IMPERFECT CAPITAL MARKETS AND PERSISTENCE OF INITIAL WEALTH INEQUALITIES

by

Thomas Piketty
Ecole Normale Supérieure – DELTA (Paris) and
London School of Economics and Political Science

Contents:

Discussion Paper
No. TE/92/255
November 1992

The Suntory Centre
Suntory and Toyota International Centres for Economics and Related Disciplines
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
Tel.: 020-7-955 6698

I am grateful to Philippe Aghion, Abhijit Banerjee, Patrick Bolton, Douglas Gale, Edward Green, Roger Guesnerie, Oliver Hart, John Moore and Kevin Roberts for their comments and encouragements.

Abstract

We consider an infinite-horizon intergenerational economy with identical agents differing only in their inherited wealth and with a constant-returns-to-scale technology using capital and labour (called "effort") and displaying a purely idiosyncratic risk. If effort is contractible, full insurance contracts make the production function deterministic and initial wealth inequalities cannot persist (just as in the neoclassical growth model). But if effort is not contractible the ability to commit is an increasing function of initial welath so that in equilibrium poorer agents face tougher credit rationing and take smaller projects (i.e. use less capital); although there is no poverty trap, the initial distribution may have long-run effects: there can be multiple long-run stationary distributions, and all are continuous and ergodic on the same interval, but have different equilibrium interest rates (and therefore different degrees of intergenerational mobility). This provides an explanation for wealth differentials within a country as well as countries, and a basis for redistributive policies with long-run effects.

Keywords: Wealth distribution, credit rationing, multiplicity.
JEL Nos.: D30, D63, D82, E43, E44.

Section 1 : Introduction.

Does the process of a market economy reduce wealth inequalities across generations, or does it merely reproduce them? Can wealth inequalities persist for ever and, in that case, does the long-run stable distribution depend on the initial distribution? This paper attempts to provide some answers to these fundamental questions in the context of a model with a standard two-factor constant-returns-to-scale technology and a fully-specified capital market imperfection of the moral-hazard variety.

We shall prove that not only wealth inequalities persist in steady-state (this is simply an incentive-compatible consequence of the moral-hazard problem) but also that the long-run wealth distribution may depend on the initial distribution; this multiplicity of long-run steady-states arises very naturally as a consequence of the dependence of credit-rationing and intergenerational mobility on the interest rate, via the following mechanism: if initially there is a large mass of agents at low wealth levels as compared to the mass of wealthier people, there will be a high demand for capital (poor people are would-be borrowers) and therefore a high interest rate, which itself, because of tougher credit-rationing, makes it more difficult for poor people to switch out from poverty, so that the wealth distribution remains as it was initially, and so does the interest rate; conversely, an economy with fewer poor people initially (as compared to the mass of potential lenders) will face a lower interest rate, social mobility will therefore be more important and inequality will remain low. With a higher long-run interest rate is associated a stationary distribution with more inequality (more exactly a distribution with a larger mass of agents at low wealth levels (i.e. potential borrowers) as compared to the mass of wealthier agents (i.e. potential lenders)) and less intergenerational social mobility.

It is worth noting that this multiplicity is not based upon a poverty trap (which would require a technological non-convexity): in this smooth constant-returns-to-scale economy, each stationary distribution is continuous and ergodic on the same interval (that is, any lineage can switch between any two wealth levels with a positive probability in a finite number of generations), but intergenerational mobility is less important in a stationary distribution with a higher interest rate (that is, the transition between a low to a high wealth level happens with a lower probability and takes longer anyway).

Such a steady-state multiplicity gives a rationale for strongly non-neutral government interventions: short-run redistributive policies can have long-run effects if they manage to shift the economy to a path converging to another stationary distribution (no matter how distortionary these short-run policies are; for example they can be subsidies to private borrowing and high wealth taxation). This contrasts heavily with usual neutrality results of macroeconomic models with a representative agent.

Also this multiplicity gives a new explanation for wealth differentials between similar economies (i.e. economies with the same technology and preferences): since the multiple stationary distributions are not Pareto-comparable (in a stationary distribution associated with a higher long-run interest rate there are at the same time some agents with higher utility levels and more agents with low utility levels; comparing them would involve the usual complications associated with poverty and inequality measurement), different countries can enjoy different steady-states depending on their distributive justice principles. This is in sharp contrast with an explanation based on a poverty trap and a technological non-convexity where the multiple steady-states can be Pareto-ranked (that is, the steady-state with nobody trapped in poverty Pareto-dominates any other steady-state; see for example Galor and Zeira(1991)), and therefore where only a (Paretian) government failure can explain why similar economies stay in different steady-states.

The rest of this paper is organized as follows: section 2 relates the paper to the recent theoretical literature on income distribution with imperfect capital markets and makes clear our modelling options; section 3 presents the model and proves that in the first-best economy (i.e. when the individual effort supply is verifiable) we get the very robust and unsurprising result of convergence to a unique single-point distribution, whatever the initial wealth distribution; section 4 defines the equilibrium of the second-best economy (i.e. when the effort supply is non-observable), analyses the set of incentive-compatible financial contracts supplied in equilibrium and proves that with an exogenously fixed interest rate the wealth distribution converges globally to a unique ergodic distribution; section 5 shows how this analysis can be extended to the case where the economy can also use a labour-consuming monitoring technology (so that safe-wage labour contracts can exist); our main result is in section 6, where we endogenize the interest rate and prove the dependence of the long-run distribution on the initial distribution with a (fairly general) example; section 7 provides concluding comments. The readers who are already familiar with moral-hazard-driven credit-rationing in an intergenerational general-equilibrium framework may want to go almost directly to section 6.

Section 2 : Relation to Existing Literature.

The idea that capital market imperfections play a crucial role in explaining the wealth distributions that we observe (and in particular their high and various degrees of inequality) is not new among economic theorists: Champernowne(1953) pointed out the role of individual income risk-bearing (even in the absence of any aggregate risk). Only recently however did formal modelling begin to reconcile economic theory with this evidence (in standard theory with first-best efficient capital markets indeed, risk-averse agents bear no risk and the wealth distribution converges to a unique single-point distribution (see the well-known Solow-Cass growth model); for a brief survey of the recent literature on income distribution with imperfect capital markets, see Aghion and Bolton(1992)); the seminal paper of Banerjee and Newman(1991) offers a formal model explaining how a stable wealth distribution can (and must) be non-Dirac: because of a moral-hazard problem in production, full insurance against idiosyncratic income risks would be worst for everybody (including the unlucky), and therefore the optimal contracts supplied by competitive capital markets involve partial insurance and the economy converges to a fully endogenous non-Dirac distribution of wealth.

There are however important limitations to Banerjee and Newman's model: firstly the (common) utility function is assumed to be unbounded below with respect to consumption, so that it is always possible for very poor agents to commit credibly to a high effort level (the bank knows that even though they do not have much to lose if their project fails, the loss in consumption utility can be made arbitrarily high), and consequently it is always possible for them to find on the market an incentive-compatible financial contract that enables them to invest in an activity requiring any initial capital investment (the unbounded below assumption is equivalent to an infinite liability in the sense of Sappington(1983)): this implies that there is in fact no credit-rationing at all. Even though Banerjee and Newman recognize that this is merely a trick that simplifies the analysis, this does not fit with the empirical fact that it is more difficult for poor people to raise capital for a project: in Banerjee and Newman(1991) exactly the contrary happens and intergenerational mobility is much higher than what empirical evidence suggests (for recent reappraisals of the extent of social mobility in the US, see Solon(1992) and Zimmerman(1992)).

Secondly, in their model the interest rate is completely exogenous (that is, the economy takes as given the world interest rate and can borrow or lend as much as it wants at that rate), so that the dynamics of the wealth distribution are Markovian (that is, the interest rate and therefore the amount bequeathed by an agent do not depend on the current distribution of wealth), and this enables Banerjee and Newman to establish global convergence to a unique distribution via ergodicity.

Aghion and Bolton(1991) goes a step further: they remove these assumptions in order to study the dual evolution of the interest rate and the wealth distribution during the development process in presence of a poverty trap: they point out that an economy starting with a low aggregate wealth may follow a (fully endogenous) Kuznets curve during its development path toward a (unique) high aggregate wealth steady-state.

As compared to these two studies, this paper has two main objectives: we keep the same basic structure of an intergenerational general-equilibrium model with moral hazard and extend it in the following directions. Firstly, a common unpleasant feature of Banerjee and Newman(1991) and Aghion and Bolton(1991) is that they consider a somewhat ad hoc technology: in both papers, the production activity requires a fixed initial investment (whose return distribution depends on the unobservable individual effort supply), so that one cannot produce anything with an initial capital investment slightly below the fixed requirement. As opposed to this strongly non-convex technology, we consider in this paper a much more usual technology, namely a smooth constant-returns-to-scale production function F(k,e) using capital and labour (called "effort"), whose only distinctive features are that F(k,e) is a random variable (the risk is purely idiosyncratic so that there is no aggregate risk) and the individual effort supply is not observable (or, more generally, not contractible). Although this makes the analysis of the equilibrium much less simple, we believe that this extension is important and not only technical: this makes much clearer the link and the contrast with usual models (if individual effort supply was contractible, full insurance contracts specifying the effort level to be taken would make the production function deterministic and the model fully equivalent to the standard neoclassical growth model; see section 3; more generally, a labour-consuming monitoring technology will make clear the gap with the zero-cost monitoring usually implicitly assumed; see section 5), in particular concerning the role of credit rationing and of initial wealth as a commitment in financial contracting (see section 4).

