RAFFAELE MATTIOLI LECTURES

Growth, Inequality, and Globalization

Theory, History, and Policy

Philippe Aghion
University College London
and
Jeffrey G. Williamson
Harvard University

CAMBRIDGE
UNIVERSITY PRESS

1 Introduction

The question of how inequality is generated and how it reproduces over time has been a major concern for social scientists for more than a century. Yet the relationship between inequality and the process of economic development is far from being well understood. In particular, for the past forty years conventional economic wisdom on inequality and growth has been dominated by two fallacies:

(a) On the effect of inequality on growth in market economies, the standard argument is that inequality is *necessarily* good for incentives and therefore good for growth, although incentive and growth considerations might (sometimes) be traded off against equity or insurance aims.

This conventional wisdom has been challenged by a number of recent empirical studies. Several papers have used cross-country regressions of GDP growth on income inequality to examine the correlation between these two variables. Alesina and Rodrik (1994), Persson and Tabellini (1994), Perotti (1996), and Hausmann and Gavin (1996b) have all found that there is a negative correlation between average growth and measures of inequality over the 1960–1985 period (although the relationship is stronger for developed than for developing countries). Persson and Tabellini (1994) also present time-series evidence for nine developed economies over the period 1830–1985: their results show that inequality has a negative impact on growth at *all* the stages of development that these economies have gone through in the past 150 years (see Benabou (1996) for a comprehensive review of the literature).

This part draws heavily from joint work with Patrick Bolton, Peter Howitt, and GianLuca Violante. We also benefitted from numerous discussions with Beatriz Armendariz, Tony Atkinson and Roland Benabou, and from the comments of Juan Antonio García, Jon Temple, and Andrea Richter. Finally, we wish to thank the "Cost of Inequality" group of the McArthur Foundation and the School of Public Policy at UCL for invaluable intellectual and financial support.

Table 1. *Korea and 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

Source: Benabou (1996).

An interesting case study is that of South Korea and the Philippines during the past thirty years, discussed by Benabou (1996). In the early 1960s, these two countries looked quite similar with regard to major macroeconomic indicators (GDP per capita, investment per capita, average saving rates, etc.), although they differed in the degree of income inequality, as we can see in table 1. In the Philippines the ratio of the income share of the top 20 percent to the bottom 40 percent of the population was twice as large as in South Korea. Over the following thirty year period, fast growth in South Korea resulted in a five-fold increase of the output level, while that of the Philippines barely doubled. That is, contrary to what the standard argument predicts, the more unequal country grew more slowly.

(b) On the reverse causal relationship from growth to inequality, the conventional wisdom is that inequality should obey the so-called Kuznets hypothesis. Based on a cross-section regression of GNP per head and income distribution across a large number of countries, Kuznets (1955) found an inverted U-shaped relation between income inequality (measured by the Gini coefficient) and GNP per head. That is, the lowest and highest levels of GNP per head were associated with low

Table 2. *Wage inequality measured as the ratio of the wages of the top to the bottom decile*

	1970	1980	1990
Germany		2.5	2.5
United States	3.2	3.8	4.5
France	3.7	3.2	3.2
Italy		2.3	2.5
Japan		2.5	2.8
United Kingdom	2.5	2.6	3.3
Sweden	2.1	2.0	2.1

Source: Piketty (1996).

inequality, while middle levels were associated with high inequality. This result, though cross-sectional, suggested a pattern of inequality along the development process. The conjecture was that inequality should necessarily increase during the early stages of development (due to urbanization and industrialization) and decrease later on as industries would attract a large fraction of the rural labor force. And indeed, in the US the share of total wealth owned by the 10 percent richest households rose from 50 percent around 1770, to 70–80 percent around 1870, and then receded back to 50 percent in 1970.

Up to the 1970s Kuznets' prediction seemed to be validated by the experience not only of the US but also of most of the OECD. However, the downward trend in inequality experienced by these economies during the twentieth century has reversed sharply in recent times. In particular, the past fifteen years have witnessed a significant increase in wage inequality both *between* and *within* groups of workers with different levels of education, as shown by figure 1 and table 2 below.

The rise in inequality shows that, as industrialization goes on, it is not necessarily the case that the income and wage distributions should become less unequal. This suggests, in turn, that the evolution of inequality may be governed by factors other than the level of GNP per capita.

The aim of this first part of the book is to challenge the conventional wisdoms on inequality and growth which, as we have just argued, cannot explain recent empirical evidence. Our analysis remains within the framework of neoclassical economics. However, the introduction of additional aspects such as credit-market imperfections, moral hazard, non-neutral technical and organizational change, and labor-market institutions, gives a more complex and, we believe, more realistic picture of the relationship between inequality and economic growth. The first half of the lecture will be concerned with the effects of inequality on growth, with a view to providing new answers to the existing questions: Does inequality matter? If so, why is excessive inequality bad for aggregate growth? Is it possible to reconcile the above aggregate findings with existing microeconomic theories of incentives? In the second half, we will discuss the Kuznets' hypothesis. We will focus on the recent upsurge in wage and income inequality in developed countries and put forward candidate explanations for it, among which technological change will come out as the most important factor.

