the Philippines' government redistributed much less, it consumed a little more: 1.66% of GDP, versus 1.58% in South Korea. These numbers are broadly consistent with the credit-constrained human capital accumulation hypothesis, and indeed Korea's educational investment skyrocketed past that of the Philippines: the enrollment rate in secondary education went from 42% in 1960 to 95% in 1985, as opposed to 50% to 65% for the Philippines. Similarly, for tertiary education Korea's 1965 enrollment rate of 6% was far behind the Philippines' 19%; by 1992 roles were reversed, with rates of 42% and 28% respectively. Another likely piece of the puzzle is the Philippines' much higher fertility rate (5.55% versus 3.60% between 1965 and 1985), although the extent to which it reflects economic as opposed to cultural factors is unclear. Turning now to theories of social conflict and property rights, political instability does not appear very relevant in this instance, as the two countries experienced similarly high levels of turmoil. The Philippines had more revolutions and political assassinations, Korea more riots, coups, government crises and constitutional changes. Their indices of political rights and

---

<sup>44</sup>Both the anecdotal evidence on the Philippines and the fact that 40% of health subsidies went to the richest 20% of the population (this number is cited by Alesina (1995)) suggest the type of path analyzed in Proposition 9, characterized by redistribution to the wealthy and low growth, or even Verdier and Ades' (1993) model of purely rent-seeking elites. Most of the Barro and Wolf (1989) data mentioned here are averages over 1960-1985; the figures given below for tertiary education are from the World Bank's 1986 and 1995 World Development Reports.

civil liberties were also virtually similar.<sup>45</sup> The more discriminating indicators of property rights protection discussed earlier (rule of law, enforceability of contracts, expropriation risk, corruption, etc.), on the other hand, put Korea distinctly ahead.<sup>46</sup> This general security of property rights was probably instrumental not only in creating a favorable climate for business investment, but also in inducing Korean households to entrust much of their considerable savings to a state-controlled banking system, which in turn channelled credit towards the industries deemed strategic by the government, without excessive dissipation along the way.

---

<sup>45</sup> Barro and Wolf's (1989) data on the average number of these events per year for Korea and the Philippines are *Strike* = (0.00, 0.00), *Revol* = (0.31, 0.46), *Riot* = (2.12, 1.00), *Assass* = (0.12, 0.73), *Coup* = (0.09, 0.00), *Crisis* = (0.38, 0.08), *Constch* = (0.22, 0.09), *Polright* = (4.80, 4.7), *Civlib* = (5.0, 4.5).

<sup>46</sup> The earliest available data are BERI's (Business Environmental Risk Intelligence) 1972 scores for bureaucratic delays, enforceability of contracts and nationalization potential. On a scale of 0 (worst) to 4 (best), Korea scored (2.1, 2.3, 2.6) and the Philippines (1.5, 1.9, 2.2). This pattern is confirmed by the additional data available for the last ten years. The averages of these three BERI scores computed by Keefer and Knack (1995) over 1986–1995 were 2.2 and 1.7 respectively. The average ICRG (International Country Risk Guide) scores over 1986–1995 were 36.7 and 20.3, on a scale of zero to 50. See Keefer and Knack (1995) for sources and definitions.

## Appendix

### Proof of Proposition 1

Let us first study the properties of the function  $V(\tau)$  defined by (12). Clearly, (13) implies that it has a unique, global maximum at  $\tau = 0$ . Moreover,

$$-\frac{V''(\tau)}{\rho\beta} = \frac{1 + \rho\beta(1 - \tau^2)}{(1 - \tau)^2(1 + \rho\beta(1 - \tau))^2} \quad (\text{A.1})$$

is positive if and only if  $\tau^2 < 1 + 1/\rho\beta$ . Thus  $V$  is strictly concave on the policy domain  $(\underline{\tau}, 1)$ , where  $\underline{\tau} \equiv -\sqrt{1 + 1/\rho\beta}$ . By (11), individual  $i$ 's utility function  $U^i(\tau)$  is of the form  $V(\tau) + (1 + \rho)\ln(A + B^i\tau)$  and therefore also strictly concave on  $(\underline{\tau}, 1)$ . The envelope theorem then yields the claims which follow equation (11), as well as Proposition 1.

### Proof of Proposition 3

The result for taxation was proved in the text. Next, note from (9) that

$$\frac{1}{\beta} \frac{\partial^2(g - \ln r)}{\partial \lambda \partial \Delta} = \frac{s'(\tau)}{s(\tau)} \frac{\partial^2 \tau}{\partial \lambda \partial \Delta} + \frac{\partial}{\partial \tau} \left( \frac{s'(\tau)}{s(\tau)} \right) \frac{\partial \tau}{\partial \Delta} \frac{\partial \tau}{\partial \lambda}.$$

For  $\Delta$  small enough,  $\partial^2 \tau / \partial \lambda \partial \Delta$  was seen to be negative, and the fact that  $\partial \tau / \partial \lambda$  is proportional to  $\Delta$  makes the second term again dominated by the first; hence the result. Similar derivations apply to intertemporal efficiency  $V$ .

### Proof of Proposition 4

The first claim follows directly from (22). Differentiating this equation and using (6) also implies that

$$g'(\tau) \geq 0 \Leftrightarrow (1 - \beta)\Delta^2 \geq \frac{1}{(1 - \tau)^2(1 + \rho\beta(1 - \tau))}. \quad (\text{A.2})$$

Since the left side is increasing in  $\tau$ ,  $g$  increases with  $\tau$  up to some  $\tau^g(\Delta) \in [\underline{\tau}, 1)$ , then decreases. This maximum is interior (and increasing in  $\Delta$ ) if and only if

$$(1 - \beta)\Delta^2 > \frac{1}{(1 - \underline{\tau})^2(1 + \rho\beta)}. \quad (\text{A.3})$$

As  $\underline{\tau} < -1$ , it suffices for instance that  $(1 - \beta)\Delta^2 > 1/(4 + 8\rho\beta)$ . The growth-maximizing tax rate is positive if:

$$\tau^g(\Delta) > 0 \Leftrightarrow (1 - \beta)\Delta^2 > \frac{1}{1 + \rho\beta}. \quad (\text{A.4})$$