Secondly (and mainly, as emphasized in section 1), we prove that if the interest rate is determined by market-clearing on the market for capital (unlike in Banerjee and Newman(1991)) and like in Aghion and Bolton(1991)), then the long-run interest rate and wealth distribution may depend on the initial distribution; we argue that this is not a technical curiosity but rather a natural consequence of credit-rationing faced by poor people, which is itself a natural consequence of the commitment role played by initial wealth when individual effort supply is not contractible (see section 1 for an informal argumentation and the next sections for a formal proof). We insist on the fact that whereas it is fairly straightforward to obtain long-run effects of the initial distribution with a technological non-convexity and a capital market imperfection (since in that context poor people can be trapped for ever in poverty by credit rationing; see for example Galor and Zeira(1991)), the path-dependence that we point out in this paper is of a completely different (and much more robust) nature: since with an endogenous interest rate individual transitions do depend on aggregate variables (i.e. the interest rate and the current distribution), we can obtain multiple continuous stationary distributions with a common support, which cannot happen with linear Markovian dynamics (as noted by Banerjee and Newman(1992); see section 6 for a more precise discussion and for the relation between our results and those of Banerjee and Newman(1992)). We believe that the main advantage of the standard smooth constant-returns-to-scale technology that we consider in this paper is that it makes clear the specificity of the path-dependence we are concerned with (for example if the interest rate was exogenously fixed the wealth distribution would converge to a unique ergodic distribution, which would not be the case in general with a non-convex technology; see section 4).

Section 3 : Verifiable effort supply: Convergence to a unique Dirac distribution.

We consider an economy with a continuum of mass one of identical agents living for one period during which they earn income by supplying labour and capital; the resulting income is divided at the end of the lifetime between their own consumption and a bequest for their offspring; there are two goods in this economy: one labour good called "effort" and one physical good that can be used as a capital good or as a consumption good; we shall consider a very simple demographic structure: each agent has exactly one offspring and generations succeed each other ad infinitum. The distribution of wealth at the beginning of period t (resulting from the bequests left by generation t-1) is represented by the distribution function $f_t(w)$ defined over the set of non-negative wealth levels $w \geq 0$; the aggregate wealth at period t W_t is given by:

$$W_t = \int_{w \geq 0} w df_t(w)$$

We assume agents to be risk-averse. Following Andreoni(1989), Banerjee and Newman(1992) and Aghion and Bolton(1991), we assume preferences to be defined directly over (consumption c,bequest b,labour e) bundles; using Andreoni's terminology this means that the motive for letting bequest is of the "warm glow" variety (see Andreoni(1989)); anyway, it is worth noting that any theory which attempts to explain the existence of bequests will eventually generate indirect preferences defined over (consumption c,bequest b,labour e) bundles (*1). Without any significant loss in generality, we also assume preferences to be quasi-linear with respect to (consumption c,bequest b) bundles and labour e (*2):

$$U = U(c,b,e) = W(c,b) - e$$

(with $W(c,b)$ concave and monotonic, and $W(0,0)=0$)

At the end of their lifetime, the agents allocate their total (capital and labour) income i between consumption c and bequest b;

their indirect utility for income i is given by:

$$U(i) = \text{Max } W(c,b) \text{ under } c+b \leq i$$

$$(U \text{ is concave and monotonic, } U(0)=0)$$

Although the results of this paper could be easily extended to more general bequest behaviour, we follow Banerjee and Newman(1992) and Aghion and Bolton(1991) and assume that the agents bequeath a fixed fraction of their total income:

$$W(c,b) = U(hc^{\alpha}b^{1-\alpha})$$

$$(0 \leq \alpha \leq 1,\ h=\alpha^{-\alpha}(1-\alpha)^{\alpha-1})$$

During each period of this economy, the agents can use capital (i.e their inherited wealth) and labour to produce output (output can be used as a consumption good and as a capital good); they can also choose to store their wealth at zero-cost and to allocate it between consumption and bequest at the end of their lifetime. The productive technology of the economy is entirely described by a stochastic gross production function F(k,e) which (as a random variable) exhibits constant returns to scale with respect to capital k and labour e:

$$F(k,e) = rk \text{ with probability } p=g^{-1}(e/k)$$

$$F(k,e) = 0 \text{ with probability } 1-p$$

$$(\text{with } g \text{ defined on } [0;1[\ (g=\infty \text{ above } 1),\ g(0)=0,\ g'>0,\ g''\geq 0)$$

We choose this very rudimentary stochastic structure because it greatly simplifies the analysis of the optimal incentive-compatible contracts of the second-best economy (since for a fixed k total production can take only two values), but in fact the only important features of the production are the following: the production is uncertain (otherwise the unobservable individual effort supply can be exactly inferred from the observable output), there is a positive probability that the net return be negative (otherwise there is no credit-rationing), for any capital investment k the probability that an effort level e was taken over the probability that an effort level e'<e was taken is an increasing function of the return (this is the usual monotone-likelihood ratio property of moral hazard literature which guarantees the monotonicity of the optimal incentive-compatible contracts; for a general analysis of optimal contracts with moral hazard, see Grossman and Hart(1983)).

Because of the constant-returns-to-scale assumption we can study the production activity at the individual level, and we assume the risk to be i.i.d. across individual projects (so that there is no aggregate risk). One can interpret the individual investments as entrepreneurial investments as well as educational investments; note also that the risk can come from pure luck (an entrepreneurial project is intrinsically risky, or uncertainty about future demands for specific kinds of human capital) as well as from differences in ability of which one is not aware ex ante (this is formally equivalent).

When the effort level taken by an individual agent is verifiable, the analysis of the competitive equilibrium of this economy is straightforward. Since the agents are risk-averse and there is no aggregate risk, in equilibrium competitive insurance companies (operating at zero administrative cost) supply complete insurance contracts to effort-suppliers; this means that if the current interest rate is $A \geq 1$ an agent putting an effort e in a project of size k will get a safe return $E(F(k,e)-Ak)=p(e,k)rk-Ak$ by signing a insurance contract specifying that the effort level e is to be taken ($E()$ is the expectation operator); an agent supplying his wealth w to the capital markets gets a safe return Aw (at the end of the period). Thus as long as individual effort supply is contractible the economy we consider is fully equivalent to an economy with a deterministic constant-returns-to-scale production function $E(F(k,e))$ (for example if $g(p)=p^{1/a}$ with $0<a<1$ we get a Cobb-Douglas production function $E(F(k,e))=k^{1-a}e^{a}$), and therefore the model degenerates to a very standard model.

Therefore at the beginning of any period t an agent with inherited wealth w facing an interest rate A chooses an effort level $e_{FB}(w,A)$ and a project size $k_{FB}(w,A)$ by solving the following maximization programm:

$$(e,k)_{FB}(w,A) = \text{ArgMax} \;\; U(i)-e \text{ under } 0 \leq i \leq rp(e,k)k+A(w-k),\ e,k \geq 0.$$

This is equivalent to choose a project size $k_{FB}(w,A)$ and a probability of success $p_{FB}(w,A)$ satisfying

$$(k,p)_{FB}(w,A) = \text{ArgMax } U(rpk+A(w-k)) - kg(p)$$
$$\text{under } k\geq 0, 0\leq p\leq 1, rpk+A(w-k)\geq 0.$$

Note that since the production function exhibits constant returns to scale $k_{FB}(w,A)/e_{FB}(w,A)$ does not depend on w: this is the unique optimal capital/labour ratio (which only depends on A, as usual). We have $e_{FB}(w,A)=k_{FB}(w,A)g(p_{FB}(w,A))$, $c_{FB}(w,A)=\alpha i_{FB}(w,A)$, $b_{FB}(w,A)=(1-\alpha)i_{FB}(w,A)$, where $i_{FB}(w,A)=i_{lFB}(w,A)+i_{kFB}(w,A)$ is total income, $i_{lFB}(w,A)=(rp_{FB}(w,A)-A)k_{FB}(w,A)$ is labour income and $i_{kFB}(w,A)=Aw$ is capital income. The equilibrium interest rate A_t must equalize the demand and the supply for funds in the economy at generation t (the supply of funds at period t being simply the aggregate wealth at the beginnig of period t):

$$W_t = \int_{w\geq 0} k_{FB}(w,A_t)\text{beginning}_t(w) \quad \text{if } A_t>1$$
$$W_t \geq \int_{w\geq 0} k_{FB}(w,A_t)\text{beginning}_t(w) \quad \text{if } A_t=1$$

(recall that it is possible to store wealth at zero-cost instead of using it as a capital input)

The demand for capital being naturally a decreasing function of the interest rate, there exists a unique competitive equilibrium. Since leisure is assumed to be a normal good, agents with a smaller initial wealth supply a higher amount of labour (or equivalently choose a higher project size) in equilibrium. This is expressed by the following proposition:

Proposition 1 : For any wealth distribution at period t f_t there exists a unique competitive equilibrium $A_t = A_{FB}(f_t)$ for the first-best economy at period t. Moreover, there exists $w_0 > 0$ such that

(a) for $w \geq w_0$ $k_{FB}(w) = e_{FB}(w) = 0$, for $w < w_0$ $k_{FB}(w), e_{FB}(w) > 0$ and $k_{FB}(w), e_{FB}(w)$ tend to 0 as w tends to w_0^-.

