
ELSEVIER
Volume 84, No. 1, September 2007
ISSN 0304-3878
Editor:
MARK ROSENZWEIG
Co-Editors:
A. CASE
E. DUFLO
G.H. HANSON
D. MOOKHERJEE
L. PRITCHETT
Book Review Editor:
CLIVE BELL
Associate Editors:
T. BESLEY
C. BRODA
R. DI TELLA
W. EASTERLY
M. FAFCHAMPS
H. JACOBY
L. KILIAN
A. KRAAY
E. LIGON
E. MIGUEL
K. MUNSHI
R. PANDE
N. PAVCNIK
D. RAY
J. TEMPLE
D. THOMAS
84(1) 1–572
(2007)
JOURNAL OF
Development
ECONOMICS
CONTENTS
D. Mckenzie and H. Rapoport, Network effects and the dynamics of migration and inequality: Theory and evidence from Mexico 1
K. Grier and R. Grier, Only income diverges: A neoclassical anomaly 25
C. Van Marrewijk and K.G. Berden, On the static and dynamic costs of trade restrictions for small developing countries 46
A. Mizala, P. Romaguera and M. Urquiola, Socioeconomic status or noise? Tradeoffs in the generation of school quality information 61
K. Yuki, Urbanization, informal sector, and development 76
P. Sauré, Revisiting the infant industry argument 104
F. Estevan and J.-M. Baland, Mortality risks, education and child labor 118
Contents continued on outside backcover

ELSEVIER

Journal of Development Economics 84 (2007) 251–270

JOURNAL OF Development ECONOMICS

www.elsevier.com/locate/econbase

Persistent inequality: An explanation based on limited parental altruism

Mausumi Das

Department of Economics, Delhi School of Economics, University of Delhi, Delhi 110007, India

Received 1 June 2003; accepted 1 October 2005

Abstract

This paper proposes an alternative theory for the observed persistence in income inequality across households, a theory based on limited parental altruism. We argue that the degree of parental altruism is 'limited' by the economic status of the parent. A poor parent not only has less ability, but also has less willingness to invest in children's human capital formation. This generates a non-linearity in such investment expenditures. As a result, initial income differences may perpetuate over time—even with convex technology and convex preferences. In this context, we also compare the efficacy of the public vis-à-vis the private education system from the perspective of long run growth.

JEL classification: O40; D31
Keywords: Income distribution; Human capital; Intergenerational mobility; Growth

1. Introduction

In recent years, the endogenous growth literature has shed new lights on issues pertaining to income distribution, human capital formation and intergenerational mobility. In two influential papers, Galor and Zeira (1993) and Banerjee and Newman (1993) have argued that inequality in income distribution might persist in the long run in the presence of credit market imperfections and some degree of technological indivisibility. In an unequal society, credit market imperfections lead to unequal opportunities to invest in the short run—resulting in polarization, and polarization is perpetuated in the long run due to the assumed indivisibility in the investment technology. In so far as the production technology exhibits non-decreasing returns to scale, such polarization negatively affects the long run growth scenario.

E-mail address: mausumi@econdse.org.

doi:10.1016/j.jdeveco.2005.10.006

An impressive volume of literature has developed subsequently that emphasizes the role of credit market imperfections in explaining intergenerational mobility or the lack of it. Some important contributions in this area include Freeman (1996), Aghion and Bolton (1997), Piketty (1997), Owen and Weil (1998), Maoz and Moav (1999), Ghatak and Jiang (2002) and Mookherjee and Ray (2002, 2003).[1] Most of these models however rely on the non-convexity of technology to generate long run persistence; credit market imperfections alone cannot generate this result. This is essentially because these models assume homothetic preferences, whereby low income only impinges upon investment from the supply side (through credit market imperfection), not from the demand side (through preferences). Since homothetic preferences imply a constant investment propensity across households (rich and poor alike), non-convexity of technology becomes important in explaining divergent long run outcomes. In this paper, we explore an alternative channel–a channel working from the preference side–through which income may impinge upon investment, and therefore future earning abilities, of the poorer households. We consider a model of human capital accumulation where the credit market is altogether missing. We show that, despite convex preferences and convex technology, inequality may perpetuate in the long run due to a complex interaction between income and preferences, which negatively affects the human capital formation decision of a poor household.

We argue that in any family the human capital formation decisions affecting the next generation (e.g., how much to investment in children's schooling, health care, etc.) are typically undertaken by the parents. Therefore, the degree of parental altruism plays an important role in determining the future earning abilities of the children. But parental perceptions about the utility of children's education seem to vary as one moves from the higher income to the lower income families. Typically in a poor family, which is close to subsistence, consumption of the family assumes more importance than the level of education of the children. Thus, the crucial assumption that we make in this paper is that the degree of parental altruism is endogenously determined and it varies with the earning ability of the parent.

The paper bestows 'warm glow' kind of altruism on the parents where parents derive direct utility by incurring expenditure on children's education and/or health, i.e., on human capital formation. However, we assume that parental altruism is 'limited' by the income status of the parent. The postulated positive relationship between the degree of parental altruism and parents' economic status has been captured by introducing a weight on the utility derived from expenditure on children's human capital formation, and the weight is assumed to be an increasing function of the parent's own consumption.[2]

This particular specification of the preference pattern spells out different outcomes for the rich vis-à-vis the poorer households. A poor parent is likely to attach less weight to children's education than her rich counterpart. As a result, not only does she have less ability to invest in children's human capital formation, but also has less *willingness*—a factor that contributes significantly to the perpetuation of lower earning abilities generation after generation.

[1] An alternative theory of persistent inequality has been put forward by Benabou (1994, 1996) and Durlauf (1996). This strand of the literature seeks to explain stratification or local segregation of communities in terms of local human capital externalities and analyses its impact on human capital formation at the community level.

[2] A somewhat similar route was followed by Cardak (1999), who introduced an "idiosyncratic weight" on education expenditure in the parents' utility function to represent heterogeneity in preferences. The crucial difference is that in our model this weight is endogenously determined. Thus, in our model, there is no inherent difference in the preference pattern of the poor and the rich households. Households are homogeneous in terms of tastes and preferences—they have exactly identical preference ordering. They only differ in terms of their earning abilities.

Moav (2002) has shown that in a credit constrained economy the convexity of the bequest function, whereby poorer parents bequeath lower *proportion* of their income to their children, could generate intergenerational persistence of inequality—even in the absence of technological indivisibilities. However, Moav simply assumes a convex bequest function. In this paper, on the other hand, we provide an economic justification as to why indeed the bequest or savings function is likely to be convex. In this sense, this paper provides a micro-foundation to Moav's argument.

There exist a number of socio-medical studies that link parental care for children to parents' socio-economic status. It has been argued in this context "poverty creates a heightened parental stress, straining or limiting the capacity of parents to provide warmth, understanding and guidance for their children".[3] These 'parental stress' theories, widely held in the fields of sociology and psychology,[4] thus postulate an indirect relationship between parental income and children's welfare: income initially affects the behaviour of parents, which in turn affects their children. This in effect implies that a poor person would be less concerned about her children's overall well-being—including their educational attainment.

A related, but slightly different argument was put forward by Irving Fisher who emphasized that poorer people are more concerned about current consumption (which may include the current consumption of their children as well), and in so far as investment in children's education necessitates a reduction in the family's current consumption, it is a luxury that they can ill-afford. This argument in effect implies that poorer households have a higher rate of time preference. In the Theory of Interest, Fisher writes, "Poverty bears down heavily on all portions of a man's expected life, both that which is immediate and that which is remote. But it enhances the utility of immediate income even more than the future income".[5] This hypothesis has found recent support in the writings of Koopmans (1986), Blanchard and Fischer (1989) and Barro and Sala-i-Martin (1995). Note that Fisher's argument does not imply a lower degree of altruism on the part of the poorer households per se (in so far as the household's current utility depends on the current consumption of the entire family—including that of the children), but it nevertheless implies that poorer households would be less concerned about children's education—since the latter essentially impacts on their future consumption. Our model attempts to capture this aspect of poverty and analyse its implication for intergenerational mobility across households and the long run pattern of development of an economy.[6]

Economic studies relating altruism to economic status is rather scarce (presumably because identification and measurement of an unobservable characteristic like parental altruism is difficult), but a very recent study indeed indicates that degree of parental altruism is an increasing function of income. On the basis of primary level data on Australia (where questionnaires were designed to assess respondents' 'attitude' towards financing children's education), Beal (2001) shows that the belief in supporting adult children during years of higher education is positively

[3] Hisnanick and Coddington (2000, p. 82).