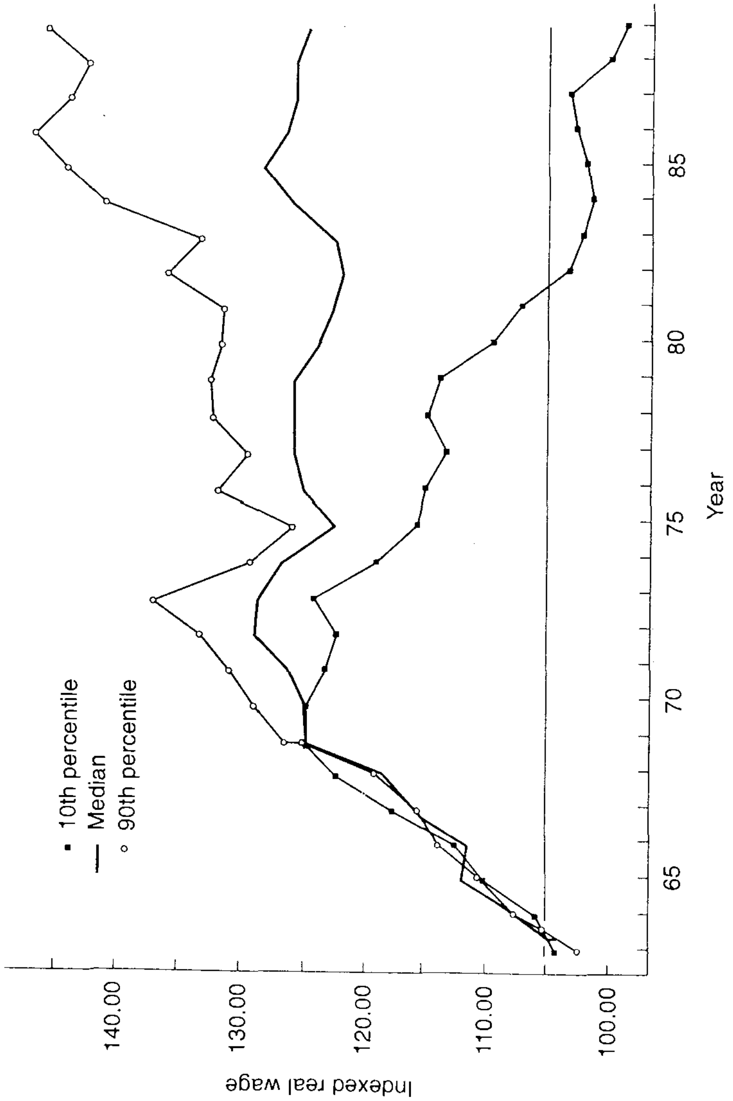

Figure 1 Indexed real weekly wages by percentile, 1963–1989
Source: Juhn, Murphy, and Pierce (1993).

2 Inequality, incentives, and growth

Until recently, most economists agreed that inequality should, if at all, have a *stimulating* effect on capital accumulation and growth. Consequently, there would be a fundamental *tradeoff* between productive efficiency (and/or growth) and social justice, as redistribution would reduce differences in income and wealth, but would also diminish the incentives to accumulate wealth.

Two main considerations appear to underlie the presupposition that inequality should be growth enhancing. The first argument has to do with *investment indivisibilities*: investment projects, in particular the setting up of new industries or the implementation of innovations, often involve large sunk costs. In the absence of a broad and well-functioning market for shares, wealth obviously needs to be sufficiently concentrated in order for an individual (or a family) to be able to cover such large sunk costs and thereby initiate a new industrial activity. This issue has been recently emphasized by policy advisers to transition economies in Central and Eastern Europe and the former Soviet Union. Corporate governance is also subject to indivisibilities as a multiplicity of owners tends to complicate the decision-making process within the firm – when it is necessary to monitor the performance and effort of the firm's manager and employees, having many (dispersed) shareholders raises the scope for free-riding, resulting in a suboptimal level of monitoring.

The second argument, based on *incentive* considerations, was first formalized by Mirrlees (1971). Namely, in a moral hazard context where output realization depends on an *unobservable* effort borne by agents (or "employees"), rewarding the employees with a constant wage independent from (the observable) output performance, will obviously discourage them from investing any effort. On the other hand, making the reward too sensitive to output performance may also be inefficient from an *insurance* point of view when output realizations are highly uncertain and the employees are *risk averse*. This insurance argument is nothing but a natural way to formalize the social justice or "equity" motive for reducing inequality.

The basic incentive argument carries over to the aggregate economy when agents are identical and/or capital markets are perfect (see Rebelo 1991). Consider the neoclassical Ramsey–Cass–Koopmans growth model. Infinitely lived agents maximize their intertemporal utility subject to their budget constraint. Each agent then solves the problem

$$\max_{c_t} \int_0^\infty u(c_t)\, e^{-\rho t} dt$$

subject to $w_t + r_t k_t = c_t$,

where ρ is the intertemporal discount rate, w_t the net wage, k_t the capital stock or wealth of the individual, c_t consumption, and r_t the *after-tax* interest rate. Solving this program we obtain the optimal rate of growth of individual consumption as a function of the after-tax real interest rate

$$g = \frac{r-\rho}{\sigma},$$

where $\sigma = -u''(c)c/u'(c)$ is the intertemporal elasticity of substitution.

When all agents are identical, the above expression gives the aggregate rate of growth of the economy. Redistribution, by making the after-tax rate of interest smaller, reduces the return to saving, thus lowering the rate of growth of consumption and of capital accumulation.

We will now challenge, by means of a simple growth model, the conventional microeconomic tradeoff between equity and incentives. In particular, we will address whether such a tradeoff still exists when we introduce wealth heterogeneity or differences in human capital endowments across individuals together with *capital-market imperfections*. There are at least three reasons why redistribution to the less endowed can be growth enhancing when capital markets are imperfect:

(a) redistribution creates opportunities,
(b) redistribution improves borrowers' incentives,
(c) redistribution reduces macroeconomic volatility.

The next subsections examine under which conditions these mechanisms reverse the conventional tradeoff.

2.1 *The opportunity-enhancing effect of redistribution*

One of the cornerstones of neoclassical economics is the assumption that there are diminishing returns to capital. It is precisely this assumption that drives the familiar convergence results, both at the cross-country level (as in the Solow growth model) and for individuals (as in Tamura 1991). The convergence results rely crucially on perfect capital markets. However, as Stiglitz (1969) first pointed out, when there are decreasing returns to capital and capital markets are imperfect, individual wealth will not converge to a common level and the aggregate level of output will be affected by its distribution. This section reconsiders Stiglitz's arguments in the context of the recent literature on "endogenous growth."

