# Poverty Traps and Human Capital Accumulation*†

Carlotta Berti Ceroni
Dipartimento di Scienze Economiche
Universita' di Bologna
and
Departament de Economia i Empresa
Universitat Pompeu Fabra

April 2, 1998

*This paper is part of my Ph.D. Dissertation at Columbia University and has been completed while visiting Universitat Pompeu Fabra as part of a Project under the Training and Mobility of Researchers (TMR) Programme, financed by the European Community (Marie Curie Research Training Grant, contract nr. ERBFMBICT961393). I am greatly indebted to Andrew Newman and Roberto Perotti for their help and encouragement. I have also benefited from comments and suggestions on previous versions of the paper by Giorgio Bellettini, Michele Boldrin and Lorenzo Garbo. All remaining errors are mine alone.

†Address: Dipartimento di Scienze Economiche, P.zza Scaravilli, 2, Bologna, 40126, Italy. Phone: +39-51258017. Fax: +39-51-221968. E-mail: bceroni@economia.unibo.it

## Abstract

In this paper, we analyze the emergence and persistence of poverty traps and study how wide-spread poverty and the unequal distribution of income can slow down the accumulation process and lead to steady-state equilibria characterized by low aggregate output levels. We define poverty as a state deriving from the lack of adequate skills and we associate income inequality with the unequal distribution of education attainments. In this context the goal of our contribution is twofold. First, we show that low asymptotic mobility and persistent income inequality can emerge as a consequence of the fact that the poor require relatively higher returns to increase expenditure on education, so that they devote to education smaller shares of their income than the rich. Second, we critically evaluate our and other related results, in order to shed light on the explanatory power of different sets of assumptions.

# 1. Introduction

In this paper, we examin the relation between income distribution, accumulation and the macroeconomic equilibrium. In particular, we analyze the emergence and persistence of poverty traps and study how wide-spread poverty and the unequal distribution of income can slow down the accumulation process and lead to steady-state equilibria characterized by low aggregate output levels.[1]

Following a considerable research stream[2], we define poverty as a state deriving from the lack of adequate skills and we associate income inequality with the unequal distribution of education attainments. According to this definition, poverty tends to be transmitted from one generation to the next, relative income rankings tend to persist over time and mobility across income classes tends to remain low, if low income leads to low human capital accumulation.[3] The fact that individual human capital accumulation paths depend on initial relative income rankings implies a relation between the initial distribution of income and the growth path followed by the economy as a whole. In this context, the goal of our contribution is twofold. On one hand, we investigate an explanation of low asymptotic mobility which has not yet been carefully explored in the literature. On the other hand, we critically evaluate our and other related results, in order to shed light on the explanatory power of different sets of assumptions.

Many recent models explain low asymptotic mobility and persistent income inequality, by assuming that the rich enjoy relatively higher net marginal returns from investment in education than the poor. This assumption has been justified by invoking the existence of incentive constraints on financial contracts and of local non-convexities in the human capital accumulation technology.[4] Instead, we

[1]Theoretical efforts devoted to understanding the relation between the distribution of resources and accumulation is motivated by recent empirical evidence suggesting a positive relation between equity and per-capita income growth (Alesina-Rodrick [2], Perotti [39], Persson-Tabellini [40]).

[2]Among others, see Becker-Tomes [12], Benabou [15] [16] [17], Durlauf [27], Galor-Zeira [29], Galor-Tsiddon [28], Lowry [35], Tamura [50].

[3]Recent estimates of intergenerational income correlation in the U.S. have shown that regression to the mean in earnings is much less rapid and mobility much lower than it was believed on the basis of previous studies (Solon [47] and Zimmerman [54]). In particular, Zimmerman shows that a consideration of sons' earnings in 1981 and fathers' data in 1965 reveals that as much as 69 (74) percent of children whose father fall in the lowest (highest) earnings quartile do not rise above (below) the second earnings quartile.

[4]If the incentives to maximize returns from investment in education are lower the larger the amount an agent needs to borrow, the poor are more likely credit-constrained than the rich

assume that low asymptotic mobility and persistent inequality are due to the fact that the poor require relatively higher marginal returns to increase expenditure on education, so that they devote to education smaller shares of their income than the rich.[5] [6] This assumption can be justified in different ways. Parents may regard the education attainment of their children as a luxury good. Educated parents may pass on a view about the consumption benefits of education. There might exist minimum consumption requirements. Risk-averse poor agents, lacking a safety net, may try to smooth income and consumption by seeking employment sooner than the rich. In this paper, we take up the first of these explanations and look at investment in education as a form of bequest of altruistic parents who derive utility from the education attainment of their children.

In our economy, a single good is produced by means of a constant returns to scale technology that utilizes human capital as the only factor of production. The intergenerational transmission of human capital occurs by means of a decreasing returns technology, which transforms private education expenditure into education attainments. External effects of parents' human capital on individual learning productivity are ruled out, as well as spillover effects of the aggregate stock of human capital at the neighborhood and economy-wide level. Education decisions are taken by altruistic parents, who value their children education attainment and allocate part of their disposable income to school them. This is equivalent to rule out the availability of education loans, since parents can not borrow on their children income.[7] The absence of spillover effects in the human

---

and are forced to take up relatively low-paid unskilled jobs (Galor-Zeira [*cit*]). If the family's (neighborhood's) human capital stock has a positive external effect in the production of human capital, relative differences in skills and income across families (neighborhoods) may be passed from one generation to the next (Benabou [*cit*], Borjas [19], Durlauf [*cit*], Galor-Tsiddon [*cit*]).

[5] As a matter of fact, convex savings have been shown to imply persistence of wealth inequality as well as a non-negative association between inequality and growth, in the context of standard neoclassical growth models, where, due to perfect capital markets, marginal net returns from investment in physical capital are equalized across classes (Bourguignon [20], Chatterjee [22], Stiglitz [48]). As we argue below, none of these implications remain necessarily true in the context of a human capital accumulation model, where a market for education loans is very likely to be missing and the poor enjoy relatively higher marginal net returns from investment in education than the rich.

[6] Several empirical papers provide microevidence consistent with the view that the poor save proportionately less than the rich (among others see Atkenson-Ogaki [4], Lawrence [33] and Ogaki-Atkenson [37]).

[7] This form of market incompleteness might be judged too extreme in other contexts but is indeed quite adherent to reality in the context of human capital accumulation models: the fact that expected human capital is not usually accepted as collateral implies that a market for

capital accumulation technology, the fact that investment in education is privately financed and linearity of the production function imply that the evolution of each dynasty's human capital can be studied in isolation. In this set up, we show that if the share of income allocated to education is increasing with the parents' human capital poverty traps may arise such that the poor are locked in the lower tail of the human capital distribution forever and differences in education attainment and income across families persist over time. Moreover, we show that inequality may undermine growth along the transition path to the steady state and that unequal economies, where a large fraction of the population is initially below the poverty line, achieve lower levels of aggregate output than more equal ones.

One of the salient features of our deterministic model is the possibility of re-interpreting it as a special case of a stochastic model where each dynasty's human capital follows a *linear* Markov process and where poverty traps are identifiable with the existence of two ergodic sets, one including states of extreme poverty and one including states of extreme wealthiness. This is due to the fact that, at any point in time, the transition functions governing each dynasty's human capital accumulation are independent of that of other dynasties so that current family's income is a sufficient statistic to determin a child's education achievement in the following period. This feature, which is common to other deterministic models in this research stream (Galor-Zeira [*cit*] and Galor-Tsiddon [*cit*]), has a nasty implication: persistence of inequality and low asymptotic mobility are not theoretically robust results. As we will show, it is sufficient to introduce an exogenous positive probability of moving back and forth the poverty trap threshold to destroy these results. In fact, convergence to a unique invariant distribution in the theory of linear Markov processes is guaranteed if there exists at least one state of positive measure that can be reached from any state in the state space in finite time. As shown by a few recent contributions analyzing the dynamics of wealth distribution (Banerjee-Newman [6] [7], Piketty [41]), persistence of inequality and low asymptotic mobility can instead be preserved in spite of the fact that each dynasty's wealth can travel all over the state space in finite time, if the stochastic process followed by each dynasty's wealth is *non linear*, that is if individual transition functions depend on the distribution of wealth across dynasties.

education loans is unlikely to develop and that the direct and opportunity costs of education are often self-financed by altruistic parents (Becker [10] [11]). Recent empirical evidence of the existence of binding borrowing constraints in financing education is provided by Bernham, Pollack and Taubman [14], who find an inverse relation between sibship size and sib schooling using U.S. data, and by Bernham et al. [13] and Cameron-Heckman [23], who find that family financial resources are a significant determinant of education levels in the U.S..