[4] Subscribers include Elder et al. (1985), Parker et al. (1988), McLoyd (1990), Cogner et al. (1992) and Huston et al. (1994). McLoyd (1990) provides an excellent survey.

[5] Fisher (1930, p. 72).

[6] The concept that poorer people might have a higher rate of time preference and its implications for economic development and growth has remained largely unexplored in economic literature, although the alternative assumption, namely that richer people have higher rate of time preference, have been applied widely in the growth literature. Some of our recent works (Das, 2000, 2003) provide complete characterization of the dynamic path under the assumption that poorer people are more impatience in the context of an infinite horizon optimal growth framework with physical capital formation.

linked to income and wealth of parents.[7] Indirect support can also be found in Jensen and Richter (2001) who observe that smoking and drinking behaviour of rich and poor women in Russia during pregnancy varies with the income status: wealthier women reduce or eliminate these behaviours, whereas women belonging to lower socio-economic status are less likely to change. Meara (2001) reports similar results for the United States.

Income-wise classification of data on households' educational expenditure could also impart valuable insights into the households' attitude towards children's education. If poor households care less for children's education, then this should get reflected in the households' expenditure on education. Poor households would not only spend less, but would also spend lower proportions of their income on children's education. In other words, education is likely to have the characteristics of a luxury commodity, exhibiting a convex Engel curve with income elasticity greater then unity.[8] The household level data on educational expenditure in India does lend support to this hypothesis. On the basis of National Accounts Statistics (NAS) data on household expenditure for the period 1950–1951 to 1996–1997, Tilak (2000) reports that that the coefficient of elasticity of household expenditure on education to total expenditure (which serves as a proxy of household income) is about 1.5. The coefficient of elasticity turns out to be even higher when these figures are adjusted for household sizes. More direct evidence is found from a study that examines the causes of non-enrollment and de-enrollment among children belonging to the age group of 6–14. Using National Council of Applied Economic Research (NCAER) 1994 survey data, Duraisamy (2002) finds that the percentage of children belonging to this age group who are currently enrolled goes up as one moves from lower income groups to higher ones. Moreover, one of the chief reasons cited in the study for not attending school by children is that "parents feel not important". This, in fact, has been cited as the single most important reason for discontinuation of schooling (see Duraisamy, 2002, Table 2, p. 12). Thus, it appears that parental preferences do play an important role in children's education decision.

At the theoretical level, our work comes close to that of Mulligan (1997), who also develops a model of endogenous parental altruism. According to Mulligan, parental altruism depends on the time spent with children, the latter being optimally decided by the parents on the basis of a utility maximization exercise. Since high wage families have higher opportunity cost of time, high-income dynasties will spend less time with children and therefore will be less altruistic.[9] Note that Mulligan's model thus predicts a concave bequest function. This however contradicts the available empirical evidence which strongly suggests a convex bequest/savings function (e.g., Menchik and David, 1983; Carroll, 2000; Dynan et al., 2004). To our mind, the postulated positive relationship between altruism and income in our model is a more realistic way to capture the inherent endogeneity in parental altruism, which is consistent with the empirical evidence and also enables us to explain the observed persistence of inequality in income across households.

In the context of human capital formation, growth and distribution, the education system plays an important role. In so far as inequality impinges upon long run growth by hindering human capital formation, it has often been argued that, instead of a direct re-distributive policy which is

[7] Note that this study directly corroborates our hypothesis. Since the questionnaires aim to identify parental *attitude* or *belief* in supporting children's education (i.e., the study is not based on ex post data about how much parents actually invest), the question of credit constraint limiting poor parents' actual choice set does not arise.

[8] I am grateful to Nobuhiro Kiyotaki for pointing out this alternative interpretation.

[9] Similar hypothesis has been put forward in the context of households' time preference in Becker and Mulligan (1997).

sometimes politically infeasible, one can operate through a public education policy to achieve similar results. The issue of public education in the context of growth and inequality was first addressed by Glomm and Ravikumar (1992) who analyzed the relative merits and demerits of the public education system in a standard overlapping generations model with human capital formation. The issue assumes special significance in the context of our model: given that poor parents at the margin are less willing to incur expenditure on children's education, in a poor economy characterized by high inequality in skill and income distribution, would public education system perform better than private education system in terms of growth? In order to address this issue, we consider an alternative structure incorporating a public education system which is financed by a uniform proportional income tax. As in Glomm and Ravikumar, we assume that the tax rate is decided by majority voting. We derive the condition under which the public education system performs better than the private education system in terms of long run growth.

The paper is organized as follows. Section 2 develops the micro-foundation for a convex bequest curve in terms of endogenous altruism. Section 3 lays out the corresponding macro-structure of the economy and discusses the intergenerational dynamics. In Section 4, we introduce a tax-financed public education system as a policy instrument and discuss its implications from the point of view of long run growth. Section 5 offers the final comments and conclusion.

2. The micro-framework of household choices

In this section, we develop the micro-foundation for a convex bequest function based on endogenous parental altruism. In the discussion that follows, we first analyse the concept of endogenous altruism in the present context. We then illustrate how allowing for endogenous altruism may generate a convex bequest function even under standard assumptions about the utility function. In this context, we also elucidate the close link between endogenous altruism and endogenous time preference.

The micro-analysis is essentially centred around a representative agent's preferences and choices. Consider a representative agent with an income y. The agent derives utility from own consumption as well as from the investment made on children's education (where investment incurred on children's education can be thought of as a form of bequest). Let $\bar{c}$ be some subsistence level of consumption which each household must maintain in order to survive. Any consumption above this subsistence level gives them positive utility, as does children's education. However we assume that for all agents $y \geq \bar{c}$, so that the subsistence consumption does not play any significant role in the optimal behaviour of the agent.[10]

That expenditure incurred on children's education generates positive utility for the agent implies the presence of 'warm glow' kind of altruism.[11] The preferences of the representative agent are denoted by the following utility function:

$$W(\hat{c}, b) = u(\hat{c}) + \beta u(b) \tag{1}$$

[10] The only role of subsistence consumption here is that of shifting the origin from zero consumption to some positive consumption level $\bar{c}$. As we shall see in the next section (when we consider the intergenerational dynamics), this provides a lower bound to the downward movement of income and consumption.

[11] An alternative way to capture altruism is to let the parental utility function depend directly on the children's utility level. This in effect would generate an infinite horizon dynastic utility function for the household. We have discussed later why this does not make any difference to our basic argument (see Remark 2).

where $\hat{c}$ denotes consumption over the subsistence level, b denotes the amount invested in children's education and β is the weight attached to utility derived from children's education, which represents the *degree* of parental altruism.

The instantaneous utility function $u(.)$ follows all the standard properties as specified below.

Assumption 1. $u(.)$ is a real valued, twice continuously differentiable function defined on $(0,\infty)$ such that $u(0)=0$, and for all $c,b \geq 0$, $u'(.)>0$; $u''(.)<0$. Further, the function $u(.)$ exhibits constant elasticity with respect to its arguments, such that $\sigma_u \equiv \frac{xu'(x)}{u(x)}$ $(x = \hat{c}, b)$ is a positive constant with values lying within (0,1).

All the assumptions about the utility function are rather standard. A constant-elasticity u function has been assumed to ensure that, when the altruism coefficient β is constant, the resulting utility function becomes homothetic in $\hat{c}$ and b.[12]

The key innovation of the paper involves endogenization of the altruism coefficient β. We postulate that the *degree* of altruism is related the parents' own consumption level. To be more specific, we assume that the degree of altruism is an increasing function of consumption over and above the subsistence level. This fact is captured by assuming that the weight β depends positively on consumption of the parent over subsistence. Accordingly, we specify the following set of assumptions which characterizes the β function:

Assumption 2. $\beta(\hat{c})$ is a real valued, twice continuously differentiable function defined on $(0,\infty)$ such that, $\beta(0)=0$, and for all $\hat{c} \geq 0$, $\beta'(.)>0$; $\beta''(.)<0$. Further $\beta(.)$ exhibit constant elasticity with respect to its argument, i.e., $\sigma_\beta \equiv \frac{\hat{c}\beta'(\hat{c})}{\beta(\hat{c})}$ is a positive constant lying within (0,1).[13]

Given Assumption 2, the utility function of the representative agent can be written as:

$$W(\hat{c}, b) = u(\hat{c}) + \beta(\hat{c})u(b) \tag{2}$$

Note that $\beta'(.)>0$ is the crucial assumption that embodies our hypothesis of limited parental altruism; the rest of Assumption 2 are standard regularity conditions. Assumptions 1 and 2 together ensure that the agent's preference schedule $W(\hat{c},b)$ is monotonic and quasi-concave in $\hat{c}$ and b. As we show below, these set of assumptions are also sufficient to generate a convex bequest function even though the instantaneous utility function $u(.)$ specified under Assumption 1 cannot generate a convex bequest function on its own (i.e., when β is exogenous). Thus, endogenous altruism, as specified under Assumption 2, has important implications for the savings/bequest decisions made by a poor family vis-à-vis a rich family.

