

# The Allocation of Public Funds in a Hierarchical Educational System\*

Xuejuan Su

Department of Economics, Indiana University  
Wylie Hall 105, Bloomington, IN 47405  
xuesu@indiana.edu

October, 2003

## Abstract

This paper studies the dynamic effects of allocating public funds between basic and advanced education. Holding the size of total public education funds fixed, I identify the effects of their composition on aggregate efficiency and (in)equality. First, since basic education is a prerequisite for attending advanced education, there exists a lower bound on funding it: allocation policies below this lower bound are strictly Pareto dominated. Whether a corresponding lower bound on funding advanced education exists or not depends on the development stage of the economy and the size of total public funds. Secondly, while an allocation policy favoring basic education generates the usual redistribution from top to bottom, one favoring advanced education may result in reverse redistribution from bottom to top. Last, through the inter-generational link, short run allocation policies may have long run effects. A simple rule-of-thumb is that for an economy in its early development stage, focusing on basic education for sufficiently long duration is the only way out of polarization and low aggregate efficiency, contrary to the actual policies pursued by many of the less developed economies.

---

\*I am indebted to Michael Kaganovich and Gerhard Glomm for many helpful discussions. I thank Peter Rangazas, participants at the BLISS Conference and Midwest Macroeconomics Conference and an anonymous referee for their comments. All remaining errors are my own.

## 1. Introduction

Any educational system is intrinsically hierarchical. Individuals have to learn basic knowledge before they can study more advanced knowledge. Consequently, within the commonly observed three-stage system - primary, secondary and tertiary education, individuals have to finish lower levels successfully before they can enter higher levels, and their academic achievements in lower levels determine their preparations and qualifications for the higher levels.

No wonder, then, that when compulsory schooling legislation is present,<sup>1</sup> it exclusively regulates school participation at lower levels, and leaves that at higher levels to private choice. Consequently, school participation is relatively uniform in lower levels, yet it exhibits pronounced difference in higher levels across individuals. Furthermore, due to the hierarchical nature of education, drop-out is much more likely than drop-in at any intermediate stage, so the participation rate shrinks when a cohort of individuals moves up along the education ladder. Table 1 provides some cross-country data revealing this pattern. While the participation rate shrinks only moderately in developed economies like the U.S., the U. K. and Japan, it shrinks drastically in some less developed economies like Chad, Lesotho and Niger.

At any educational stage, the quality of education depends on the resource availability, and it both affects and is affected by individuals' participation decisions. Our focus is on the public education funds, the major component of total education expenditure for almost all countries. Adjusting the aggregate funding level by participation rate, expenditure per pupil is a strong indicator of schooling quality at a given stage. Data in Table 2 reveal that countries differ in their allocation policies concerning public education funds, and the relative schooling quality across educational stages exhibits prominent variations. While the relative schooling quality between higher stage and lower stage remains close to 1 in developed economies, it rises beyond 10 in some less developed economies. This raises the main issue of the current paper.

Given the hierarchical nature of the educational system, different stages of education are not perfectly substitutable, so a change in the schooling quality at one stage may not be readily offset by an opposite change at another stage. Consequently, the resource allocation policy may have profound impact on the final outcomes. This paper conducts positive analysis of the dynamic effects of the allocation of public funds across educational stages. Given a fixed size of public

---

<sup>1</sup>UNESCO (1999) report shows that over 95% of all countries have such legislations.

education funds, how does their allocation across different stages affect individuals' participation decisions and hence human capital and income? What are the policy effects on (in)equality? In particular, does the difference in participation decisions at lower and higher stages lead to reverse redistribution of resources from low-income to high-income families? And to what extent can this reverse redistribution be offset through income tax? What are the policy effects on aggregate efficiency? Is there a Pareto dominant allocation policy, or does a policy change necessarily benefit some individuals at the expense of others? Finally what are the long run effects of an allocation policy?

These issues cannot be analyzed without looking into the internal structure of the educational system. Instead of viewing the entire education process as a black-box, I introduce the hierarchical structure. Namely, in this model, the educational system consists of two stages. The human capital output produced at the lower stage acts as an input into the education process at the higher stage. This approach enables us to identify the composition effect free from the size effect of public education funds.

Economists have studied the size effect of public education funds from various perspectives; but very few have studied the composition effect based on the hierarchical structure of the educational system. One exception is Driskill and Horowitz (2002), who explicitly model the education hierarchy. They consider a social planner's problem in a standard growth framework and study the optimal investment in hierarchically produced human capital as well as in physical capital. They show that the optimal investment program, part of which is the investment plan at different stages of education, depends on the relative scarcity of physical capital and the differential levels of human capital, and hence on the development stage of an economy. However, that paper sheds no light on the policy effects in a decentralized decision-making setup, and it cannot be used to study the distributional effects across individuals. As will be outlined below, individuals differ in their participation decisions, and the distributional effects constitute an important aspect of the allocation policy, hence should not be ignored. The current paper fills these gaps.

Within a dynamic framework of successive generations, this paper models education as a hierarchical two-stage system with two distinct technologies. The lower and higher stages are also referred to as the basic and advanced stages. The technology at the basic level is standard, where human capital is produced with conventional inputs such as an individual's pre-school preparation (also called initial qualification), schooling time and the relevant schooling quality. The tech-

nology at the advanced level uses the human capital output produced from basic education as an input. This factor determines individuals' preparation for advanced education (also called augmented qualification). Poorly prepared individuals (with augmented qualifications below some threshold level) simply cannot benefit from advanced education even if they do attend.<sup>2</sup> Only sufficiently prepared individuals (with augmented qualifications above the threshold level) can benefit from advanced education, and they make their own participation decisions at the advanced stage.

Individuals' participation decisions at the advanced stage vary according to their augmented qualifications, which in turn depend on the schooling quality at the basic stage. Consequently, their benefits from public education funds depend on the allocation policy. There are two opposite forces at work. On the one hand, individuals with low qualifications, not attending advanced education themselves, help finance it through income tax, so there is reverse redistribution of resources from the poor to the rich.<sup>3</sup> On the other hand, individuals with high qualifications earn more income and pay more taxes, so the reverse redistribution may be partly or even fully offset. (See Bevia and Iturbe-Ormaetxe (2002)). Given a fixed size of public education funds, different allocation policies elicit quite different individual responses in school participation decisions, and therefore have different distributional effects.

With a fixed size of public funds, different allocation policies also differ in their effects on aggregate efficiency. Some policies lead to equilibrium outcomes that are strictly Pareto dominated. Since basic education is essential for producing human capital output in its own right and providing the necessary input, the augmented qualification, in advanced education, there exists a lower bound on its funding level. Allocation policies that fund basic education below this lower bound are strictly Pareto dominated by the one meeting this lower bound. On the other hand, whether a corresponding lower bound on funding advanced education exists or not depends on the initial distribution and the size of total public funds. When a fraction of the population has extremely low initial qualifications, and/or when the size of the total public funds is not big enough, any allocation of funding

---

<sup>2</sup>Think that students who can barely do algebra have nothing to gain to sit in an advanced analysis course.

<sup>3</sup>In the United States data, Hansen (1970), Radner and Miller (1970), Peltzman (1973) and Bishop (1977) find some evidence that students from high income families are more likely to attend higher education. Le Grand (1982) documents similar phenomenon for the United Kingdom. For a recent political economy analysis on this issue, see Fernandez and Rogerson (1995).

to advanced education always hurts those individuals who are unprepared for the advanced stage, and hence the allocation policies may not be Pareto ranked. Thus in this model, policy effects depend crucially on the development stage of an economy.

Outside the Pareto improvable range where there is no conflict of interests across individuals, there exists a range of allocation policies that benefit some individuals at the expense of others. Consequently there may be a trade-off between aggregate efficiency and equality. This trade-off, again, depends crucially on the development stage of an economy, i.e., the distribution of individuals' initial qualifications. Since basic education is universally beneficial and advanced education may be exclusive, the policy effects are essentially determined by how exclusive advanced education is in an economy. In a less developed economy, the majority of the population has extremely low initial qualifications, so advanced education is extremely exclusive. Then an allocation policy favoring basic education is most likely to improve both the aggregate efficiency and equality. In a developed economy, a substantial fraction of the population has high initial qualifications and can benefit more from advanced education. Therefore an allocation policy favoring basic education is likely to improve equality but hurt the aggregate efficiency. This intra-generational trade-off is an important aspect of the policy effects.