For this purpose, we will consider a discrete-time version of the so-called AK growth model. This is a model in which, although there are diminishing returns to individual investments, there are constant returns to the aggregate capital stock,[1] so that the level of output can be expressed as $Y=AK$, where A is a constant and K the aggregate capital stock.[2]

There is only one good in the economy that serves both as a *capital* and *consumption* good. There is a continuum of overlapping-generations families, indexed by $i \in [0,1]$. Each individual lives for two periods. The intertemporal utility of an individual i born at date t is given by

$$U_t^i = \ln c_t^i + \rho \cdot \ln d_t^i, \tag{1}$$

where c_t^i and d_t^i denote current and future consumption respectively. Individuals differ in their initial endowments of human capital. Let the endowment of individual i upon birth at date t be given by

$$w_t^i = \epsilon_t^i \cdot A_t,$$

where ϵ_t^i is an identically and independently distributed random shock that measures individual i's access to general knowledge. We normalize the mean of ϵ_t^i at one, so that $\int_0^1 w_t^i di = A_t$.

Individual i can either use the efficiency units of labor he is endowed with in order to produce the *current* consumption good, according to a linear "one-for-one" technology, or invest it into the production of the future consumption good. Production of the future consumption good (i.e., of the good

[1] See Aghion and Howitt (1998), chapter 1.

[2] The particular formulation we use in this subsection is taken from Benabou (1996).

available at date $(t+1)$) takes place at date t according to the AK technology

$$y_t^i = (k_t^i)^\alpha (A_t)^{1-\alpha}, \tag{2}$$

where k_t^i denotes the amount of investment by individual i at date t, A_t is the average level of human capital or knowledge available in period t, and $0<\alpha<1$.

We assume that the economy exhibits learning-by-doing: the more an economy produces in one period, the more it learns, and hence the greater the level of knowledge available in the next period. Formally

$$A_t = \int_0^1 y_{t-1}^i di = y_{t-1}. \tag{3}$$

That is, the accumulation of knowledge results from past production activities.

The interesting aspect of this section will result from the presence of heterogeneity or *inequality* among individuals of the same generation, and more specifically from the interplay between *capital-market imperfections* and the effect of *redistribution policies.*

The rate of growth between period $t-1$ and t is given by:

$$g_t = \ln \frac{y_t}{y_{t-1}}$$

that is

$$g_t = \ln \int_0^1 \left(\frac{k_t^i}{A_t}\right)^\alpha di,$$

where k_t^i is determined by intertemporal optimization. It then can be expressed simply as

$$g_t = \ln \frac{E_t(k^\alpha)}{A_t^\alpha},$$

where $E_t(k^\alpha)$ is the mathematical expectation over the output generated by individual investment levels k at date t.

Because of decreasing returns with respect to individual capital investments k^i (in other words, the fact that $\alpha<1$ and therefore the function $k \mapsto k^\alpha$ is concave) greater inequality between individual investments for a given aggregate capital stock will reduce aggregate output.[3] Therefore the more unequal the distribution of individual investments k_t^i, the smaller current aggregate output and therefore the lower the growth rate g in the above AK model.

Is this sufficient for redistribution to the less endowed to be growth enhancing? Not unless capital markets are imperfect. In the absence of capital-market imperfections all individuals choose to invest the same amount of capital $k^i \equiv k^*$, no matter what the initial distribution of human capital or "wealth" across individuals (see Aghion and Howitt 1998, chapter 9). The reason is that the opportunity cost of investing is the rate of interest, both for lenders and borrowers. Hence all individuals wish to invest up to the point where the marginal product of capital is equal to the rate of interest. Those whose wealth is above this level lend, those whose wealth is below it borrow. As a result, aggregate output and growth cannot be positively affected by wealth distribution policies.

Conversely, when capital markets are highly imperfect and therefore credit is scarce and costly, equilibrium investments under laissez-faire will remain unequal across individuals with heterogenous human-capital endowments. Consider the extreme situation in which borrowing is simply not possible and agents are constrained by their wealth, $k_t^i \leq w_t^i$. In this case, individual investments are simply a constant fraction of their wealth $k_t^i = s \cdot w_t^i$. Thus, in contrast to the perfect capital-market case, when credit is unavailable equilibrium investments will differ across individuals (being an increasing function of their initial endowments in human capital), and the rate of growth is given by the distribution of endowments

[3] This, in turn, follows from the following standard theorem in expected utility theory:

Theorem: Let u be a concave function on the non-negative reals. Let X and Y be two random variables, such that the expectations $Eu(X)$ and $Eu(Y)$ exist and are finite, and such that Y is obtained from X through a sequence of mean-preserving spreads. Then $Eu(Y) \leq Eu(X)$. Because a convex function is the negative of a concave function, the opposite inequality holds for a convex function. Then, since

$$E_t(k^\alpha) = \int_0^\infty k^\alpha \cdot f_t(k)dk,$$

where $f_t(k)$ is the density function over individual investments at date t, $E_t(k^\alpha)$ is reduced by a mean-preserving spread.

$$g_t = \alpha \ln s + \ln \int_0^1 (\epsilon_t^i)^\alpha \, di.$$

More inequality is therefore bad for growth when capital markets are highly imperfect.