### Proof of Proposition 5

Because  $W(\tau)$  is the sum of  $V(\tau)$  and a concave, quadratic function of  $\tau$ , it is strictly concave on  $(\underline{\tau}, 1)$ . The same statements hold for individual preferences  $U^i(\tau)$ , now given by (24). The social and (interior) private optima are thus respectively defined by  $W'(\tau) = 0$  and  $U'^i(\tau) = 0$ . In the planner's case, (23) implies that  $V'(\tau^p) = -\rho\beta(1 - \beta)(1 - \tau^p)\Delta^2$ . The fact that  $V$  is strictly

concave and maximized at zero then yields *Claim 1*, while *Claim 2* follows from setting  $\ln w^i = m$  in the first-order condition  $U^i(\tau) = 0$  derived from (25). Next, let  $\tau^a$  be the preferred tax of the agent with average wealth,  $\ln w = m + \Delta^2/2$ . Clearly,  $\tau^a < \tau^p$  if and only if  $U^a(\tau^p) < 0$ , which by (25) is equivalent to  $W'(\tau^p) + [\rho\beta^2(1 - \tau^p) - \rho\beta/2]\Delta^2 < 0$ , that is, to  $\beta(1 - \tau^p) < 1/2$ . Equivalently,  $W'(1 - (2\beta)^{-1}) > 0$ , so by (23) and (13),

$$\tau^p > 1 - (2\beta)^{-1} \Leftrightarrow \frac{(1 - \beta)\Delta^2}{2\beta} > \frac{1 - (2\beta)^{-1}}{(2\beta)^{-1}(1 + \rho\beta(2\beta)^{-1})} \Leftrightarrow (1 - \beta)\Delta^2 > \frac{4\beta(2\beta - 1)}{2 + \rho},$$

hence *Claim 3*. Finally, let us show that when  $\beta < 1/4$  there exists an interval where the tax rate  $\tau^m$  set by the median voter is less inefficient than that set by the average agent,  $\tau^a$ . As  $\Delta$  goes to zero the median converges to the mean, and both  $\tau^m$  and  $\tau^a$  converge to the planers' optimum, which is  $\tau^p(0) = 0$ . Therefore, from (25):  $W'(\tau^a)/\rho\beta = (1/2 - \beta(1 - \tau^a))\Delta^2 = (1/2 - \beta)\Delta^2 + o(\Delta^2)$  and  $W'(\tau^m)/\rho\beta = -\beta(1 - \tau^m)\Delta^2 = -\beta\Delta^2 + o(\Delta^2)$ . Thus if  $\beta < 1/4$  and  $\Delta$  is small enough,  $W'(\tau^a) > -W'(\tau^m)$ , which implies (using Taylor expansions) that  $\tau^p - \tau^a > \tau^m - \tau^p$  and therefore  $W(\tau^a) < W(\tau^m)$ .

### Proof of Proposition 6

Strict concavity of individual preferences  $U^i(\tau)$  implies that the equilibrium tax rate is uniquely defined by the first-order condition (26), or

$$\Psi(\tau, \Delta, \lambda) \equiv (\rho\beta)^{-1}V'(\tau) + (1 - \tau)\Delta^2 - \lambda\Delta = 0, \quad (\text{A.5})$$

as long as there is an interior solution  $\tau^*(\Delta, \lambda) > \underline{\tau}$ . Since  $\Psi_\lambda < 0 < -\Psi_\tau$  it is clear that  $\tau^*$  is strictly increasing in  $\lambda$ . Proposition 5 showed that maximizing  $W$  requires  $\lambda_W > 0$ ; conversely  $\lambda_W < \Delta$  because for  $\lambda > \Delta$ ,  $\Psi(0, \Delta, \lambda) < 0$  so  $\tau^*$  is negative, implying  $W'(\tau^*) > 0$  since  $\tau^p(\Delta) > 0$ . The fact that growth is maximized at  $\lambda_G > \lambda_W$  is obvious; conversely, as  $\lambda$  increases towards the value which makes  $\tau^*(\Delta, \lambda)$  equal to  $\underline{\tau}$ ,  $\tau^*(\Delta, \lambda)$  falls below  $\tau^g(\Delta)$ , implying  $g'(\tau^*) > 0$ . Therefore  $\lambda_G$  is interior, as long as (A.3) is satisfied. This concludes the proof of *Claim 1*, which concerns the effects of  $\lambda$  on equilibrium taxes  $\tau^*(\Delta, \lambda)$ .

I now turn to the effects of  $\Delta$  on  $\tau^*$ , and the resulting contributions to  $W$  and  $g$ . The solution to (A.5) is interior for all  $\Delta$  if and only if  $\max_\Delta \{(\rho\beta)^{-1}V'(\underline{\tau}) + (1 - \underline{\tau})\Delta^2 - \lambda\Delta\} > 0$ , or

$$\frac{\lambda^2}{4} \leq \frac{-\underline{\tau}}{1 + \rho\beta(1 - \underline{\tau})}, \quad (\text{A.6})$$

which will be assumed from here on (since  $\underline{\tau} < -1$ , a sufficient condition is  $\lambda^2 \leq 4/(1 + 2\rho\beta)$ ). Equations (A.5) and (25) then imply:

$$W'(\tau^*) = \rho\beta\Delta(\lambda - \beta(1 - \tau^*)\Delta), \quad (\text{A.7})$$

$$\frac{\partial \tau^*}{\partial \Delta} = \frac{2\Delta(1 - \tau^*) - \lambda}{-V''(\tau^*)/\rho\beta + \Delta^2}. \quad (\text{A.8})$$

Therefore, whenever rising inequality induces tax cuts, these are always inefficient:

$$\frac{\partial \tau^*}{\partial \Delta} \leq 0 \Rightarrow W'(\tau^*) > 0. \quad (\text{A.9})$$

For  $\lambda \leq 0$  equations (A.7) and (A.8) imply  $\partial \tau^* / \partial \Delta > 0$  and  $W'(\tau^*) < 0$ , hence  $g'(\tau^*) < 0$  a fortiori. This proves *Claim 2*. For  $\lambda > 0$ ,  $\partial \tau^* / \partial \Delta > 0$  if and only if  $\tau^* \leq 1 - \lambda / (2\Delta)$ , or equivalently  $\Psi(1 - \lambda / (2\Delta), \Delta, \lambda) < 0$ . Substituting (13) into (A.5), this means

$$\frac{\lambda / (2\Delta) - 1}{(\lambda / 2\Delta)(1 + \rho\beta\lambda / 2\Delta)} + \frac{\lambda\Delta}{2} - \lambda\Delta < 0 \Leftrightarrow \Delta > \frac{\lambda}{2} \left( \frac{1 - \rho\beta\lambda^2 / 4}{1 + \lambda^2 / 4} \right) \equiv \tilde{\Delta}. \quad (\text{A.10})$$