Through the inter-generational link, higher human capital levels of parents imply better pre-school preparations and hence higher initial qualifications of children, so an allocation policy that affects the current generation's decisions on human capital accumulation has long lasting impacts over all future generations. Consequently when there are multiple steady states, convergence to which steady state is path-dependent. Numerical examples are used to illustrate the simple rule-of-thumb. For an economy in its early stage of development, allocating most of the public funds on basic education for a sufficiently long time span is the only way out of the "underdevelopment trap" and polarization in the long run.

The rest of the paper is structured as follows. Section 2 lays out the model. Section 3 characterizes the equilibrium and analyzes the policy effects on equality and aggregate efficiency. When analytical results cannot be obtained, some numerical examples are provided in Section 4. Section 5 analyzes the long term policy effects. Section 6 concludes and discusses directions for future work. All proofs are relegated to Appendix A, as is the algorithm used in the numerical analysis. In Appendix B, the original model is extended to include both physical capital and private education expenditure. It is shown that the extended model shares the same qualitative features of the original model.

## 2. Two-stage education model

This is a successive generations model with heterogeneous individuals.<sup>4</sup> Within a generation, individuals differ in their pre-school preparations (initial qualifications), which is assumed to be an increasing function of the human capital levels of their parents.<sup>5</sup>

Individuals accumulate human capital through tax-financed public education. In the following analysis, I deliberately abstract from private education expenditure to focus on the response of individuals' school participation decisions to the allocation policy of public education funds.<sup>6</sup>

The hierarchical structure of the educational system is modeled as a two-stage process. Given the prevalence of compulsory schooling legislations, basic education at the lower stage is assumed fully compulsory and uniform in duration for all individuals. The human capital output from basic education determines individuals' preparations for advanced education (augmented qualifications) and is an input thereof. Poorly prepared individuals (with augmented qualifications below some threshold level) simply cannot benefit from advanced education even if they do attend, so optimally they forego that stage. Only sufficiently prepared individuals (with augmented qualifications above the threshold level) can benefit from advanced education, and they make their own participation decisions at that stage.

After finishing education, individuals join the labor force and work, get wage

---

<sup>4</sup>The successive generations modeling approach to the study of distributional dynamics follows Banerjee and Newman (1991) and Perotti (1993).

<sup>5</sup>That children's pre-school preparations depend on the human capital levels of parents is a commonly adopted specification in the literature. Just to mention a few, see Glomm and Ravikumar (1992), Eckstein and Zilcha (1994), Benabou (1996) and Kaganovich and Zilcha (1999).

<sup>6</sup>As pointed out by an anonymous referee, private education expenditure is also an important factor affecting education outcomes in many countries. A substantial body of literature focuses on the interaction between public and private education expenditures, in the form of public/private school choice and/or education vouchers (Glomm and Ravikumar (1992), Epple and Romano (1998), Caucutt (2002), Ladd(2002)). However, within the current framework, if private education expenditure is introduced, its allocation across education stages poses the same problem as the allocation of public education funds. In Appendix B, I show that if the allocation of private education expenditure follows some fixed rule, it is subsumed in the interpretation of the current education technology. However, if the allocation of private education expenditure adjusts in response to the public allocation policy, together with the adjustment of private schooling time decision, the model becomes insolvable for analytical results. The necessary simplification is dictated by the focus of the present study.

income, pay income tax and consume the remaining disposable income. At the end of life, each individual is replaced by one offspring, so the population remains constant.

It is assumed that a long-lived government collects tax revenue from one generation, and allocates the public funds between the basic and advanced stages of education for the next generation. Theoretically the government also sets the compulsory schooling duration at the basic stage. However, in reality, compulsory schooling legislations change infrequently compared to the annual budget allocation decisions. So our analysis treats the compulsory duration in basic education as exogenous and fixed, and focuses on the effects of different allocation policies of a fixed size of public education funds.

Below I describe the formal model. There are  $I$  distinct groups of individuals with measures  $\lambda_i$ ,  $i = 1, 2, \dots, I$ . Each group consists of a continuum of homogeneous agents, so an individual ignores the impact of his own decision on the aggregate. Total population is normalized to 1. Generation 0 individuals are endowed with group-specific levels of human capital  $h_0^i$ ,  $i = 1, 2, \dots, I$ , where  $h_0^1 < h_0^2 < \dots < h_0^I$ .

In variables describing a generation  $t$  individual of group  $i$ , subscript  $t$  indexes the generation and superscript  $i$  indexes the group. An individual starts with basic education according to the following technology:

$$h_t^{b,i} = Bl(h_{t-1}^i)^\delta (G_t/L)^\theta \quad (1)$$

This technology is fairly standard. There are three inputs: the initial qualification (pre-school preparation) captured by the term  $h_{t-1}^i$ ,<sup>7</sup> the compulsory schooling time  $l$ , and the relevant school quality  $G_t/L$ . Here  $G_t$  is the public funds allocated to the basic level,  $L$  is the aggregate time input in this level, which equals  $l$  since population is 1, so  $G_t/L$  measures annual public education

---

<sup>7</sup>Another interpretation of the term  $h_{t-1}^i$  would be that it captures the effect of private education expenditure, as long as we make the assumption that private educational expenditure is a normal good and its allocation across stages follows a fixed rule. See Appendix B for the detailed analysis. However, at the advanced stage, the non-hierarchical nature of private educational expenditure stands in sharp contrast to the hierarchical nature of preparations and qualifications for different stages. In other words, a dollar can be freely shifted between the basic and the advanced stages without change in value, yet the preparation for advanced education (augmented qualification) is affected by but does not affect the pre-school preparation for basic education (initial qualification). For reasons discussed before, I simplify from the issue of private education expenditure, and adopt the interpretation of pre-school preparation (initial qualification) associated with  $h_{t-1}^i$ .

expenditure per pupil at the basic stage. The human capital output is  $h_t^{b,i}$ . Since there is no private choice involved at the basic level, the distribution of  $h_t^{b,i}$ , the augmented qualification, can be treated as exogenously determined by the distribution of  $h_{t-1}^i$ , the initial qualification, together with public policy parameters. Namely, lack of preparation for basic education (low initial qualification  $h_{t-1}^i$ ) is directly translated into lack of preparation for advanced education (low augmented qualification  $h_t^{b,i}$ ), and good preparation for basic education (high initial qualification  $h_{t-1}^i$ ) implies good preparation for advanced education (high augmented qualification  $h_t^{b,i}$ ).

After basic education, an individual makes his own participation decision at the advanced stage according to the following technology:

$$h_t^{a,i} = An_t^i (h_t^{b,i} - \widehat{h})^\delta (g_t/N_t)^\theta \quad (2)$$

Again there are three inputs. An individual's preparation for advanced education (augmented qualification) depends on his human capital output from basic education,  $h_t^{b,i}$ ,<sup>8</sup> with the threshold  $\widehat{h}$  capturing the hierarchical nature: An individual with low augmented qualification is too poorly prepared to benefit from advanced education. Strong evidence in support of this threshold concept exists and is almost self-evident. For example, a child has to develop certain familiarity with the alphabet before he can start learning words and writing compositions; a student has to learn algebra to a certain degree before he can study calculus. It is worth pointing out that this threshold is intrinsic to the learning technology at the advanced stage. It is not subject to policy manipulation.<sup>9</sup>

Two other inputs are standard and parallel to those at the basic stage: the privately chosen schooling time  $n_t^i$ , and the relevant school quality  $g_t/N_t$ . Here  $g_t$  is the public funds allocated to the advanced level,  $N_t$  is the aggregate time input in this level, so  $g_t/N_t$  measures annual public education expenditure per pupil at the advanced stage. The human capital output is  $h_t^{a,i}$ .

For simplicity, it is assumed that human capital from basic and advanced education are perfect substitutes in goods production,<sup>10</sup> so an individual's total human capital  $h_t^i$  is given by:

---

<sup>8</sup>It follows directly from Lucas (1988) that existing human capital improves the learning efficiency in subsequent education.

<sup>9</sup>For analysis of the effects of policy-chosen threshold educational standards, see Costrell (1994) and Betts (1998). Notice that if both the technology-intrinsic and policy-chosen thresholds coexist, the latter has to be higher than the former and hence more exclusive.