There is now a role for suitably designed redistribution policies in enhancing aggregate productive efficiency and growth. We will analyze the effects of an *ex-ante* redistribution of human-capital endowments. Consider a lump-sum transfer policy which consists of taxing highly endowed individuals directly on their endowments, and then using the revenues from this tax in order to subsidize human-capital improvements by the less endowed. Thus, the post-tax endowment of individual i can be simply defined by

$$\hat{w}_t^i = w_t^i + \beta(A - w_t^i), \quad 0 < \beta < 1. \tag{4}$$

Recall that A is the average endowment. Those with above-average wealth pay a tax of $\beta(w_t^i - A)$, while those with below-average receive a net subsidy, $\beta(A - w_t^i)$. Because it is a lump-sum tax it does not change the returns to k_t^i, and hence it only affects the incentives to invest in so far as it changes the current wealth of the individual. As the tax rate β increases and the distribution of disposable endowments becomes more equal across individuals, investments by the poorly endowed will increase while investments by the rich will decrease. However, as we already argued, because the production technology exhibits *decreasing* returns with respect to individual capital investments, we should expect redistribution to have an overall *positive* effect on aggregate output and growth. The rate of growth becomes:

$$g = \alpha \ln s + \ln \int_0^1 (\epsilon_t^i + \beta(1 - \epsilon_t^i))^\alpha di. \tag{5}$$

Now consider the term under the integral sign. As β increases, the heterogeneity among individual investment levels (which are proportional to $[\epsilon_t^i + \beta(1 - \epsilon_t^i)]$) decreases, and therefore so does the aggregate efficiency loss due to the unequal distribution of w^i. In the limiting case where $\beta = 1$, the term under the integral sign is constant across individuals i, and the highest possible growth rate is achieved.

The implication of the foregoing analysis is that, when credit is unavailable, redistribution to the poorly endowed, that is, to those individuals who exhibit the higher marginal returns to investment, will be growth enhancing.

2.2 The positive incentive effect of redistribution: questioning the traditional argument

Our modeling of capital-market imperfections in the previous subsection was somewhat extreme, as we were simply assuming away all possibilities of borrowing and lending. Using such a reduced-form representation of credit-market imperfections, we were able to show that redistributing wealth from the rich (whose marginal productivity of investment is relatively low, due to decreasing returns to individual capital investments) to the poor (whose marginal productivity of investment is relatively high, but who cannot invest more than their limited endowments w_i), would enhance aggregate productivity and therefore growth in the preceding AK model. In other words, *redistribution creates investment opportunities* in the absence of well-functioning capital markets, which in turn increases aggregate productivity and growth. Note that this "opportunity creation effect" of redistribution does not rely on incentive considerations: even if one could force the poor to invest *all* their initial endowments rather than maximize intertemporal utility as in the preceding analysis, redistributing wealth from the richest to the poorest individuals would still have an overall positive effect on aggregate productivity and growth, again because of decreasing returns to individual investments.

In this subsection we want to push the analysis one step further and introduce incentives as the microeconomic source of capital-market imperfections. This will enable us to challenge the view that the incentive effect of redistribution should always be negative. In fact, as we will now illustrate, *redistribution may sometimes be growth enhancing as a result of incentive effects only!*

Following Aghion and Bolton (1997), we introduce moral-hazard considerations as the explicit source of credit-market imperfections into the AK with overlapping-generations framework developed above. Specifically, we again assume the existence of a continuum of non-altruistic, overlapping-generations families, indexed by $i \in [0,1]$. The utility of individual i in generation t is

$$U_t^i = d_t^i - c(e_t^i), \tag{6}$$

where d_t^i denotes individual i's second-period consumption (for simplicity we assume that individuals consume only when old), e^i is the non-monetary effort incurred by individual i when young, and $c(e^i) = A(e^i)^2/2$ denotes the non-monetary cost of effort. The parameter A still measures productivity on the current technology. As before, the human-capital endowment of individual i is taken to be an idiosyncratic proportion of average knowledge at date t, that is, $w_t^i = \epsilon_t^i \cdot A_t$.

The production technology involves an extreme form of U-shaped average cost curve with respect to capital investments, namely:

(a) the production activity requires a *fixed* and indivisible capital outlay equal to $k_t^i = \varphi \cdot A_t$;
(b) conditional upon the required investment $\varphi \cdot A_t$ being made at date t, the output from investment in this technology is uncertain and given by

$$y_t^i = \begin{cases} \sigma \cdot A_t \text{ with probability } e_t^i \\ 0 \text{ with probability } 1 - e_t^i, \end{cases}$$

where e_t^i is individual i's effort at date t. We assume that second-period outcomes y_t^i are independently identically distributed across individuals of the same generation.

The source of capital-market imperfection will be moral hazard with limited wealth constraints (or limited liability), in other words, the assumption that:

(a) efforts e^i are not observable;
(b) a borrower's repayment to his lenders cannot exceed his second period output y_t^i.

Consider the effort decision of an individual who does not need to borrow, that is, for whom $w^i \geq \varphi A$. The problem he faces is

$$\max_e \{e \cdot \sigma A - c(e)\},$$

which gives the first-best level of effort, $e^* = \sigma$.

An agent with initial endowment $w^i < \varphi A$ needs to borrow $b^i = \varphi A - w^i$ in order to invest. Let ρ be the unit repayment rate owed by individual w^i. Hence, he chooses effort e^i to maximize the expected second-period revenue net of both repayment to the lenders and effort cost, namely

$$e^i = \max_e \{e(\sigma A - \rho(\varphi \cdot A - w^i)) - c(e)\}$$

$$= e(\rho, w^i), \qquad (7)$$

where $e(\rho, w^i) = \sigma - \rho(\varphi - w^i/A)$ is less than the first-best effort e^*, and is decreasing in ρ and increasing in w^i.

What is important in order to find moral hazard is that effort be increasing in the wealth of the individual. That is, for given ρ, the lower a borrower's initial wealth, the *less* effort he will devote to increasing the probability of success of his project. The more an individual needs to borrow in order to get production started, the less incentives he has to supply effort, in that he must share a larger fraction of the marginal returns from his effort with lenders. An immediate consequence of this result is that redistributing wealth toward borrowers will have a *positive* effect on their effort *incentives*. Whenever this positive incentive effect more than compensates the potentially negative incentive effect on lenders' efforts, then such a redistribution will indeed be growth enhancing based on incentive considerations only.

Before turning to the analysis of redistribution, let us make two important remarks. First, individuals with initial wealth $w^i \geq \varphi A$ (in other words the lenders), will systematically supply the first-best level of effort because they remain residual claimants on all returns from such effort: $e^i(w^i \geq \varphi A) = e^*$.