Note now that (A.6) implies that  $\rho\beta\lambda^2 / 4 < 1$ , hence  $\tilde{\Delta} > 0$ . Therefore  $\tau^*$  is *U-shaped* with respect to  $\Delta$ : it declines from  $\tau^*(0, \lambda) = 0$  to a minimum  $\tilde{\tau}^* < 0$  at  $\tilde{\Delta}$ , then rises again towards  $\tau^*(\infty, \lambda) = 1$ , cutting the  $\tau = 0$  axis at  $\Delta = \lambda$ . Finally, (A.5) shows that as  $\Delta \rightarrow +\infty$ ,  $\tau^*$  tends to one with  $1 - \tau^* \approx \theta / \Delta$ , where

$$-\frac{1}{1 - \tau} + (1 - \tau)\Delta^2 - \lambda\Delta \approx 0 \Rightarrow -\Delta / \theta + \Delta\theta - \lambda\Delta = 0 \Leftrightarrow \theta - 1/\theta = \lambda, \quad (\text{A.11})$$

which has a unique solution  $\theta > 0$  for each  $\lambda$ . For further reference, note also that this limit is reached from below:

$$\forall (\Delta, \lambda) \in \mathfrak{R}_+^2, (1 - \tau^*(\Delta, \lambda))\Delta < \theta. \quad (\text{A.12})$$

Let us now turn to the efficiency consequences of these variations in  $\tau^*$ . By (A.8),  $W'(\tau^*) \geq 0$  if and only if  $\tau^* \geq 1 - \lambda / (\beta\Delta)$ . Equivalently,  $\Psi(1 - \lambda / (\beta\Delta), \Delta, \lambda) \geq 0$ , or:

$$\frac{\lambda\Delta^2}{\beta\Delta} - \lambda\Delta \geq \frac{1 - \lambda / (\beta\Delta)}{(\lambda / \beta\Delta)(1 + \rho\beta\lambda / \beta\Delta)} \Leftrightarrow \lambda^2(1/\beta - 1) \geq \frac{\beta - \lambda/\Delta}{1 + \rho\lambda/\Delta} \Leftrightarrow \frac{\lambda}{\Delta} \geq \frac{\beta - \lambda^2(1/\beta - 1)}{1 + \rho\lambda^2(1/\beta - 1)}.$$

Two cases are thus possible:

*Case (i):* if  $\lambda^2 > \beta^2 / (1 - \beta)$  then  $W'(\tau^*) > 0$  for all  $\Delta$ . Therefore on the interval  $[0, \tilde{\Delta}]$  where taxes decrease with inequality, these reductions in  $\tau^*$  are always inefficient. On  $[\tilde{\Delta}, +\infty)$  taxes increases with inequality, although never sufficiently fast. Indeed, one can show that  $\lim_{\Delta \rightarrow \infty} [\Delta(1 - \tau^p(\Delta))] = (1 - \beta)^{-1/2}$ , which is less than  $\theta$  defined in (A.11) if and only if  $\lambda^2 > \beta^2 / (1 - \beta)$ .

*Case (ii):* if  $\lambda^2 < \beta^2 / (1 - \beta)$  then  $W'(\tau^*) > 0$  on  $[0, \bar{\Delta})$  and  $W'(\tau^*) < 0$  on  $(\bar{\Delta}, +\infty)$ , where

$$\bar{\Delta} \equiv \frac{\lambda(1 + \rho\lambda^2(1/\beta - 1))}{\beta - \lambda^2(1/\beta - 1)}. \quad (\text{A.13})$$

Observe now that  $\bar{\Delta} > \lambda > \tilde{\Delta}$ . This implies that  $W'(\tau^*) > 0$  over the whole interval  $[0, \tilde{\Delta}]$  where  $\tau^*$  decreases with inequality, and even over a strict superset of the interval  $[0, \lambda]$  where capital is subsidized ( $\tau < 0$ ). This is intuitive, as the planner always wants to set  $\tau^p(\Delta) > 0$ . Only once

inequality has reached  $\bar{\Delta}$  does it drive equilibrium taxes  $\tau^*(\Delta, \lambda) > 0$  to excessively high levels. This concludes the proof of *Claim 3*, as far as intertemporal efficiency is concerned.

It only remains to consider the growth effects of  $\partial\tau^*/\partial\Delta$ , that is, the sign of  $g'(\tau^*)$ . Using (A.2) and (A.5) one can write, as long as  $\tau^* \neq 0$ ,

$$g'(\tau^*) > 0 \Leftrightarrow \frac{(1 - \tau^*)\Delta^2 - \lambda\Delta}{\tau^*(1 - \tau^*)} < (1 - \beta)\Delta^2 \Leftrightarrow \tau^*[\lambda/\Delta - (1 - \tau^*)(1 - (1 - \beta)\tau^*)] > 0. \quad (\text{A.14})$$

Note first that in the neighborhood of  $\Delta = 0^+$ , where  $\tau \approx 0^-$ , this condition is never satisfied. This is intuitive since the losses from heterogeneity are of second order in  $\Delta^2$ , whereas the induced tax effects on savings are of first order. Next, consider the condition in the neighborhood of  $\bar{\Delta}$  defined by (A.10), where  $\tau^* = 1 - \lambda/(2\bar{\Delta}) < 0$ :

$$\begin{aligned} g'(\tau^*(\bar{\Delta}, \lambda)) > 0 &\Leftrightarrow 2 < 1 - (1 - \beta)(1 - \lambda/(2\bar{\Delta})) \Leftrightarrow \frac{2 - \beta}{1 - \beta} < \frac{1 + \lambda^2/4}{1 - \rho\beta\lambda^2/4} \Leftrightarrow \\ \frac{\lambda^2}{4} &> \frac{1}{\rho\beta + (1 - \beta)(1 + \rho\beta)}. \end{aligned} \quad (\text{A.15})$$

It is easily shown that the right-hand side is smaller than that of (A.6) if and only if  $(1 - \beta)|\underline{\tau}| > 1$ , or equivalently  $(1 - \beta)^2(1 + 1/\rho\beta) > 1$ . This inequality holds if and only if  $\beta$  is not too close to one, in which case there exists a range of  $\lambda$ 's satisfying (A.6) such that  $g'(\tau^*(\bar{\Delta}, \lambda)) > 0$ . In this parameter configuration, tax cuts induced by rising inequality just below  $\bar{\Delta}$  reduce growth, while tax increases which occur as  $\Delta$  rises just above  $\bar{\Delta}$  tend to improve it. To show that this last result is not specific to the region regressive of capital subsidies ( $\tau^* < 0$ ), let us next examine  $g'(\tau^*)$  at  $\Delta = \lambda$ . Since  $\tau^*(\lambda, \lambda) = 0$ , (A.2) shows that:

$$g'(\tau^*(\lambda, \lambda)) > 0 \Leftrightarrow \lambda^2 > \frac{1}{(1 - \beta)(1 + \rho\beta)}.$$

Again it can be shown that the right-hand side is greater than of (A.6) as long as  $\beta$  is below some fixed upper bound. In that case there exists a range of  $\lambda$ 's such that increases in inequality in the neighborhood of  $\Delta = \lambda^+$  generate tax increases to positive levels which still improve the growth rate. Finally, let us consider (A.14) as  $\Delta \rightarrow \infty$  and  $1 - \tau^* \approx \theta/\Delta$ . The condition becomes  $\lambda \geq \beta\theta$ , which by (A.11) is equivalent to  $\lambda^2 \geq \beta^2/(1 - \beta)$ . When this condition (which once more is compatible with (A.6) if and only if  $\beta$  is not too close to one) holds,  $g'(\tau^*(\Delta, \lambda)) > 0$  even asymptotically: equilibrium taxes increase with inequality, but too slowly from the point of view of maximizing not only  $W$  but even  $g$ . When  $\lambda^2 < \beta^2/(1 - \beta)$ , on the other hand, taxes eventually increase too fast with rising inequality, whether from the point of view of welfare or from that of growth.