Although models involving individual linear dymnamics are weak from a theoretical point of view, the mechanisms underlying the poverty traps they describe may matter observationally and empirically. In this paper, we will explore the implications of assuming that the exogenous probabilities of moving back and forth the poverty trap threshold are very low, so that the economy is observationally equivalent to one where poverty traps exist for most of the time, and we will devote some attention to the study of how mobility evolves over time. Moreover, we will discuss some casual evidence supporting the view that extreme poverty and cultural backwardness are states of social and economic isolation, such that the link from aggregate to individual behavior is not very effective. Note that most part of this analysis can be applied to Galor-Zeira's [*cit*], Galor-Tsiddon's [*cit*] and "linear" models in general, as well as to our model.

The plan of the paper is as follows. In section 2, we present our benchmark deterministic human capital accumulation model and show how poverty traps can arise when the poor require relatively higher returns to increase investment in education. Section 3 extends the analysis to a stochastic framework and shows that poverty traps obtained in the context of linear models tend to be extremely fragile. In section 4, we show that linear models may still be relevant both on observational and on empirical grounds.

## 2. Human Capital Accumulation and Poverty Traps

In this section, we present our benchmark deterministic human capital accumulation model and show that poverty traps such that a fraction of the population is stuck in the lower tail of the income distribution forever, while a fraction of the population is always away from it, can derive from the fact that the poor require relatively higher returns to increase investment in education, so that they devote to education relatively smaller shares of their income.

*A. The Model*

We consider an economy populated in each period by a unitary mass of families composed by two individuals, a parent and a child. Each agent lives two periods and is endowed with a unit of time in each period. Agents can only make decisions in the second period of their lives. When young, agents may get educated if their parents decide so. In this case they spend their unit of time at school. Young agents not attending school simply remain idle. When old, all agents supply inelastically to the labor market their unit of time, earning an income proportional to their human capital stock, and decide how to allocate it between family consumption and offsprings education in order to maximize their utility. The fact that education decisions are taken by parents amounts to assume that a market for education loans does not exist, since parents can not borrow on children's future income to finance education. Investment in education is thus constrained by parents' income.

In order to obtain close form solutions for individual transition functions and to keep the aggregation problem as simple as possible, we choose a specification of preferences and technology such that investment in education is linear in the parent's human capital. Qualitatively similar results can be obtained with more general specifications, provided that the salient features of the functions we choose are maintained.

The utility function of the $i$-th parent at time $t$ is defined over consumption at time $t$ and the $i$-th child's human capital stock at time $t+1$.[8] We assume that the utility function is quasi-linear in the child's human capital stock and takes the form:

$$U^i\left(c_t^i, h_{t+1}^i\right) = \log\left(c_t^i\right) + \delta h_{t+1}^i \tag{2.1}$$

where $\delta$ is a parameter measuring the strenght of the altruistic motive. This utility function is obviously non-homothetic: instead, the marginal rate of substitution between parent's consumption and child's education attainment calculated for a given ratio of the former to the latter is decreasing with the parent's human capital stock. Moreover, the marginal rate of substitution between parent's consumption and child's education attainment calculated at zero investment is decreasing with the parent's human capital stock, so that the poor require relatively higher returns on education to start investing.

The technology for the production of human capital takes the form:

$$h_{t+1}^i = F\left(e_t^i\right) = \log\left(\gamma e_t^i + \nu\right) + \upsilon \tag{2.2}$$

where $e_t^i$ is parent $i$'s expenditure on child $i$'s education at time $t$, $\nu \geq 1$, $\gamma > 0$, $\upsilon > 0$ and $\mu = \log(\nu) + \upsilon$. Note that when no investment in human capital is made by parents, children nevertheless achieve a minimum level of human capital in adulthood, $\mu$, which is independent of family's income. Returns to investment in education are bounded at zero investment, they are everywhere decreasing and

[8] We choose a "warm glow" specification of altruism, in order to simplify the analysis, but our results do not depend on this assumption.

they do not depend on the parents' human capital stock, nor on the aggregate human capital stock.

The economy produces a single good by means of a constant returns to scale technology, which utilizes human capital as the only factor of production. The aggregate production function is given by:

$$Y_t = A \int_I h_t^i di = AH_t \tag{2.3}$$

where $h_t^i$ is the human capital stock of the $i$-th parent at time $t$, $H_t$ is the aggregate stock of human capital at time $t$ and average human capital productivity, $A$, is constant. Note that this specification of the production function implicitly assumes that different types of human capital are perfect substitutes in production. Also, the real wage per efficiency unit of labor will be constant and equal to $A$ in equilibrium.

The human capital of the initial generation of parents is distributed according to an exogenously given distribution $g_0(h_0^i)$, such that:

$$g_0(h_0^i) \geq 0 \qquad \int_\alpha^\beta g_0(h_0^i)dh_0^i = 1 \quad h_0^i \in (\alpha, \beta) \quad \beta > \alpha \geq \mu$$

Note that we assume that all parents of the first generation have human capital at least equal to $\mu$.

*B. Individual Income Dynamics*

The individual maximization problem solved by parent $i$ at time $t$ is given by:

$$\max_{e_t^i} \log\left(c_t^i\right) + \delta h_{t+1}^i \tag{2.4}$$

$$\begin{aligned} s.t. \quad c_t^i &= Ah_t^i - e_t^i \\ h_{t+1}^i &= \log\left(\gamma e_t^i + \nu\right) + v \\ \left(e_t^i, c_t^i\right) &\geq (0, 0) \end{aligned}$$

Note that $e_t^i$ is constrained to be non-negative.[9]

[9]By a simple change of variable, expenditure on education can be interpreted in terms of foregone labor income. Let $z_t^i$ denote the share of time that parents devote to their children

The solution of problem 2.4 implies that investment in education is given by:

$$e_t^{i*}\left(h_t^i\right)=\begin{cases}0 & h_t^i\leq \underline{h}\\[2ex] \dfrac{\delta}{1+\delta}Ah_t^i-\dfrac{\nu}{\gamma\left(1+\delta\right)} & h_t^i>\underline{h}\end{cases}\tag{2.5}$$

where $\underline{h}=\dfrac{\nu}{\delta\gamma A}$. Note that, the share of income allocated to education is increasing with the parent's human capital stock, whenever investment in education is positive. This is due to the non-homotheticity of the utility function. Also, note that corner solutions can arise at low income levels as a consequence of the fact that the poor require relatively higher returns to start investing in education and that returns from education are bounded at zero investment. Without loss of generality, let $A=1$ from now on. With this normalization, income and human capital are always equal both at the individual and at the aggregate level and can be used interchangeably.

Constancy of the real wage and independence of the productivity of individual investment in education of the aggregate stock of human capital in the economy and of its distribution imply that the dynamics of each dynasty's human capital is independent of the evolution of the aggregate stock of human capital and of its distribution and can be studied in isolation. In particular, the evolution of dynasty $i$'s human capital is described by the transition equation:

$$h_{t+1}^i=F\left[e_t^{i*}\left(h_t^i\right)\right]=\Phi(h_t^i)\tag{2.6}$$

where:

$$\Phi(h_t^i)=\begin{cases}\mu & h_t^i\leq \underline{h}\\[2ex] \log\left[\dfrac{\delta\left(\gamma h_t^i+\nu\right)}{1+\delta}\right]+\upsilon & h_t^i>\underline{h}\end{cases}\tag{2.7}$$

Obviously, the transition function $\Phi\left(h_t^i\right)$ has positive slope and is concave for $h_t^i>\underline{h}$ . This is due to the fact the marginal propensity to invest in educa-

education. Then expenditure on education and consumption can be rewritten as $e_t^i=z_t^i\omega_t h_t^i$ and $c_t^i=\left(1-z_t^i\right)\omega_t h_t^i$. By using these definitions, the decision of how much money to spend on education can be transformed into a decision of how much time to devote to education. This is important, since, as we noted above, the education process involves mostly opportunity costs rather than direct costs.

tion is constant at interior solutions and returns to investment in education are decreasing.

*C. The Dynamics of Income Distribution*

At any point in time, the current distribution of income determines tomorrow's income distribution. In particular, we have:

$$g_t\left(h_t^i\right) = g_{t-1}\left(\Phi^{-1}\left(h_t^i\right)\right) \quad h_t^i \in [\alpha, \beta] \tag{2.8}$$

Given the initial distribution of human capital and the individual transition equation, it is then possible to analyze the behavior of the income distribution over time. In particular, we are interested in establishing conditions under which initial inequality persists over time. In the context we are analyzing, persistence of income inequality can be characterized as the emergence of poverty traps such that a fraction of dynasties is stuck in the lower tail of the income distribution forever while a fraction of dynasties is permanently away from it. This is equivalent to requiring the individual transition equation to exhibit multiple steady states.