<sup>10</sup>This is solely for analytical simplicity. Main results are robust to alternative specifications.



$$h_t^i = h_t^{b,i} + h_t^{a,i} \quad (3)$$

After finishing education, an individual spends the rest of his lifetime working. His total life time is normalized to 1, so the remaining working time is  $1 - l - n_t^i$ .<sup>11</sup> Individuals do not value leisure, so the flat-rate income tax  $\tau_t$  is not distortionary, and individuals consume all their disposable income. Without loss of generality, I abstract from life cycle consumption profile, so utility maximization is equivalent to after-tax income maximization, or simply income maximization. Consequently, without saving decision, there is no physical capital.<sup>12</sup> Goods production is linear in effective labor, so the wage rate can be normalized to 1 and wage payment exhausts output.<sup>13</sup> An individual's after-tax income is given by:

$$(1 - \tau_t)h_t^i(1 - l - n_t^i) \quad (4)$$

An individual makes his school participation decision  $n_t^i$  to maximize utility (4), subject to the constraints (1), (2) and (3), taking public policies and aggregate variables as given.

### 3. Equilibrium analysis

*The economy's equilibrium is defined as the collection of the sequences of exogenous public policy parameters  $(\tau_{t-1}, G_t, g_t)_{t=1}^\infty$ , privately chosen advanced schooling time  $\{\{n_t^i\}_{i=1}^I\}_{t=1}^\infty$ , and the aggregate time input in advanced education  $\{N_t\}_{t=1}^\infty$ , such that:*

1. For each  $t \geq 1$ , given  $G_t, g_t$  and  $N_t$ ,  $\{n_t^i\}_{i=1}^I$  solve private individuals' problems of maximizing (4) subject to (1), (2) and (3),;

---

<sup>11</sup>It is assumed that individuals reap the human capital after their time investment in education within the same period. See Galor and Tsiddon (1997) for a similar treatment.

<sup>12</sup>Even when life cycle consumption profile is introduced, with perfect credit market, life-time utility maximization is equivalent to (properly discounted) life-time income maximization. Within the current framework of fully publicly funded education, the credit market is non-operative since no individual faces liquidity constraint. See Appendix B for a detailed analysis.

<sup>13</sup>Or this can be thought as partial equilibrium analysis, where individuals take the market wage rate as given, and income is linear in effective labor. Since our focus is on the composition of human capital from basic and advanced education, we abstract from the general equilibrium effect arising from the interaction between aggregate human capital and aggregate physical capital. The two approaches yield equivalent results in intra-generational comparison across individuals. See Appendix B for a detailed analysis.

2. The consistency condition is satisfied:  $N_t = \sum_{i=1}^I \lambda_i n_t^i$ , i.e., individuals' expectation of the aggregate time input in advanced education is consistent with its realization;

3. The allocation policy respects the government budget in each period:

$$G_t + g_t = \tau_{t-1} \sum_{i=1}^I \lambda_i h_{t-1}^i (1 - l - n_{t-1}^i) \text{ for } t > 1 \quad (5)$$

$$G_1 + g_1 = \tau_0 \sum_{i=1}^I \lambda_i h_0^i \quad (6)$$

The first-order condition for individual  $i$ 's optimization problem is given by:

$$-(h_t^{b,i} + h_t^{a,i}) + h_t^{a,i} (1 - l - n_t^i) / n_t^i \leq 0, \quad n_t^i \geq 0, \text{ with at least one equality} \quad (7)$$

In other words,

$$n_t^i = 0 \text{ if } h_t^{b,i} \leq \widehat{h} \quad (8)$$

$$n_t^i = \max\left\{0, \frac{1-l}{2} - \frac{h_t^{b,i}}{2A(h_t^{b,i} - \widehat{h})^\delta (g_t/N_t)^\theta}\right\} \text{ if } h_t^{b,i} > \widehat{h} \quad (9)$$

Aggregation of (8) and (9) over all individuals yields the following equation:

$$N_t = \sum_{i=1}^I \lambda_i n_t^i = \sum_{h_t^{b,i} > \widehat{h}} \lambda_i \max\left\{0, \frac{1-l}{2} - \frac{h_t^{b,i}}{2A(h_t^{b,i} - \widehat{h})^\delta (g_t/N_t)^\theta}\right\} \quad (10)$$

**Proposition 1.** *In period  $t$ , given the distribution of initial qualifications  $h_{t-1}^1 < h_{t-1}^2 < \dots < h_{t-1}^I$ , if the allocation policy  $G_t$  and  $g_t$  is such that  $h_t^{b,I} > \widehat{h}$  after basic education, i.e., some individuals are qualified for advanced education, and  $g_t > 0$ , i.e., there is publicly funded advanced education, then there exists a unique aggregate advanced schooling time  $N_t$  that satisfies the equilibrium conditions. Also all private advanced schooling decisions  $\{n_t^i\}_{i=1}^I$  are uniquely determined.*

This proposition establishes the existence and uniqueness of the equilibrium under a given allocation policy within a period. Given the initial distribution of human capital for generation 0, public policies at period 1 uniquely determine the

distribution of human capital for generation 1. By induction, the human capital distribution for all future generations will be uniquely determined, given the sequence of allocation policies. The result is robust so long as there is negative externality in the education production technology, in the sense that when an individual chooses more schooling time at the advanced level, he ignores the effect of his decision on the schooling quality. From a private perspective, the education technology is linear in individual schooling time.<sup>14</sup> Since the size of public education funds is fixed, there are economy-wide diminishing returns to the aggregate time input in the education process.

**Proposition 2.** *In the equilibrium, there exist  $\underline{h}$  and  $\bar{h}$ , where  $\hat{h} < \underline{h} < \bar{h} < \infty$ , such that  $n_t^i > 0$  if and only if  $h_t^{b,i} \in (\underline{h}, \bar{h})$ . Moreover, between individual  $i$  and individual  $j$ , if  $\underline{h} < h_t^{b,i} < h_t^{b,j} \leq \frac{\hat{h}}{1-\delta}$ , then  $n_t^i < n_t^j$ ; if  $\frac{\hat{h}}{1-\delta} \leq h_t^{b,i} < h_t^{b,j} < \bar{h}$ , then  $n_t^i > n_t^j$ .*

This proposition characterizes how individuals base their advanced schooling decisions on their augmented qualifications (and hence initial qualifications), taking the allocation policy as given. It is not comparative statics analysis, but an intra-generational comparison across individuals. Given the hierarchical nature of the education system, after basic education, individuals with low augmented qualifications ( $h_t^{b,i} \leq \hat{h}$ ) are too poorly prepared to benefit from advanced education, so they forego this stage. For individuals with medium to high augmented qualifications ( $h_t^{b,i} > \hat{h}$ ), their advanced schooling decisions are not even monotonically increasing. Those barely qualified (with  $\hat{h} < h_t^{b,i} \leq \underline{h}$ ) may find their learning efficiencies too low for advanced study, and those super qualified (with  $h_t^{b,i} \geq \bar{h}$ ), if exist, may find their opportunity costs of foregone wage incomes too high for additional schooling, and both voluntarily opt out of the advanced level.<sup>15</sup> Last, for those individuals who find it optimal to attend advanced education (with  $\underline{h} < h_t^{b,i} < \bar{h}$ ), their advanced schooling time decisions increase with their augmented qualifications up to a point ( $\frac{\hat{h}}{1-\delta}$ ), and then decrease.

---

<sup>14</sup>This is consistent with the stylized fact that (log) earnings to schooling are essentially linear. See Mincer (1974) and Heckman and Solomon (1974).

<sup>15</sup>That the super qualified individuals may opt out of advanced education is robust to alternative specifications. As basic education is of fixed compulsory duration, individuals with extremely high initial qualifications may have accumulated enough human capital after basic education, so their opportunity costs of foregone wage incomes for further study may very well exceed the additional benefit. Then it is optimal to opt out of advanced education.

Within the hierarchical educational system, individuals' benefits from basic education funds are uniform; yet their benefits from advanced education funds differ due to their advanced schooling time decisions. More specifically, if there are individuals with extremely low initial qualifications (and hence extremely low augmented qualifications after basic education), they do not benefit from advanced education but help finance it through income tax, then publicly funded advanced education involves reverse distribution of resources away from the bottom of the population. On the other hand, who benefit most from this reverse redistribution depends on how far right the upper tail of the distribution stretches. Since economies at different stages of development are typically characterized by different distributions of initial qualifications (pre-school preparations), the distributional effects of an allocation policy depend critically on the development stage of an economy.