(b) for $w < w_0$ $k_{FB}(w)$ and $e_{FB}(w)$ are decreasing functions of initial wealth w.

(c) for $w < w_0$ $i_{FB}(w)$ does not depend on w.

Proof : see the appendix.

Note that (a) and (b) of proposition 1 do not depend on the particular assumptions we made on preferences: only the normality of leisure is used. One can be sure that w_0 is finite by assuming that U'(i) tends to 0 as i tends to infinity: if the marginal indirect utility of income tends to be arbitrarily small then agents above a finite initial wealth level w_0 will choose not to work at all and to become rentier. Part (c) of proposition 1 however relies on the quasi-linear form of the utility function;

this has the only effect of reinforcing the convergence result.

We now turn to the dynamics of the wealth distribution. For the first-best economy as for the second-best economy, wealth distribution dynamics are completely given by the study of the mapping $b(w,A_t)$ which gives the offspring's initial wealth as a function of the parent's initial wealth. Proposition 1 gives us that $b_{FB}(w,A_t)$ is horizontal until w_0 and then linear (see figure 1); therefore if $(f_t(w),w\geq 0)$ is such that $f_t(w)=0$ for $w>w_0$, then $(f_{t+1}(w),w\geq 0)$ is a Dirac distribution; more generally, it is straightforward that f_t converges to the Dirac distribution 1_{w*} as t tends to infinite, w* being the (necessarily unique) wealth level such that $b_{FB}(w*,A_{FB}(1_{w*}))=w*$:

Proposition 2 : For any initial distribution f_0 , the wealth distribution f_t converges to the Dirac distribution 1_{w*} as t tends to infinite.

Proof : follows directly from proposition 1 and the remarks above.

Therefore as long as the effort level supplied by individual agents is observable initial wealth inequalities cannot persist (every lineage of every country should converge to w*). It is worth noting that this idea that market forces imply convergence of wealth levels is exactly the same intuition as that of the standard Solow-Cass "growth" model (for which it is well-known that initial wealth inequalities cannot persist). Needless to say, there is very little empirical evidence to support such a view of the development process (for LDCs as well as for developed countries).

Section 4 : Non-verifiable effort supply: Optimal Second-Best Contracts and Credit Rationing.

If the individual effort supply is not observable (more precisely, if a contract specifying an effort level to be taken is not enforceable), then the economy is no longer equivalent to an economy with a deterministic production function; indeed, if an insurance company offers a full insurance contract to entrepreneurs (i.e. an income that does not depend on whether the project succeeds or fails), then it is a dominant strategy for any of these entrepreneurs to choose an effort level $e=0$, and consequently the insurance company will suffer a loss (the probability of success of a project $p(e,k)$ is equal to 0 for $e=0$) once the fixed payment offered to the entrepreneurs is positive. Therefore insurance companies will choose to offer financial contracts involving partial insurance in order to create incentives for effort and to make non-negative profits.

What will the competitive equilibrium look like in the presence of this moral hazard problem? Even though this is obviously a very important question, our objective in this paper is not to deal with the difficult issue of competition with moral hazard, and we shall just note that the (second-best) efficiency of the competitive equilibrium in such a world is known to be very problematic (see Arnott and Stiglitz(1990); roughly speaking, one has to assume that an exclusivity clause is enforceable by the financial institutions in order to get equilibrium efficiency). In this paper we assume that competitive markets are second-best Pareto-optimal, i.e. that the optimal incentive-compatible contracts are supplied.

Consider an individual with initial wealth w who is willing to undertake a project of size k and who is facing an interest rate A. A financial contract is a couple (i_s, i_f), where i_s is the payment made to the entrepreneur if the project is successful and i_f the payment if the project fails. In the first-best economy we had:

$$i_s = i_f = rp(e,k)k - Ak$$

where e was the effort level written in the contract and to be taken by the entrepreneur. But in the second-best economy the entrepreneur cannot commit to an effort level by writing it in the contract, and financial markets anticipate that facing the contract (i_s,i_f) the entrepreneur will choose a probability of success $p(i_s,i_f,k,w,A)$ solving the following programm:

$$p(i_s,i_f,k,w,A) = \text{Argmax } pU(i_s+Aw) + (1-p)U(i_f+Aw) - kg(p)$$
$$\text{under } 0 \leq p \leq 1.$$

Note that this programm is well-defined only if $\min(i_s,i_f)+Aw \geq 0$ (this limited liability constraint comes simply from the requirement that neither bequest nor consumption can be strictly negative: one cannot credibly commit for one's offspring to pay back a loan); by convention we put $p(i_s,i_f,k,w,A)=0$ when $\min(i_s,i_f)+Aw<0$.

A financial contract (i_s,i_f) is said to be incentive-compatible if an insurance company offering it makes a non-negative expected profit, i.e. if

$$rp(i_s,i_f,k,w,A)k - Ak \geq p(i_s,i_f,k,w,A)i_s + (1-p(i_s,i_f,k,w,A))i_f$$

We define $C(k,w,A)$ as the set of incentive-compatible financial contracts for an agent with initial wealth w who is willing to undertake a project of size k when the market interest rate is A; note that $C(k,w,A)$ can be empty, in particular because of the limited liability constraint. Obviously for A>A' we have

$$C(k,w,A) \subset C(k,w,A')$$

Following the discussion above, in equilibrium only the contracts of $C(k,w,A)$ maximizing the utility of the agents will be traded; we note $k_{SB}(w,A)$ the size of the project chosen by an agent with initial wealth w and $(i_s,i_f)(w,A)$ the associated contract (which is also the (stochastic) labour income $i_{lSB}(w,A)$):

$$(k_{SB},i_s,i_f)(w,A) = \text{ArgMax } p(i_s,i_f,k,w,A)U(i_s+Aw) + (1-p(i_s,i_f,k,w,A))U(i_f+Aw) - kg(p(i_s,i_f,k,w,A))$$
$$\text{under } (i_s,i_f) \in C(k,w,A)$$

We also define $p_{SB}(w,A)=p(i_s(w,A),i_f(w,A),k_{SB}(w,A),w)$. The wealth bequeathed by this agent to his offspring is no longer deterministic; we note $b_{SBs}(w,A)$ the bequest left by this agent if his project succeeds and $b_{SBf}(w,A)$ if his project fails:

$$b_{SBs}(w) = i_s(w,A)+Aw$$
$$b_{SBf}(w) = i_f(w,A)+Aw$$

The equilibrium interest rate $A_t=A_{SB}(f_t)$ is given by the same equilibrium condition as in the first best economy:

$$W_t = \int_{w\geq 0} k_{SB}(w,A_t)\ wdf_t(w) \quad \text{if } A_t>1$$
$$W_t \geq \int_{w\geq 0} k_{SB}(w,A_t)\ wdf_t(w) \quad \text{if } A_t=1$$

Note that although the amount bequeathed by an individual agent is stochastic, the wealth distribution at period t+1 f_{t+1} is a deterministic function of the wealth distribution at period t f_t since there is a continuum of agents (there is no aggregate uncertainty); the crucial difference between linear and non-linear dynamics referred to in section 1 is the following: if the interest rate A is fixed exogenously the mapping $f_{t+1}(f_t)$ is linear (and the dynamics can be analyzed at the level of a single lineage), whereas if A is determined by market-clearing the mapping $f_{t+1}(f_t)$ is no longer linear (that is, the transition rules for a single lineage depend on the current distribution). Since non-linear dynamics with an infinite-dimensional state space become very quickly intractable, we begin the analysis with an exogenously fixed interest rate (say that until section 5 we consider a small open economy taking the world interest rate as given).

Before that we start the analysis of this equilibrium and of the wealth distribution dynamics, it is worth noting that these financial contracts have a very intuitive interpretation in terms of corporate finance and incentives: risk-aversion and limited initial wealth imply that for Pareto-efficiency reasons an entrepreneur would always prefer to use only outside finance for his project, i.e. to go to the equity market and issue shares; but in order to create incentives for effort-taking the market will accept to buy his shares only if he purchases himself a sufficient part of them. Therefore poor agents will be unable to raise capital for a large project, since by definition they are not able to purchase the shares necessary to convince the market that they will take a sufficiently high effort level.

The first step to characterize the equilibrium of this economy is to study the structure of the sets $C(k,w,A)$ of incentive-compatible contracts. As it was suggested above, there exists for every project size a minimal initial wealth below which there does not exist any incentive-compatible contract:

Proposition 3 : There exists a mapping $w_0(k,A)$ such that $C(k,w,A)$ is non-empty if and only if $w \geq w_0(k)$. Moreover $w_0(k)$ is given by

$$w_0(k) = k - (\ \mathrm{Max}_{x\geq 0}\ g'^{-1}(U(x)/k)(rk-x)\ )/A$$

Proof : see the appendix.

In order to show with a simple example what sort of credit-rationing curve proposition 3 implies, we give the following corollary of proposition 1:

Corollary : For $U(i)=(1-\exp(-qi))/q$ $(q>0)$, $g(p)=ap^2/2$ and a,A,r such that $r/2a<1$ and $r^2/4aA<1$, there exists a concave and increasing mapping $k_0(w,A)$ $(=w_0^{-1}(w,A))$ such that

(a) $C(k,w,A)$ non-empty iff $k\leq k_0(w,A)$

(b) $k_0(0,A)=0$

(c) $k_0'(0,A)=1/(1-r^2/4aA)$

Proof : see the appendix.