### Proof of Proposition 7

*Claim 1* immediately results from Propositions 4 and 6, or directly from (A.7)-(A.8). For  $\lambda > 0$ , let us write out:

$$(\rho\beta\Delta)^{-1} \frac{dW}{d\Delta} \Big|_{\tau=\tau^*} = -(1-\beta)(1-\tau^*)^2 + \frac{(\lambda - \beta\Delta(1-\tau^*))(2\Delta(1-\tau^*) - \lambda)}{-V''(\tau^*)/\rho\beta + \Delta^2}. \quad (\text{A.16})$$

The numerator of the second term is a quadratic polynomial in  $\lambda$  whose maximal value is  $[(2-\beta)(1-\tau^*)\Delta]^2/4$ . Therefore,  $(dW/d\Delta)_{\tau=\tau^*} < 0$  as long as:

$$(1-\beta) \left(1 - \frac{V''(\tau^*)}{\rho\beta\Delta^2}\right) > \frac{(2-\beta)^2}{4} \Leftrightarrow \frac{\beta^2}{4(1-\beta)} < \frac{-V''(\tau^*)}{\rho\beta\Delta^2} = \frac{1 + \rho\beta(1-\tau^{*2})}{\Delta^2(1-\tau^*)^2(1 + \rho\beta(1-\tau^*))^2}, \quad (\text{A.17})$$

by (A.1). Now, recall from (A.12) that  $\Delta(1-\tau^*) < \theta$  for all  $(\Delta, \lambda, \beta)$ , so the right-hand side is greater than  $\theta^{-2}(1 + \rho\beta(1-\tau^*))^{-2}$ . Provided  $\beta$  is below this value,  $dW/d\Delta$  is therefore negative for all  $(\Delta, \lambda)$ . For  $\beta$  close to 1, on the other hand, it is clear that the second term in (A.16) dominates. By (A.7) and (A.8) this second term is positive if and only if  $\partial\tau^*/\partial\Delta > 0$  and  $W'(\tau^*) > 0$ , which occurs when  $\Delta \in (\tilde{\Delta}, \bar{\Delta})$ , where these two bounds are defined by (A.10) and (A.13) (in which  $\beta$  can be set equal to 1). This concludes the proof of *Claim 2*.

Consider now growth. First, we showed earlier that if  $\lambda^2 < \beta^2/(1-\beta)$ ,  $g'(\tau^*)$  becomes negative as  $\Delta$  becomes large enough. Since  $\partial\tau^*/\partial\Delta > 0$  for  $\Delta > \tilde{\Delta}$  and  $\partial g/\partial\Delta < 0$  always, this proves that  $dg/d\Delta < 0$  for  $\Delta$  large enough in this case. To obtain a result which is independent of  $\lambda$  and to make clear the role played by  $\beta < 1$ , let us use (A.2) and (A.5) to compute:

$$(\beta\Delta)^{-1} \frac{dg}{d\Delta} \Big|_{\tau=\tau^*} = -(1-\beta)(1-\tau^*)^2 + \frac{((1-\beta)\tau^*(1-\tau^*)\Delta - (1-\tau^*)\Delta + \lambda)(2\Delta(1-\tau^*) - \lambda)}{\tau^*(-V''(\tau^*)/\rho\beta + \Delta^2)} \quad (\text{A.18})$$

Focus now on values  $\Delta > \lambda$ , for which  $\tau^* > 0$ . The numerator of the second term is a quadratic polynomial in  $\lambda$  whose maximal value is  $(1-\tau^*)^2(1+(1-\beta)\tau^*)^2\Delta^2/4$ . Therefore,  $(dg/d\Delta)_{\tau=\tau^*} < 0$  for  $\Delta > \lambda$  as long as:

$$(1-\beta) \left(1 - \frac{V''(\tau^*)}{\rho\beta\Delta^2}\right) > \frac{(1+(1-\beta)\tau^*)^2}{4\tau^*}.$$

As  $\Delta$  tends to infinity,  $1-\tau^* \approx \theta/\Delta$  and  $-V''(\tau^*)/(\rho\beta\Delta^2) \approx \theta^{-2}$ , so the condition becomes  $1 + \theta^{-2} > (1-\beta/2)^2/(1-\beta)$ . This holds if and only if  $\beta$  is below some critical value in  $(0, 1)$ , hence the first part of *Claim 3*. Conversely, one can show from (A.18) that  $\lim_{\Delta \rightarrow 0} [(dg/d\Delta)_{\tau=\tau^*}] = \lambda^2\beta \lim_{\Delta \rightarrow 0} [-\Delta/\tau^*] > 0$ . Indeed as  $\Delta \rightarrow 0$  (or also as  $\beta \rightarrow 1$ ) the growth rate (22) reduces to the complete markets case, where  $\partial g/\partial\tau < 0$ .

### Proof of Proposition 8

The asymptotic variance  $\Delta_\infty^2 = s^2/(1 - \beta^2(1 - \tau)^2)$  associated to a constant tax rate  $\tau$  decreases on the domain  $(\underline{\tau}, 1)$ ; note in particular that  $\beta(1 - \underline{\tau}) < 1$ . Let  $T(\Delta)$  denote the inverse, decreasing function, which maps  $(s^2, s^2/(1 - \beta^2(1 - \underline{\tau})^2))$  into  $(\underline{\tau}, 1)$ . A steady-state is an intersection of this decreasing curve with the equilibrium tax function  $\tau^*(\Delta, \lambda)$  defined by (26) or (A.5). When  $\lambda \leq 0$  this function is upward-sloping, by Proposition 6; hence there can be at most one equilibrium. Moreover,  $\tau^*(s^2, \lambda) < \tau^*(\infty, \lambda) = 1 = T(s^2)$  and  $\tau^*(s^2/(1 - \beta^2), \lambda) > \tau^*(0, \lambda) = 0 = T(s^2/(1 - \beta^2))$ , so there is a unique steady-state, with  $\tau_\infty \in (0, 1)$ . Finally, a decrease in  $\lambda$  shifts up the  $\tau^*(\Delta, \lambda)$  function, which then cuts  $T(\Delta)$  at a higher  $\tau$  and a lower  $\Delta$ .