**Proposition 3.** *In the equilibrium, individuals with better pre-school preparations have higher human capital and higher wage income, i.e., if  $h_{t-1}^i < h_{t-1}^j$ , then  $h_t^i < h_t^j$  and  $h_t^i(1 - l - n_t^i) < h_t^j(1 - l - n_t^j)$ .*

This proposition shows that individuals with higher initial qualifications always earn more income and hence pay more taxes, i.e., the flat-rate income tax involves the usual redistribution from top to bottom. On the other hand, Proposition 2 shows that individuals with higher initial qualifications may benefit more from advanced education, i.e., there is reverse redistribution from bottom to top. The net distributional effect is the composition of the two opposite forces. In the next section, numerical examples are used to show that allocation policies may indeed differ quite drastically in their distributional effects.<sup>16</sup>

Now let's turn to comparative statics to study the policy effects on aggregate efficiency. In period  $t$ , the total tax revenue  $R_t$  is predetermined by the aggregate output and tax rate in the previous period. So the size of public education funds is fixed, and the allocation choice involves a one-for-one trade-off between funding basic and advanced education.

**Proposition 4.** *In period  $t$ , if the size of total public education funds is not too small, i.e.,  $R_t > \underline{G}_t$ , where  $\underline{G}_t$  is determined by the equation  $\hat{h} = Bl(h_{t-1}^I)^\delta (\underline{G}_t/L)^\theta$ ,*

---

<sup>16</sup>Another implication of this proposition is that there is no social mobility over time. This is only an artifact of our assumptions on identical preferences across individuals and lack of uncertainty. Without uncertainty that directly shuffles the social ranking, and without preference heterogeneity with respect to one's offsprings, overtake incurs too much cost and generates too little benefit for the current generation, so there is no social mobility.

so that after basic education with funding  $\underline{G}_t$ , individuals with the best initial qualifications are just qualified for advanced education, then there exists  $\varepsilon > 0$  such that the allocation policy with  $G_t = (1+\varepsilon)\underline{G}_t$  Pareto dominates any allocation policy with  $G_t < (1 + \varepsilon)\underline{G}_t$ .

This proposition establishes a lower bound on funding basic education, namely  $(1+\varepsilon)\underline{G}_t$ , for Pareto dominance. It arises because basic education is the foundation of the entire hierarchical educational system. Basic education produces human capital output in its own right, and it provides the necessary input, the augmented qualification, for advanced education. Thus any allocation policy that poorly funds basic education is strictly Pareto dominated.

On the other hand, whether a corresponding lower bound on funding advanced education exists or not depends on the development stage of an economy and the size of total public funds. When a big fraction of the population has extremely low initial qualifications, as usually in a less developed economy, or when the size of total public funds is not big enough, funding advanced education always hurts those who are unprepared for the advanced stage, so Pareto ranking may not be established. This asymmetry of policy implications roots from the asymmetry between the basic and advanced stages in the educational hierarchy.

Pareto ranking is desirable but extremely difficult to obtain in a setup of heterogeneous individuals. Given a fixed size of public funds, there exists a wide range of allocation policies that benefit some individuals at the expense of others. Hence a policy change may involve a trade-off between aggregate efficiency and equality. The highly nonlinear equation (10) prevents us from characterizing the trade-off analytically, so let's turn to numerical examples next.

## 4. Numerical examples

In this section the model is solved numerically to demonstrate the policy effects. The algorithm for the numerical procedure is provided in Appendix A. Given a fixed size of public education funds, this exercise compares the entire range of feasible allocation policies and their effects on equality and aggregate efficiency. So when the effects differ between two allocation policies, that difference is solely due to the composition rather than the size of public education funds. This enables us to identify the composition effect free from the size effect.

In the following examples, the size of total public education funds is set at  $R_t = 4$ . An allocation policy is uniquely specified by its funding level for basic

education,  $G_t$ , with the remaining funds for advanced education. All feasible allocation policies must satisfy the condition  $G_t > \underline{G}_t$ , with  $\underline{G}_t$  defined in Proposition 4. Otherwise, no individual qualifies for advanced education, and the advanced education funds are obviously wasted.

The distributions of initial qualifications are chosen to characterize different development stages of an economy. Other parameters are picked solely for algebraic simplicity and presented in Table 3.

**Example 1. A less developed economy**

Let the distribution of initial qualification be as follows:  $h_{t-1}^1 = 1$ ,  $\lambda_1 = 0.7$ ,  $h_{t-1}^2 = 6$ ,  $\lambda_2 = 0.2$ , and  $h_{t-1}^3 = 20$ ,  $\lambda_3 = 0.1$ . It is characteristic of a less developed economy, where a big fraction at the bottom of the population has extremely low initial qualifications, the middle class is small, and a very small portion at the top has high initial qualifications. The equilibrium outcomes associated with different allocation policies are presented in Figure 1.

The entire range of allocation policies are represented on the horizontal axis, where funding for basic education ranges from  $\underline{G}_t$  up to  $R_t$ , and consequently, funding for advanced education ranges from  $R_t - \underline{G}_t$  down to 0. The corresponding equilibrium outcomes are represented on the vertical axis, where different graphs give both the group-specific outcomes and the economy-wide aggregates.

The first two graphs demonstrate the policy effects on group-specific human capital and income levels. In this less developed economy, the bottom fraction of the population cannot qualify for advanced education regardless of the allocation policy, the middle class can attend advanced education only when  $G_t$  is sufficiently high, and the top fraction benefits most from advanced education when it is highly exclusive, i.e., when  $G_t$  is moderate. Consequently, allocation policies differ tremendously in their effects on group-specific outcomes. Once outside the Pareto improving range where all three groups benefit from higher  $G_t$  given by Proposition 4, a further increase in  $G_t$  may benefit both the bottom and the middle class yet hurt the top fraction up to 50% of its maximal human capital and 20% of its maximal income. The composition effects of public education funds are prominent.

The second two graphs illustrate the policy effects on individuals' net benefits from public education funds. The net benefits are computed by deducting individuals' tax payments from their benefits (which in turn depend on their schooling time decisions), holding the size of public education funds constant overtime. When  $G_t$  is moderate and hence advanced education is highly exclusive, the tax payment of the top fraction is not sufficient to cover its high benefit from advanced

education, so the allocation policies result in reverse redistribution of resources from bottom to top. When  $G_t$  is sufficiently big, this reverse redistribution is more than fully offset, and the allocation policy generates the usual redistributive effect from top to bottom. In this example, the small middle class always has negative net benefit regardless of the allocation policies. Moreover, once we restrict our attention to allocation policies outside the Pareto improving range, the reverse redistribution effect disappears. All allocation policies funding basic education at sufficiently high levels generate the usual redistribution from top and the middle to bottom.

The last two graphs focus on the macroeconomic issue of aggregate efficiency and equality. Aggregate efficiency is measured as the economy-wide aggregates, and (in)equality is measured by the Gini coefficients. Within the Pareto improving range of allocation policies, Proposition 4 tells us that the efficiency unambiguously improves with higher  $G_t$ . The Gini coefficient increases first and then levels off. So the conflict between efficiency and equality starts severe but mitigates when the policy puts more and more focus on basic education. The moderate (or even lack of) trade-off between efficiency and equality continues with higher and higher  $G_t$ . Overall in a less developed economy, allocation policies focusing on basic education benefit the majority of the population and are likely to improve both aggregate efficiency and equality.

**Example 2. A developed economy**

Now let the distribution of initial qualifications be as follows:  $h_{t-1}^1 = 10$ ,  $\lambda_1 = 0.2$ ,  $h_{t-1}^2 = 15$ ,  $\lambda_2 = 0.7$ , and  $h_{t-1}^3 = 50$ ,  $\lambda_3 = 0.1$ . This distribution becomes characteristic of one in a developed economy. The entire population has at least medium qualifications, with a small bottom group, a big middle class, and an elite top group. The equilibrium outcomes associated with different allocation policies are reported in Figure 2. The policy effects in a developed economy is quite different from those in a less developed economy.