In order to prove that Assumptions 1 and 2 are sufficient to generate a convex educational bequest function, first consider the following lemma.

Lemma 1. *Under Assumptions 1 and 2, the marginal rate of substitution between consumption (above subsistence) and educational expenditure decreases along a ray from the origin.*

Proof. The marginal rate of substitution between consumption and educational expenditure is given by $\mathrm{MRS}_{\hat{c},b} \equiv \frac{W\hat{c}}{W_b} = \frac{u'(\hat{c})+\beta'(\hat{c})u(b)}{\beta(\hat{c})u'(b)}$. Now consider a ray from the origin (with $\hat{c}$ along the horizontal axis and b along the vertical axis) such that $\frac{b}{\hat{c}} = \alpha$ (some constant). Take any point on

[12] Note that, when the elasticity of $u(.)$ with respect to its argument is a constant, the elasticity of $u'(.)$ with respect to its argument is also a constant This follows from the relation that $u'(x) = \sigma_u \frac{u(x)}{x}$ for all x. Hence, differentiating both sides and simplifying, $\frac{-xu''(x)}{u'(x)} = 1-\sigma_u$, a constant.

[13] Notice that the two elasticities σ_u and σ_β will have to be less than unity in order to be consistent with the rest of the assumptions.

this line, say $\left(\frac{K}{\alpha}, K\right)$, where K is any real number. Note that, for a particular value of K, we can thus identify one particular point on the $b=\alpha\hat{c}$ locus, and as we increase the value of K we move on to higher and higher points lying on the same locus. The MRS at this point is given by:

$$\mathrm{MRS}|_{b=\alpha\hat{c}=K} = \frac{u'(K/\alpha)+\beta'(K/\alpha)u(K)}{\beta(K/\alpha)u'(K)} = \frac{u'(K/\alpha)}{\beta(K/\alpha)u'(K)} + \frac{\sigma_\beta}{\alpha\sigma_u}$$

Since σ_β and σ_u are constants, it follows that

$$\frac{\mathrm{dMRS}|_{b=\alpha\hat{c}=K}}{\mathrm{d}K} = \frac{[\beta(K/\alpha)\cdot u'(K)]u''(K/\alpha)\cdot\dfrac{1}{\alpha} - u'(K/\alpha)\left[\beta'(K/\alpha)\cdot u'(K)\cdot\dfrac{1}{\alpha} + \beta(K/\alpha)\cdot u''(K)\right]}{[\beta(K/\alpha)\cdot u'(K)]^2}$$

$$= \frac{-\left(\dfrac{1}{\alpha}\right)u'(K/\alpha)u'(K)\beta'(K/\alpha) + \beta(K/\alpha)\cdot u'(K/\alpha)\cdot u'(K)\cdot\dfrac{1}{K}\left[\dfrac{(K/\alpha)u''(K/\alpha)}{u'(K/\alpha)} - \dfrac{K\cdot u''(K)}{u'(K)}\right]}{[\beta(K/\alpha)\cdot u'(K)]^2}$$

Using the fact that the elasticity of $u'(.)$ with respect to its argument is a constant, it can be easily shown the last term in the numerator is zero, so that $\frac{\mathrm{dMRS}|_{b=\alpha\hat{c}=K}}{\mathrm{d}K} < 0$. Thus, along a ray from the origin as we move to higher indifference curves (i.e., as K rises), the MRS falls and the indifference curves become flatter. □

The rationale behind this result becomes clear once we examine the MRS expression more closely: $\frac{W_c}{W_b} = \frac{u'(\hat{c})+\beta'(\hat{c})u(b)}{\beta(\hat{c})u'(b)} = \frac{u'(\hat{c})}{\beta(\hat{c})u'(b)} + \frac{\beta'(\hat{c})u(b)}{\beta(\hat{c})u'(b)}$. When elasticities of u and β are constants, a proportional increase in $\hat{c}$ and b along a ray from the origin leaves the second term above unchanged. Hence, any change in the MRS comes entirely due to the change in the first term and here our assumption of endogenous altruism (whereby $\beta'(\hat{c})>0$) comes into play: as $\hat{c}$ and b increase proportionately, the numerator ($u'(\hat{c})$) falls, while the denominator ($\beta(\hat{c})u'(b)$) does not necessarily fall and even if it falls it only falls less than proportionately. Hence, the ratio must decline, implying a fall in the MRS. One could easily verify that constancy of the elasticity of the β function is not necessary for this result; the result will go through if, for example, σ_β is decreasing in its argument. The important point is incorporating endogenous altruism makes an otherwise homothetic utility function inherently *non-homothetic* and it is essentially this non-homotheticity property that drives the result here. Of course an *arbitrarily specified* non-homothetic utility function could also generate the same result; however, our assumption of endogenous altruism provides a *justification* as to why the utility function would indeed be non-homothetic and thus provides an economic foundation to any such arbitrary characterization of the utility function.

Given Lemma 1, it is now easy to prove that educational bequest function would be convex in income. Optimal educational expenditure of the representative agent can be obtained by maximizing (2) subject to the budget constraint: $c+b=y$. Subtracting the subsistence consumption, $\bar{c}$, from both sides, one can rewrite the budget constraint as: $\hat{c}+b=\hat{y}$, where $\hat{c}$ and $\hat{y}$ denote consumption and income over and above the subsistence level, respectively. (Recall that income of every household is assumed to cover at least the subsistence consumption; so $\hat{y}$ is non-negative.) Then, the necessary and sufficient conditions for interior optima are given by:

$$u'(\hat{c}) + u(b)\cdot\beta'(\hat{c}) = \beta(\hat{c})\cdot u'(b) \tag{3}$$

$$\hat{c} = \hat{y} - b \tag{4}$$

From (3) and (4), the expenditure on children's education, b, can be expressed as a function of $\hat{y}$. The standard assumption regarding $\hat{c}$ and b–that both are normal goods–ensures that

$0 < \frac{\mathrm{d}b(\hat{y})}{\mathrm{d}\hat{y}} < 1$. In tracing the curvature of $b(\hat{y})$, first note that, under Lemma 1, the marginal rate of substitution between $\hat{c}$ and b decreases along a ray from the origin. This implies that as income increases, in the $(\hat{c},b)$ plane, the successive equilibrium points (where the marginal rate of substitution is equal to the slope of the budget line) shift to the left (see Fig. 1 in this context, which traces the income–consumption curve for the household). Thus, as income increases the households invests *higher proportion* of their income on children's education, implying a convex $b(\hat{y})$ curve such that $\frac{\mathrm{d}^2 b(\hat{y})}{\mathrm{d}\hat{y}^2} > 0$.

It is important to emphasize at this point that our formulation of endogenous altruism, whereby parents enjoying higher consumption place greater weight on children's education, is *not* synonymous with the assumption that children's education is a luxury good. In other words, we do not begin by assuming that education expenditure is convex in income; we simply propose a mechanism for endogenizing altruism (which, to our mind, is intuitively plausible and is also supported by empirical evidence) and then show that *under reasonable assumptions* this mechanism could generate convex bequest/education functions.