Second, when analyzing the relationship between initial wealth and effort, we have treated the repayment schedule ρ as given. However, because the risk of default on a loan increases with the size of the loan (the probability of success $e(\rho, w)$ decreases when

w decreases), the unit repayment rate ρ may vary with w to reflect the change in default risk. Aghion and Bolton (1997) show that even once this effect is taken into account, effort is increasing in w^i.

The growth rate of the economy is given by

$$g = \ln \frac{\sigma A \cdot \int e^i di}{A}$$

$$= \ln \sigma + \ln \int_0^1 e^i di, \qquad (8)$$

with efforts $e^i \leq \sigma$. If either (a) or (b) were violated, then the first-best effort would automatically be elicited from *all* individuals no matter what their human-capital endowments were. The growth rate would then be unaffected by redistribution and always be equal to $g = \ln \sigma^2$. This corresponds to nothing but the case of *perfect* capital markets, that is of capital markets that do *not* suffer from incentive problems. When there are incentive problems, the more unequal the distribution of wealth is, that is, the larger the number of individuals with wealth below the threshold level φA, the lower the aggregate level of effort will be. Consequently, inequality has a negative effect on both the income level and the growth rate.

We now have all the elements we need to analyze the incentive effects of redistribution. Because individuals with initial wealth $w^i \geq \varphi A$ supply the first-best effort $e^* = \sigma$, raising a lump-sum tax $t^i < w^i - \varphi A$ on the endowment of each such individual and then distributing the total proceeds among borrowers:

(i) will not affect the effort e^* supplied by the wealthy, whose *after-tax* endowments remain strictly above the required fixed cost φA;
(ii) will increase the effort supplied by any subsidized borrower.

The above redistribution scheme will then have an unambiguously positive *incentive* effect on growth, as efforts e^i either increase or remain constant as a result of redistribution.

We have just put the traditional incentive-distribution tradeoff upside-down, since we have shown that in the context of an imperfect credit market with moral hazard, redistribution enhances growth. For quite similar reasons inequality will tend to discourage cooperation between uneven equity holders engaged in the same venture or partnership. This lack of cooperation may typically take the form of *free-riding* by the poor on the rich's effort.[4] The effect on (long-run) growth will obviously be negative.

To see how inequality induces free-riding consider the following set up. Suppose that the economy gives birth to only two individuals each period, and that these two individuals (who both live for two periods) need to join forces (that is, to pool their initial resources) in order to produce. Let $\bar{w}_t = \bar{w}A_t$ and $w_t = wA_t$ denote the initial endowment of the richer and the poorer of these two individuals. As above, we denote by $\varphi \cdot A_t$ the fixed cost of the project initiated at date t, and we assume that

$$\bar{w} + \underline{w} \geq \varphi > \bar{w} > \underline{w}.$$

In other words, the project requires the financial participation of both individuals in order to be implemented at all.

Once the fixed cost φA_t has been sunk, the project yields $\sigma \cdot A_t$ with probability $(\bar{e} + \underline{e})/2$ and zero with probability $(1 - \bar{e} + \underline{e})/2)$, where $\bar{e}$ and $\underline{e}$ denote the effort of the richer and the poorer individuals. The return of the project is then distributed between the two individuals according to their shares in the total investment. They can choose whether to exert one unit of effort or no effort at all. There is a "moral hazard in team" problem between the two individuals.

Suppose that there is a non-zero effort cost for each individual, and let us assume, as before, that individuals only care for expected second-period output net of their effort cost. Then, the resulting Nash equilibrium depends on the degree of inequality. In particular, when the discrepancy between the rich and the poor is sufficiently large relative to the cost of effort, full cooperation between both individuals (i.e., $\bar{e} = \underline{e} = 1$) will not be sustainable in equilibrium. Rather, the poor individual will free-ride on the rich one, as part of the (unique) equilibrium $\bar{e} = 1, \underline{e} = 0$. Moving toward a more egalitarian distribution of wealth (i.e., toward $\bar{w} = \underline{w} = 1/2$) between the two individuals, will favor their cooperation and thereby increase the level of output and the growth rate.

[4] Legros and Newman (1994) have also emphasized the idea that a high degree of inequality between the rich and the poor may induce the rich to choose inefficient organizational structures in order to better take advantage of their bargaining power *vis-à-vis* poor partners within the same firms.

Table 3. *Income inequality is increased by high macroeconomic volatility*

	Index of inequality	Percentage of difference
Income inequality		
Latin America	6.284	
Industrial countries	2.270	
Difference	4.014	100.0
Impact of		
Initial income inequality	2.047	51.0
Growth in per capita income	0.067	1.7
Average inflation	0.029	0.7
Volatility of real GDP	0.912	22.7
Unexplained	0.959	23.9

Source: Gavin and Hausman (1996b).

2.3 Macroeconomic volatility

Another reason why excessive inequality may be bad for growth is that it generates macroeconomic volatility. The idea that macroeconomic instability is fundamentally detrimental to growth has been pointed out by various authors, especially Alesina and Perotti (1996). It also emerges quite clearly from the cross-country regression for Latin America performed by Hausmann and Gavin (1996a,b). Interestingly for our purpose in this chapter, Hausmann and Gavin find (a) a *positive* correlation between macroeconomic volatility and both income inequality and financial underdevelopment (table 3), and (b) a *negative* correlation between volatility and growth (figure 2).