Again the first two graphs report the policy effects on group-specific human capital and income levels. In a developed economy, advanced education is less exclusive, so the Pareto improving range of allocation policies established in Proposition 4 has greatly shrunk, i.e., the lower bound is only slightly above  $\underline{G}_t$ . A moderately higher  $G_t$  is likely to expand the fraction of qualified individuals and induce participation in advanced education, so the elite top group is hurt due to diluted schooling quality at the advanced stage. Overall policy effects involve sharp conflict of interests across groups. Moreover, in a developed economy, since the entire population can benefit from advanced education so long as the size of

total funds is big enough, there may exist a lower bound on funding advanced education for Pareto criterion, which can never exist in a less developed economy.

In the second two graphs, it is obvious that catch-up may take place between the middle class and the top group, and policies differ in their effects on individuals net benefits from public education. When the middle class invests more schooling time in advanced education than the top group, its net benefit from public education becomes positive, which never happens in the previous example of a less developed economy. Similar to the less developed economy case, reverse redistribution of resources exists when  $G_t$  is low and hence advanced education is highly exclusive, and this reverse redistribution is more than fully offset when  $G_t$  becomes high.

The last two graphs report the policy effects on aggregate efficiency and equality. Interesting enough, they share the same pattern as in a less developed economy. Within the Pareto improving range of allocation policies, aggregate efficiency improves with higher  $G_t$ , and inequality increases sharply first and quickly levels off. Outside this range, efficiency keeps increasing with higher  $G_t$ , and so does equality. Even in a developed economy, allocation policies favoring basic education may achieve both aggregate efficiency and equality, especially when the size of total public funds is not too big.

In summary, the two numerical examples illustrate the following points. First, there is a lower bound on funding basic education for Pareto dominance as suggested in Proposition 4, and this lower bound is much higher for a less developed economy than a developed economy. Promoting basic education involves less conflict of interests across groups in the former than in the latter case. Second, a corresponding lower bound on funding advanced education for Pareto dominance fails to exist when a fraction of the population has extremely low initial qualifications, and/or when the size of total public funds is not too large. Third, an allocation policy is more likely to induce reverse redistribution of resources when advanced education is highly exclusive, or equivalently, when its funding level for basic education is only moderate. Last but not the least, the trade-off between efficiency and equality may be quite moderate under allocation policies favoring basic education for a less developed economy, or for a developed economy with relatively small size of total public funds.



## 5. Long run policy effects

In this model, parents' decisions on human capital accumulation directly affect children's pre-school preparations, so policy effects are transmitted across generations and long-lasting. As before, to single out the composition effect, the total tax revenue is held constant over time. The experiment is to change the short run allocation policies while holding the long run allocation policies constant. Due to the non-convexity in the advanced education technology, there may be multiple steady states, and convergence to which steady states is path dependent.<sup>17</sup>

With a fixed size of public funds, in a steady state, the allocation policy consists of the time-invariant funding levels for basic and advanced education,  $G$  and  $g$ , respectively. Omitting the time subscript  $t$ , a steady state is governed by the following system of equations:

$$h^{b,i} = Bl(h^{b,i} + h^{a,i})^\delta (G/L)^\theta \quad (11)$$

$$h^{a,i} = An^i(h^{b,i} - \hat{h})^\delta (g/N)^\theta \quad (12)$$

$$n^i = 0 \text{ if } h^{b,i} \leq \hat{h} \quad (13)$$

$$n^i = \max\left\{0, \frac{1-l}{2} - \frac{h^{b,i}}{2A(h^{b,i} - \hat{h})^\delta (g/N)^\theta}\right\} \text{ if } h^{b,i} > \hat{h} \quad (14)$$

$$N = \sum_{h^{b,i} > \hat{h}} \lambda_i \max\left\{0, \frac{1-l}{2} - \frac{h^{b,i}}{2A(h^{b,i} - \hat{h})^\delta (g/N)^\theta}\right\} \quad (15)$$

This is a system of  $3I+1$  non-linear equations with  $3I+1$  unknown variables. It may have multiple solutions, corresponding to multiple steady states supportable under the same long run allocation policies. To which steady state the economy actually converges depends on the short run policies. Below is a numerical example to illustrate this path-dependent convergence.

### Example 3 Path-dependent convergence

Let's return to the less developed economy example, and adopt the initial distribution thereof. The size of public education funds is fixed at  $R = 4$  over

---

<sup>17</sup>See Azariadis and Drazen (1990) for their discussion on threshold effects, where small policy changes may have decisive effects on the dynamics of the economy.

time. From the analysis of the previous section, it is known that focusing on basic education improves both aggregate efficiency and equality, so the short run allocation policy is set to allocate all the public funds on basic education, i.e.,  $G_t = 4$  and  $g_t = 0$ . The time-invariant long run allocation policy is set at  $G = 1.5$  and  $g = 2.5$ . This exercise considers six distinct policy sequences, where the effective duration of the short run allocation policy ranges from 1 to 6 periods, before switching to the long run allocation policy. Hence any difference in the steady states of the economy is solely due to the duration of the short run policies. The results are in figure 3.

It is interesting to see that slight difference in the short run allocation policy may have tremendous impact on the long run outcomes of the economy. The six policy sequences can be roughly divided into three categories. Even with the basic education oriented short run allocation policy, it takes time for poorly-endowed dynasties to catch up; and the lower are their initial qualifications, the longer duration it takes. In the first category, when the short run allocation policy is effective for only 1 or 2 periods, the middle class and the bottom population remain stuck with low human capital, and their poor qualifications prevent them from benefiting from advanced education in the long run. The small top group enjoys the highly exclusive advanced education. There is polarization in the economy in the long run. Also the long run aggregate output is 10.99, which is the lowest among all three categories.

In the second category, the short run allocation policy is effective for 3 or 4 periods. While it enables group 2 to catch up and benefit from advanced education in the long run, it is still insufficient for the extremely poorly-endowed group 1. Consequently in the long run, the middle class joins the top group in advanced education (which necessarily split the benefit that originally solely accrues to the top group), yet the bottom of the population remains outside advanced education. Especially interesting is the case when  $t = 4$ , where the bottom group does once become qualified for advanced education under the short run allocation policy, yet gets crowded out under the long run allocation policy that is much less favorable to them. The long run inequality is still prominent, but it is much improved compared to the polarization outcome in the first category. The long run aggregate output is 12.16, 11% higher than that in the first category and the highest among all three.

In the third category, the short run allocation policy is effective for 5 or 6 periods, and it is sufficient to promote both the middle class and the bottom group to qualify for advanced education. Then when the economy finally switch

to its long run allocation policy, the entire population benefits from advanced education and inequality is minimal. The long run aggregate output is 11.54, 5% higher than that in the first category.

This exercise is by no means exhaustive, but it does reveal the simple rule-of-thumb. In a less developed economy, or equivalently, when an economy is at its early development stage, the short run allocation policy may have great impact on the long run outcomes. More specifically, when a big fraction of the population has extremely low initial qualifications, funding advanced education at the expense of basic education will keep the bottom population in the low human capital trap permanently. Inequality will be persistent and substantially high, and long run aggregate efficiency is low. Instead, heavily funding basic education for sufficiently long duration enables the entire population to qualify for and benefit from advanced education, and the long run outcomes are much more equal. This suggests a negative correlation between current inequality and the relative focus the allocation policy puts on basic education in an economy's early development stage, which can be tested by the data.

## 6. Conclusion

This paper studies the dynamic effects of the allocation of public funds between basic and advanced education. We identify the composition effects of public education funds by holding the size fixed. The allocation policy has big impact on both aggregate efficiency and equality.

On the efficiency side, there exists a lower bound on funding basic education. Any allocation policy that funds basic education below this lower bound is strictly Pareto dominated. On the other hand, whether a corresponding lower bound on funding advanced education exists or not depends on the development stage of the economy and the size of total public funds. Only when the entire population has at least medium initial qualifications, and the size of total public education funds is big enough, will funding advanced education be Pareto improving.

On the equality side, an allocation policy has profound distributional effect. The distributional effect arises when the policy induces individuals to adjust their advanced schooling time decisions, and hence their net benefits from public education funds differ. It is shown that while an allocation policy favoring basic education generates the usual redistribution associated with public education from top to bottom, one favoring advanced education renders it highly exclusive and

results in reverse redistribution from bottom to top.