The convexity of the educational bequest function has important implications for the intergenerational mobility, which we discuss in the next section. At this juncture, however, it seems appropriate to draw attention to the close link between our formalization of endogenous altruism with the concept of endogenous time preference. The rate of time preference typically reflects an agent's preference between the present and the future. Thus, in a two-period set up, if $W(c_1,c_2)$ represents the total utility of an agent (as a function of her current consumption (c_1) and her future consumption (c_2)), then the rate of time preference is defined as $\log \mathrm{MRS}_{c_1,c_2|_{c_1=c_2}}$.[14] In other words, the degree of time preference is measured by the slope of an indifference curve along a 45° line through the origin. If along the 45° line the slope of the indifference curve remains constant, that would imply a constant (exogenous) rate of time preference; if it falls, then that would imply decreasing time preference (or 'decreasing marginal impatience'); and if it rises, that would imply increasing time preference (or 'increasing marginal impatience'). It is easy to see that in our utility function defined in (2) if we replace $\hat{c}$ by c_1 and b by c_2, then by virtue of Assumptions 1 and 2 and Lemma 1, the indifference map would imply decreasing time preference whereby poorer people exhibit a higher rate of time preference, i.e., lower patience (see Fig. 2 which depicts this case in terms of the households' indifference map). Indeed, such negative relationship between households' income and time preference finds empirical support in a number of studies (see, for example, Lawrence, 1991; Ogaki and Atkenson, 1997; Samwick, 1998). In the context of a standard Diamond-type two-period overlapping generations framework, where the only source of second period consumption is out of first period savings, the abovementioned endogeneity of time preference would once again generate a convex savings function. In a world where the intergenerational link appears through parental savings, such convex savings functions could be instrumental in generating persistent inequality. It seems plausible to argue that endogenous altruism (as reflected not only in the children's schooling decision, but also in the care and help that the parents provide during childhood) plays a more important role during the formative years of a child's development which affects his future abilities, while endogenous time preference is likely to affect the progeny's overall life-time earnings/wealth through inheritance. In fact, the two mechanisms could work in tandem to accentuate the wealth inequality across households. In this paper, however, we focus on parental altruism working through children's education, as we believe that this mechanism has a more direct bearing on the question of intergenerational

[14] See Obstfeld (1990).

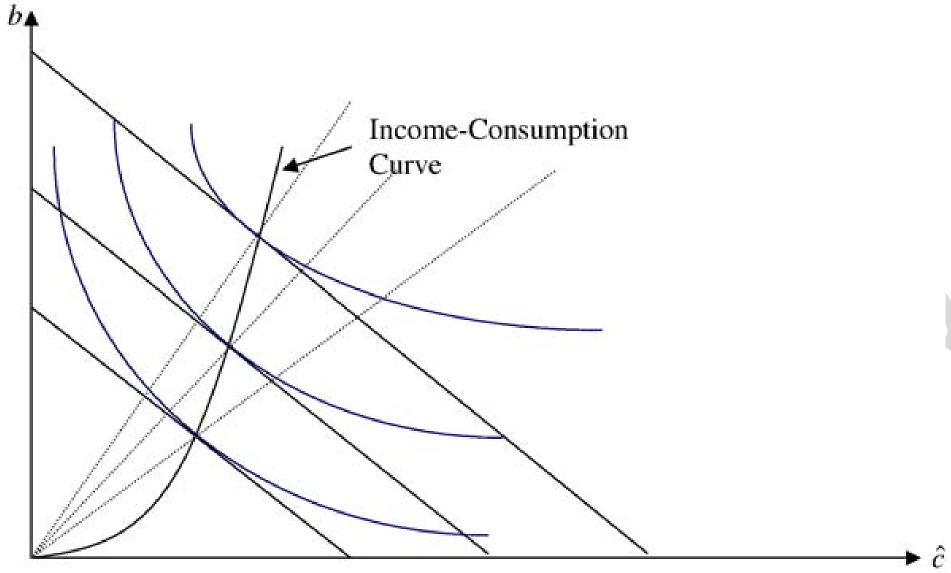

Fig. 1. Households' income–consumption curve.

mobility across different occupations, and therefore has important implications for the overall growth prospect of the macro-economy.

3. The macro-economy and intergenerational dynamics

Having described the households' decision making process at the micro-level, we now turn to the description of the macro-economy. The macro-economy is represented by an overlapping generations framework with constant population. Time is discrete with $t=0, 1, 2, \ldots, \infty$.

There exist a finite N number of households in the economy. Every individual member of a household lives for two periods and has one child born to her in the second period of her life. Thus, at any point of time t, every household in the economy has exactly two members: one child

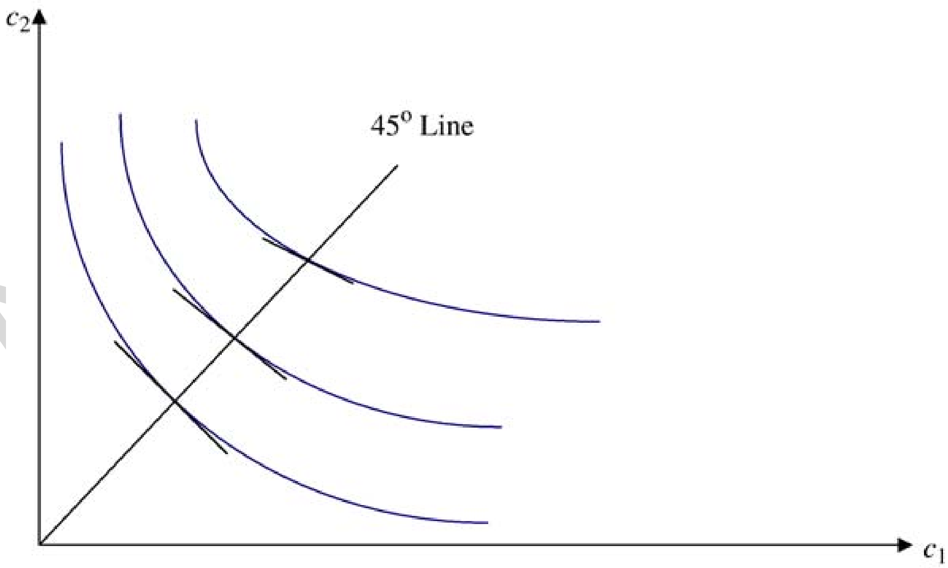

Fig. 2. Households' indifference map with decreasing marginal impatience.

and one adult. The size of newborns in each cohort is constant, given by N. Individuals are identical in terms of tastes and preferences—both across households and across generations.

The life cycle of an agent born at period t (generation t) is as follows. Each individual is born with some basic natural skill level, denoted by $h_{\min}$. Additional skills (or human capital) can be acquired during childhood through investment in education. In the first period of her life, as a child, the agent consumes nothing and devotes the entire time in acquiring skill, the acquired skill level being a function of the amount of investment made by her parent on her education. In the next period, as an adult, she works and earns a certain income y_t on the basis of the skills that she has acquired in the previous period. Out of this income she consumes a part and invests the rest in educating her child. She dies at the end of this period.

The preference structure of the agent, as already described in the previous section, is represented by Eq. (2). An agent belonging to generation t with income y_t maximizes (2) subject to her budget constraint.

3.1. Production technology

A single commodity is produced which may be either consumed or invested. Following the endogenous growth literature, production technology is assumed to be linear, exhibiting constant returns to human capital or skill level. Thus, aggregate output produced at time t is,

$$Y_t = AH_t,\ A > 0 \tag{5}$$

where H_t is the total amount of human capital available in the economy at time t and A is the (constant) return per unit of skilled labour. A worker with human capital h^i earns an income Ah^i. Thus, there exists a continuum of occupations in the economy, requiring different skill levels and generating different income levels.

Note that the minimum possible income level in this economy is given by $y_{\min} \equiv Ah_{\min}$. For convenience, let us assume that the minimum income level just enables the household to maintain the subsistence consumption, i.e.,

$$Ah_{\min} = \bar{c} \tag{6}$$

This assumption has the straightforward implication that, among all the different occupational choices that are available in the economy, there exists a subsistence sector that offers lowest income, requires minimum skill and merely fulfils the subsistence consumption requirement of the household.

Given (6), we define $\hat{y}_t(\equiv y_t - y_{\min})$ as the earnings above subsistence and write the budget constraint of the representative agent of generation t as,

$$\hat{c}_t + b_t = \hat{y}_t \tag{7}$$

3.2. Human capital formation

As we have stated before, each individual is born with some basic minimum level of human capital $h_{\min}$. Additional human capital formation is a function of the investment expenditure on her education incurred by her parent in the previous period. Thus,

$$h_{t+1} = h_{\min} + g(b_t) \tag{8}$$

where the human capital formation technology is characterized below.

Assumption 3. $g(b)$ is a real valued, continuous function defined on $(0,\infty)$ such that $g(0)=0$; for all $b \geq 0$, $g'(b)>0$, $g''(b) \leq 0$ and $\lim_{b\to\infty} g(b) = \bar{g}$ (a positive constant).