It is worth noting that if U( ) was unbounded below (i.e. U(i) tends to minus infinity as i tends to 0^+)the sets C(k,w,A) would always be non-empty: by giving an arbitrarily small consumption to the entrepreneur when the project fails it is always possible to induce an arbitrarily high effort level; however this seems extremely unrealistic, and assuming U( ) to be bounded below seems more appropriate.

In that case the limited liability constraint puts a finite upper bound on the project size that an agent can undertake: for example the corollary says that with a CARA utility function and a quadratic cost function and for a large subset of parameters arbitrarily poor agents have the opportunity to take only arbitrarily small projects, so that (at least below a certain wealth level) the competitive equilibrium will exhibit exactly the opposite property than when the effort level to be taken could be written in a contract: the agents with a higher initial wealth will take a larger project and supply more labour, even though leisure is a normal good and technology is constant-returns-to-scale. Note that this does not occur because of differences in relative risk aversion between poor and rich agents; this occurs because of the limited liability constraint: the agents with a low initial wealth have very limited commitment possibilities, and consequently the competitive financial markets lend them funds only for small projects. Note also that $k_{SB}(w,A)$ cannot in general be an increasing function of w over $[0;\infty[$: for very rich agents, it is even less profitable to work in the second-best economy than in the first-best economy, so that those with an initial wealth above w_0 (see proposition 1) still choose to become rentiers; in other words for very wealthy agents the normality of leisure continues to play its role in the second-best economy, whereas for less wealthy agents the credit rationing due to the limited liability constraint reverse dramatically the effect of the normality of leisure by giving more earning opportunities to wealthier agents(see figure 3).

We now turn to the analysis of the dynamics of the wealth distribution (we are still assuming a fixed interest rate). In the same manner as for the first-best economy this can be done simply by looking at the mappings $(b_{SBs},b_{SBf})(w,A)$. The most striking difference with the first-best case is that now a stationary distribution can no longer be a single-point distribution. Moreover, given the non-convexity of the incentive-compatibility constraints, ergodicity is a priori not guaranteed (see figure 3 for an example of non-convex mappings implying the existence of multiple ergodic sets, noted $[w_1,w_2]$ and $[w_3,w_4]$ on the figure). However, these perverse outcomes are ruled out by the following proposition, which is essentially a consequence of normality of leisure (that is, richer agents have the opportunity to take larger projects only if they accept to pay a higher amount out of their capital income than poorer agent in the contingent state where they are unlucky; otherwise they have no more opportunities than poorer agents and the normality of leisure implies that they supply less labour):

Proposition 4 : For any $w \geq w' \geq 0$, $i_s(w',A) \geq i_f(w,A)$.

Proof : see the appendix.

If we restrict our attention to the utility and cost functions considered in the corrolary of proposition 3 (*3), proposition 4 implies that the mappings $(b_{SBs},b_{SBf})(w)$ look as depicted on figure 4; therefore the dynamics are ergodic, and the distribution of wealth converges to a unique invariant distribution, whatever the initial distribution:

Proposition 5: There exists $w^{**}>0$ and a unique wealth and continuous distribution f^* defined on the interval $[0;w^*]$ such that whatever the initial distribution f_0, the wealth distribution at period t f_t converges to f^* as t goes to infinity.

Proof : see the appendix.

Therefore with a fixed interest rate and a constant-returns-to-scale technology, the main effect of moral-hazard is that the long-run wealth distribution is non-Dirac, but the latter is still independent of the initial distribution. By investing in larger and larger projects, any lineage can escape poverty with a positive probability. Note however that this is not true in general if there is a technological non-convexity: if there is a minimal project size lineage who are too poor initially remain poor for ever with probability one (see Aghion and Bolton(1991) and, with a more brutal kind of capital market imperfection, Galor and Zeira(1991)).

However, the fact that the long-run distribution of wealth is ergodic must not hide the fact that in such an economy, in spite of the constant-returns-to-scale technology, getting out from poverty takes a long time and is highly unlikely, even though it happens with a positive probability (by definition of ergodicity); proposition 6 makes precise this idea:

Proposition 6 : For any wealth level $w>0$ and integer T, there exists a wealth level $w(w,T)>0$ such that with probability one a dynasty starting with an initial wealth below $w(w,T)$ will still be below w after T generations.

Proof : see the appendix.

Section 5 : Monitoring Technology: Safe Wages and Poor Workers.

A somewhat unpleasant feature of the technology considered so far is that there is no room for safe wages: since the individual labour supply is not observable at all, such a contract would be worst for everybody. An extreme consequence is that, via credit-rationing, the individual effort supply in equilibrium tends to zero as the individual wealth tends to zero (since arbitrarily poor agents can only borrow arbitrarily poor amounts and therefore invest in arbitrarily small projects); note that if one further assumes that there exists an indivisibility in labour supply (no matter how small), this gives an easy way to obtain involuntary unemployment. However, this does not seem to fit at all with the fact that safe wages do exist in the real world. A natural way to extend our model such that safe wage contracts can exist is to assume that there exists a labour-consuming monitoring technology: note that this is simply a median way between the first-best economy (where individual labour supply is implicitly assumed to be observable at no cost) and the second-best economy (where the cost of observing individual labour supply is implicitly assumed to be infinite). In this section, we briefly sketch how the previous analysis can be extended to this case.

The monitoring technology is assumed to be the following: by spending e units of effort, an agent can monitor h(e) units of effort spent in the productive activity (for simplicity, we assume that it is not possible to monitor the monitors). The function h is assumed to be concave, and h(0)=0. An agent supplying some labour which is monitored by another agent is called a worker. Monitored labour is paid at a safe wage rate v>0 (since agents are risk-averse and there is no aggregate risk it would be costly and useless to make v dependent on whether the worker is lucky or not).

Facing an interest rate A and a wage rate v, an agent with initial wealth w must now choose the size k(w,A,v) of the project on which he wants to be self-employed, the quantity of labour $D(w,A,v)$ that he wants to monitor and the quantity of labour $S(w,A,v)$ that he supplies on the market for monitored labour; the equilibrium interest rate $A_t=A(f_t)$ and wage rate $v_t=v(f_t)$ are given by market-clearings on the market for capital and monitored labour (the formal definitions are left to the reader).

Naturally, whether in equilibrium some agents choose to become a worker (i.e. to supply a quantity $S(w,A,v)$ of monitored labour) depends on the profitability of monitoring over self-employment, i.e. on the function $h(e)$. However, we can be sure that if the monitoring technology is profitable at all (that is, a positive quantity of monitored labour is supplied in equilibrium) then sufficiently poor agents will become workers (since with self-employment they can only supply an arbitrarily small quantity of labour):

Proposition 7: If monitored labour is traded in equilibrium then there exists $w'>0$ such that $D(w,A_t,v_t)>0$ for any $w<w'$.

Proof : see the appendix.

Note that the way we obtain that poor people become workers monitored by wealthier people is very similar to the approach first developed by Newman(1991) and Banerjee and Newman(1992): because of informational asymmetries in financial contracting, there is some credit-rationing so that only wealthy people can supply the initial investment required to start up a business, whereas poor people can only work for a safe wage (which does not require any initial investment). This view of occupational choice is substantially different from that based on decreasing absolute risk aversion (developed by Kihlstrom and Laffont(1979) and challenged by Newman(1991)).

If we assume that both A and v are exogenously fixed (say that we are in a small open economy and capital as well as monitored labour are perfectly mobile), then the dynamics of the wealth distribution with the monitoring technology look very similar to the case analyzed in the previous section: individual transitions do not depend on aggregate variables, and by using the same argument as proposition 4 one can prove that there cannot exist multiple ergodic sets, so that the wealth distribution converges to a unique distribution, whatever the initial conditions.

Section 6 : Long-run Effects of the Initial Distribution.

We have established that when the interest rate and the wage rate are fixed then the initial distribution has no long-run effects. We now show why this is no longer the case when they are determined by market forces.

The case where the interest rate is fixed and the wage rate is endogenous is that considered by Banerjee and Newman(1992); they prove that in that case the long-run wage rate (and therefore the long-run distribution) may depend on the initial distribution via the following mechanism: an initial distribution with many poor agents will lead to a high supply of monitored labour, and therefore to a low wage rate, which in turn implies that a large mass of poor people will remain wage-earners (i.e. monitored-labour suppliers); conversely, a less unequal initial distribution leads to a higher wage rate and more social mobility.

In this section, we shall consider the other polar case where the wage rate is exogenous and the interest rate is endogenous, and prove that in the same manner the long-run interest rate (and therefore the long-run ergodic distribution) may depend on the initial distribution (naturally, both cases could be combined to obtain a dependence of both the long-run wage rate and the long-run interest rate on the initial distribution). In order to make the path dependence as clear as possible, we make the following assumptions:

$$U(i) = i \qquad (1)$$

$$\text{For } e>e_0,\ U(i,e)=-\infty \qquad (2)$$

$$g(p)=ap \text{ for } 0\leq p\leq q,\ b(p-q)+aq \text{ for } q\leq p\leq q',\ \text{and } +\infty \text{ for } p>q' \ (b>a) \qquad (3)$$

$$r>b, \ (q'r-1)/g(q')>1 \qquad (4)$$

$$v<(q'r-1)/g(q') \qquad (5)$$

The roles of the different assumptions are the following: (1) makes the optimal incentive-compatible contracts computable; (2) implies that agents will not supply arbitrarily high quantity of labour (so that the distribution has a compact support); (3) and (4) imply that in equilibrium every agent chooses the same probability of success on his project $p=q'$ and that for any $A\geq 1$ there is some credit-rationing (see the appendix); (5) makes sure than being self-employed is more profitable than being a worker (as long as the funds are available). Note that none of these assumptions is really restrictive (in particular the piecewise-linearity of g could be easily removed), but they make the model analytically more tractable.