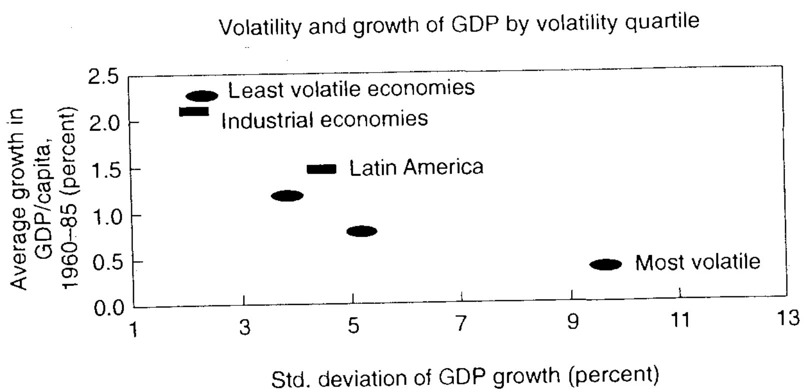

Figure 2 Volatility and growth of real GDP per capita
Source: Gavin and Haussman (1996)

Several explanations have been put forward to account for the correlation between [high] inequality and macroeconomic volatility. Alesina and Perotti (1996) maintain that causality runs from high inequality to political and institutional instability, which in turn results in macroeconomic volatility. The approach we take in this subsection, based on Aghion, Banerjee, and Piketty (1997) (ABP from now onwards), postulates a direct effect of inequality on macroeconomic fluctuations. Inequality, however, takes the form of unequal access to investment opportunities across individuals, which, together with a high degree of capital-market imperfection, can generate persistent credit cycles. Beyond its theoretical appeal, we believe that the ABP set-up summarized below can be useful in understanding the kind of financial crises recently experienced by the growing economies of South-East Asia.

Specifically we consider a dynamic economy in which only a fraction of the active population has access to high-yield investment opportunities. There are a number of reasons why access to investment opportunities may be restricted. Particular skills, ideas, or connections may be required, and often there may be crucial information that can only be acquired by those already in the business. Investment indivisibilities are another potential cause. Individuals may also differ in their attitudes toward risk, hence only those with little risk aversion will be willing to under-

take risky projects rather than work under a riskless employment contract. It is this inequality of access to investments and the consequent separation of investors and savers that will give rise to volatility.

Consider an economy where there are two production technologies: a traditional technology and a high-yield technology. Two crucial assumptions are needed for inequality to affect volatility:

1 Inequality of access to investment: Only a fraction of savers can directly invest in high-yield projects, whereas all individuals can invest in the low-yield technology.
2 Credit-market imperfections: Because of incentive compatibility considerations, an investor with wealth w can borrow only a limited amount, νw, where $\nu < \infty$.

Now assume that all individuals in the economy save a constant fraction of their wealth, *s*. What do the saving and the investment functions look like? The total supply of funds in period t is a fraction s of the aggregate level of wealth in period $t-1$. Savings at t are therefore independent of any variable in that period. The total demand for investment in the high-return project at time t is proportional to the wealth of those who have access to the high-yield investment, and thus is also completely determined by the previous period's income and by the (exogenously given) credit multiplier. There is therefore no market-clearing mechanism that will equalize the supply of funds and the demand for investment in the more productive technology. Consequently the economy will experience either "idle" savings (i.e., a fraction of savings are not invested in high-yield projects) or unrealized investment opportunities.

The link between inequality and volatility hence stems from the fact that those who invest and those who save are not the same individuals. Slumps are periods of idle savings, in which funds are invested in the low-return technology therefore generating a loss of potential output. If everybody had the possibility of investing in the high-yield technology, all agents would choose to invest all their savings, and there would be no slumps. Similarly, if investors were not credit constrained they could absorb all savings.

More precisely, during booms investors' net wealth increases and therefore so does their borrowing capacity, νw. Investors can thus accumulate debt during booms, thereby increasing the demand for investable funds. The interest rate is given by the marginal product of capital. Since all funds are invested in the high-yield technology, interest rates are high during booms. Eventually, the accelerated increase in their debt repayment obligations ends up squeezing the investors' borrowing capacity, up to a point where a positive fraction of savings becomes idle. At this point the economy experiences a slump: some funds have to be invested in the traditional technology, therefore the marginal product of capital falls and interest rates drop. This in turn allows the investors to progressively reconstitute their borrowing capacity, and so eventually the economy will re-enter a boom. If the fraction of the population with high-yield investment possibilities is small enough and/or the credit multiplier low enough, there will be continuous oscillations of the investment level. Such volatility of investment in turn implies that there are unexploited production possibilities and hence the long-run growth rate is lower than it could be.

The government has two structural policy options to try to move the economy out of the above cyclical equilibrium into a situation in which all savings are invested in the high-return production technology. One is to reduce the borrowing constraints, thus increasing the credit multiplier and ensuring that there is sufficient demand for funds. This is, however, a hard policy to implement unless the government is willing to lend to individuals itself. Moreover, if the credit constraint is the result of a moral hazard problem, such as that examined in subsection 2.2, it would not be possible to increase the credit multiplier without generating adverse incentive effects. A second structural policy consists in reducing the degree of inequality of access to investment. By increasing the fraction of savers that can directly invest in high-yield projects, the economy can move to a permanent-boom situation and thus increase its growth rate. Structural reforms such as investing in infrastructure or in human capital, or reducing the bureaucratic obstacles faced by entrepreneurs that wish to set up a firm, would reduce entry barriers and promote growth.

Structural policies may be hard to implement though, especially in the short run. An alternative would be to transfer the idle funds from savers to investors. This policy ensures that all savings are invested in the high-yield technology. However, it transfers resources from those that are worse off to those that are better off. Yet ABP show that this policy does not entail negative distributive consequences for savers. The higher level of income trickles down to savers for two reasons: first, the interest rate is higher, so (poor) lenders are better off; second, as more capital is invested in the high-yield technology, the productivity of labor and thus the wage rate is also higher.