Greater investment in education produces higher human capital or skill. However, infinitely large investment in education does not result in infinitely high skill level; there exists a finite upper bound on human capital formation.

Note that, according to our specification, children's human capital formation does not depend directly on the parent's skill level. This is a simplification. The purpose here is to capture the non-linearity in human capital formation of poor vis-à-vis rich households that arises out of parental investment decisions. Incorporating parental skill level in the human capital formation technology directly would only accentuate any such non-linearity.

The stock of human capital of an individual fully determines her income and therefore her decisions as to how much to consume and how much to invest in her children's education. As before, we can define a new variable $\hat{h}$ which denotes the skill level above the basic minimum (i.e., $\hat{h}_t = h_t - h_{\min}$) and write the above function as,

$$\hat{h}_{t+1} = g(b_t) \tag{9}$$

For expositional purposes, it is convenient to work with a linear human capital formation technology. Therefore, for the rest of the paper, we shall assume a linear functional form of g, given by

$$\left.\begin{aligned} g(b) &= \gamma b (\gamma > 0) \text{ for } b \leq \bar{b} \\ &= \bar{g} \text{ for } b \geq \bar{b} \end{aligned}\right\} \tag{10}$$

where $\bar{b} = \frac{\bar{g}}{\gamma}$.

3.3. Distribution of human capital

Let $f_t(h)$ denote the cumulative distribution of human capital (skill) across agents belonging to generation t. Then total human capital available in the economy at time t is,

$$H_t = \int_{h_{\min}}^{\infty} h_t \mathrm{d} f_t(h_t). \tag{11}$$

Starting with an initial distribution f_0, the distribution of human capital evolves endogenously over time in accordance with the investment on children's education made by each family. Since at every point of time, the aggregate stock of human capital in the economy is determined by the distribution of human capital across households, the initial distribution of human capital also determines the growth path of the economy.

3.4. Intergenerational dynamics

Intergenerational mobility in this economy is determined by the investment in education by each family. Note that investment in education is zero when $h_t = h_{\min}$. For all individuals with skill level higher than $h_{\min}$ (and therefore with income level greater than $y_{\min}$), optimal investment in education can be obtained by maximizing (2) subject to (7). As we have already seen in Section 2, from the necessary and sufficient conditions for optima, one derives the optimal education expenditure function $b(\hat{y}_t)$ which in increasing and convex in $\hat{y}_t$. Noting that

$\hat{y}_t = A\hbar_t$, we can then characterize the intergenerational dynamics in terms of $\hat{h}$ (or h) in the following way.

From (10), $g(b)=\bar{g}$ when $b \geq \frac{\bar{g}}{\gamma}$. We can define the upper bound on g in terms of $\hat{h}$ as: $g=\bar{g}$ when $\hat{h} \geq \frac{1}{A} b^{-1}\left(\frac{\bar{g}}{\gamma}\right) \equiv \tilde{h}$. Thus, from (9) and (10), intergenerational mobility is described by the following dynamic equation:

$$\left.\begin{array}{ll} \hat{h}_{t+1} & = 0 \text{ for } \hat{h}_t = 0 \\ & = \gamma b(A\hat{h}_t) \text{ for } 0 < \hat{h}_t \leq \hat{h} \\ & = \bar{g} \text{ for } \hat{h}_t \geq \hat{h} \end{array}\right\}. \tag{12}$$

From (12), it is easy to see that $\hbar = 0$ (i.e., $h = h_{\min}$) represents a steady state. For skill level above $h_{\min}$, the dynamics of $\hbar$ (and therefore h) depends crucially on the convexity of $b(A\hbar_t)$.

Lemma 2 below establishes a set of sufficient conditions for the existence of multiple equilibria in this case.

Lemma 2. *Let* $\lim_{\hat{h}\to 0} \frac{db(A\hat{h}_t)}{d\hat{h}_t} < 1$ *and* $b\left(A\hat{h}_t\right) > \frac{\hat{h}_t}{\gamma}$ *for some* $\hbar_t \in (0, \tilde{h})$. *Then there exists a steady state value of* $\hbar$ *lying between 0 and* $\tilde{h}$, *say* $\hbar^*$, *which is locally unstable. Moreover, this unstable steady state will be unique.*

Proof. Since $b(A\hbar_t)=0$ for $\hbar_t = 0$ and $b(A\hbar_t)$ is continuous in $\hbar_t$, the existence of $\hbar^*$ can be guaranteed by the intermediate value theorem. Uniqueness and instability of $\hbar^*$ follows from the convexity of $b(A\hbar_t)$. □

Given Lemma 2, the complete dynamic system represented by (12) has three steady states given by 0, $\hbar^*$ and $\bar{g}$, respectively. Fig. 3 depicts this scenario.[15] The dark line in the diagram traces the path of $\hbar_{t+1}$, which intersects the 45° line at three points, indicating three steady states.

The dynamics of h determines how the income of successive generations changes over time across households. The following proposition specifies the condition for intergeneration mobility in terms of the initial skill level of a household:

Proposition 1. *Under Assumptions 1, 2 and 3, and Lemma 2, the dynamic system for the variable* $\hbar$ *(represented by (12)) has three non-trivial steady states, given respectively by 0,* $\hbar^*$ *and* $\bar{g}$*, where the first and the third ones are locally stable and the middle one is locally unstable. Likewise, the dynamic system for the variable h will have three non-trivial steady states given by* h_{min}, $\hbar^* + h_{min}$ *and* $\bar{g} + h_{min}$, *of which the first and the third are stable and the second one unstable. Accordingly, the evolution of a household with initial skill level* h_0^i *will be governed by the following condition:*

$$\lim_{t\to\infty} h_t^i = h_{\min} \text{ if } h_0^i \in [h_{\min}, \hat{h}^* + h_{\min})$$

$$= \bar{g} + h_{\min} \text{ if } h_0^i \in (\hat{h}^* + h_{\min}, \infty)$$

Proof. Follows from Lemma 2. □

From the above proposition, it is evident that intergenerational mobility depends crucially on the initial level of human capital of the adult member of the households. All those households which start with an initial skill level lying below the critical value $\hbar^* + h_{\min}$ will invest less and

[15] Fig. 3 has been drawn in terms of $\hat{h}$, plotting $\hat{h}_t$ and $\hat{h}_{t+1}$ along the horizontal and the vertical axes, respectively, and (0,0) representing the subsistence point. Alternatively, one can read this diagram in terms of h, interpreting the origin as ($h_{\min}$, $h_{\min}$) and measuring h_t and h_{t+1} along the horizontal and the vertical axes, respectively.

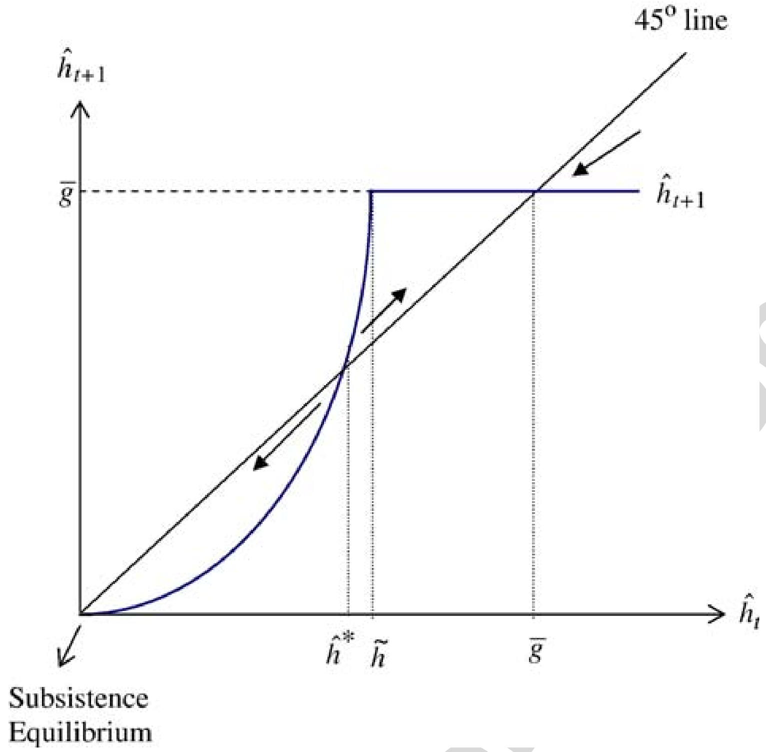

Fig. 3. Intergenerational dynamics.

less in children's education, and eventually reach $h_{\min}$, surviving only at the subsistence level. On the other hand, households starting above $\hbar^*+h_{\min}$ will invest more on children's education, and will eventually reach the higher steady state, earning and consuming at the maximum possible level. Thus, $\hbar^*+h_{\min}$ constitutes the threshold level of human capital (and correspondingly $A(\hbar^*+h_{\min})$ constitutes the threshold income level), which determines whether a household in the long run approaches the higher stable steady state or the lower one.