Under these assumptions, an agent with initial wealth w cannot borrow more than $k_0(w,A)=w/(1-q'(r-b)/A)$ for any interest rate $A\geq 1$ (see the appendix): we are therefore in the case of fully binding credit-rationing. We prove the dependence of the long-run distribution on the initial distribution for a large subset of parameters:

Proposition 8 : There exists $p_1<p_2$, $\alpha_1<\alpha_2$, $w_1<w_2$ such that if $0<q<p_1<q'<p_2<1$ and $\alpha_1<\alpha<\alpha_2$ then there exists two interest rates $A_1^*<A_2^*$ and two distributions f_1^* and f_2^* such that:

(a) if $f_0(w_2)<1-q'^2$, then (f_t,A_t) converges to (f_1^*,A_1^*).

(b) if $f_0(w_1)>1-q'^3$, then (f_t,A_t) converges to (f_2^*,A_2^*).

Proof : see the appendix.

Note that in general there may exist more than two possible long-run interest rates, although we only prove that if in the initial distribution there are too many poor people as compared to the mass of wealthy people then the economy will converge to a high-interest rate stationary distribution, and conversely; the intuition of the result can be understood very simply by looking at the figures 5 and 6: if the interest rate is initially high (because of a too large mass of borrowers as compared to the mass of potential lenders), credit-rationing is very tough and the transition from a low wealth level to a high wealth takes time (at least three generations; see figure 6) and happens with a low probability (being lucky during three consecutive generations happens with probability q'^3); this implies that the mass of borrowers will remain high as compared to the mass of lenders so that the interest rate will remain high; conversely, if the interest rate is initially low, intergenerational mobility is more important (in two generations and with probability q'^2 it is possible to switch from a low wealth level to a high wealth level; see figure 5), and the interest rate will remain low. It is worth noting that what matters is really the initial distribution and not the initial aggregate wealth: an economy with a high aggregate wealth can very well converge to the high-interest rate steady state if the initial distribution is sufficiently unequal, and conversely. Note also that although in our example the more unequal distribution has a smaller aggregate wealth, in general it is unclear which of the different stationary distributions displays the highest aggregate wealth (there are wealthier people in a stationary distribution associated to a higher interest rate but there are also more poor people): the example we point out in section 8 is only one particular case of the general phenomenon of steady-states multiplicity due to the dependence of individual transitions on aggregate variables (that is, on the current wealth distribution, via the interest rate), although we believe this example is particularly intuitive.

Although the link and the complementarity between our results and Banerjee and Newman(1992)'s results are clear enough, there are other important differences. Firstly, a (minor) unpleasant feature of Banerjee and Newman(1992)'s model is the (somewhat crude) capital market imperfection they consider: a borrower can flee and succeed in escaping the lender with an exogenous probability $1-\pi$; if the borrower is caught he incurs an exogenous non-monetary punishment F (note that if π is equal to zero there is no credit market at all); the problem with this capital market imperfection is that it is not enough to generate the kind of steady-states multiplicity we are concerned with: one needs also uncertainty in individual income (otherwise with a convex technology the distribution converges to a unique Dirac distribution), and the only robust way to obtain it is a moral-hazard problem in production (Banerjee and Newman(1992) obtains it by assuming that agents are risk-neutral and choose not to insure against their idiosyncratic income risk); although one may think that these are merely technical details, we believe that it makes sense to try to see what is needed and what is implied by such or such precise kind of capital market imperfection (in particular as far as the analysis of possible intervention in the banking system is concerned).

Secondly, as we argued in section 2, an important difference between this paper and Banerjee and Newman(1992) is that we consider a standard convex technology whereas their technology is based on multiple non-convexities; it seems to us that this is essential to make clear that the intuition behind the long-run effects of initial conditions that this paper and Banerjee and Newman(1992) point out does not rely at all on any technological non-convexity (as opposed to other studies such as Galor and Zeira(1991)), nor on a "pecuniary" non-convexity (Banerjee and Newman suggest that what they need for their path-dependence is some "increasing returns to wealth", as opposed to technological increasing returns; but in fact this is not necessary: in our model the credit-rationing curve $k_0(w)$ is perfectly smooth and concave because of the smooth constant-returns-to-scale technology).

Section 7 : Concluding Comments.

In the context of an inter-generational general-equilibrium model with a standard constant-returns-to-scale production function, we have shown how removing the usual implicit assumption of zero-cost monitoring of individual labour supply modifies dramatically the conclusions regarding the long-run distribution of wealth.

Firstly, the necessity to provide incentives implies partial insurance contracts and credit-rationing, so that the long-run wealth distribution is non-degenerate. Because the initial wealth plays an essential commitment role in financial contracting, and in spite of the ergodicity of the long-run distribution, a lineage can stay for very long periods at very low wealth levels.

Secondly and mainly, the extent of social mobility being dependent on the toughness of credit-rationing, and therefore on the interest rate which itself depends on the relative mass of potential borrowers and lenders (that is, on the distribution), the long-run interest rate and the long-run distribution of wealth may depend on the initial distribution. Roughly speaking, a higher long-run interest rate is associated with a less mobile and more unequal distribution.

With the constant-returns-to-scale technology considered in this paper there is no room for self-sustained growth (exactly as in the Solow-Cass model). The impact of growth on the distributional issues discussed so far depends obviously on the mechanism that makes long-run growth sustainable. If long-run growth is due to an economy-wide externality, then this will certainly be neutral with respect to the long-run relative inequality. But if the externality is more sectoral (say, a sector is an interval of project sizes), then the dynamics of the wealth distribution may become even less trickle-down. More precise analysis must await further research.

Appendix.

Proof of proposition 1 (section 3):

The maximization program of the agents is:

$$(k,p)_{FB}(w,A) = \text{ArgMax } U(rpk + A(w-k)) - kg(p)$$
$$\text{under } rpk+A(w-k)\geq 0, k\geq 0, 0\leq p\leq 1.$$

This is a standard concave maximization program with a unique and continuous solution (except at the corner solution $k=0$, for which every p is equally optimal as $e=kg(p)=0$ anyway), the first-order conditions of which are:

$$(rp-A)U'(rpk+A(w-k)) = g(p) \qquad (1)$$
$$rkU'(rpk+A(w-k)) = kg'(p) \qquad (2)$$

If $k_{FB}(w,A)>0$, (2) becomes:

$$rU'(rpk+A(w-k)) = g'(p) \qquad (2')$$

(1) and (2') give:

$$p - g(p)/g'(p) = A/r \qquad (3)$$

(3) has a unique solution, and therefore $p_{FB}(w,A)$ does not depend on w as long as $k_{FB}(w,A)>0$. Coming back to (2') this implies that $i_{FB}(w,A)=rp_{FB}(w,A)k_{FB}(w,A)+A(w-k_{FB}(w,A))$ does not depend on w once $k_{FB}(w,A)>0$ (this proves part (c) of proposition 1), and consequently that $k_{FB}(w,A)$ is a strictly decreasing function of w, and $e_{FB}(w,A)$ a strictly increasing function of w (this proves part (b)). Moreover $k_{FB}(w,A)=0$ for $w\geq w_0$ with

$$(rp^*-A)U'(w_0) = g'(p^*)$$

where p^* is the unique solution of (3) (this proves part(a)). Finally there exists a unique equilibrium interest rate since $k_{FB}(w,A)$ is a decreasing function of A (and the supply of funds is inelastic for $A\geq 1$). CQFD.

Proof of proposition 3 (section 4):

$p(w,k,i_s,i_f,A)$ solves the following concave program:

$$p(w,k,i_s,i_f,A) = \text{ArgMax } pU(i_s+Aw)+(1-p)U(i_f+Aw)-kg(p)$$
$$\text{under } 0 \le p \le 1.$$

The first-order condition is:

$$U(I+Aw+srk)-U(I+Aw) = kg'(p) \qquad (4)$$

There is no aggregate risk in this continuum economy, therefore what matters is the expected return of a contract and $C(k,w)$ is empty if and only if for any (i_s,i_f)

$$R(w,k,i_s,i_f,A) = p(w,k,i_s i_f,A)(rk-i_s+i_f)-i_f-Ak+Aw < 0$$

Equation (4) and $U''<0$ imply that the maximum of $R(w,k,i_s,i_f,A)$ with respect to i_f is reached for $i_f=-Aw$. We have:

$$R(w,k,i_s,i_f=-Aw,A) = g'^{-1}(U(i_s)/k)(rk-i_s)-Ak+Aw$$

Therefore we have

$$(\ C(k,w,A) \text{ empty }) \text{ iff } (\ w < w_0(k,A)\ )$$
$$\text{with } w_0(k,A) = k - (\ \text{Max}_{x\ge 0} g'^{-1}(U(x)/k)(rk-x))/A$$

This completes the proof. CQFD.