### Proof of Proposition 9

By Proposition 6, the equilibrium tax function  $\tau^*(\Delta, \lambda)$  is now  $U$ -shaped in  $\Delta$ : it decreases from  $\tau^*(0, \lambda) = 0$  to a minimum at  $\Delta = \tilde{\Delta}$ , then rises towards  $\tau^*(\infty, \lambda) = 1$ , cutting the  $\tau = 0$  axis at  $\Delta = \lambda$ . The fact that  $\tau^*(\Delta, \lambda)$  now has a decreasing segment implies that it can intersect  $T(\Delta)$  at several points. Here I shall only *characterize* the set of equilibria, rather than derive sufficient conditions for multiplicity (this is done in Bénabou (1995a), for a closely related model). It is easy to see that such multiple intersection can only happen in the range  $\Delta < \lambda$  where  $\tau^*(\Delta, \lambda) < 0$ , and that this scenario requires  $s^2/(1 - \beta^2) < \tilde{\Delta}$ . The number of intersections must then be odd, as  $\tau^*(s^2/(1 - \beta^2), \lambda) < 0 = T(s^2/(1 - \beta^2))$  and  $\tau^*(s^2/(1 - \beta^2(1 - \underline{\tau})^2), \lambda) > \underline{\tau} = T(s^2/(1 - \beta^2(1 - \underline{\tau})^2))$ . Denote  $\{(\Delta_j, \tau_j)\}_{j=1}^n$  these intersections, with  $\Delta_j < \Delta_{j+1}$  and  $0 > \tau_j > \tau_{j+1}$ , and let  $W(\tau_j, \Delta_j)$  be the corresponding levels of intertemporal efficiency. By Proposition 5, each  $W(\cdot, \Delta_j)$  is strictly concave and maximized at some  $\tau^p(\Delta_j) > 0$ . Since  $W(\tau, \Delta_j)$  is decreasing in its second argument, this allows us to write:

$$W(\tau_{j+1}, \Delta_{j+1}) < W(\tau_{j+1}, \Delta_j) < W(\tau_j, \Delta_j), \quad (\text{A.19})$$

hence the result that intertemporal efficiency declines as one moves to a more regressive steady-state. The same holds a fortiori for ex-ante welfare, since the insurance motive makes it increasingly costly to reduce income-sharing from  $\tau_j$  to  $\tau_{j+1}$  as the variability of income rises from  $\Delta_j$  to  $\Delta_{j+1}$ . A ranking similar to (A.19) obtains for growth rates as long as  $\tau_j = \tau^*(\Delta_j, \lambda) < \tau^g(\Delta_j)$ ; such is the case in particular if the  $\Delta_j$ 's satisfy (A.4), ensuring that  $\tau^g(\Delta_j) > 0$ . Since  $\tau_j < 0$  for all  $j$ , implying  $\Delta_j^2 > s^2/(1 - \beta^2)$ , it suffices that  $s^2 > (1 + \beta)/(1 + \rho\beta)$ .

## References

- Acemoglu, D. (1995) "Reward Structures and the Allocation of Talent," *European Economic Review*, 39, 17-34.
- Aghion, P. and Bolton, P. (1996) "A Trickle-Down Theory of Growth and Development with Debt Overhang," *Review of Economics Studies*, forthcoming.
- Alesina, A. (1995) "The Political Economy of Macroeconomic Stabilizations and Income Inequality: Myths and Reality," Harvard University mimeo, March.
- Alesina, A. and Drazen, A. (1991) "Why Are Stabilizations Delayed?" *American Economic Review*, 81(5), 1170-1189.
- Alesina, A. Özler, S., Roubini, N. and Swagel, P. (1996) "Political Instability and Economic Growth," *Journal of Economic Growth*, forthcoming.
- Alesina, A. and Perotti, R. (1996) "Income Distribution, Political Instability, and Investment," *European Economic Review*, forthcoming.
- Alesina, A. and Rodrik, D. (1994) "Distributive Politics and Economic Growth," *Quarterly Journal of Economics*, 109, 465-490.
- Ayagari, R. (1994) "Uninsured Idiosyncratic Risk and Aggregate Savings," *Quarterly Journal of Economics*, 109, 659-684.
- 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.
- Barbosa, S., Jovanovic, B. and Spiegel, M. (1996) "Inequality and Stability," New York University mimeo, June.
- Barro, R. (1990) "Government Spending in a Simple Model of Endogenous Growth," *Journal of Political Economy*, 98(2), 103-125.
- Barro, R. (1991) "Economic Growth in a Cross-Section of Countries," *Quarterly Journal of Economics*, 106, 407-443.
- Barro, R. (1996) "Democracy and Growth," *Journal of Economic Growth*, 1(1), 1-28.
- Barro, R. and Sala-i-Martin, X. (1992) "Convergence," *Journal of Political Economy*, 100(2), 223-251.
- Barro, R. and Wolf, H. (1989) "Data Appendix for Economic Growth in a Cross-Section of Countries," NBER mimeo, November.
- Becker, G. and Tomes, N. (1979) "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *Journal of Political Economy*, 87(6), 1153-1189.
- Bénabou, R. (1993) "Workings of a City: Location, Education and Production," *Quarterly Journal of Economics*, 108, 619-652.
- Bénabou, R. (1996a) "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance," *American Economic Review*, 86(3), 584-609.

Bénabou, R. (1996b) "Equity and Efficiency in Human Capital Investment: The Local Connection," *Review of Economic Studies*, 62(2), 237-264.

Bénabou, R. (1995a) "Unequal Societies," New York University mimeo, April. Revised as NBER Working Paper 5583, May 1996.

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

Benhabib J. and Spiegel, M. (1994), "The Role of Human Capital in Economic Development: Evidence from Aggregate Cross-Country Data," *Journal of Monetary Economics* 24, 143-173.

Benhabib, J. and Rustichini, A. (1996) "Social Conflict and Growth," *Journal of Economic Growth*, 1, 129-146.

Bernanke, B. and Gertler, M. (1989) "Agency Costs, Net Worth, and Business Fluctuations," *American Economic Review*, 79, 14-31.

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

Bertola, G. (1995) "Accumulation and the Extent of Inequality," CEPR Discussion Paper 1187, May.

Binswanger, H. and Deininger, K. (1995) "Towards a Political Economy of Agriculture and Agrarian Relations," World Bank mimeo, September.

Binswanger, H. and Rosenzweig, M. (1995) "Are Small Farmers Too Small to be Efficient?," World Bank mimeo, October.