Given our assumptions on preferences and technology, multiple steady states are possible only for parameters' configurations such that corner solutions arise at low income levels. In fact, if all dynasties invest in education, concavity of the individual transition function implies that there exists a unique steady state level of human capital to which all dynasties eventually converge.[10] Since all dynasties have human capital at least equal to $\mu$, a necessary and sufficient condition for corner solutions to emerge at low income levels is $\underline{h} > \mu$. This condition also guarantees that once a dynasty reaches the human capital stock $\mu$, it remains stuck there forever, that is $\mu$ is a fixed point of the individual transition equation. For poverty traps to emerge, the individual transition equation must exhibit two more steady states, besides $\mu$. In particular, there must exist an unstable steady state $h_u > \underline{h}$ and a stable steady state $h^* > h_u$. Since the individual transition function is concave for $h_t^i > \underline{h}$, this is possible only if the slope of the transition function tends to a number larger than one, as human capital tends to $\underline{h}$ from above. It is immediate to verify that a necessary and sufficient condition for this

[10] Under more general assumptions on preferences and the technology for human capital accumulation, yet satisfying the afore-mentioned features, poverty traps can emerge due to the fact that poor parents invest relatively smaller shares of their income in children's education, even if all dynasties always invest positive amounts in education.

to hold is $\gamma > \nu \dfrac{1+\delta}{\delta}$. Continuity of $\Phi'(h_t^i)$ then ensures that $\Phi'(h_t^i) > 1$ for $h_t^i > \underline{h}$, for $h_t^i$ close enough to $\underline{h}$.

Therefore, for choices of parameters such that $\gamma > \nu$ and $\underline{h} > \mu$, with $\underline{h} \simeq \mu$, the individual transition equation will exhibit three steady states, at human capital levels $\mu$, $h_u$ and $h^*$, and poverty traps will emerge, for some initial distribution of income.[11] Dynasties whose human capital is initially below the poverty threshold $h_u$ permanently remain below such threshold. Parents belonging to these "poor" dynasties may initially invest in their children's education, but the dynasty's human capital stock decays generation after generation until it reaches the steady state level $\mu$ and no further investment in education is ever made. Dynasties whose human capital is initially above the poverty threshold $h_u$ permanently remain above such threshold. Parents belonging to these "rich" dynasties keep investing in their children's education and the dynasty's human capital stock grows generation after generation until it reaches the steady state level $h^*$. In the long run, dynasties are concentrated in two groups, the poor who do not invest in education and thus remain poor and the rich who invest in education and thus remain rich. Note that no asymptotic mobility across income classes is possible in our model. Initial inequality persists over time in the sense that the long run distribution of income, that is the relative size of the two income groups, is determined by the initial one. In particular, let the degree of inequality in the distribution of income at any point in time be defined by the fraction of dynasties whose human capital is below the poverty threshold $h_u$. Then, the larger the fraction of dynasties initially below the poverty threshold $h_u$, the more unequal the long run distribution of income.

### *D. Income Inequality and the Macroeconomic Equilibrium*

At any point in time, the current distribution of income determines current aggregate investment in education and tomorrow's aggregate human capital and ouput. In particular, we have:

$$E_t = \frac{\delta}{1+\delta} \int_{\underline{h}}^{\beta} h_t^i g_t\left(h_t^i\right) - \frac{\nu}{\gamma(1+\delta)} \left[1 - G_t\left(\underline{h}\right)\right] \tag{2.9}$$

[11] Figure 1.1 in the Appendix provides a graphical representation of a choice of parameters such that poverty trap arises.

$$H_t = Y_t = \int_{\underline{h}}^{\beta} \log \left[ \frac{\delta \left( \gamma h_{t-1}^i + \nu \right)}{1+\delta} \right] g_{t-1} \left( h_{t-1}^i \right) + (v - \mu) G_{t-1} \left( \underline{h} \right) - \mu \qquad (2.10)$$

By observation of 2.9 and 2.10 it is clear that the current distribution of income affects aggregate human capital accumulation and output growth along the transition path to the steady state. Moreover, since the current distribution of income depends on the initial one, there also exists a relation between the initial distribution of income and current aggregate output level.

As for the relation between inequality and aggregate output growth along the transition path toward the steady state, a clearcut conclusion regarding the sign of the association can not be reached, in presence of poverty traps. In fact, if we consider a redistribution of income from dynasties above the threshold $h_u$ to dynasties below such threshold, that is a reduction in the degree of income inequality, two opposite effects are at work. On one hand, redistributing income from dynasties experiencing positive human capital growth rates $(h_t^i > h_u)$ to dynasties experiencing negative growth rates $(h_u > h_t^i > \underline{h})$ tends to reduce aggregate human capital and output growth rates. On the other hand, redistributing income in favor of dynasties who would otherwise be unable to invest in education $(h_t^i \leq \underline{h})$ can foster aggregate output growth, if these dynasties are put in conditions to start investing in education and to outgrow poverty.

The relation between the initial degree of inequality and current aggregate output is obviously negative along the transition path toward the steady state, since the human capital stock of initially poor dynasties is always below that of initially rich ones. Such negative relation will persist in the long run, in presence of poverty traps. In particular, in presence of poverty traps, aggregate output at steady state is given by:

$$Y_\infty = \left[ h^* + h_u g_0 \left( h_u \right) \right] - \left( h^* - \mu \right) G_0 \left( h_u \right) \qquad (2.11)$$

and is obviously decreasing with the fraction of dynasties whose human capital is initially below the threshold $h_u$, that is with the initial degree of income inequality.

### *E. A Brief Discussion of Our Results*

As we have already pointed out, the mechanism underlying the poverty trap we derived in this section is grounded on the assumption that the poor require relatively higher returns to increase investment in education than the rich. Qualitatively similar results have been obtained in the context of deterministic human capital accumulation models by Galor-Zeira [*cit*] and Galor-Tsiddon [*cit*] by assuming instead that, due to some non-convexity in the technology for human capital production, the poor enjoy relatively lower net marginal returns from investment in education than the rich. In particular, Galor-Zeira [*cit*] assume that, present capital market imperfections, education involves the payment of a fixed cost (tuition), while Galor-Tsiddon [*cit*] assume that the parents' human capital has a positive spillover effect on the children productivity at learning (home environment). These non-convexities are crucial to obtain asymptotic persistence of inequality. In fact, absent non-convexities, these models imply that the poor enjoy marginal net returns from investment in education that are not lower than those enjoyed by the rich so that their human capital stock tends to catch up in the long run. However, none of the assumptions introduced to obtain non-convexities is completely satisfactory from our point of view. On one hand, in most countries education involves variable opportunity costs in terms of foregone income rather than fixed direct costs. On the other hand, available empirical evidence seems to suggest that, although a positive correlation between parent's human capital and children's earnings seems to exist, much of it would be intermediated through the children education attainment: increasing the average skill level of the father seems to have only small effects on the son's income, given the son's schooling (see Schultz [46] and more recently Altonnji and Dunn [3]). The assumptions of our model, instead, are in line with the above observations.

Before concluding this section, it might be useful to shed further light on some features of our model. We showed above that both the extent of inequality at steady state and the steady state aggregate ouput level are determined by the initial extent of poverty. So far we have assumed that, for given choices of the parameters $\delta$, $\gamma$, $\nu$, $\upsilon$ and $A$, which determine the position of the poverty threshold $h_u$, the initial extent of poverty is determined by the initial distribution of human capital across dynasties. However, our reasoning might be reversed to argue that, given the initial distribution of human capital, the initial extent of poverty and therefore the long run behavior of the economy, is determined by the parameters which influence the position of the poverty threshold $h_u$.

For example, regarding $\gamma$, which determines the productivity of investment in education, the following remark can be made.

Remark 1. As $\gamma$ decreases, the fraction of dynasties who end up with human capital $\mu$ rather than $h^*$increases.

This is due to the fact that corner solutions become more relevant as $\gamma$ decreases, since $\underline{h}$ is decreasing with $\gamma$, and to the fact that the existence of poverty traps itself depends on the magnitude of $\gamma$. In particular, as $\gamma$ decreases, the following situations prevail: (a) if $\gamma > \nu\dfrac{1+\delta}{\delta}$ and $\underline{h} < \mu$ all dynasties eventually converge to the human capital level $h^*$; (b) if $\gamma > \nu\dfrac{1+\delta}{\delta}$ and $\underline{h} > \mu$, with $\underline{h} \simeq \mu$ poverty traps emerge such that initially poor dynasties end up at the low steady state $\mu$, while initially rich dynasties eventually converge to the high steady state $h^*$; (c) if $\gamma > \nu\dfrac{1+\delta}{\delta}$ and $\underline{h} >> \mu$, corner solutions emerge at low income levels but poverty traps do not, so that all dynasties eventually converge to the low steady state $\mu$; (d) if $\gamma < \nu\dfrac{1+\delta}{\delta}$ and $\underline{h} >> \mu$, corner solutions emerge at low income levels but poverty traps do not, so that all dynasties eventually converge to the low steady state $\mu$.[12]

A similar remark can be made regarding the real wage level. So far, we have assumed $A = 1$, for simplicity. Instead, let $A \neq 1$. Then we have:

Remark 2. As $A$ decreases, the fraction of dynasties who end up with human capital $\mu$ rather than $h^*$increases.