Interesting enough, the trade-off between efficiency and equality seems quite moderate within the relevant range of allocation policies. In particular, for a less developed economy, an allocation policy that focuses on basic education achieves both aggregate efficiency and equality. This is in contrast to the actual policies pursued by many less developed economies, which tend to focus more on advanced education as compared to developed economies. This exciting political economy question is left for future research.

Through the inter-generational link, an allocation policy that affects parents' human capital levels also has impact on children's pre-school preparations, so its effects are long lasting and transmitted across generations. Different short run policies may lead an economy to different steady states in the long run, which differ sharply in their efficiency and equality. Numerical examples are provided to demonstrate the simple rule-of-thumb, that for an economy in its early development stage, focusing on basic education for sufficiently long duration is the only way out of polarization and low aggregate efficiency.

This paper is only the first step toward understanding the policy effects of allocating public funds within a hierarchical educational system. Many issues remain open and to be taken in future research. First, adding private education expenditure and allowing flexibility in its allocation across educational stages will enrich the resources-time interactions, and more complicated computational techniques are needed to characterize the equilibrium and analyze the policy effects. Second, based on the positive analysis of the policy effects, we can take the model one step forward and conduct normative analysis on the optimal policy issue. Since short run policies may have long run effects, it is necessary to reduce the infinite dimension problem into some finite dimension summary indexes for policy comparison criteria. Third, the hierarchical structure of the educational system can be put into a political economy framework and used to study how individuals' policy preferences are aggregated into the public allocation policy. Last but not the least, introducing randomness into the inter-generational transmission of human capital may generate some value-added.

## References

- Azariadis, Costas and Drazen, Allan. "Threshold Externalities in Economic Development", *Quarterly journal of Economics*, May 1990, 105(2), pp. 501-526.
- Banerjee, Abhijit V. And Newman, Andrew. "Risk-bearing and the Theory of Income Distribution", *Review of Economics Studies*, April 1991, 58(2), pp. 211-235.
- Benabou, Roland. "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance", June 1996, 86(3), pp. 584-609.
- Betts, Julian R. "The Impact of Educational Standards on the Level and Distribution of Earnings", *American Economic Review*, March 1998, 88(1), pp. 266-275.
- Bevia, Carmen and Iturbe-Ormaetxe, Inigo. "Redistribution and Subsidies for Higher Education". *Scandinavian Journal of Economics*, June 2002, 104(2), pp. 321-340.
- Bishop, John. "The Effect of Public Policies on the Demand for Higher Education", *Journal of Human Resource*, 1977, 12(3), pp. 285-307.
- Caucutt, Elizabeth M. "Educational Vouchers When There Are Peer Group Effects - Size Matters". *International Economic Review*, Feb. 2002, 43(1), pp. 195-222.
- Costrell, Robert M. "A Simple Model of Educational Standards", *American Economic Review*, Sep. 1994, 84(4), pp. 956-971.
- Driskill, Robert A. and Horowitz, Andrew. "Investment in Hierarchical Human Capital", *Review of Development Economics*, Feb. 2002, 6(1), pp. 48-58.
- Eckstein, Zvi and Zilcha, Itzhak. "The Effects of Compulsory Schooling on Growth, Income Distribution and Welfare", *Journal of Public Economics*, July 1994, 54(3), pp. 339-359.
- Epple, Dennis and Romano, Richard E. "Competition between Private and Public Schools, Vouchers, and Peer Group Effects". *American Economic Review*, March 1998, 88(1), pp. 33-62.
- Fernandez, Raquel and Rogerson, Richard. "On the Political Economy of Education Subsidies", *Review of Economic Studies*, Apr. 1995, 62(2), pp. 249-262.
- Galor, Oded and Tsiddon, Daniel. "Technological Progress, Mobility, and Endogenous Growth", *American Economic Review*, June 1997, 87(3), pp. 363-382.
- Glomm, Gerhard and Ravikumar, B. "Public versus Private Investment in

Human Capital: Endogenous Growth and Income Inequality”, *Journal of Political Economy*, Aug. 1992, 100(4), pp. 818-834.

Hansen, W. Lee. “Income Distribution Effects of Higher Education”, *American Economic Review*, May 1970, 60(2), pp. 335-340.

Heckman, James J. and Solomon, Polachek. “Empirical Evidence on the Functional Form of the earnings-schooling relationship”, *Journal of the American Statistical Association*, June 1974, 69(346), pp. 350-354.

Kaganovich, Michael and Zilcha, Itzhak. “Education, Social Security and Growth”, *Journal of Public Economics*, Feb. 1999, 71(2), pp. 289-309.

Ladd, Helen F. "School Vouchers: A Critical Review". *Journal of Economic Perspective*, Fall 2002, 16(4), pp. 3-24.

Le Grand, J. “The Strategy of Equality: Redistribution and the Social Services”, George Allen and Unwin (Publishers) Ltd, 1982: 54-81

Lucas, Robert E. Jr. “ On the Mechanics of Economic Development”, *Journal of Monetary Economics*, July 1988, 22(1), pp. 3-42.

Mincer, Jacob. “Schooling, Experience, and Earnings”, New York, Columbia University Press, 1974.

Peltzman, Sam. “The Effect of Government Subsidies-in-kind on Private Expenditures: the Case of Higher Education”, *Journal of Political Economy*, Feb. 1973, pp. 1-27.

Perotti, Roberto. “Political Equilibrium, Income Distribution, and Growth”, *Review of Economic Studies*, October 1993, 60(4), pp. 755-776.

Radner, R. and Miller, L. S. “Demand and Supply in U. S. Higher Education: a Progress Report”, *American Economic Review*, May 1970, 60(2), pp. 326-334.

UNESCO, '99 Statistical Yearbook.

## Appendix A

### Proposition 1.

**Proof.** The equilibrium solution for  $N_t$  satisfies equation (10). The left hand side of (10) is continuous and strictly increasing in  $N_t$ . The right hand side, denoted as  $f(N_t)$ , is continuous and decreasing in  $N_t$ , with  $f(0) > 0$ , and  $f(\frac{1-l}{2}) < \frac{1-l}{2}$ . The single crossing point of the two curves is the equilibrium solution for  $N_t$ , and  $N_t \in (0, \frac{1-l}{2})$ . Then  $\{n_t^i\}_{i=1}^I$  can be determined through (8) and (9). ■

### Proposition 2.

**Proof.** Individuals take the equilibrium solution for  $N_t$  parametrically, and solve for their own advanced schooling time  $n_t^i$  according to (8) and (9). So  $n_t^i > 0$  if and only if for some  $h_t^{b,i} > \hat{h}$ ,  $\frac{1-l}{2} - \frac{h_t^{b,i}}{2A(h_t^{b,i} - \hat{h})^\delta (g_t/N_t)^\theta} > 0$ ; or equivalently,  $A(1-l)(h_t^{b,i} - \hat{h})^\delta (g_t/N_t)^\theta > h_t^{b,i}$ . Denote  $f(h) = A(1-l)(h - \hat{h})^\delta (g_t/N_t)^\theta$  for  $h > \hat{h}$  and  $\delta \in (0, 1)$ , then  $f'(h) > 0$ ,  $f''(h) < 0$ ,  $\lim_{h \rightarrow \hat{h}+} f(h) = 0 < \hat{h}$ ,  $\lim_{h \rightarrow \hat{h}+} f'(h) = +\infty$ , and  $\lim_{h \rightarrow +\infty} f'(h) = 0$ . Suppose  $f(h) \leq h$  for all  $h > \hat{h}$ , then in the economy,  $n_t^i = 0$  for all groups, which is contradictory to the positive equilibrium solution  $N_t$ . When  $f(h) > h$  for some  $h > \hat{h}$ , by continuity, there exist  $\underline{h}$  and  $\bar{h}$ , with  $\hat{h} < \underline{h} < \bar{h} < \infty$ , such that  $f(h) > h$  if and only if  $h \in (\underline{h}, \bar{h})$ . So in the economy,  $n_t^i > 0$  if and only if  $h_t^{b,i} \in (\underline{h}, \bar{h})$ . Now for the function  $n = \frac{1-l}{2} - \frac{h}{2A(h - \hat{h})^\delta (g_t/N_t)^\theta}$  with  $h \in (\underline{h}, \bar{h})$ , we have  $\text{sign}(\frac{dn}{dh}) = \text{sign}(-(1-\delta)h + \hat{h})$ , so  $\frac{dn}{dh} > 0$  when  $h \in (\underline{h}, \frac{\hat{h}}{1-\delta})$ ,  $\frac{dn}{dh} = 0$  when  $h = \frac{\hat{h}}{1-\delta}$ , and  $\frac{dn}{dh} < 0$  when  $h \in (\frac{\hat{h}}{1-\delta}, \bar{h})$ . It then follows that if  $\underline{h} < h_t^{b,i} < h_t^{b,j} \leq \frac{\hat{h}}{1-\delta}$ , then  $n_t^i < n_t^j$ ; if  $\frac{\hat{h}}{1-\delta} \leq h_t^{b,i} < h_t^{b,j} < \bar{h}$ , then  $n_t^i > n_t^j$ . ■