Bourguignon, F. (1994) "Growth, Distribution, and Human Resources," in *En Route to Modern Growth*, G. Ranis ed., Baltimore, John Hopkins University Press.

Brandolini, A. and Rossi, N. (1995) "Income Distribution and Sustainable Growth in Industrial Countries," Syracuse University Working Paper 130, September.

Chatterjee, S. (1994) "Transitional Dynamics and the Distribution of Wealth in a Neoclassical Growth Model," *Journal of Public Economics*, 54, 97-119.

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

Cooper, S. (1996) "Redistribution and the Persistence of Income Inequality," Kennedy School of Government mimeo, Harvard University, January.

Deininger, K. and Squire, L. (1995a) "Measuring Income Inequality: A New Data-Base," World Bank mimeo, December.

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

Devarajan, S., Swaroop, V. and Zou, H.-F., (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*, 1(1), 75-94.

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.

Durlauf, S. and Johnson, P. (1992) "Multiple Regimes and Cross-Country Growth Behavior," *Journal of Applied Econometrics*, 10(4), 365-384.

Easterly, W. and Rebello, S. (1993) "Fiscal Policy and 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. (1996) "Income Distribution, Communities and the Quality of Public Education," *Quarterly Journal of Economics*, 111, 135-164.

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.

Fields, G. and Jakubson, G. (1994) "New Evidence on the Kuznets Curve," Cornell University mimeo.

Galor, O. and Zang, H. (1993) "Fertility, Income Distribution and Economic Growth: Theory and Cross-Country Evidence," Brown University mimeo.

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

Galor, O. and Tsiddon, D. (1993) "Income Distribution and Growth: the Kuznets Hypothesis Revisited," Brown University Working Paper 93-1.

Gastil, R. (1982) "Freedom in the World," Westport, CT, Greenwood Press.

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.

Gottschalk, P. and Smeeding, T. (1995) "Cross-National Comparisons of Levels and Trends in Inequality," Syracuse University Working Paper 126, July.

Grossman, H. (1991) "A General Model of Insurrections," *American Economic Review*, 81, 912-921.

Grossman, H. (1994) "Production, Appropriation and Land Reform," *American Economic Review*, 84, 705-712.

Grossman, H. (1995) "Robin Hood and the Redistribution of Property Income," *European Journal of Political Economy*, 11, 399-410.

Grossman, H. and Kim, M. (1996) "Predation and Accumulation," *Journal of Economic Growth*, forthcoming.

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

Kremer, M. and Maskin, E. (1994) "Segregation by Skills and The Rise in Inequality," MIT mimeo.

- Krusell, P., Quadrini, V. and Ríos-Rull, J.-V. (1996) "Politico-Economic Equilibrium and Economic Growth," *Journal of Economic Dynamics and Controls*, forthcoming.
- Krusell, P., Quadrini, V. and Ríos-Rull, J.-V. (1994) "Are Consumption Taxes Really Better than Income Taxes?," University of Rochester mimeo, April.
- Kuznets, S. (1955) "Economic Growth and Income Inequality," *American Economic Review*, 45, 1-28.
- Levine, R. and Renelt, D. (1992) "A Sensitivity Analysis of Cross-Country Growth Regressions," *American Economic Review*, 82, 942-963.
- Lindert, P. (1996) "What Limits Social Spending?" *Explorations in Economic History*, 33, 1-34.
- Londregan, J., and Poole, K. (1990) "Poverty, the Coup Trap, and the Seizure of Executive Power," *World Politics*, 1-24.
- Loury, G. (1981) "Intergenerational Transfers and the Distribution of Earnings," *Econometrica*, 49, 843-867.
- Lucas, R. (1993) "Making a Miracle," *Econometrica*, 61, 251-272.
- Meltzer, A., and Richards, A. (1981) "A Rational Theory of the Size of Government," *Journal of Political Economy*, 89, 914-927.
- McCallum, J. and Blais, A. (1987) "Government, Special Interest Groups, and Economic Growth," *Public Choice*, 54, 3-18.
- Mulligan, C. (1995) "Some Evidence on the Role of Imperfect Markets for the Transmission of Inequality," University of Chicago mimeo, July.
- Murphy, K., Shleifer, A. and Vishny, R. (1989) "Income Distribution, Market Size and Industrialization," *Quarterly Journal of Economics*, 104, 537-564.
- OECD (1996) "Income Distribution in OECD Countries," Social Policy Studies 18. Report prepared by T. Atkinson, L. Rainwater and T. Smeeding.
- 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. (1992) "Growth, Distribution, and Politics," in A. Cuckierman, Z. Hercowitz and L. Lederman, eds., *Political Economy, Growth, and Business Cycles*. MIT Press, Cambridge, MA.
- Persson, T. and Tabellini, G. (1996) "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*.

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

Psacharopoulos, G. (1993) "Returns to Education: A Global Update," World Bank Policy Research Paper 1067, January.

Quah, D. (1993) "Empirical Cross-Section Dynamics for Economic Growth," *European Economic Review*, 37, 426-434.

Rodrik, D. (1995) "Getting Interventions Right: How South Korea and Taiwan Grew Rich," *Economic Policy*, 55-107.

Rosenzweig, M. and Wolpin, K. (1993) "Credit Market Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low Income Countries: Investments in Bullocks in India," *Journal of Political Economy* 101(2), 223-244.

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.

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

Svensson, J. (1993) "Investment, Property Rights and Political Instability: Theory and Evidence", Stockholm University, Institute for International Economic Studies mimeo, December.

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

Taylor, C. and Hudson, M. (1972) *World Handbook of Political and Social Indicators*, New Haven and London, Yale University Press.

Tornell, A. (1993) "Economic Growth and Decline with Endogenous Property Rights," NBER Working Paper 4354, May.

Tornell, A. and Velasco, A. (1992) "The Tragedy of the Commons and Economic Growth: Why Does Capital Flow from Rich to Poor Countries?" *Journal of Political Economy*, 100(6), 1208-1231.

Townsend, R. (1995) "Financial Systems in Northern Thai Villages," *Quarterly Journal of Economics*, 110, 1011-1046.

Venieris, Y. and Gupta, D. (1986) "Income Distribution and Sociopolitical Instability as Determinants of Savings: A Cross-Sectional Model," *Journal of Political Economy*, 94, 873-883.