This is due to the fact that corner solutions become more relevant as $A$ decreases, since $\underline{h}$ is decreasing with $A$ and to the fact that the existence of poverty traps itself depends on the magnitude of $A$. In fact, when $A \neq 1$, the necessary and sufficient condition for the slope of the transition function to tend to a number larger than one as human capital tends to $\underline{h}$ from above becomes $\gamma A > \nu\dfrac{1+\delta}{\delta}$. Therefore, as $A$ decreases, the following situations prevail: (a) if $\gamma A > \nu\dfrac{1+\delta}{\delta}$ and $\underline{h} < \mu$ all dynasties eventually converge to the human capital level $h^*$; (b) if $\gamma A > \nu\dfrac{1+\delta}{\delta}$ and $\underline{h} > \mu$, with $\underline{h} \simeq \mu$ poverty traps emerge such that initially poor dynasties end up at the low steady state $\mu$, while initially rich dynasties

[12] Figure 1.2-1.5 in the Appendix provide a graphical representation of cases (a)-(d), for convenient choices of parameters.

eventually converge to the high steady $h^*$; (c) if $\gamma A > \nu \dfrac{1+\delta}{\delta}$ and $\underline{h} >> \mu$,corner solutions emerge at low income levels but poverty traps do not, so that all dynasties eventually converge to the low steady state $\mu$; (d) if $\gamma A < \nu \dfrac{1+\delta}{\delta}$ and $\underline{h} >> \mu$, corner solutions emerge at low income levels but poverty traps do not, so that all dynasties eventually converge to the low steady state $\mu$.[13]

Remarks 1 and 2 make clear that, for given initial inequality in the distribution of human capital, the lower productivity of investment in education and the lower the real wage for efficiency unit of labor, the more pervasive is poverty in the long run, the poorer is the economy on aggregate and the more likely is income inequality to persist in the long run. However, as we noted before, neither of these variables is endogenous in our model and, in particular, neither of them depends on the distribution of human capital in the economy. As the following sections shows, this is indeed quite a relevant feature of our model.

## 3. Robust Poverty Traps: Linear versus non Linear Models

In this section we show that poverty traps of the kind we analyzed so far are not very robust theoretically. In fact, it is sufficient to introduce exogenous probabilities, no matter how small, of moving back and forth the poverty threshold to destroy them.

Consider a stochastic generalization of our model where, given parents' investment in education, the children education attainment is a stochastic variable, due to children's innate ability which becomes known only in adulthood. In particular, assume that, for all dynasties at all points in time, there exists a probability $\pi_1$ that, due to his poor comprehension, the child does not profit from education, no matter how extensive, and a probability $\pi_2$ that the child is a genius and reaches the highest possible level of skills independently of the education he receives. Let these shocks to innate ability be independently distributed across dynasties. Then, we have:

$$h_{t+1}^i = \begin{cases} \mu & w.\ p.\ \pi_1 \\ F\left(e_t^i\right) & w.\ p.\ 1-\pi_1-\pi_2 \\ \beta & w.\ p.\ \pi_2 \end{cases} \tag{3.1}$$

[13]Figure 1.2-1.5 in the Appendix provide a graphical representation of cases (a)-(d), for convenient choices of parameters.

The individual maximization problem solved by parent $i$ at time $t$ is now:[14]

$$\begin{aligned} &\max_{e_t^i} \log\left(h_t^i - e_t^i\right) + \delta\left[\pi_1\mu + (1-\pi_1-\pi_2) F\left(e_t^i\right) + \pi_2\beta\right] \\ &s.t. \quad e_t^i \geq 0 \\ &\qquad F\left(e_t^i\right) = \log\left(\gamma e_t^i + \nu\right) + \upsilon \end{aligned} \tag{3.2}$$

Note that our assumptions on preferences imply risk-neutrality so that parents have no incentives to insure their children against the risk they incur.[15] The solution of problem 3.2 then implies that the evolution of dynasty $i$'s human capital is described by:

$$h_{t+1}^i = \begin{cases} \mu & w.\ p. \quad \pi_1 \\ \Phi\left(h_t^i\right) & w.\ p. \quad 1-\pi_1-\pi_2 \\ \beta & w.\ p. \quad \pi_2 \end{cases} \tag{3.3}$$

where $\Phi\left(h_t^i\right) = F\left[e_t^{i*}\left(h_t^i\right)\right]$. Note that, when positive, optimal investment in education is given by:

$$e_t^{i*}\left(h_t^i\right) = \frac{\delta\left(1-\pi_1-\pi_2\right)}{1+\delta\left(1-\pi_1-\pi_2\right)} h_t^i - \frac{\nu}{\gamma\left[1+\delta\left(1-\pi_1-\pi_2\right)\right]} \tag{3.4}$$

and that corner solutions arise for $h_t^i \leq \widetilde{h} = \dfrac{\nu}{\delta\gamma\left(1-\pi_1-\pi_2\right)}$.

Let $\gamma > \nu\dfrac{1+\delta}{\delta}$ and $\widetilde{h} > \mu$, $\widetilde{h} \simeq \mu$, so that the function $\Phi\left(h_t^i\right)$ exhibits three steady states, $\mu$, $h_u$, $h^*$, with $\mu < h_u < h^*$, of which two are stable ($\mu$ and $h^*$) and one is unstable ($h_u$), and poverty traps would emerge, absent ability shocks.

[14] As before, we set $A = 1$.

[15] With risk-aversion, the same situation emerges if no institution exists to insure against this type of risk or if access to full insurance is not available to some part of the population. One might think of special education as offering at least partial cover against poor comprehension. If full cover through special education is possible for all dynasties, the introduction of individual risk does not change the dynamics of the model. However, it is quite easy to imagine situations where achieving complete insurance through special education is not possible for all dynasties. For example, if access to special education requires the payment of a fixed cost, poor parents will be unable to afford it, given their inability to borrow on future children's income. In this case, and in any other case where at least a fraction of population has no access to full insurance, the introduction of risk will have implications for the dynamics of income distribution that are qualitatively similar to the ones we reach by ruling out the incentive to insure altogether.

Constancy of the real wage and independence of the productivity of individual investment in education of the aggregate stock of human capital in the economy and of its distribution imply that, at any point in time, the position of a dynasty's human capital in the infinite state space $S = [\alpha, \beta]$ depends only on her position in the state space in the previous period and on the exogenous probabilities $\pi_1$ and $\pi_2$ of being hitten by a shock. In other words, dynastic human capital accumulation follows a *linear* Markov process.

It should be noted that an infinite state space representation is not necessary in this case. In fact, under our assumptions, any dynasty lying in the interval $R = (h_u, \beta]$ at time $t$ either remains there in the following period -this event happens with probability $1 - \pi_1$ independently of the exact position of the dynasty in $R$- or is sent by the individual transition function to $\mu$, which lies in the complement of $R$, $P = [\alpha, h_u]$, -this event happens with probability $\pi_1$, independently of the exact position of the dynasty in $R$. Similarly, any dynasty lying in $P$ at time $t$ either remains there in the following period -this event happens with probability $1 - \pi_2$ independently of the exact position of the dynasty in $P$- or is sent by 3.3 to $\beta$, which lies in the complement of $P$, that is $R$ -this event happens with probability $\pi_2$, independently of the exact position of the dynasty in $P$. Therefore, the properties of our stochastic process can be studied by anlyzing those of a Markov chain defined over the two-states space, $S = (P, R)$. In particular, dynastic human capital accumulation is governed by transition functions which are represented by the $2 \times 2$ stochastic matrix:

$$\Pi = \begin{pmatrix} 1 - \pi_2 & \pi_2 \\ \pi_1 & 1 - \pi_1 \end{pmatrix} \tag{3.5}$$

Since at any point in time we have an infinite mass of dynasties in each state, the probability that a dynasty is in state $s$ at time $t$ is equal to the mass of dynasties which are in state $s$ at time $t$. The dynamics of the income distribution is therefore described by:

$$p_t = p_{t-1}\Pi = p_0\Pi^t \tag{3.6}$$

where $p_t = [G_t(h_u), 1 - G_t(h_u)]$. In this context, we call invariant a distribution of income such that $p^* = p^*\Pi$.