### Proposition 3.

**Proof.** It is obvious that  $h_{t-1}^i < h_{t-1}^j$  implies  $h_t^{b,i} < h_t^{b,j}$ . When  $h_t^{b,i} \leq \hat{h}$ ,  $h_t^{a,i} = 0 \leq h_t^{a,j}$ , then it follows that  $h_t^i < h_t^j$ . When  $\hat{h} < h_t^{b,i} < h_t^{b,j}$ , both individual  $i$  and individual  $j$  chooses their advanced schooling time  $n_t^i$  and  $n_t^j$  optimally according to (7). Denote  $\bar{n}_t^i$  as the advanced schooling time needed for individual  $i$  so that  $h_t^i = h_t^j$ , and it is obvious that  $\bar{n}_t^i > n_t^j$ . Suppose  $n_t^i \geq \bar{n}_t^i$ , it is easy to check that  $-(h_t^{b,i} + h_t^{a,i}) + h_t^{a,i}(1-l-n_t^i)/n_t^i < -(h_t^{b,j} + h_t^{a,j}) + h_t^{a,j}(1-l-n_t^j)/n_t^j \leq 0$ , which cannot be optimal for individual  $i$ . So  $n_t^i < \bar{n}_t^i$ , and then it follows that  $h_t^i < h_t^j$ . Similar argument shows that  $h_t^i(1-l-n_t^i) < h_t^j(1-l-n_t^j)$ . ■

**Proposition 4.**

**Proof.** For  $0 < x < y \leq 1$ , denote the equilibrium outcomes under  $G_t = (1+y\varepsilon)\underline{G}_t$  as variables with a tilde, and those under  $G_t = (1+x\varepsilon)\underline{G}_t$  without. There exists  $\varepsilon > 0$  such that no group but group  $I$  is qualified for advanced education.

For group  $i$ , where  $i = 1, 2, \dots, I-1$ , it is obvious that  $\tilde{h}_t^{b,i} > h_t^{b,i}$  and  $\tilde{h}_t^{b,i}(1-l) > h_t^{b,i}(1-l)$ , so both human capital and gross income are higher when  $G_t = (1+\varepsilon)\underline{G}_t$  than when  $G_t < (1+\varepsilon)\underline{G}_t$ .

For group  $I$ , the following equations hold:

$$h_t^{b,I} = Bl(h_{t-1}^I)^\delta((1+x\varepsilon)\underline{G}_t/L)^\theta \approx Bl(h_{t-1}^I)^\delta(\underline{G}_t/L)^\theta(1+\theta x\varepsilon) = (1+\theta x\varepsilon)\hat{h} \quad (16)$$

$$\tilde{h}_t^{b,I} = Bl(h_{t-1}^I)^\delta((1+y\varepsilon)\underline{G}_t/L)^\theta \approx Bl(h_{t-1}^I)^\delta(\underline{G}_t/L)^\theta(1+\theta y\varepsilon) = (1+\theta y\varepsilon)\hat{h} \quad (17)$$

$$N_t = \lambda_I n_t^I = \lambda_I \left( \frac{1-l}{2} - \frac{(1+\theta x\varepsilon)\hat{h}^{1-\delta} N_t^\theta}{2A(\theta x\varepsilon)^\delta(R_t - (1+x\varepsilon)\underline{G}_t)^\theta} \right) \quad (18)$$

$$\tilde{N}_t = \lambda_I \tilde{n}_t^I = \lambda_I \left( \frac{1-l}{2} - \frac{(1+\theta y\varepsilon)\hat{h}^{1-\delta} \tilde{N}_t^\theta}{2A(\theta y\varepsilon)^\delta(R_t - (1+y\varepsilon)\underline{G}_t)^\theta} \right) \quad (19)$$

Furthermore, we know that:

$$\lim_{\varepsilon \rightarrow 0^+} \frac{(1+\theta x\varepsilon)}{2A(\theta x\varepsilon)^\delta(R_t - (1+x\varepsilon)\underline{G}_t)^\theta} / \frac{(1+\theta y\varepsilon)}{2A(\theta y\varepsilon)^\delta(R_t - (1+y\varepsilon)\underline{G}_t)^\theta} = \left(\frac{y}{x}\right)^\delta > 1 \quad (20)$$

Suppose  $\tilde{N}_t \leq N_t$ , comparing the right hand side of (18) and (19), we have  $\tilde{N}_t > N_t$ , which leads to contradiction. It then follows that  $\tilde{N}_t > N_t$ , and then  $\tilde{n}_t^I > n_t^I$ . Thus:

$$\lim_{\varepsilon \rightarrow 0^+} \frac{\tilde{h}_t^{a,I}}{h_t^{a,I}} = \lim_{\varepsilon \rightarrow 0^+} \frac{A\tilde{n}_t^I(\theta y\varepsilon\hat{h})^\delta(R_t - (1+y\varepsilon)\underline{G}_t)^\theta}{An_t^I(\theta x\varepsilon\hat{h})^\delta(R_t - (1+x\varepsilon)\underline{G}_t)^\theta} > \left(\frac{y}{x}\right)^\delta > 1 \quad (21)$$

So for group  $I$  individuals, the human capital level is higher when  $G_t = (1+\varepsilon)\underline{G}_t$  than when  $G_t < (1+\varepsilon)\underline{G}_t$ .

As for the income level, when  $G_t = (1+y\varepsilon)\underline{G}_t$ , it is feasible but not optimal to choose  $n_t^I$  instead of  $\tilde{n}_t^I$ ; denote the corresponding outcomes as variables with double tilde. Then we have:

$$\lim_{\varepsilon \rightarrow 0^+} \frac{\tilde{\tilde{h}}_t^{a,I}}{h_t^{a,I}} = \lim_{\varepsilon \rightarrow 0^+} \frac{A\tilde{\tilde{n}}_t^I(\theta y\varepsilon\hat{h})^\delta(R_t - (1+y\varepsilon)\underline{G}_t)^\theta}{An_t^I(\theta x\varepsilon\hat{h})^\delta(R_t - (1+x\varepsilon)\underline{G}_t)^\theta} = \left(\frac{y}{x}\right)^\delta > 1 \quad (22)$$



So  $\tilde{h}_t^I(1-l-\tilde{n}_t^I) > \tilde{h}_t^I(1-l-n_t^I) > h_t^I(1-l-n_t^I)$ , where the first inequality follows from the optimality of the choice  $\tilde{n}_t^I$  and the second inequality follows since  $\tilde{h}_t^I > h_t^I$ . Consequently the income level for group  $I$  is also higher when  $G_t = (1+\varepsilon)\underline{G}_t$  than when  $G_t < (1+\varepsilon)\underline{G}_t$ .

Through the inter-generational link, all future generations have higher pre-school preparations when  $G_t = (1+\varepsilon)\underline{G}_t$ . So the allocation policy with  $G_t = (1+\varepsilon)\underline{G}_t$  Pareto dominates any allocation policy with  $G_t < (1+\varepsilon)\underline{G}_t$ . ■

### Equilibrium solution algorithm

By Proposition 1, we know that there is a unique solution for  $N_t \in (0, \frac{1-l}{2})$ .

Step 1. Set  $a = 0$  and  $b = \frac{1-l}{2}$ .

Step 2. Let  $N = \frac{a+b}{2}$ , substitute the value into (8) or (9) to compute  $n_t^i$  for all group  $i$ , and then aggregate  $n_t^i$  as in (10) to obtain  $N_t$ .