Proof of the corrolary of proposition 3 (section 4) :

For $U(i)=(1-\exp(-qi))/q$ and $g(p)=ap^2/2$, $x \dashrightarrow g'^{-1}(U(x)/k)(rk-x)$ is a concave function, the maximum of which is reached for $x(k)$ given by the following equation (as long as the constraint $p\le 1$ is not binding):

$$rk-x(k) = (\exp(qx(k)-1)/q$$

Since $\exp(qx)-1>qx$, the assumption $r/2a<1$ guarantees that the probability $g'^{-1}(U(x(k))/k)$ is always smaller than 1.
For k sufficiently small the first-order approximation of $x(k)$ is $rk/2$, so that the first-order approximation of $\text{Max}_{x\ge 0} g'^{-1}(U(x)/k)(rk-x)$ is $r^2k/4a$.
Moreover $k \dashrightarrow \text{Max}_{x\ge 0} g'^{-1}(U(x)/k)(rk-x)$ is concave, so that $w_0(k,A)$ is convex. The assumption $r^2/4aA<1$ guarantees that $w_0'(0,A)>0$, and therefore $w_0(k,A)$ is a convex and increasing function. One can define $k_0(w,A)=w_0^{-1}(w,A)$ and it is straightforward that $k_0(w,A)$ verifies the properties (a), (b) and (c) of the corollary of proposition 3.

Proof of proposition 4 (section 4):

Assume that there exists $w>w'$ such that $i_f(w)>i_s(w')$. There exists a unique $i_f>-Aw$ such that

$$U(i_s(w')+Aw)-U(i_f+Aw)=U(i_s(w')+Aw')-U(i_f(w')+Aw')$$

Since for all k $p(w,k,i_s(w'),i_f)=p(w',k,i_s(w'),i_f(w'))$ and since by definition we have $(i_s(w'),i_f(w'))\in C(k_1,w')$ for some k_1, we have:

$$(i_s(w'),i_f')\in C(k_1,w,A)$$

Since w did not choose this incentive-conpatible contract we must have:

$$U_w(i_s(w),i_f(w))> U_w(i_s(w'),i_f')$$
(with obvious notations)

But this is contradictory with the fact that w' chose $(i_s,i_f)(w)$, i.e. that $U_{w'}(i_s(w),i_f(w))< U_w(i_s(w'),i_f(w'))$ since by assumption

$$U(i_s(w')+Aw)-U(i_s(w')+Aw')> U(i_f(w)+Aw)-U(i_f(w)+Aw')$$

This gives proposition 4. CQFD.

Proof of proposition 5 (section 4) :

The corrolary of proposition 3 implies that

$$k_{SB}(w,A)< w/(1-r^2/4aA)$$

Therefore $i_s(w,A)$ tends to 0 as w tends to 0^+. Proposition 4 then implies that $i_f(w,A)\leq 0$ for all $w>0$. Therefore $b_{SBf}(w,A)<w$ for all $w>0$ (as long as $(1-\alpha)A<1$).
Two cases can happen: if $b_{SBs}(w,A)<w$ for w smaller than some w', then it is staightforward that every wealth level converges to 0 with probability one; the other case is that represented in figure 4: if $b_{SBs}(w,A)>w$ for $w<w^{**}$ (we know that such a w^{**} exists since there exists $w_0'\leq w_0$ such that for $w\geq w_0'$ $k_{SB}(w,A)=0$), then it is clear that the open interval $]0;w^{**}[$ is the unique ergodic set of the wealth Markov process, and standard convergence results give that whatever the initial wealth distribution, the wealth distribution converges to a unique continuous wealth distribution concentrated on the interval $]0,w^{**}[$.
The conditions $r/a<1$ and $r^2/2aA>1$ guarantee that we are in the second case since they imply that the "stand-alone" utility level obtained by an agent who invests all his initial wealth in a project without seeking for outside finance is higher than $U(Aw)$:

For $r/a<1$ and $r^2/2aA>1$, $\mathrm{Max}_{0\leq p\leq 1}\, pU(rw)-wg(p) > U(Aw)$

(for w sufficiently small).

Therefore if $(1-\alpha)r$ is sufficiently high, $b_{SBs}(w,A)>w$ for w sufficiently small. CQFD.

Proof of proposition 6 (section 4):

Proposition 6 is a direct implication of the corrolary of proposition 3: we know that $b_{SBs}(w,A)/w$ is bounded above by some positive real z. Therfore for any integer T and positive real $w>0$, an agent with initial wealth smaller than w/z^T can be sure that his successor T generations later will have a wealth smaller than w, no matters how lucky his dynasty can be. CQFD.

Proof of proposition 7 (section 5) :

The assumption that a positive quantity of monitored labour is traded in equilibrium implies that $vU'(0)>1$. But since arbitrarily poor agents can borrow only arbitrarily small amounts, there exists $w'>0$ such that for $w<w'$ $U'(i_f(w,A))<1/v$, and therefore $S(w,A,v)>0$. CQFD.

Proof of proposition 8 (section 6) :

Since agents are risk-neutral, the optimal contracts are always such that $i_f=-Aw$ once $k\geq w$ (and for $k<w$, no insurance contract is traded). We first compute the credit-rationing curve: assumptions (3) and (4) imply that for any $A\geq 1$, an agent with initial wealth w facing a loan of size k-w for a project of size k>w will choose the probability of success $p=q'$ if and only if

$$k<k_0(w,A)=w/(1-q'(r-b)/A)$$

As long as $1\leq A\leq q'r-g(q')$, any agent would like to invest in a project as large as possible (modulo $e\leq e_0$). Therefore for $w<w(A)=e_0(1-q'(r-b)/g(q'))/g(q')$ the credit-rationing is fully binding and these agents choose to supply the rest of their labour force as monitored labour (in particular an agent with an initial wealth w=0 cannot borrow at all and earn the safe wage ve_0); for any $1\leq A\leq q'r-g(q')$ the individual transitions are therefore given by:

for $w\leq w(A)$, $b(w,A)=(1-\alpha)(ve_0+(b-vg(q'))w/(1-q'(r-b)/A)$ with probability q'

$b(w,A)=(1-\alpha)(ve_0-vg(q'))w/(1-q'(r-b)/A)$ with probability $1-q'$

for $w(A)\leq w\leq e_0/g(q')$, $b(w,A)=(1-\alpha)((r-A/q')e_0/g(q')+Aw/q')$ with probability q'

$b(w,A)=0$ with probability $1-q'$

for $w\geq e_0/g(q')$, $b(w,A)=(1-\alpha)((r-A)e_0/g(q')+Aw)$ with probability q'

$b(w,A)=(1-\alpha)Aw$ with probability $1-q'$

$$(1-\alpha)vg(q') < 1-q'(r-b)/(q'r-g(q')) \quad \text{(b)}$$
$$(1-\alpha)(q'r-g(q')) < 1 \quad \text{(c)}$$

If the interest rate was fixed and equal to $A_1^*=1$, then (a) implies that the wealth distribution would converge to a (unique) distribution f_1^* such that $1-f_1^*(e_0/g(q'))>q'^2$ (see figure 5); therefore there exists p_1 such that if $q'>p_1$, the equilibrium interest rate for f_1^* $A(f_1^*)$ is equal to one. This also implies that there exists $w_2>e_0/g(q')$ such that if initially $1-f_0(w_2)>q'^2$ then $A(f_t)=1$ for all $t\geq 0$. This proves part (a) of proposition 8.

In the same manner, if the interest rate was fixed and equal to $A_2^*=rq'-g(q')$, then (b) implies that (whatever the way we choose to solve agents' indifferences) the wealth distribution would converge to a (unique) distribution f_2^* such that $1-f_2^*(e_0/g(q'))<q'^3$ (see figure 6); therefore there exists p_2 ($>p_1$) such that if $q'<p_2$, the equilibrium interest rate for f_2^* $A(f_2^*)$ is equal to $rq'-g(q')$. Together with (c), this also implies that there exists $w_1<e_0/g(q')$ such that if initially $1-f_0(w_1)<q'^3$ then $A(f_t)=rq'-g(q')$ for all $t\geq 0$. This proves part (b) of proposition 8. CQFD.