It is immediate to prove that the stochastic process we defined is such that, for all $p_0$ in the one-dimensional unit simplex, there exists a unique invariant distribution $p^*$ in the one-dimensional unit simplex to which the sequence $\{p_0\Pi^t\}$ converges as $t \to \infty$. All we have to do is to verify that the necessary and sufficient condition for convergence of distribution of probabilities over finite state spaces to a unique limit, as stated in Theorem 11.4 in Stockey-Lucas [49], applies to our process. This condition requires that there exists a time $T \geq 1$ such that there exists at least one state $j$ such that the probability of moving from any other state in the state space to $j$ in $T$ periods is positive and is obviously satisfied in our case, since all elements of the the transition matrix $\Pi$ are positive. By Theorem 11.4 in Stockey-Lucas, this implies that $\lim_{t\to\infty} \{\Pi^t\} = Q$, where $Q$ is a stochastic matrix with rows identically equal to $p^*$, which is the unique invariant distribution of $\Pi$, and that $\lim_{t\to\infty} \{p_0\Pi^t\} = p_0Q = p^*$.[16]

By solving the system of linear difference equations defined by ??, it is immediate to characterize the exact path followed by the distribution of income, as well as the limit distribution, $p^*$. In particular, we have:

$$p_t' = \begin{bmatrix} \dfrac{\pi_1}{\pi_1+\pi_2} + \left(G_0\left(h_u\right) - \dfrac{\pi_1}{\pi_1+\pi_2}\right)\left(1-\pi_1-\pi_2\right)^t \\ \dfrac{\pi_2}{\pi_1+\pi_2} + \left(1 - G_0\left(h_u\right) - \dfrac{\pi_2}{\pi_1+\pi_2}\right)\left(1-\pi_1-\pi_2\right)^t \end{bmatrix} \tag{3.7}$$

and $p^* = \left[\dfrac{\pi_1}{\pi_1+\pi_2}, \dfrac{\pi_2}{\pi_1+\pi_2}\right]$.

The introduction of shocks to individual abilities such that at any point in time there exists a positive probability to move back and forth the poverty threshold $h_u$ is sufficient to destroy the poverty trap we derived in the previous section. In fact, the introduction of such shocks has two implications. First, since each row of $Q$ is equal to $p^*$, all dynasties have the same probability of belonging to either income class in the long run, no matter which class they initially belonged to, that is there is complete asymptotic mobility. Second, since the limit distribution of income does not depend on the initial one, initial inequality does not persist in the long run and there is no relation between the initial degree of income inequality and the steady state aggregate level of output.

Before concluding this section a few observations are in order.

First, although the qualitative properties of the shocks we introduced are quite peculiar, in that their support is sufficiently large to allow a reduction of the state space to a two-states set. non-robustness of the poverty traps derived in the context of linear models must be considered a general result. In fact, in such context, conditions for convergence of distribution of probabilities to a unique

[16] Convergence results qualitatively similar to the one we invoked here hold in the infinite state space case.

limit are quite weak and essentially have to do with the possibility of moving from any recurrent state to any other in finite time. In particular, multiplicity of the invariant distribution is possible only if the transition matrix exhibits at least one zero in each column. Such condition is obviously quite easily perturbed by the introduction of shocks to individual income, even with smaller support than the one we assumed above.

Second, non-robustness of poverty traps to the introduction of uninsurable individual risk is by no means a peculiarity of our model, but rather applies to any other linear model where individual transitions depend only on the individual position relative to some poverty threshold and notably to the models of Galor-Zeira [*cit*] and Galor-Tsiddon [*cit*], which both have this feature.

Third, although asymptotic mobility is complete in presence of uninsurable individual risk in the context of linear models, convergence to the unique invariant distribution may be very slow if the probability of being hit by ability shocks is low. This implies that mobility will be low along the transition path to the steady state and that the economy will behave as if there were two separate income classes most of the time. Therefore, though not robust from a theoretical point of view, poverty traps of the kind we analyzed in the previous section may be important observationally. In the following section, we will characterize further the properties of our model in terms of mobility along the transition path to the steady state and study how mobility evolves over time.

Finally, as shown by a few recent papers in a related literature studying the relation between wealth inequality and physical capital accumulation, robust poverty traps can be restored if the evolution of individual wealth is made dependent on the evolution of aggregate wealth and in particular if the evolution of individual wealth is made dependent on wealth distribution, that is if the stochastic process governing individual wealth accumulation is assumed to be non-linear. In fact, under these assumptions, multiplicity of the invariant distribution is possible even if all recurrent states communicate in finite time.[17] In the context of human capital accumulation models, stochastic non-linear Markov processes can naturally be obtained by making the productivity of investment in education and the real wage endogenous to the distribution of income. It might then be possible to characterize situations where the unequal distribution of human capital gives rise to low returns from investment in education and low wages, which, as we showed in the previous section, in turn translate into pervasive poverty traps and persistent inequality. Although, from a theoretical point of view, this kind of

[17] See Banerjee-Newman [*cit*] and Piketty [*cit*].

poverty traps would fare better than those we obtained in the previous section, the latter may still be relevant on empirical grounds. In the next section, we will discuss some casual evidence supporting the view that extreme poverty and cultural backwardness effectively are states of social and economic isolation, such that the link from aggregate to individual behavior is not very effective.

## 4. Relevance of Linear Models

In this section we argue that poverty traps obtained in the context of linear models, though weak from a theoretical point of view, may nevertheless be relevant for at least two reasons. First, we show that how fast the initial income distribution converges to the unique invariant distribution depends on the properties of the shocks: if ability shocks are not very likely, mobility across income classes will be low and income inequality will persist for a long time. Linear models will then be observationally equivalent to non-linear models exhibiting robust poverty traps. Second, on the basis of casual empirical evidence, we document the existence of situations where the links from the distribution of income to individual behavior are indeed quite feeble, as far as private education decisions and individual human capital accumulation are concerned. In these situations, models involving linear dynamics will provide a more accurate description of reality than those involving non-linear dynamics.

### 4.1. Mobility along the Transition to the Steady State

Consider the stochastic generalization of our human capital accumulation model. As we know from the previous section, the distribution of dynasties over the two income classes $P$ (poor) and $R$ (rich) at time $t$ is given by:

$$p_t = p_0 \Pi^t$$

Therefore by setting $p_0$ alternatively equal to the vectors $(1, 0)$ and $(0, 1)$, we obtain information concerning the position in the state space of any dynasty after $t$ generations. In other words, the generic element of $\Pi^t$, $\pi_{ij}^t$, gives the probability that a dynasty initially belonging to class $i$ will belong to class $j$ after $t$ generations.

Mobility in our economy can be gauged by comparing the elements of $\Pi^t$. In particular, we will say that mobility is low at time $t$ if an initially poor (rich) dynasty is more likely to still be poor (rich) after $t$ generations than an initially rich (poor) dynasty is to become poor (rich) after $t$ generations, that is if $\pi_{PP}^t - \pi_{RP}^t > 0$

(or equivalently if $\pi^t_{RR} - \pi^t_{PR} > 0$). Accordingly, mobility will be higher the closer to zero are $\pi^t_{PP} - \pi^t_{RP}$ and $\pi^t_{RR} - \pi^t_{PR}$. Mobility will be complete at time $t$ if the probability that a dynasty belongs to either class after $t$ generations does not depend on which class the dynasty initially belonged to, that is if $\pi^t_{PP} - \pi^t_{RP} = \pi^t_{RR} - \pi^t_{PR} = 0$. Also, we will say that our economy exhibits more upward than downward mobility if the probability that an initially poor dynasty has become rich after $t$ generations is higher than the probability that an initially rich dynasty has become poor after $t$ generations, that is if $\pi^t_{PR} - \pi^t_{RP} > 0$.

In order to derive expressions for the elements of $\Pi^t$, recall that $\Pi = PDP^{-1}$, where $D$ is a diagonal matrix with elements the eigenvalues of $\Pi$ and $P$ is a matrix with columns equal to the corresponding eigenvectors. This implies that $\Pi^t = PD^tP^{-1}$. By making use of this transformation, it is easily verified that:

$$\begin{aligned}
\pi^t_{PP} &= \frac{\pi_1}{\pi_1+\pi_2} + \frac{\pi_2\left(1-\pi_1-\pi_2\right)^t}{\pi_1+\pi_2} \\
\pi^t_{PR} &= \frac{\pi_2}{\pi_1+\pi_2} - \frac{\pi_2\left(1-\pi_1-\pi_2\right)^t}{\pi_1+\pi_2} \\
\pi^t_{RP} &= \frac{\pi_1}{\pi_1+\pi_2} - \frac{\pi_1\left(1-\pi_1-\pi_2\right)^t}{\pi_1+\pi_2} \\
\pi^t_{RR} &= \frac{\pi_2}{\pi_1+\pi_2} + \frac{\pi_1\left(1-\pi_1-\pi_2\right)^t}{\pi_1+\pi_2}
\end{aligned} \tag{4.1}$$

By observation of 4.1, it is immediate to check that, $\forall t < \infty$, $\pi^t_{PP} - \pi^t_{RP} > 0$, $\pi^t_{RR} - \pi^t_{PR} > 0$, $\frac{d\pi^t_{PP}}{dt} < 0$, $\frac{d\pi^t_{RR}}{dt} < 0$, $\frac{d\pi^t_{RP}}{dt} > 0$, $\frac{d\pi^t_{PR}}{dt} > 0$ and $\pi^t_{PR} - \pi^t_{RP} > 0 \Longleftrightarrow \pi_2 > \pi_1$. Moreover, $\lim_{t\to\infty} \pi^t_{iP} = p^*_P = \frac{\pi_1}{\pi_1+\pi_2}$ and $\lim_{t\to\infty} \pi^t_{iR} = p^*_R = \frac{\pi_2}{\pi_1+\pi_2}$, $i = P, R$. It is then possible to make the following remarks.