Our analysis so far has concentrated on the case of a closed economy and much of the output cycle appeared to be driven by movements in the real interest rate. However, in more recent work with P. Bacchetta and A. Banerjee, we are considering a small open economy extension of the same framework, where real interest rates remain fixed at the international market-clearing level and the transmission variable becomes the price of non-tradeable goods in terms of the tradeable good. More precisely, high-yield investments in the domestic economy require the use of non-tradeable goods (such as real estate) as inputs to produce tradeable goods. Then, the story goes as follows: during a boom the domestic demand for non-tradeable goods keeps going up as high-yield investments build up, and thus so does the price of non-tradeables relative to that of tradeables. This, together with the accumulation of debt that still goes on during booms, will eventually squeeze investors' borrowing capacity and therefore the demand for non-tradeable goods. At this point, the economy experiences a slump and two things occur: the price of non-tradeables collapses to the level where it is equal to the real rate of return of the asset (i.e., it falls relative to that of tradeables), while a fraction of the assets on offer is not purchased as there are not enough investment funds. This second effect has real consequences, as those individuals who cannot undertake tradeable production have to move into the backyard technology. The collapse in the price of non-tradeables thus results in a contraction of the tradeable-goods sector and of the level of real output.

Unlike in Krugman (1998), the argument that we have just presented does not rely on any regime or policy change.[5] Investors are constrained in their borrowing at any point in time. The increase in the price of non-tradeable goods relative to tradeables and the accumulation of debt, make the credit-market constraint bind at a certain moment in time, and bring about the collapse in the price of non-tradeables. The effect of credit-market imperfections would, clearly, be worsened if production were risky and if there were moral hazard on the part of investors. What is new about this approach is that the financial slump is the consequence of rapid growth. Growth is financed by the accumulation of debt. The debt build up and the consequent increase in the price of non-tradeables is what causes the crisis. This raises the question of what is sustainable growth. If periods of fast growth are followed by slumps due to excessive debt build-up, it may be a better long-run strategy to allow the economy to develop at a slower but steady pace.

2.4 *Political economy*

Economic conflicts surface through the political process, especially when the society as a whole must decide about redistribution or public-good investments such as education or health. By affecting the outcome of the political game, inequality will directly influence the extent of redistribution and thereby the rate of growth. Interestingly, the *direction* in which inequality affects growth through the political process turns out to depend heavily on the importance of credit constraints, as we will now illustrate.

As has been argued in the previous subsections, redistribution affects the rate of growth in an AK-model. If inequality determines the extent of redistribution, it will then have an indirect effect on the rate of growth of the economy. Several authors, such as Alesina and Rodrik (1994), Persson and Tabellini (1994), and Benabou (1996), maintain that inequality affects taxation through the political process when individuals are allowed to vote in order to choose the tax rate (or, equivalently, vote to elect a government whose program includes a certain redistributive policy). In general, we would expect that in very unequal societies, a majority of voters prefer high redistribution than in more equal societies. If redistribution is harmful for growth, then more unequal societies would grow faster.

[5] Krugman (1998) argues that the Asian crisis has been caused by moral hazard on the part of financial intermediaries whose liabilities were guaranteed by the government. The resulting overinvestment and excessive risk-taking made asset prices rise. Eventually, a "change in regime" has implied that liabilities are no longer guaranteed and asset prices have collapsed.

To illustrate this argument suppose that individuals are, as before, endowed with different amounts of human capital, given by $w_t^i = \epsilon_t^i \cdot A_t$, and that production of the future consumption good takes place according to the AK-technology. The government now introduces redistributive taxation that takes the following form: there is a proportional tax on individual investments and the revenue is used to distribute a lump-sum subsidy which is proportional to the average investment. This is, an individual i with pre-tax investment k^i ends up with the post-tax investment

$$k^i(\tau) = (1-\tau)\cdot k^i + \tau \cdot k, \quad 0 < \tau < 1,$$

where k is the average investment. Clearly, those with above-average investments pay a net tax, while those with below-average k_i receive a net subsidy.

When capital markets are perfect so that all agents can borrow at the risk-free interest rate, all individuals choose to invest the same amount $k^i = s(\tau)\cdot w$, where $s(\tau)$ is the saving rate. Individual investments depend on the *average* endowment and on the saving rate. In the absence of moral hazard, the saving rate is the same for all individuals and is decreasing in the tax rate due to the standard negative incentive effect (see appendix 2). Moreover, in the AK-model, the growth rate is a function of the saving rate, $g_t = \alpha \ln s(\tau)$. The standard incentive argument then implies that a high tax rate, by reducing the net return to investment, reduces the fraction of wealth that is invested and the growth rate.

The tax rate affects individuals differently depending on their initial income, as it has two distinct effects: on the one hand, it affects an agent's current income through the net tax/subsidy; on the other hand, it affects his future income through the changes in the growth rate. In fact, we can express the indirect utility function of individual i as a function of his relative wealth and the tax rate

$$U^i(\tau) = V(\tau) + G(w^i/w,\tau).$$

The term $V(\tau)$ captures the *incentive* effect of redistribution and is the same for all agents: redistribution affects the growth rate and hence utility. In our particular example, $V(\tau)$ is decreasing in τ as a result of the negative incentive effect. $G(w^i/w,\tau)$ is an individual-specific term that reflects the *redistribution* effect of the tax. Those agents with wealth above w pay a net tax, as they are taxed more than they receive in subsidies. Hence for them the term $G(w^i/w,\tau)$ is negative. For those agents with $w^i < w$, this term is positive as they receive a net subsidy. The net investment of the individual with average wealth, $w^i = w$, is unaffected by the tax, i.e., $G(1,\tau)$ is zero. Moreover, the impact of an increase in the tax rate on $G(w^i/w,\tau)$ depends on the relative income position of the individual: a higher tax reduces the utility of agents with above-average wealth through the redistribution effect, and increases that of agents with below-average w^i. We have

$$\frac{\partial G(w^i/w,\tau)}{\partial \tau} \begin{cases} < 0 & \text{for } w^i > w \\ = 0 & \text{for } w^i = w \\ > 0 & \text{for } w^i < w. \end{cases}$$

An individual will prefer the tax rate at which the marginal increase in utility of redistribution equals the marginal loss due to the incentive effect. The preferred tax rate of individual i is given by the first-order condition $\partial U^i(\tau)/\partial\tau = 0$

$$\frac{\partial V(\tau)}{\partial \tau} = -\frac{\partial G(w^i/w,\tau)}{\partial \tau}.$$

Individuals with initial wealth equal to w or greater will prefer a zero tax rate, as a higher tax reduces their utility through both the incentive effect and the redistribution effect. Individuals with initial wealth $w^i < w$ will prefer a positive tax rate $\tau(w^i)$. The resulting preferred tax rate $\tau(w^i)$ is decreasing in w^i. Not surprisingly, poorer individuals will prefer a higher tax rate τ^i, as the redistribution effect is stronger the lower w^i.