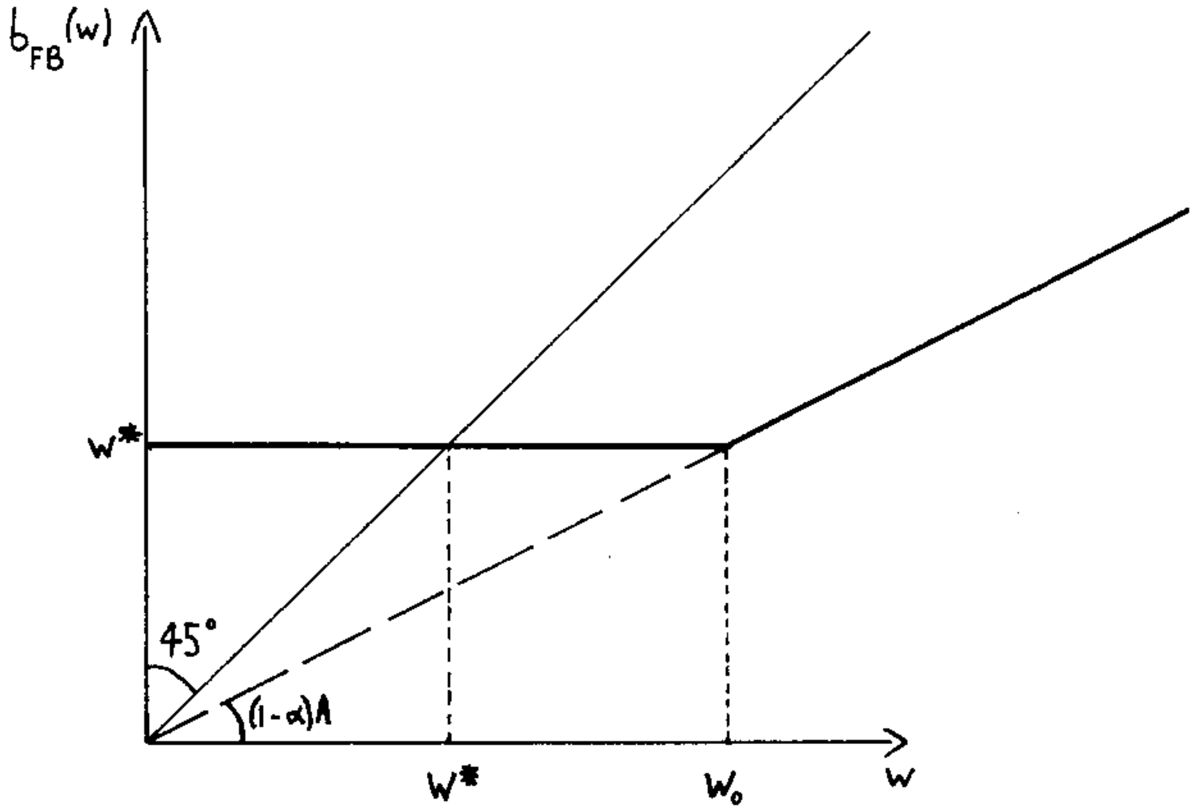

Figure 1

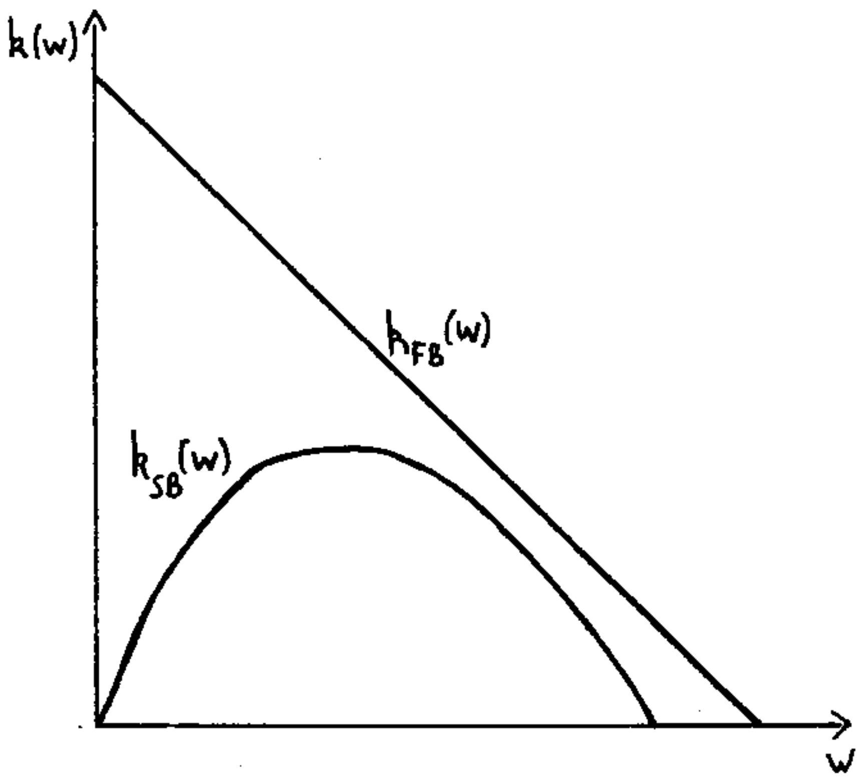

Figure 2

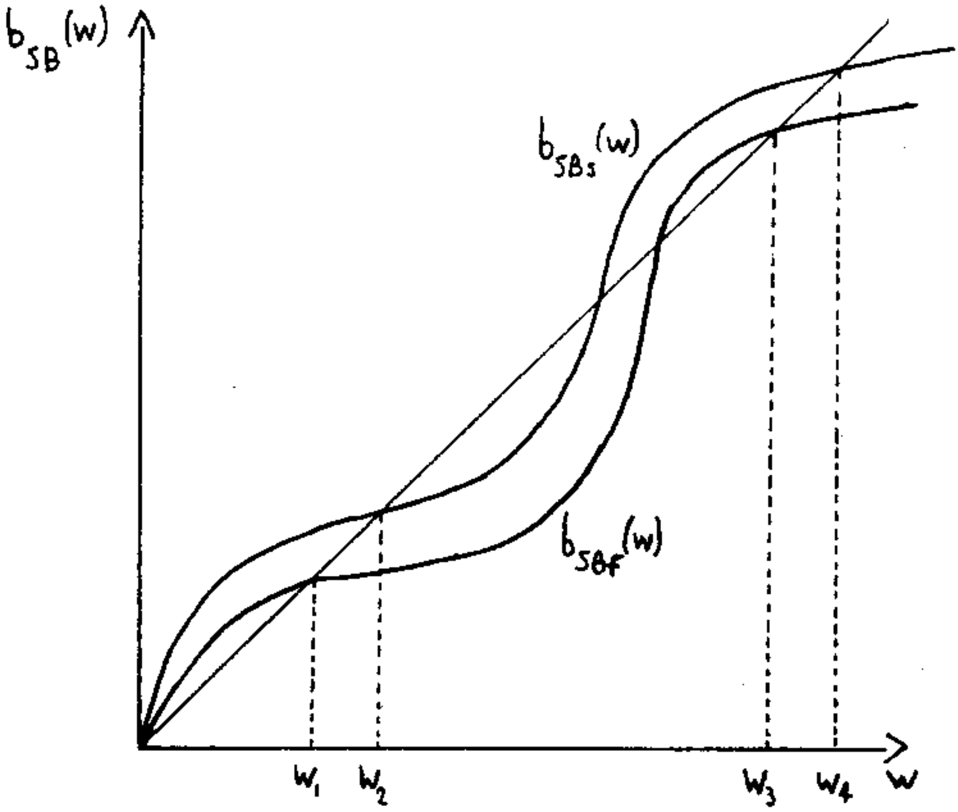

Figure 3

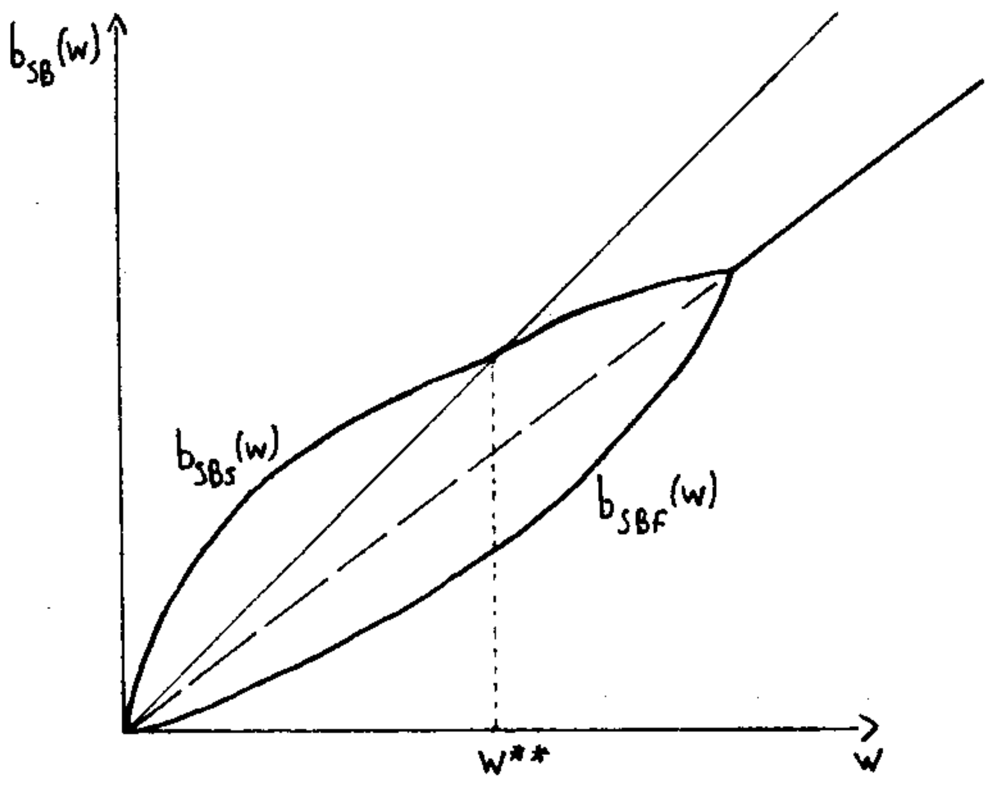

Figure 4

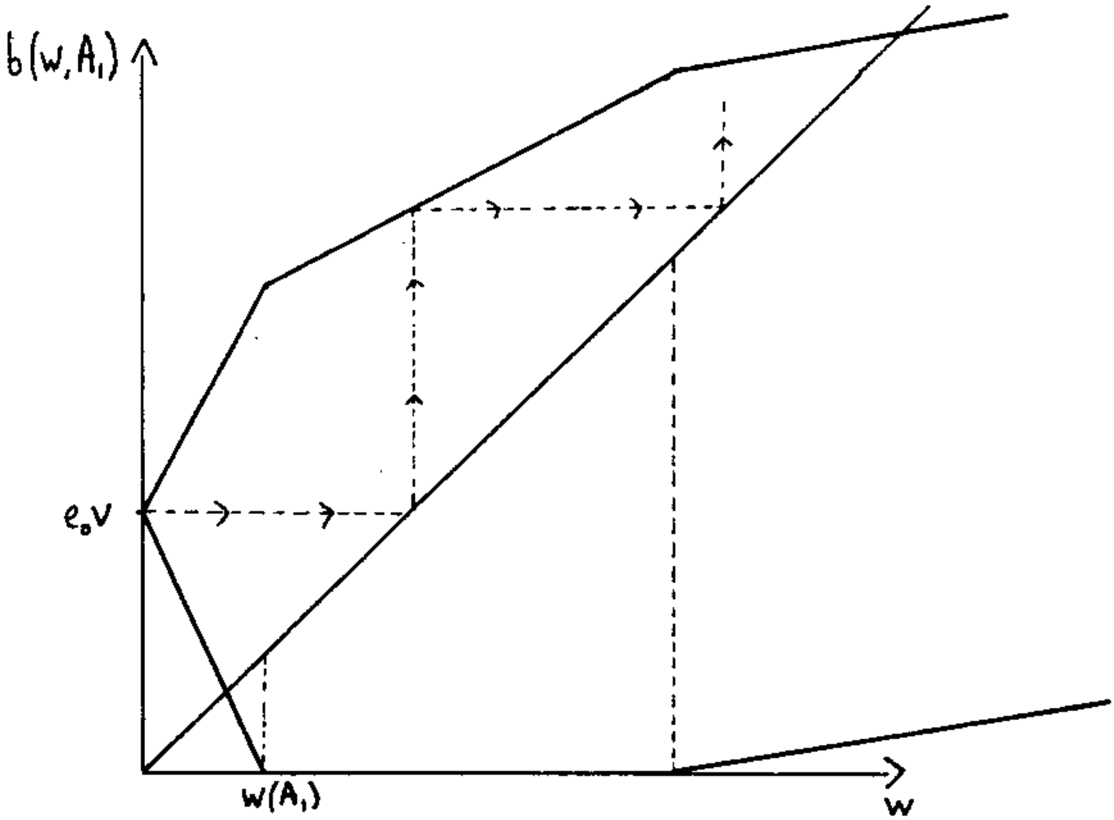

Figure 5

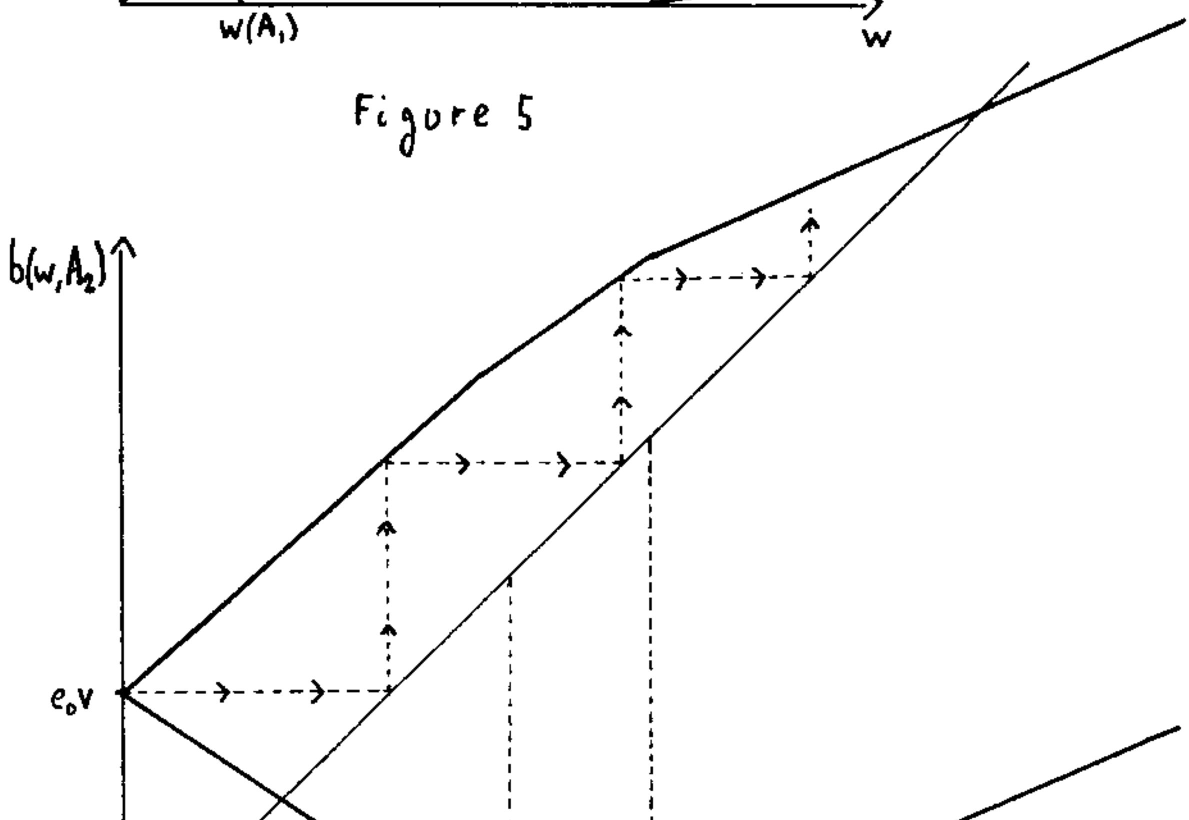

Figure 6

Footnotes:

(*1) : Note however that if one interprets the utility function as being a dynamic-programming value function, then the latter may not be concave in the second-best economy because of the non-convex incentive-compatibility constraints. Also this value function may not be easy to compute: for example in the case of a dynastic behaviour (i.e. maximization of the discounted sum of future utilities; note that this does not seem to be particularly realistic: see Bernheim(1988) and Abel and Bernheim(1991)), the set of incentive-compatible contracts supplied in equilibrium can only be defined recursively. Therefore there is much gain in tractability and little loss in generality by assuming preferences to be defined directly over (c,b,e) bundles.

(*2) : The rest of the paper would be virtually unchanged if we assume only preferences to be separable (i.e. U(c.b.e)=W(c,b)-V(e)), or more generally labour to be a normal good.

(*3) : Since proposition 4 holds for any utility and cost function, the ergodicity of the wealth distribution dynamics can be established for the general case; however, we are primarily concerned with the case where credit rationing is fully binding for poor agents.

References.

-Abel and Bernheim(1991), "Fiscal policy with impure intergenerationnal altruism", Econometrica.

-Aghion and Bolton(1991), "A trickle-down theory of growth and development with debt-overhang", mimeo.

-Aghion and Bolton(1992), "Distribution and growth in models of imperfect capital markets", European Economic Review 36.

-Arnott and Stiglitz(1990), "The welfare economics of moral hazard", WP NBER no 3316.