Remark 3. Along the transition path to the steady state mobility is low in our economy, though increasing with time. Mobility becomes complete only at steady state.

Remark 4. At any point in time, upward mobility is higher than downward mobility in our economy if and only if the probability that a positive ability shock occurs is higher than the probability that a negative ability shock occurs.

In presence of individual uninsurable risk, it is impossible to maintain in the long run a distinction across dynasties in terms of their initial position in the state space. In fact, as we know from the previous section, the introduction of uninsurable risk in our linear human capital accumulation model implies that all dynasties become asymptotically equally likely to end up poor or rich so that differences across classes are overcome in the long run. Yet, along the transition path to the steady state, dynasties face very different opportunities depending on their initial position in the state space, that is, along the transition path to the steady state, our economy behaves as if there existed two income classes. Since mobility across classes is entirely driven by exogenous ability shocks in our model, it is quite intuitive that this situation will last longer, the lower the probability that such shocks occur. The following remark makes this intuitive observation more precise.

Remark 5. Let $k$, $1 \geq k \geq 0$, denote a certain degree of mobility and $t_k$ the instant in time when such degree of mobility is reached, that is:

$$\pi_{PP}^{t_k} - \pi_{RP}^{t_K} = \pi_{RR}^{t_k} - \pi_{PR}^{t_k} = k$$

It is immediate to verify that $t_k = \dfrac{\log(k)}{\log(1 - \pi_1 - \pi_2)}$ and that $t_k$ is decreasing with $\pi_1$ and $\pi_2$. Moreover, $\lim\limits_{(\pi_1+\pi_2)\to 0} t_k = \infty$.[18]

Therefore, if ability shocks occur with very low probability, our economy will be observationally equivalent to one where income inequality persists over time and where no mobility exists across income classes, that is an economy where poverty traps exist.

It might also be interesting to calculate after how long a dynasty can expect to change class with given probability and how this depends on the probability that ability shocks occur. The following remark sheds some light on this issue, showing that the lower the probability that ability shocks occur, the longer dynasties should expect to wait before changing class with given probability.

Remark 6. Let $\rho$ denote a given probability of changing class, $\rho < \dfrac{\pi_z}{\pi_1 + \pi_2}$, $z = 1, 2$, and let $u_\rho$ and $d_\rho$ be defined by $\pi_{PR}^{u_\rho} = \rho = \pi_{RP}^{d_\rho}$. It is immediate to

[18] Obviously, there is no reason to expect that $t_k$ is an entire number. When it is not, the number of generations required to reach the degree of mobility $k$ will be the smallest entire number containing $t_k$.

verify that $u_\rho = \dfrac{\log\left(1 - \dfrac{\rho(\pi_1 + \pi_2)}{\pi_1}\right)}{\log(1 - \pi_1 - \pi_2)}$ and $d_\rho = \dfrac{\log\left(1 - \dfrac{\rho(\pi_1 + \pi_2)}{\pi_2}\right)}{\log(1 - \pi_1 - \pi_2)}$, so that $u_\rho - d_\rho > 0$ if and only if $\pi_2 > \pi_1$. It is also immediate to verify that $\dfrac{\delta u_\rho}{\delta \pi_s} < 0$ and $\dfrac{\delta d_\rho}{\delta \pi_s} < 0$, $z = 1, 2$.[19]

Our discussion so far makes clear that in order to gauge the extent of mobility in our model information is needed about the magnitudes of $\pi_1$ and $\pi_2$. Since direct estimates of $\pi_1$ and $\pi_2$ are quite difficult to obtain, in order to have at least an approximate idea of the extent of mobility in our model, we calculate the number of generations that are required for convergence to the unique invariant distribution for given values of $\pi_1$ and $\pi_2$. Such calculations show that if $\pi_1 = \pi_2 = 0.3$,[20] convergence already requires eight generations (slightly more than one century). Moreover, if $\pi_1 = \pi_2 = 0.03$, the number of generations required for convergence rises over one hundred (more than sixteen centuries) and if $\pi_1 = \pi_2 = 0.003$, convergence requires more than one million generations! On the grounds of these calculations, we are led to conclude that poverty traps of the kind we analyzed in the previous section may well be important observationally.

## 4.2. Some Casual Evidence

As we argued above, robust poverty traps can be restored by assuming that individual transitions depend on the distribution of income so that they define non-linear dynamics of human capital accumulation. The idea is that initially unequal distributions of income are more likely to generate poverty traps so that inequality is more likely to persist in initially unequal economies. Here, we highlight some of the channels through which the distribution of income should matter for individual transitions in the context of human capital accumulation models. Then, we argue that these channels might not be effective in some situations and, on the

[19] As before, there is no reason to expect that $u_\rho$ and $d_\rho$ are integers. When they are not, the number of generations required to reach the degree of upward and downward mobility $\rho$ will be the smallest integer respectively containing $u_\rho$ and $d_\rho$.

[20] These are approximately the probabilities that a 1965 born child attains the top (lowest) earnings quartile in 1981 conditional upon his father's earnings in 1965 lying in the two bottom (top) quartiles as measured by Zimmerman [*cit.*] using U.S. data. These probabilities provide a good measure of the elements of $\Pi^t$ at a given point in time, but are likely to largely overestimate the "true" values of $\pi_1$ and $\pi_2$.

basis of casual empirical evidence, we document the existence of such situations. This provides an empirical argument in favor of linear models.

Non-linearity of individual human capital dynamics can be obtained in the context of our human capital accumulation model by removing the assumption that the productivity of individual investment in education and the wage rate, and more generally factor prices, do not depend on the distribution of human capital.

One compelling reason why the productivity of individual investment in education should depend on the distribution of human capital are spillover effects in the learning process. In particular, it is reasonable to expect that the productivity of individual investment in education is lower the larger the fraction of children coming from poor and uneducated families in the learning environment (peer-effects, role models, norms of behavior, crime, etc.). Since low productivity of individual investment in human capital causes wide-spread poverty traps, this implies that inequality is more likely to persist when the distribution of income is initially unequal. Recent empirical analysis on externalities in the learning process shows that social capital in effect has an important role in human capital formation, at least at the local level. Brooks-Gunn et al. [21], Case and Katz [24] and Crane [25] provide evidence that variables having to do with the distribution of income at the neighborhood level, such as the fraction of high income families in the neighborhood, the fraction of youths involved in crime in the neighborhood and the fraction of workers in the neighborhood holding professional or managerial jobs, affect drop out rates in ways that are consistent with the existence of significant spillover effects at the neighborhood level. Also, Downes and Pogue [26] show that communities with a higher fraction of students from disadvantaged backgrounds have higher per-student costs, for a given level of educational achievement. Yet, evidence of the existence of spillover effects at the local level is not sufficient to dismiss linear models altogether. In fact, if spillover effects only matter at the local community level (e.g. neighborhood, village, city) and are related to the clustering of poor families in poor neighborhoods and to the inability of poor families from poor villages and cities to migrate to rich cities[21], one might argue that linear models still provide an acceptable representation of the country-wide

[21] One can think of different explanations for why very poor families from poor villages face disproportionally high mobility costs. Banerjee and Newman [8] argue that, in presence of imperfect information, informal credit institutions in rural villages can provide very poor agents with insurance against risks that would be unavailable to them in the city's formal credit market. This prevents migration of very poor agents from villages to cities, where they might earn higher wages.

reality, while non-linear models better describe local dynamics.[22] In particular, if the chance that a family escapes poverty once in a poor community does not depend on the distribution of income across communities, linear models can still apply if we take representative families of each community as units of observation.

A link from the distribution of human capital to the wage rate is quite naturally obtained by removing the assumption that different types of human capital are perfect substitutes in production. In particular, if skilled labor (supplied by educated agents) is complement in production with unskilled labor (supplied by uneducated agents), the smaller the ratio of educated to uneducated agents, that is the more unequal the distribution of human capital, the lower (higher) unskilled (skilled) labor's productivity and the lower (higher) unskilled (skilled) labor's wage in a competitive labor market. Although a large gap between skilled and unskilled labor's wages obviously raises the rate of return of investment in education and might provide an incentive to increase it, such incentive is likely to be ineffective at very low income levels.[23] On the contrary, the more unequal the distribution of human capital and the lower the unskilled labor's wage, the more wide-spread is poverty and the more persistent is inequality. This link will be enfeebled if complementarities in production between skilled and unskilled labor are only relevant at the local level and if labor markets are spatially segmented due to the inability of poor families from poor villages and cities to migrate to rich cities. If this is the case, again linear models can provide an acceptable description of reality, if we take representative families of each community as units of observation. Recent empirical work on U.S. data indeed lends some support to these presumptions: Rauch [44] finds that average productivity is higher in metropolitan areas where average human capital is higher. Montgomery [36] provides evidence of wage differentials across metropolitan areas. Finally, Topel [51] finds that distinctly local factors affect relative wages across regions and suggests that the extent of the labor market is limited by geography. On the other hand, if spatial labor markets segmentation does not sound appealing, another argu-

[22] Benabou [15], [16], [17] and Durlauf [27] construct theoretical models where local spillover effects in the production of human capital endogenously lead to the segregation of poor families in poor neighborhoods and consequently to permanent poverty and persistent income inequality in cities where income is initially unequally distributed.