In order to convince the reader that the dynamics described in Proposition 1 is indeed a possibility under reasonable assumptions about the utility function, an example might be appropriate here.

Example. Let $u(\hat{c}) = \beta(\hat{c}) = \sqrt{\hat{c}}$ and $u(b) = \sqrt{b}$. Then, the utility function of the representative member of generation t can be written as

$$W(\hat{c}, b) = \sqrt{\hat{c}}\left(1 + \sqrt{b}\right) \tag{13}$$

These specific forms of the u(.) and the β(.) satisfy Assumptions 1, 2 and 3. Maximizing (13) subject to the budget constraint of the household, we get the following set of first order conditions:

$$\sqrt{b}(1 + \sqrt{b}) = \hat{c} \tag{14}$$

$$\hat{c} = \hat{y} - b \tag{15}$$

Simplifying, we get a quadratic equation of the form:

$$2(\sqrt{b})^2 + \sqrt{b} - \hat{y} = 0 \tag{16}$$

Ruling out the negative solution for $\sqrt{b}$, we get the optimal expenditure on children's education as:

$$b = \left(\frac{-1 + \sqrt{1 + 8\hat{y}}}{4}\right)^2 \tag{17}$$

It is easy to see that $\frac{\mathrm{d}b_t}{\mathrm{d}\hat{h}_t} = \frac{A}{2}\left(1 - \frac{1}{\sqrt{1 + 8A\hat{h}_t}}\right) > 0$ and $\frac{\mathrm{d}b_t^2}{\mathrm{d}\hat{h}_t^2} = \frac{A}{4}\left(\frac{8}{(1+8A\hat{h}_t)^{3/2}}\right) > 0$. Thus, $b(A\hbar)$ is a convex function of h with $\lim_{\hat{h}\to 0} \frac{\mathrm{d}b_t}{\mathrm{d}\hat{h}_t} = 0$. Also as $\hat{h}$ approaches infinity, so does $b(A\hat{h})$, with the limiting value of the slope given by: $\lim_{\hat{h}\to\infty} \frac{\mathrm{d}b_t}{\mathrm{d}\hat{h}_t} = \frac{A}{2}$. Therefore, for a wide range of parameter values, the conditions specified in Lemma 1 will indeed be satisfied and the depending on its initial value, the skill level of the agents will converge to either $h^{\min}$ or $\bar{g} + h^{\min}$ in the long run.[16]

Remark 1. Note that the threshold effect elaborated in this section arises purely due to the preference structure of the households which relates the degree of altruism to the parental income level. While the credit market for human capital (whereby parents can borrow against their children's future earnings) is missing here, we have not assumed any non-convexity either in the technology or in the preferences. In our model, inequality perpetuates in the long run because the poor parents, being more concerned about their current consumption, choose to invest a lower proportion of their income in children's human capital formation. It is this convexity of the savings behaviour that generates threshold effects and persistence of inequality in our model. Of course our model predicts strict polarization, while in reality one observes some tendency towards regression to the mean over very long run.[17] Such tendencies could be explained in our model by incorporating a random ability factor that works independent of the income factor. Thus, while the income factor, working through endogenous altruism, may generate tendencies towards persistence, a randomly distributed ability factor may offset any such tendency over the very long run.

Remark 2. As we have mentioned before, the altruism coefficient β can alternatively be interpreted as the discount factor, which is negatively related the households' rate of time preference. Thus, one can interpret our model as a model of variable time preference. This second interpretation would be more relevant, were we to model household preferences over infinite horizon, allowing savings to adjust optimally over time. In the context of persistence inequality, it has been argued that 'warm glow' kind of altruism in an overlapping generations framework precludes strategic or optimal savings behaviour, thereby limiting the scope of the poorer households to escape poverty by continuous up-gradation of skills over generations. (See Mookherjee and Ray, 2001, 2002 for an elaboration of this argument.) In fact, in an important contribution, Loury (1981) has shown that, even if capital markets are missing altogether, strategic savings behaviour on the part of the households over infinite horizon and convexity of technology ensure that households' incomes converge in the long run. Loury however assumes that households' rate of time preference is constant. Loury's convergence result need not hold when households' rate of time preference is endogenously determined. If poorer people have a higher rate of time preference (which is analogous to the our assumption of β' being positive here), then facing a constant return on investment, it is likely that poorer people would invest less

[16] The value of $\hat{h}^*$ in this example turns out to be $\frac{\gamma}{(A\gamma-2)^2}$. Therefore, for any parameter values such that $\frac{\gamma}{(A\gamma-2)^2} < \frac{(1+4\sqrt{\bar{g}/\gamma})^2}{8A} \equiv \tilde{h}$, the economy will be characterized by three steady states.

[17] Recent empirical studies (e.g., Zimmerman, 1992; Mulligan, 1997; Solon, 1992, 2002) estimate the persistence coefficient to lie between 0.4 and 0.6, which though high compared to the earlier estimates is nonetheless less than unity.

and would therefore be perpetually stuck in a poverty trap—even in a infinite horizon model which allows for strategic savings behaviour. Such possibilities have been discussed in Das (2000, 2003) in the context of an infinite horizon optimal growth model with physical capital formation.

Remark 3. Finally, note that the intergenerational dynamics specified in Proposition 1 has important implications for long run growth of the economy. According to Proposition 1, the income of the adult member converges either to $h_{\min}$ or to $\overline{g}+h_{\min}$ depending on whether her ancestors started with an initial skill level lying below or above $\hat{h}^*+h_{\min}$. Therefore, as in Galor and Zeira, in the long run, there will be complete polarization of the households with the number of households surviving at the subsistence level given by

$$N_{\text{sub}} = \int_{h_{\min}}^{\hat{h}^*+h_{\min}} \mathrm{d}f_0(h_0) \tag{18}$$

It then follows that the long run average income of the economy will depend on the initial distribution of human capital (or equivalently on the initial distribution of income). The higher is the proportion of people with initial skill levels below $h^*+h_{\min}$, the lower is the long run average income. Hence, the Galor-Zeira conclusion that inequality and growth are negatively correlated is reinstated here—though the dynamics works here through a different channel. It is however important to note here that a convex savings function may also imply a positive relationship between inequality and growth because it generates higher aggregate savings which boosts growth (as in Bourguignon, 1981). In fact, as has been argued by Galor and Moav (2004), this positive effect through higher aggregate savings would dominate if the marginal return to capital formation is sufficiently non-diminishing. In our model, however, the marginal return to human capital formation falls sharply to zero beyond a certain point, which discourages human capital accumulation by the rich beyond a level. As a result, the growth enhancing impact of inequality gets thwarted.

4. Public education system

In our analyses so far we have assumed that each parent finances education of her child privately. In this section, we discuss the consequence of introducing a public education system which is financed by a proportional income tax. The choice of the tax rate is endogenous: it is decided by majority voting. The issue that we consider here is under what circumstances the public education system will perform better than the private education system in terms of long run growth?

Let us assume that government imposes a proportional income tax at the rate τ on any income above the subsistence level and invests the entire tax proceeds in a public education programme. The households do not have the option of pursuing an independent private education program.

In every period, the adults vote to decide the tax rate and the tax rate chosen by the majority is accepted by the government.[18] In this case, an individual's decision-making is a two-step procedure. She first chooses the level of consumption that maximizes her utility for a given post tax income, treating the expenditure on children's education as given. Since marginal utility from consumption is positive, this would imply that she consumes her entire post tax income. Thus, we

[18] This of course presupposes that such a political equilibrium exits. However, given the assumptions about the utility function in our model, existence of a political equilibrium can be easily proved.

can derive the indirect utility function of the household as a function of the tax rate. In the next step, she chooses the tax rate τ so as to maximize this indirect utility function.

It follows trivially that, when everybody has identical skill level and identical income (which is also the average income of the economy), the tax rate chosen under majority voting will coincide with the proportion of income spent on children's education under the private education regime. Hence, the investment in human capital formation per child will be exactly the same under the private and the public education system, and the long run growth path will also be identical. The public education system will generate a different growth path for the economy than the private education system if and only if the initial distribution is non-degenerate.