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

World Bank, *World Development Report*, 1986 and 1995, Oxford University Press, New York.

|             | Sample 1 (N=33)  |                  | Sample 2 (N=24)   |                  | Sample 3 (N=18)  |                  | Sample 4 (N=9)   |                  |
|-------------|------------------|------------------|-------------------|------------------|------------------|------------------|------------------|------------------|
|             | dgini8070        | dginires8070     | dgini8070         | dginires8070     | dgini8070        | dginires8070     | dgini8070        | dginires8070     |
| c           | 9.02<br>(2.49)   | 13.93<br>(2.39)  | 7.66<br>(1.93)    | 17.36<br>(2.44)  | 1.39<br>(1.21)   | 18.65<br>(2.34)  | 18.42<br>(1.38)  | -1.29<br>(-0.71) |
| gini70      | -.231<br>(-2.69) | -.320<br>(-2.64) | -.190<br>(-1.98)  | -.364<br>(-2.59) | -.337<br>(-2.63) | -.411<br>(-2.37) | -.496<br>(-1.39) | -.743<br>(-1.93) |
| expdum70    | 7.98<br>(2.49)   |                  | 5.37<br>(2.47)    |                  |                  |                  |                  |                  |
| perdum70    | 4.10<br>(0.73)   |                  | 2.26<br>(0.30)    |                  |                  |                  |                  |                  |
| perdum80    | -4.68<br>(-.93)  |                  | -5.80<br>(-.77)   |                  |                  |                  |                  |                  |
| netdum70    | 2.65<br>(.72)    |                  | 6.72<br>(2.27)    |                  |                  |                  |                  |                  |
| netdum80    | -9.56<br>(-2.28) |                  | -13.28<br>(-2.91) |                  |                  |                  |                  |                  |
| $\bar{R}^2$ | .101             | .155             | .053              | .142             | .191             | .249             | -.020            | .247             |
| sdev70      | 10.23            | 8.98             | 10.18             |                  | 8.09             | 8.45             | 5.45             | 6.67             |
| sdev80      | 9.98             | 8.10             | 10.20             |                  | 8.27             | 7.34             | 8.28             | 7.09             |

Table 5a: convergence tests for Ginis, 1970 to 1980  
(heteroskedasticity-consistent t-statistics in parenthesis)

Sample 1: countries for which 1970 and 1980 data were available or interpolated  
Sample 2: subset of countries for which 1970, 1980 and 1990 data were available or interpolated  
Sample 3: subset of countries using gross household income surveys  
Sample 4: subset of available OECD countries.

|             | Sample 1 (N=38) |                   | Sample 2 (N=24) |                  | Sample 3 (N=16) |                  | Sample 4 (N=15)  |              |
|-------------|-----------------|-------------------|-----------------|------------------|-----------------|------------------|------------------|--------------|
|             | dgini9080       | dginires9080      | dgini9080       | dginires9080     | dgini9080       | dginires9080     | dgini9080        | dginires9080 |
| c           | 10.93<br>(2.14) | 14.11<br>(2.13)   | 10.51<br>(1.64) | 15.72<br>(1.35)  | 15.25<br>(1.41) | 31.49<br>(3.73)  | -1.94<br>(-1.10) |              |
| gini80      | -275<br>(-1.83) | -0.334<br>(-1.73) | -252<br>(-1.39) | -377<br>(-1.21)  | -358<br>(-1.26) | -.934<br>(-3.48) | -.903<br>(-2.55) |              |
| expdum80    |                 | .04<br>(.04)      |                 | 15.72<br>(1.35)  |                 |                  |                  |              |
| perdum80    |                 | -.342<br>(-1.84)  |                 | -377<br>(-1.21)  |                 |                  |                  |              |
| perdum90    |                 | -2.32<br>(-1.73)  |                 | -1.0.2<br>(-.29) |                 |                  |                  |              |
| netdum80    |                 | -2.32<br>(-.65)   |                 | -2.26<br>(-.41)  |                 |                  |                  |              |
| netdum90    |                 | 4.23<br>(1.27)    |                 | 5.05<br>(.86)    |                 |                  |                  |              |
|             |                 | 6.44<br>(2.11)    |                 | 4.30<br>(0.81)   |                 |                  |                  |              |
|             |                 | -9.32<br>(-4.26)  |                 | -8.91<br>(-2.18) |                 |                  |                  |              |
| $\bar{R}^2$ | .120            | .175              | .070            | -.040            | .123            | .081             | .682             | .579         |
| sdev80      | 9.08            | 7.55              | 10.20           |                  | 8.27            | 8.18             | 6.80             | 6.66         |
| sdev90      | 8.97            | 7.61              | 10.55           |                  | 8.74            | 8.89             | 4.14             | 4.86         |

Table 5b: convergence tests for Ginis, 1980 to 1990  
(heteroskedasticity—consistent t-statistics in parenthesis)

Sample 1: countries for which 1980 and 1990 data were available or interpolated  
Sample 2: subset of countries for which 1970, 1980 and 1990 data were available or interpolated  
Sample 3: subset of countries using gross household income surveys  
Sample 4: subset of available OECD countries.

|             | Sample 1 (N=25)  |                  | Sample 2 (N=24)  |                  | Sample 3 (N=14) |                  | Sample 4 (N=7)   |                  |
|-------------|------------------|------------------|------------------|------------------|-----------------|------------------|------------------|------------------|
|             | dgini9070        | dginires9070     | dgini9070        | dginires9070     | dgini9070       | dginires9070     | dgini9070        | dginires9070     |
| c           | 2.72<br>(1.08)   | 7.32<br>(1.54)   | 2.38<br>(.92)    | 7.02<br>(1.44)   | 3.50<br>(.58)   | 19.01<br>(1.99)  | 19.01<br>(1.99)  | -.74<br>(-.41)   |
| gini70      | -.039<br>(-.64)  | -.141<br>(-1.37) | -.028<br>(-.44)  | -.130<br>(-1.20) | -.46<br>(-.37)  | -.542<br>(-1.87) | -.542<br>(-1.87) | -.576<br>(-2.64) |
| expdum70    | -.54<br>(-.29)   |                  | -.62<br>(-.32)   |                  |                 |                  |                  |                  |
| perdum70    | -2.96<br>(-1.07) |                  | -2.78<br>(-.95)  |                  |                 |                  |                  |                  |
| perdum90    | 1.90<br>(.88)    |                  | 1.69<br>(.74)    |                  |                 |                  |                  |                  |
| netdum70    | 1.63<br>(.80)    |                  | 1.56<br>(.74)    |                  |                 |                  |                  |                  |
| netdum90    | -3.23<br>(-1.42) |                  | -3.13<br>(-1.36) |                  |                 |                  |                  |                  |
| $\bar{R}^2$ | -.030            | -.139            | -.039            | -.171            | -.069           | .024             | .024             | .243             |
| sdev70      | 10.37            |                  | 10.18            |                  | 9.08            | 4.26             | 4.26             | 6.30             |
| sdev90      | 10.60            |                  | 10.55            |                  | 9.38            | 5.20             | 5.20             | 5.45             |

Table 5c: convergence tests for Ginis, 1970 to 1990  
(heteroskedasticity-consistent t-statistics in parenthesis)