[23] In our model, the skilled-unskilled labor's wage differential is not a key determinant of the existence of poverty traps. In principle, however, it is not hard to generalize the model so as to allow for this possibility. In fact, it is sufficient to assume that parents care about their children's income rather than about their human capital. *Coeteris paribus*, a higher wage differential would then make investment in education more profitable.

ment in favor of linear models rests on quite the opposite view: in an integrated world, a country's distribution of human capital is not an important determinant of the skilled-unskilled labor's wage differential, which tends to converge to common world-wide values due to free trade. The relevant question here is whether recent economic integration and globalization has in fact pushed in the direction of factor price equalization. Available empirical evidence suggests that there has been a sustantial tendence to factor price equalization across industrialized countries in the nineteenth and twentieth centuries[24] and confirms that international trade played an important role in explaining such tendence[25]. Also, recent empirical evidence shows that large trade deficits with developing low-wage countries contribute to explain the upsurge of wage inequality in the US during the 1980s[26].

The assumption that a market for education loans does not exist rules out the possibility that the interest rate channel generates non-linearity of individual human capital dynamics in our model. Instead, such non-linearities might emerge if the credit market imperfection took less extreme forms.[27] However, as we argued above, the assumption that a loan market does not exist reflects an important feature of reality in the context of human capital accumulation model. Therefore, the linearity of human capital accumulation models in this respect is quite justified.

Our discussion makes clear that there exist situations where the links from the distribution of income to individual behavior are indeed quite feeble as far as human capital accumulation is concerned. In these situations, linear models, though not robust from a strictly theoretical point of view, may provide a better description of reality than non-linear models and this may justify their use.

# 5. Concluding Remarks

In this paper, we analyze the emergence and persistence of poverty traps and study how wide-spread poverty and the unequal distribution of income can slow down the accumulation process and lead to steady-state equilibria characterized by low aggregate output levels. We define poverty as a state deriving from the lack of adequate skills and we associate income inequality with the unequal distribution of education attainments. In this context, we show that low asymptotic mobility

[24] See Williamson [52].

[25] See Williamson [53].

[26] See Katz and Murphy [32], Leamer [35] and Sachs and Shatz [45].

[27] See Aghion-Bolton [1], Banerjee-Newman [*cit.*] and Piketty [*cit.*].

and persistent income inequality can emerge as a consequence of the fact that the poor require relatively higher returns to increase expenditure on education, so that they devote to education smaller shares of their income than the rich.

Our contribution shares with other models the feature that individual human capital dynamics do not depend on the distribution of human capital, that is the Markov process governing individual human capital accumulation is linear. Although, as we show, this feature makes the poverty trap we derive theoretically not robust to the introduction of individual uninsurable risk, still we believe that our approach can provide useful insights on the mechanisms that generate persistence of income inequality and low intergenerational mobility. In fact, on one hand, we believe that mobility associated to uninsurable risk is likely to be very low. This implies that our model can be considered observationally equivalent to one where robust poverty traps exist. On the other hand, we believe that there exist situations where extreme poverty and cultural backwardness can be considered states of social and economic isolation. In these situations the link from aggregate to individual behavior is not very effective and the use of a linear model is empirically justified.

## References


[1] Aghion, P. and P. Bolton (1993), "A Theory of Trickle-Down Growth and Development", Nuffield College and LSE, mimeo.

[2] Alesina, A. and D. Rodrick (1994), "Distributive Politics and Economic Growth", *Quarterly Journal of Economics*, 109: 465-490.

[3] Altonji, J. and T. Dunn (1995), "The Effects of School and Family Characteristics on the Return to Education", NBER W.P. 5072.

[4] Atkenson, A. and M. Ogaki (1992), "Wealth-varying Intertemporal Elasticities of Substitution: Evidence form Panel and Aggregate Data", University of Rochester and University of Chicago, mimeo.

[5] Banerjee, A. and A. Newman (1991), "Risk-Bearing and the Theory of Income Distribution", *Review of Economic Studies*, 58: 211-235.

[6] ____ and ____ (1993), "Occupational Choice and the Process of Economic Development", *Journal of Political Economy*, 101: 274-299.

[7] ____ and ____ (1994), "Poverty, Incentives and Development", *American Economic Review Papers and Proceedings, 101: 211-216.*

[8] ____ and ____ (1995) "Migration, Integration and Development", Columbia D.P. 734, June 1995.

[9] Barham, V., Boadway, R., Marchand, M. and P. Pestieau (1995), "Education and the Poverty Trap", *European Economic Review*, 39: 1257-1275.

[10] Becker, G. (1975), *Human Capital*, University of Chicago Press.

[11] ___ (1991), *A Treatise on the Family*, Harvard University Press.

[12] ___ and N. Tomes (1979), "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility", *Journal of Political Economy*, 87: 1153-1189.

[13] Berhman, J, Kletzer, L., McPherson, M. and O. Morton (1992), "The College Investment Decision: Direct and Indirect Effect of Family Background on the Choice of Post-Secondary Enrollment and Quality", Williams College, mimeo.

[14] _____, Pollack R. and P. Taubman (1989), "Family Resources, Family Size and Access to Financing for College Education", *Journal of Political Economy*, 97.

[15] Benabou, R. (1992), "Heterogeneity, Stratification and Growth", MIT, W.P. 93-4.

[16] _____ (1993) "Workings of a city: Location, Education and Production", *Quarterly Journal of Economics*, 108: 123-150.

[17] _____ (1994) "Education, Income Distribution and Growth: the Local Connection", NBER W.P. 4798.

[18] Blinder, A.S. (1975), "Distribution Effects and the Aggregate Consumption Function", *Journal of Political Economy*, 83: 447-476.

[19] Borjas, G.J. (1991), "Ethnic Capital and Intergenerational Mobility", *Quarterly Journal of Economics*, 108: 123-150.

[20] Bourguignon, F. (1981), "Pareto Superiority of Unegualitarian Equilibria in Stiglitz' Model of Wealth Distribution with Convex Saving Function", *Econometrica*, 49: 1469-1475.

[21] Brooks-Gunn, J., Duncan, G., Kato Klebanov, P. and N. Sealand (1993), "Do Neighborhoods Influence Child and Adolescent Development?", *American Journal of Sociology*, 99: 353-395.

[22] Chatterje, S. (1994), "Transitional Dynamics and the Distribution of Wealth in a Neoclassical Growth Model", *Journal of PUblic Economics*, 54: 97-119.

[23] Cameron, S. and J. Heckman (1992), "The Dynamics of Educational Attainment for Blacks, Whites and Hispanics", University of Chicago, mimeo.

[24] Case, A. and L. Katz (1991), "The Company You Keep: the Effects of Familiy and Neighborhood on Disadvantaged Youths", NBER W.P. 3705.

[25] Crane. J. (1991), "The Epidemic Theory of Ghettos and Neighborhood Effects on Dropping Out and Teenage Childbearing", *American Journal of Sociology*, 96: 1226-1259.

[26] Downes, T. and T. Pogue (1993), "Adjusting School Aid Formulas for the Higher Cost of Educating Disadvantaged Students", Northwestern University, mimeo.

[27] Durlauf, S. (1996), "A Theory of Persistent Income Inequality", *Journal of Economic Growth*, 1: 75-93.

[28] Galor, O and D. Tsiddon (1992), "Income Distribution and Output Growth: the Kuznets' Hypothesis Revisited", Brown University and the Hebrew University of Jerusalem, mimeo.

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

[30] Glennerster, H. (1993), "The Economics of Education: Changing Fortunes", in *Current Issues in the Economics of Welfare"*, Barr, N. and D. Whynes, eds., MacMillan.

[31] Guaitoli, D. (1994), "Income Distribution, Growth Effects and Convergence", University of Chicago, mimeo.

[32] Katz, L. and K. Murphy (1992), "Changes in Relative Wages, 1963-1987: Supply and Demand Factors", *Quarterly Journal of Economics*, 107: 34-78.

[33] Lawrence, E.C. (1991), "Poverty and the Rate of Time Preference: Evidence from Panel Data", *Journal of Political Economy*, 99: 54-77.

[34] Leamer, E. (1996), "Wage Inequality from International Competition and Technological Change: Theory and Country Experiences", *American Economic Review, Papers and Proceedings*, 86: 309-314.