If the public education system is introduced at time 0, then the initial expenditure on education is given by

$$E_0 = \tau_0 \left[A \int_{h_{\min}}^{\infty} (h_0 - h_{\min}) \mathrm{d}f_0(h_0) \right] \tag{19}$$

Hence, education expenditure per child under public education system is given by,

$$\begin{aligned} e_0 &= \frac{E_0}{N} \\ &= \frac{\tau_0 \left[A \int_{h_{\min}}^{\infty} h_0 \mathrm{d}f_0(h_0) - A h_{\min} \int_{h_{\min}}^{\infty} \mathrm{d}f_0(h_0) \right]}{N} \\ &= \frac{\tau_0 [A h_0 - A h_{\min} N]}{N} \\ &= \tau_0 (\bar{y}_0 - y_{\min}) \end{aligned} \tag{20}$$

where $\bar{y}_0$ is the average income at time 0.

Let the variable h^{PUB} denote human capital formation under the public investment regime. Then,

$$h_1^{\mathrm{PUB}} = \gamma e_0 \tag{21}$$

Thus, under public education system, the income of every household in the next period will be $y_1^{\mathrm{PUB}} = A h_1^{\mathrm{PUB}}$.

One important implication of the public education system is that it removes the difference in the skill level across households from the next period onwards. Thus, the first round impact of the public education system would determine the subsequent pattern of development for the entire economy.

It is easy to see that, if $h_1^{\mathrm{PUB}} < \hat{h}^* + h_{\min}$, then the economy would not be better off in the long run under public education system. This is so because in the next period all the households will earn an income $y_1^{\mathrm{PUB}} < A(\hat{h}^* + h_{\min})$ and will choose a tax rate such that $e_1 = b(y_1^{\mathrm{PUB}}) < b(A(\hat{h}^* + h_{\min}))$. From our analysis in Section 3, we know that $\hat{h}^* + h_{\min}$ is an unstable equilibrium. Thus, in every subsequent period, the skill level and therefore the income level of the households will fall until it reaches $h_{\min}$ (the subsistence skill level).

The public education system will unambiguously improve the position of the economy in the long run if the initial tax rate $\tau_0 > \frac{\hat{h}^* + h_{\min}}{\gamma(\bar{y}_0 - y_{\min})}$. If the chosen tax rate satisfies this condition, then under the public education system every household of the economy in the long run will reach the highest possible income level $A(h_{\min} + \bar{g})$. Under private education system on the other hand only those households with initial skill level above $(\hat{h}^* + h_{\min})$ will attain the maximum possible income level. Thus, for any economy with an initial distribution of skill such that there is at least one household with initial skill level below $(\hat{h}^*+h_{\min})$, the economy will attain a higher average income under public education system than under private education system in the long run if the chosen initial tax rate is greater than $\frac{\hat{h}*+h_{\min}}{\gamma(\bar{y}_0-y_{\min})}$. The following proposition summarizes this result.

Proposition 2. *An economy will be better off in the long run under public education system in the sense that it will attain a higher level of per capita income compared to its initial position if the chosen initial tax rate is such that*

$$\tau_0 > \frac{\hat{h}^* + h_{\min}}{\gamma(\bar{y}_0 - y_{\min})}.$$

Moreover, under this condition, the public education system will perform better than the private education system in terms of long run growth provided there is at least one household in the economy with initial skill level below $(\hat{h}^*+h_{min})$.

It is important to note here that the chosen tax rate itself is likely to be a function of the average income in the economy. Therefore, Proposition 2 in effect imposes a condition on the initial average income, and thus on the initial distribution. To see this more clearly, let us consider an example.

The specific utility function that we consider here is the same one that we discussed as an example at the end of Section 3. Since we already know that this utility function generates multiple equilibria under a wide range of parameter values, it is convenient to illustrate our point with the help of this specific example. Thus, let (13) represent the utility function of the representative household with income $\hat{y}$. Under public education system, the indirect utility function of the household can then be written as,

$$\tilde{W} = \sqrt{(1-\tau_0)(y_0-y_{\min})}\left(1+\sqrt{\tau_0(\bar{y}_0-y_{\min})}\right) \tag{22}$$

The household maximizes (22) with respect to τ_0 to choose its optimal tax rate. Solving the maximization problem, we find that the optimal tax rate is give by,

$$\tau_0^* = \left\{\frac{-1+\sqrt{1+8(\bar{y}_0-y_{\min})}}{4\sqrt{(\bar{y}_0-y_{\min})}}\right\}^2 \tag{23}$$

The optimal tax rate derived in (23) is independent of the household's level of income and therefore all households will choose the same tax rate. Thus, by Proposition 2, the economy will be better off in the long run under public education system if

$$\left\{\frac{-1+\sqrt{1+8(\bar{y}_0-y_{\min})}}{4\sqrt{\bar{y}_0-y_{\min}}}\right\}^2 > \frac{\hat{h}^*+h_{\min}}{\gamma(\bar{y}_0-y_{\min})}$$

Simplifying, we can write this condition as,

$$\bar{h}_0 > \frac{\left(\sqrt{\frac{16(\hat{h}^*+h_{\min})}{\gamma}}+1\right)^2 - 1}{8A} + h_{\min} \tag{24}$$

where $\bar{h}_0$ denote the average initial skill level in the economy. Whenever the initial distribution is such that the average skill level in the economy is greater than the term specified in the RHS of

(24) and there is at least one household with initial skill level below $(\hat{h}^{*}+h_{\min})$, public education system will definitely be growth enhancing compared to the private education system.

Finally, in the context of this example, consider two economies which are identical in terms of technology and preferences, but differ in terms of initial distribution of skill (and income). Let $f_0^{\mathrm{I}}(h_0)$ and $f_0^{\mathrm{II}}(h_0)$ represent the initial distribution of skills in the two economies such that $\bar{h}_0^{\mathrm{I}}=\bar{h}_0^{\mathrm{II}}$, but $\sigma_0^{\mathrm{I}}(h_0)>\sigma_0^{\mathrm{II}}(h_0)$, where $\bar{h}$ denotes the average skill level and σ denotes the variance. Thus, the two economies have the same average income at the initial point, but the income is more unequally distributed in the former. Assuming that there is at least one household even in country II (i.e., the one with better distribution of income) with initial skill level below $\hat{h}^{*}+h_{\min}$, public education system will definitely perform better in economy I (where distribution of income is worse, and therefore there are more households with initial skill level below $\hat{h}^{*}+h_{\min}$), in the sense that under public education system average income of the economy in the long run will improve more in the first economy than the second one. Thus, when conditions under Proposition 2 are satisfied, the more unequal is the society, the better is the performance of the public education system from the point of view of long run growth.

Remark 4. In this section, we have only endogenized the choice of the tax rate under the public education system. The type of education system itself could be also an endogenous choice variable decided by majority voting. Needless to say, even when the public education system performs better than the private education system in terms of long run growth, it need not be the preferred system unless it improves the welfare of the majority of the voting population. However, it is easy to show that, for households with income level below the mean income, the cost of educating children under the public education regime would be lower, and therefore for these households the public education system will be necessarily welfare improving.[19] The opposite holds for households with income level above the mean income. It then follows that, if the initial the distribution of income is positively skewed such the median < mean, public education will always be welfare-improving for the majority of the voting population.

5. Conclusion

In this paper, we seek to explain the persistence of income inequality in terms of a model based on limited parental altruism. With the missing market for human capital, endogenous altruism implies that parental income determines not only the parent's ability but also her willingness to invest in children's human capital. This gives rise to a non-linearity in the investment expenditure on human capital formation, and with a constant returns to scale (convex) technology, initial low earning abilities of the parent gets translated into low earning abilities of the subsequent generations as well. A direct consequence of this is a long run polarization of skill levels and income levels. Initial distribution therefore becomes an important determinant of the long run development pattern. Hence, a one shot re-distributive policy of the government that aims at

[19] To see this note that for a household with initial income y_0 the utility maximization problem under private education system is given by max $W(\hat{c}_0,b_0)=u(\hat{c}_0)+\beta(\hat{c}_0)u(b_0)$ subject to $\hat{c}_0+b_0=\hat{y}_0$. On the other hand, noting that $\tau_0=\frac{e_0}{\bar{y}_0-y_{\min}}$, under the public education system the maximization exercise of the household can be written as max $W(\hat{c}^0,e^0)=u(\hat{c}^0)+\beta(\hat{c}^0)u(e^0)$ subject to $\hat{c}_0+\left(\frac{\hat{y}_0}{\bar{y}-y_{\min}}\right)b_0=\hat{y}_0$. The only difference between these two problems are in the budget equations: while under the private education system the relative price associated with the expenditure on children's education is unity, under the public education system, the relative price is given by $\frac{\hat{y}_0}{\bar{y}_0-y_{\min}}$. Clearly, for households with below mean income level, the public education system is cheaper and generates higher utility.

reducing inequality by shifting people from the two tail ends towards the middle could enhance the growth performance of an economy.