Assume now that the tax rate is endogenously determined each period through majority voting. Given that the intertemporal utilities $U^i(\tau)$ are single-peaked for $w^i < w$, the equilibrium tax rate τ will be that chosen by the median voter. Inequality therefore affects the degree of redistribution: the higher the degree of wealth equality, as measured by the ratio of the median voter's wealth to average wealth, the higher the tax rate τ will be. Hence *in the absence of credit-market imperfections, more inequality* (in the sense of a lower ratio of median to average wealth) *will lead to more redistribution and therefore to lower growth.*

As we noted above, appreciating the effect of capital-market imperfections is crucial to understanding the relationship between inequality and growth. Is the result that greater inequality is harmful for growth robust to the introduction of capital-market imperfections? To address this question we should couple the political economy arguments just presented with the models developed in previous subsections. Consider, in particular, the opportunity-creation effect. Suppose that a lump-sum tax β is introduced, and that the tax rate is chosen by majority voting. As we already argued in subsection 2.1, this tax has no incentive effect. It, however, affects the individual's utility in two ways. There is a redistribution effect, that implies that those with wealth below average benefit from redistribution, those with average wealth are unaffected, while individuals for whom $w^i > w$ experience a reduction in their net wealth. There is a second effect that reflects the aggregate loss from investment inequality, which arises in the no credit-market case. This loss is itself a consequence of the assumption of decreasing returns to individual capital investments; to the extent that it affects aggregate knowledge A at any point in time, this cost of inequality is to be borne by *all* individuals, the poor *and* the rich, in the economy. In terms of the indirect utility function, $U^i(\tau) = V(\tau) + G(w^i/w,\tau)$, this means that now $V(\tau)$ is increasing in the tax rate. In particular, the individual with average wealth w will now vote for a positive tax rate because (1) the redistribution effect leaves his wealth unchanged, and (2) redistribution creates investment opportunities, increasing aggregate knowledge and therefore his income. The larger the degree of inequality, the more the median voter will benefit from the direct redistribution effect, and the higher his preferred tax rate will be. The overall impact of greater inequality on the growth rate is now ambiguous: on the one hand, it reduces growth, as seen in subsection 2.1, on the other, it results in a greater degree of redistribution and therefore faster growth.

A similar point is made by Saint-Paul and Verdier (1993), Glomm and Ravikumar (1992), and Perotti (1993) who analyze the voting process over public education spending aimed at circumventing wealth constraints on private education investments. In these papers redistribution takes the form of public education or education subsidies, while revenue is raised through a tax on the returns to investment. Consequently, redistribution has both a negative incentive effect and a positive *opportunity creation* effect. The rate of growth, and hence the term $V(\tau)$ in the indirect utility function, are a nonmonotonic function of the tax rate. When inequality is great, so that a large fraction of the population is constrained in their investments, the opportunity creation effect dominates; for more equal distributions, public education only slightly increases the number of agents that have access to education while it reduces the investment of a large part of the population through the incentive effect, resulting in a reduction in the growth rate. Since the tax rate is strictly increasing in the degree of inequality, the resulting relationship between wealth distribution and growth is U-shaped.

2.5 Discussion

The main conclusion we can draw from this section is that when we allow for heterogeneity among agents along with capital-market imperfections, the traditional argument that inequality has a positive impact on growth is strongly challenged. Consider, for example, the opportunity-enhancing effect. Our argument relies on three assumptions: first, that agents are heterogenous; second, that capital markets are highly imperfect; third, that the production technology exhibits diminishing returns to capital. These may look quite strong. However, there is, at least, one particular type of investment for which these assumptions clearly

hold: education.[6] Investments in human capital are characterized by strong diminishing returns. Moreover, borrowing in order to make such an intangible investment is usually expensive (if not impossible) and hence family wealth becomes a major determinant of the size of the investment. If we view k_t^i as an investment in education and g as the rate of growth of human capital (which in turn determines the rate of output growth, as argued by Lucas (1988)), then our analysis predicts a negative relationship between wealth inequality and the rate of growth.

The importance of moral hazard in determining individual actions is well-known, and subsection 2.2 has examined its consequences for the aggregate level of investment. We saw how a lump-sum tax and transfer system results in faster growth. Consider now a transfer system in which revenue is raised through distortionary (*ex-post*) taxation. In this case there are two incentive effects: the standard effect whereby taxation reduces net returns and hence lenders' incentive to invest, and moral hazard with wealth constraints which decreases the effort exerted by entrepreneurs whose projects are largely financed by borrowing. Whether redistribution increases or decreases the rate of growth then depends on whether the standard effect of taxes on those with high wealth is smaller or greater than the positive impact on the effort of those with low wealth levels.

The third aspect we have dealt with introduces a different, and much neglected, concept of inequality. It is not the distribution of wealth that we look at, but rather the social and institutional environment that affects access to investment projects. As we have seen, this institutional source of inequality will affect both the distribution of wealth and the rate of growth of the economy.

Overall, inequality actually proves *bad* for growth in several circumstances. Redistribution is then growth enhancing because it creates opportunities, improves borrowers' incentives and/or because it reduces macroeconomic volatility. In such instances, there is no longer a tradeoff between equity and efficiency goals, and policies designed to tackle one then have a

[6] A growing literature addresses how inequality affects growth through the possibilities of agents to invest in education. See Saint-Paul and Verdier (1993), Galor and Zeira (1993), Perotti (1993), and García-Peñalosa (1995).