[35] Lowry, G.C. (1981), "Integenerational Transfers and the Distribution of Earnings", *Econometrica*, 49: 843-867.

[36] Montgomery, (1992), "Metropolitan Wage Differentials", *Journal of Urban Economics"*,

[37] Ogaki, M. and Atkenson. A., (1991), "Estimating Subsistence Levels with Euler Equations in Indian Panel Data", University of Rochester and University of Chicago, mimeo.

[38] Perotti, R. (1994), "Income Distribution and Investment", *European Economic Review*, 38: 827-836.

[39] Perotti, R. (1996), "Growth, Income Distribution and Democracy: What the Data Say", *Journal of Economic Growth*, 1.

[40] Persson, T. and G. Tabellini (1994), "Is Inequality Harmful for Growth? Theory and Evidence", *American Economic Review*, 84: 600-621.

[41] Piketty, T. (1992), "The Dynamic of Wealth Distribution and the Interest Rate with Credit Rationing", MIT, mimeo.

[42] Psacharopoulos, G. (1973), *Returns to Education: an International Comparison*, Elsevier Scientific Publishing Company, Amsterdam.

[43] ____ (1987), *Economics of Education: Research and Studies*, Pergamon Press, Oxford.

[44] Rauch, (1993), "Productivity Gains from Human Capital Concentration", *Journal of Urban Economics*

[45] Sachs, J. and H. Shatz (1996), "U.S. Trade with Developing Countries and Wage Inequality", *American Economic Review, Papers and Proceedings*, 86: 234-239.

[46] Schultz, P.T. (1988), "Education Ivestments and Returns", in *Handbook of Development Economics*, Chenery, H. and T.N. Srinivasan eds., Elsevier Science Publishers.

[47] Solon, G. (1992), "Intergenerational Income Mobility in the U.S.", *American Economic Review*, 82: 393-408.

[48] Stiglitz, J. (1969), "Distribution of Income and Wealth among Individuals", *Econometrica*, 37: 382-397.

[49] Stockey, N. and R. Lucas (1989), "*Recursive Methods in Economic Dynamics*", Harvard University Press.

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

[51] Topel, (1994), "Wage Inequality and Regional Labor Market Performance in the US". *American Economic Review, Papers and Proceedings*, 17-22.

[52] Williamson, (1995), "The Evolution of Global Labor Markets since 1830: Background Evidence and Hypotheses", *Explorations in Economic History*.

[53] _____ (1995), "Globalization, Convergence and History", *Journal of Economic History*.

[54] Zimmerman, D. (1992), "Regression Toward Mediocrity in Economic Stature", *American Economic Review*, 82: 409-429.

# 6. Appendix

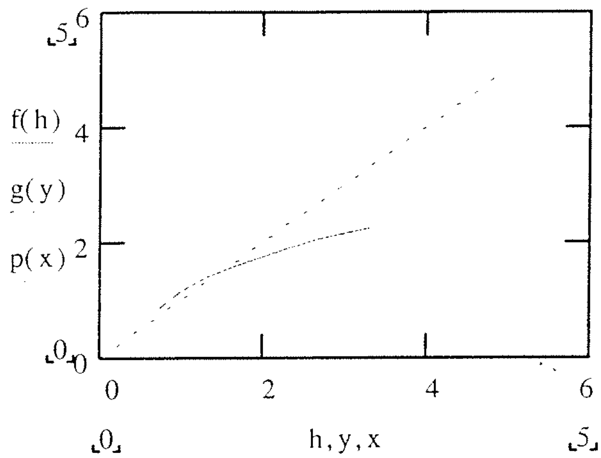


Figure 6.1:

$p(x) = \mu$; $g(y) = y$; $f(h) = \Theta\left(h_t^i\right)$, $h_t^i > \underline{h}$
$A = 1$, $\delta = 0.5$, $\gamma = 12$, $\nu = 1$, $\upsilon = 0.3$

Case (a)

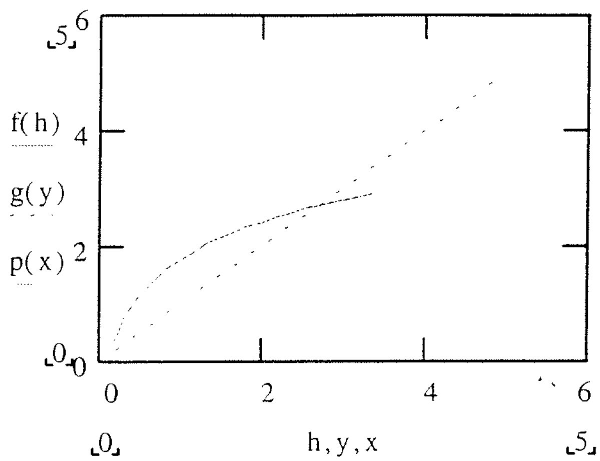


Figure 6.2:

$p(x) = \mu$; $g(y) = y$; $f(h) = \Theta\left(h_t^i\right)$, $h_t^i > \underline{h}$
$\delta = 0.5$, $\nu = 1$, $v = 0.3$
$A = 1$ and $\gamma = 12$ or $A = 6$ and $\gamma = 2$

Case (b)

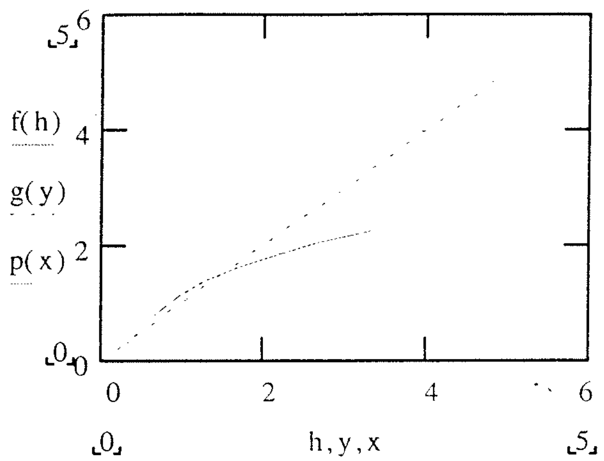


Figure 6.3:

$p(x) = \mu$; $g(y) = y$; $f(h) = \Theta(h_t^i)$, $h_t^i > \underline{h}$
$\delta = 0.5$, $\nu = 1$, $\upsilon = 0.3$
$A = 1$ and $\gamma = 6$ or $A = 3$ and $\gamma = 2$

Case (c)

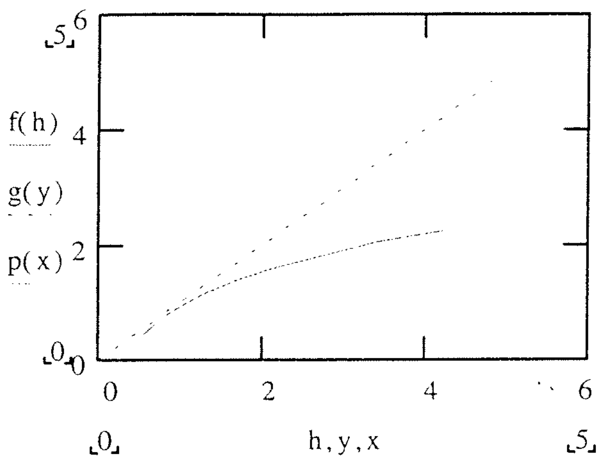


Figure 6.4:

$p(x) = \mu$; $g(y) = y$; $f(h) = \Theta\left(h_t^i\right)$, $h_t^i > \underline{h}$
$\delta = 0.5$, $\nu = 1$, $\upsilon = 0.3$
$A = 1$ and $\gamma = 4.7$ or $A = 2.4$ and $\gamma = 2$

Case (d)

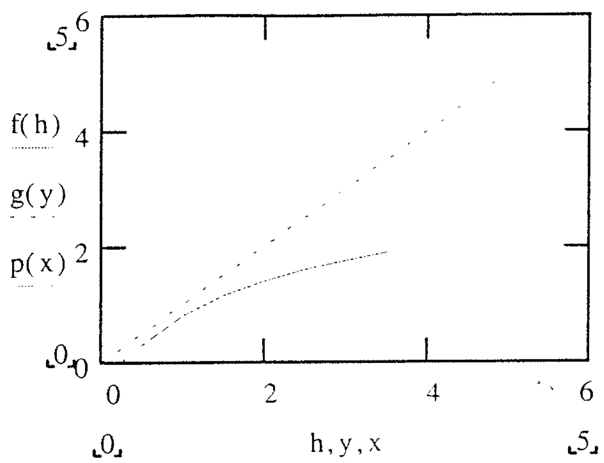


Figure 6.5:

$p(x) = \mu$; $g(y) = y$; $f(h) = \Theta(h_t^i)$, $h_t^i > \underline{h}$
$\delta = 0.5$, $\nu = 1$, $\upsilon = 0.3$
$A = 1$ and $\gamma = 4$ or $A = 2$ and $\gamma = 2$