Sample 1: countries for which 1970 and 1990 data were available or interpolated  
Sample 2: subset of countries for which 1970, 1980 and 1990 data were available or interpolated  
Sample 3: subset of countries using gross household income surveys  
Sample 4: subset of available OECD countries.

|                           | (1)<br>INEQ<br>on GR, INV | (2)<br>DEM<br>on GR, INV | (3)<br>DEM*INEQ<br>on GR, INV | (4)<br>INEQ<br>on REDIST | (5)<br>REDIST<br>on GR, INV | (6)<br>HUMCAP<br>on GR, INV | (7)<br>CRED<br>on GR, INV | (8)<br>CRED*INEQ<br>on GR, INV | (9)<br>INEQ<br>on INSTAB | (10)<br>INSTAB<br>on GR, INV |
|---------------------------|---------------------------|--------------------------|-------------------------------|--------------------------|-----------------------------|-----------------------------|---------------------------|--------------------------------|--------------------------|------------------------------|
| 1 Alesina-Rodrik (94)     | ⊖                         | 0                        | 0                             |                          |                             | ⊕                           |                           |                                |                          |                              |
| 2 Alesina-Perotti (96)    |                           |                          |                               |                          |                             | ⊕                           |                           |                                | ⊕                        | -                            |
| 3 Alesina et al (96)      |                           | 0                        |                               |                          |                             | ⊕                           |                           |                                | ⊕                        | ⊖                            |
| 4 Barro (96)              |                           | ~                        |                               |                          | +                           | M=⊕ F=⊖                     |                           |                                |                          | ⊖                            |
| 5 Benhabib-Spiegel (96)   | (-)                       |                          |                               |                          |                             | ⊕                           |                           |                                |                          | (-)                          |
| 6 Bourguignon (94)        | ⊖                         |                          |                               |                          |                             | ⊕                           |                           |                                |                          |                              |
| 7 Brandolini-Rossi (95)   | 0                         |                          |                               |                          |                             |                             |                           |                                |                          |                              |
| 8 Clarke (92)             | ⊖                         |                          | 0                             |                          |                             | ⊕                           |                           |                                |                          |                              |
| 9 Deininger-Squire (95)   | (-)                       | (±)                      | +                             |                          |                             | +                           |                           |                                |                          |                              |
| 10 Devarajan et al (93)   |                           |                          |                               |                          | +                           |                             |                           |                                |                          |                              |
| 11 Easterly-Rebello (93)  |                           |                          |                               | +                        | (-)                         | ⊕                           |                           |                                |                          | (±)                          |
| 12 Keefer-Knack (95)      | ⊖                         | (-)                      | (±)                           | (-)                      |                             | ⊕                           |                           |                                | ⊕                        | ⊖                            |
| 13 Levine-Renelt (92)     |                           |                          |                               | ⊖                        | (±)                         | ⊕                           |                           |                                |                          | -                            |
| 14 Lindert (96)           |                           |                          |                               |                          |                             |                             |                           |                                |                          |                              |
| 15 McCallum-Biais (87)    |                           |                          |                               |                          | ⊕                           |                             |                           |                                |                          |                              |
| 16 Perotti (92)           | ⊖                         |                          |                               | (-)                      | +                           | +                           |                           |                                | ⊕                        | ⊖                            |
| 17 Perotti (94)           | ⊖                         |                          |                               | (-)                      | (+)                         | 0                           | ⊖                         | -                              |                          |                              |
| 18 Perotti (96)           | ⊖                         | 0                        | (-)                           | (+)                      | ⊕                           | M=⊕ F=⊖                     |                           |                                | ⊕                        | ⊖                            |
| 19 Persson-Tabellini (92) | ⊖                         |                          | ⊖                             |                          |                             | ⊕                           |                           |                                |                          |                              |
| 20 Persson-Tabellini (94) | ⊖                         | -                        | ⊖                             | (+)                      | (-)                         | ⊕                           |                           |                                |                          |                              |
| 21 Sala-i-Martin (92)     |                           |                          |                               |                          | ⊕                           |                             |                           |                                | +                        | ⊖                            |
| 22 Svensson (93)          |                           |                          |                               |                          |                             | ⊕                           |                           |                                |                          | ⊖                            |
| 23 Venieris-Gupta (86)    | ⊖                         |                          |                               |                          |                             | ⊕                           |                           |                                |                          | ⊖                            |

Table 2: Effects of inequality on growth or investment and some of their determinants

Symbols: ⊕, ⊖: consistent sign and generally significant; +, -: consistent sign, sometimes significant; (±), (-): consistent sign but generally not significant; (±): inconsistent sign with significant coefficients; 0: inconsistent sign or close to zero, and not significant; ~: inverse U-shaped, significant.

INEQ Measures of inequality: [1] uses income and land Gini; [5], [16] and [23] use -Q3; [6] uses -Q1, -(Q1+Q2) and Q5; [7], [9] use income Gini; [8] and [11] use income Gini, Theil, coefficient of variation and Q5/(Q1+Q2); [12] uses Gini, land Gini and -Q3; [14] uses  $\ln(Q5/Q3) + \ln(Q3/Q1)$ ; [17] uses -Q3 and -(Q1+Q2); [18] uses -(Q3+Q4); [19] uses -Q3 and land Gini; [20] uses Q5 and -Q3; [22] uses -(Q1+Q2)/Q5.

DEM Measures of political rights and degree of democracy; see each study.

REDIST Measures of redistribution: EDEXP is the share of education expenditures in GDP; TRAN and TAX are various transfers and tax rates, as detailed below.

[10] uses current govt. expenditure /GDP (⊕), health and education (±); [11] uses 12 different average and marginal tax rates; only one has a significant effect; [12] uses shares in GDP of social security, welfare, government transfers, tax revenues, government expenditures and government consumption, as well as the share of employment in the state sector; [14] uses the shares in GDP of payments for: social security, welfare, unemployment, health (as well as the sum of these four transfers), and education; [16], [17] and [20] use total transfers/GDP; [18] uses GDP shares of: social security, health plus housing, and education; also the labor tax rate, average and marginal income tax rates; [15], [21] use GDP share of social security.

CRED Credit market imperfections, measured in [16] by (minus) the loan-to-value ratio for mortgages.

HUMCAP [4], [5], [9] and [15] use the initial stock of human capital (primary and /or secondary education); M=males, F=females; all others proxy for the stock with enrollment ratios.

INSTAB [1], [2], [3], [5], [13], [16] and [18] focus on socio-political instability, measured by various combinations of protests, strikes, government turnover, political violence, coups, revolutions etc. [4], [12] and [22] focus on the security of property rights, measured by indicators of country risk (default, expropriation, nationalization), contract enforceability, corruption, etc.