-Andreoni(1989), "Giving with impure altruism: applications to charity and Ricardian equivalence", Journal of Political Economy 97, 1447-1458.

-Banerjee and Newman(1992), "Occupational choice and the process of development", forthcoming in Journal of Political Economy.

-Banerjee and Newman(1991), "Risk-bearing and the theory of income distribution", Review of Economic Studies 58, 211-235.

-Bernheim(1988), "Is everything neutral?", Journal of Political Economy.

-Champernowne(1953), "A model of income distribution", Economic Journal 63, 318-351.

-Galor and Zeira(1991), "Income distribution and macoeconomics: the human capital connection", mimeo.

-Greenwood and Jovanovic(1990), "Financial development, growth and the distribution of income", Journal of Political Economy 98, 1076-1107.

-Grossman and Hart(1983), "An analysis of the principal-agent problem", Econometrica 51, 7-45.

-Kessler and Masson(1988), "On five hot issues on wealth distribution", European Economic Review 32, 644-653.

-Kihlstrom and Laffont(1979), "A general equilibrium entrepreneurial theory of firm formation based on risk aversion", Journal of Political Economy 87, 719-748.

-Loury(1981), "Intergenerationnal transfers and the distribution of earnings", Econometrica 49, 843-867.

-Lucas and Stokey(1989), "Recursive methods in economic dynamics", Harvard University Press.

-Newman(1991), "The capital market, the wealth distribution and the employment relation", mimeo.

-Persson and Tabellini(1991), "Is inequality harmful for growth: theory and evidence", mimeo.

-Persson and Tabellini(1992), "Growth, distribution and politics", European Economic Review 36, 593-602.

-Sappington(1983), "Limited liability contracts between principal and agent", Journal of Economic Theory 29, 1-21.

-Solon(1992), "Intergenerationnal income mobility int he US", American Economic Review, june 1992.

-Varian(1980), "Redistributive taxation as a social insurance", Journal of Public Economics.

-Zimmerman(1992), "Regression towards mediocrity in economic
stature", American Economic Review, june 1992.