It has often been argued that in many developing countries direct re-distributive policies are politically difficult to implement. In so far as the inequality affects the human capital formation and thus the growth prospect of a country, it is sometimes recommended that a public education programme could attain the same objective at a lower political cost. Given the fact that poorer households are less willing to invest in children's education, a public education system that reduces the cost of education for the poor households at the expense of the richer households may improve the long run growth prospect of the economy. We have shown here that a public education system does not unambiguously improve the growth prospect of an economy. However, the effectiveness of the public education system depends on the degree of inequality: the more unequal the society is, the greater is the scope of public education system in improving the long run growth scenario.

Acknowledgement

I thank Aditya Bhattacharya, Shankha Chakraborty, Gerhard Glomm, Subrata Guha, Nobuhiro Kiyotaki, Tridip Ray and Partha Sen for their valuable suggestions on an earlier version of this paper. I also gratefully acknowledge the insightful comments of two anonymous referees of this journal, which has greatly improved the exposition in the current version. The paper benefited from comments received during presentations at the Delhi School of Economics and at the Econometric Society Meeting, India and South East Asia Chapter, Delhi, December 2001. The usual disclaimer applies.

References

Aghion, P., Bolton, P., 1997. A theory of trickle-down growth and development. Review of Economic Studies 64, 151–172.

Banerjee, A.V., Newman, A., 1993. Occupational choice and the process of development. Journal of Political Economy 101, 274–298.

Barro, R.J., Sala-i-Martin, X., 1995. Economic Growth. McGraw Hill, New York.

Beal, D.J., 2001. Financial support for adult children in Australia. Economic Analysis and Policy 31, 13–24.

Becker, G.S., Mulligan, C.B., 1997. The endogenous determination of time preference. Quarterly Journal of Economics 112, 729–758.

Benabou, R., 1994. Human capital, inequality, and growth: a local perspective. European Economic Review 38, 817–826.

Benabou, R., 1996. Heterogeneity, stratification and growth: macroeconomic implications of community structure and school finance. American Economic Review 86, 584–609.

Blanchard, O.J., Fischer, S., 1989. Lectures in Macroeconomics. MIT Press, Cambridge, MA.

Bourguignon, F., 1981. Pareto superiority of unegalitarian equilibria in Stiglitz's model of wealth distribution with convex savings function. Econometrica 49, 1469–1475.

Cardak, B.A., 1999. Heterogeneous preferences, education expenditures and income distribution. The Economic Record 75, 63–76.

Carroll, C.D., 2000. Why do the rich save so much. In: Slemrod, J.B. (Ed.), Does Atlas Shrug? The Economic Consequences of Taxing the Rich. Harvard University Press.

Cogner, R.D., Cogner, K.J., Elder, G.H., Lorenz, F.O., Simons, R.L., Whitbeck, L.B., 1992. A family process model of economic hardship and adjustment of early adolescent boys. Child Development 63, 526–541.

Das, M., 2000. Optimal growth with variable rate of time preference. Working Paper. Department of Economicsm, Delhi School of Economics, Delhi University, India.

Das, M., 2003. Optimal growth with decreasing marginal impatience. Journal of Economic Dynamics and Control 27, 1881–1898.

Duraisamy, M., 2002. Child schooling and child work in India. Working Paper No. 84. National Council of Applied Economic Research, New Delhi. April.
Durlauf, S.N., 1996. A theory of persistent inequality. Journal of Economic Growth 1, 75–95.
Dynan, K., Skinner, J., Zeldes, S.P., 2004. Do the rich save more? Journal of Political Economy 112, 397–444.
Elder, G., Nguyen, T.V., Caspi, A., 1985. Linking family hardship to children's lives. Child Development 56, 361–375.
Fisher, I., 1930. The Theory of Interest. Macmillan, New York.
Freeman, S., 1996. Equilibrium income inequality among identical agents. Journal of Political Economy 104, 1047–1064.
Koopmans, T.C., 1986. Representation of preference ordering over time. In: Mcguire, C.B., Radner, R. (Eds.), Decision and Organisation, 2nd ed. University of Minnesota Press, Minneapolis.
Galor, O., Moav, O., 2004. From physical to human capital accumulation: inequality and the process of development. Review of Economic Studies 71, 1001–1026.
Galor, O., Zeira, J., 1993. Income distribution and macroeconomics. Review of Economic Studies 60, 35–52.
Ghatak, M., Jiang, N.N., 2002. A simple model of inequality, occupational choice, and development. Journal of Development Economics 69, 205–226.
Glomm, G., Ravikumar, B., 1992. Public versus private investment in human capital: endogenous growth and income inequality. Journal of Political Economy 100, 818–834.
Hisnanick, J.J., Coddington, D.A., 2000. The immunisation status of poor children: an analysis of parental altruism and child well-being. Review of Social Economy 58, 81–107.
Huston, A.C., McLyod, V.C., Coll, C.G., 1994. Children and poverty: issues in contemporary research. Child Development 65, 275–282.
Jensen, R.T., Richter, K., 2001. Understanding the relationship between poverty and children's health. European Economic Review 45, 1031–1039.
Lawrence, E.C., 1991. Poverty and the rate of time preference: evidence from panel data. Journal of Political Economy 99, 54–75.
Loury, G.C., 1981. Intergenerational transfers and the distribution of earnings. Econometrica 49, 843–867.
Maoz, Y.D., Moav, O., 1999. Intergenerational mobility and the process of development. Economic Journal 109, 677–697.
Moav, O., 2002. Income distribution and macroeconomics: the persistence of inequality in a convex technology framework. Economics Letters 75, 187–192.
McLoyd, V., 1990. The impact of economic hardship on black families and children: psychological distress, parenting, and socioemotional development. Child Development 61, 311–346.
Meara, E.., 2001. Why is health related to socioeconomic status? NBER Working Paper 8231.
Menchik, P., David, M., 1983. Income distribution, lifetime savings and bequests. American Economic Review 73, 672–690.
Mookherjee, D., Ray, D., 2001. Introduction. In: Mookherjee, D., Ray, D. (Eds.), Readings in the Theory of Economic Development. Blackwell, Massachusetts, USA.
Mookherjee, D., Ray, D., 2002. Contractual structure and wealth accumulation. American Economic Review 92, 818–849.
Mookherjee, D., Ray, D., 2003. Persistent inequality. Review of Economic Studies 70, 369–393.
Mulligan, C., 1997. Parental Priorities and Economic Inequality. University of Chicago Press, Chicago.
Obstfeld, M., 1990. Intertemporal dependence, impatience and dynamics. Journal of Monetary Economics 26, 45–75.
Ogaki, M., Atkenson, A., 1997. Rate of time preference, intertemporal elasticity of substitution, and level of wealth. Review of Economics and Statistics 79, 564–572.
Owen, A.L., Weil, D.N., 1998. Intergenerational earnings mobility, inequality, and growth. Journal of Monetary Economics 41, 71–104.
Parker, S., Greer, S., Zuckerman, B., 1988. Double jeopardy: the impact of poverty on early Child Development. Pediatric Clinics of North America 35, 1227–1240.
Piketty, T., 1997. The dynamics of the wealth distribution and the interest rate with credit rationing. Review of Economic Studies 64, 173–189.
Samwick, A.A., 1998. Discount rate heterogeneity and social security reform. Journal of Development Economics 57, 117–146.
Solon, G., 1992. Intergenerational income mobility in United States. American Economic Review 82, 393–408.
Solon, G., 2002. Cross-country differences in intergenerational earnings mobility. Journal of Economic Perspective 949–956 Summer.
Tilak, J.B.G., 2000. Household Expenditure on Education in India: a Preliminary Examination of the 52nd Round of the National Sample Survey. National Institute of Educational Planning and Administration, New Delhi. July.
Zimmerman, D.J., 1992. Regression toward mediocrity in economic stature. American Economic Review 82, 409–429.