Step 3. If  $N_t > N$ , then  $a = N$ ; if  $N_t < N$ , then  $b = N$ . Repeat step 2 till convergence occurs according to a pre-specified error band.

Step 4. Take  $N_t$  as the equilibrium solution, and compute  $n_t^i$  for all group  $i$ . All other variables can be subsequently determined.

## Appendix B

The Extended Model with Physical Capital and Private Educational Expenditure

Now the original model is extended to include both physical capital (hence life cycle saving decisions) and private educational expenditure. The extended model differs with the original model in the following aspects. Now a generation  $t$  group  $i$  individual is active for two periods  $t$  and  $t + 1$ : the first period is the same as that in the original model, and the second period is retirement. The individual allocates his after-tax wage income to the following four components: consumption when young  $c_t^i$ , consumption when old  $d_{t+1}^i$ , private education expenditure for his child at the basic stage  $e_{t+1}^{b,i}$ , and that at the advanced stage  $e_{t+1}^{a,i}$ . So his utility function, corresponding to equation (4), is given by:

$$\ln c_t^i + \beta \ln d_{t+1}^i + \gamma_1 \ln e_{t+1}^{b,i} + \gamma_2 \ln e_{t+1}^{a,i} \quad (23)$$

An individual  $i$  maximizes his utility according to the following constraints:

$$h_t^{b,i} = Bl(h_{t-1}^i)^\delta (G_t/L)^\theta (e_t^{b,i})^\kappa \quad (24)$$

$$h_t^{a,i} = An_t^i (h_t^{b,i} - \widehat{h})^\delta (g_t/N_t)^\theta (e_t^{a,i})^\kappa \quad (25)$$

$$h_t^i = h_t^{b,i} + h_t^{a,i} \quad (26)$$

$$c_t^i + s_{t+1}^i = (1 - \tau_t) w_t h_t^i (1 - l - n_t^i) \quad (27)$$

$$d_{t+1}^i + e_{t+1}^{b,i} + e_{t+1}^{a,i} = s_{t+1}^i R_{t+1} \quad (28)$$

Here  $s_{t+1}^i$  is individual  $i$ 's saving from period  $t$  to period  $t + 1$ ,  $R_{t+1}$  is the rate of return on physical capital in period  $t + 1$ , and  $w_t$  is the wage rate in period  $t$ . Goods production technology is given by the following:

$$Y = K_t^\alpha L_t^{1-\alpha} \quad (29)$$

*The equilibrium is defined as the collection of the sequences of exogenous policy parameters  $\{\tau_{t-1}, G_t, g_t\}_{t=1}^\infty$ , private choices  $\{\{n_t^i, c_t^i, s_t^i, d_t^i, e_t^{b,i}, e_t^{a,i}\}_{i=1}^I\}_{t=1}^\infty$ , the aggregate variables  $\{N_t, K_t, L_t\}_{t=1}^\infty$ , and market prices  $\{w_t, R_t\}_{t=1}^\infty$ , such that:*

1. For  $t \geq 1$ , given  $G_t, g_t, N_t, w_t$  and  $R_{t+1}$ ,  $\{n_t^i, c_t^i, s_{t+1}^i, d_{t+1}^i, e_{t+1}^{b,i}, e_{t+1}^{a,i}\}_{i=1}^I$  solve private individuals' problems;

2. The consistency condition is satisfied:  $N_t = \sum_{i=1}^I \lambda_i n_t^i$ ;
3. Competitive factors markets:  $R_t = \alpha K_t^{\alpha-1} L_t^{1-\alpha}$  and  $w_t = (1-\alpha) K_t^\alpha L_t^{-\alpha}$ ;
4. Market clearing:  $K_t = \sum_{i=1}^I \lambda_i s_t^i$  and  $L_t = \sum_{i=1}^I \lambda_i h_t^i (1-l-n_t^i)$ ;
5. The allocation policy respects the government budget in each period (with  $w_0 = 1$  and  $L_0 = \sum_{i=1}^I \lambda_i h_0^i$ ):

$$G_t + g_t = \tau_{t-1} w_{t-1} L_{t-1} \quad (30)$$

When solving the extended model, observe that the solution procedure can be divided into two steps: the first step involves the schooling time allocation decision  $n_t^i$ , optimally chosen to maximize the wage income; and the second step involves the life cycle consumption decisions  $\{c_t^i, s_t^i, d_t^i, e_t^{b,i}, e_t^{a,i}\}$ , optimally chosen to maximize utility, take the wage income as given. The solution generated by this sequential maximization problem is equivalent to the solution given by solving the model simultaneously.

Taking the maximized wage income as given, the life cycle consumption decisions are given by:

$$c_t^i = \frac{R_{t+1}}{R_{t+1} + \beta + \gamma_1 + \gamma_2} (1 - \tau_t) w_t h_t^i (1 - l - n_t^i) \quad (31)$$

$$d_{t+1}^i = \frac{\beta R_{t+1}}{R_{t+1} + \beta + \gamma_1 + \gamma_2} (1 - \tau_t) w_t h_t^i (1 - l - n_t^i) \quad (32)$$

$$e_{t+1}^{b,i} = \frac{\gamma_1 R_{t+1}}{R_{t+1} + \beta + \gamma_1 + \gamma_2} (1 - \tau_t) w_t h_t^i (1 - l - n_t^i) \quad (33)$$

$$e_{t+1}^{a,i} = \frac{\gamma_2 R_{t+1}}{R_{t+1} + \beta + \gamma_1 + \gamma_2} (1 - \tau_t) w_t h_t^i (1 - l - n_t^i) \quad (34)$$

In the extended model, since private education expenditure in basic education  $e_t^{b,i}$  is an increasing function of parental human capital level  $h_{t-1}^i$  (shifting (33) 1 period backward), its interpretation is already subsumed in the basic education technology specified in the original model. Adding private education expenditure in advanced education  $e_t^{a,i}$  in the learning technology (shifting (34) 1 period backward), the first order condition on the optimal advanced schooling time is again given by:

$$-(h_t^{b,i} + h_t^{a,i}) + h_t^{a,i} (1 - l - n_t^i) / n_t^i \leq 0, \quad n_t^i \geq 0, \quad \text{with at least one equality} \quad (35)$$

In other words,

$$n_t^i = 0 \text{ if } h_t^{b,i} \leq \hat{h} \quad (36)$$

$$n_t^i = \max\left\{0, \frac{1-l}{2} - \frac{(h_t^{b,i})^{1-\kappa}}{2\tilde{A}(h_t^{b,i} - \hat{h})^\delta (g_t/N_t)^\theta}\right\} \text{ if } h_t^{b,i} > \hat{h} \quad (37)$$

Here  $\tilde{A}$  is a function of the technology parameter  $A$ , the preference parameters  $\beta$ ,  $\gamma_1$  and  $\gamma_2$ , exogenous policy parameters  $\tau_{t-1}$  and  $l$ , predetermined parental choice  $n_{t-1}^i$ , predetermined market wage rate  $w_{t-1}$ , and the current market rate of return on physical capital  $R_t$ : non of which is subjective to the influence of the generation  $t$  group  $i$  individual when he makes his schooling decision in advanced education and utilizes the private education expenditure prepared by his parent for the advanced stage. For the purpose of intra-generational comparison across individuals,  $\tilde{A}$  is only a scale parameter and can be normalized. Since solution (37) differs from solution (9) in the original model only by the exponents on  $h_t^{b,i}$ ,  $1 - \kappa$  versus 1, it is clear that Proposition 1, 2, 3 and 4 hold under both specifications.

Table 1 School Participation Rate and Survival Rate

	CS Years (1)	Enrolment Ratio (%)			Survival Rate (%)	
		Prim. (2)	Second. (3)	Tert. (4)	(5)= $\frac{(3)}{(2)}$	(6)= $\frac{(4)}{(2)}$
Bangladesh **	5	89	43	7	48	8
Brazil **	8	97	71	17	73	18
Chad *	6	57	8	1	14	2
Japan **	9	100	100	48	100	48
Lesotho **	7	78	21	3	27	4
Niger **	8	30	5	1	17	3
United Kingdom *	11	99	94	58	95	59
United States *	10	95	87	72	92	76

Countries with 1999/2000 data are indicated with \* and those with 2000/2001 data are indicated with \*\*.

Table 2 Allocation of Public Educational Expenditure

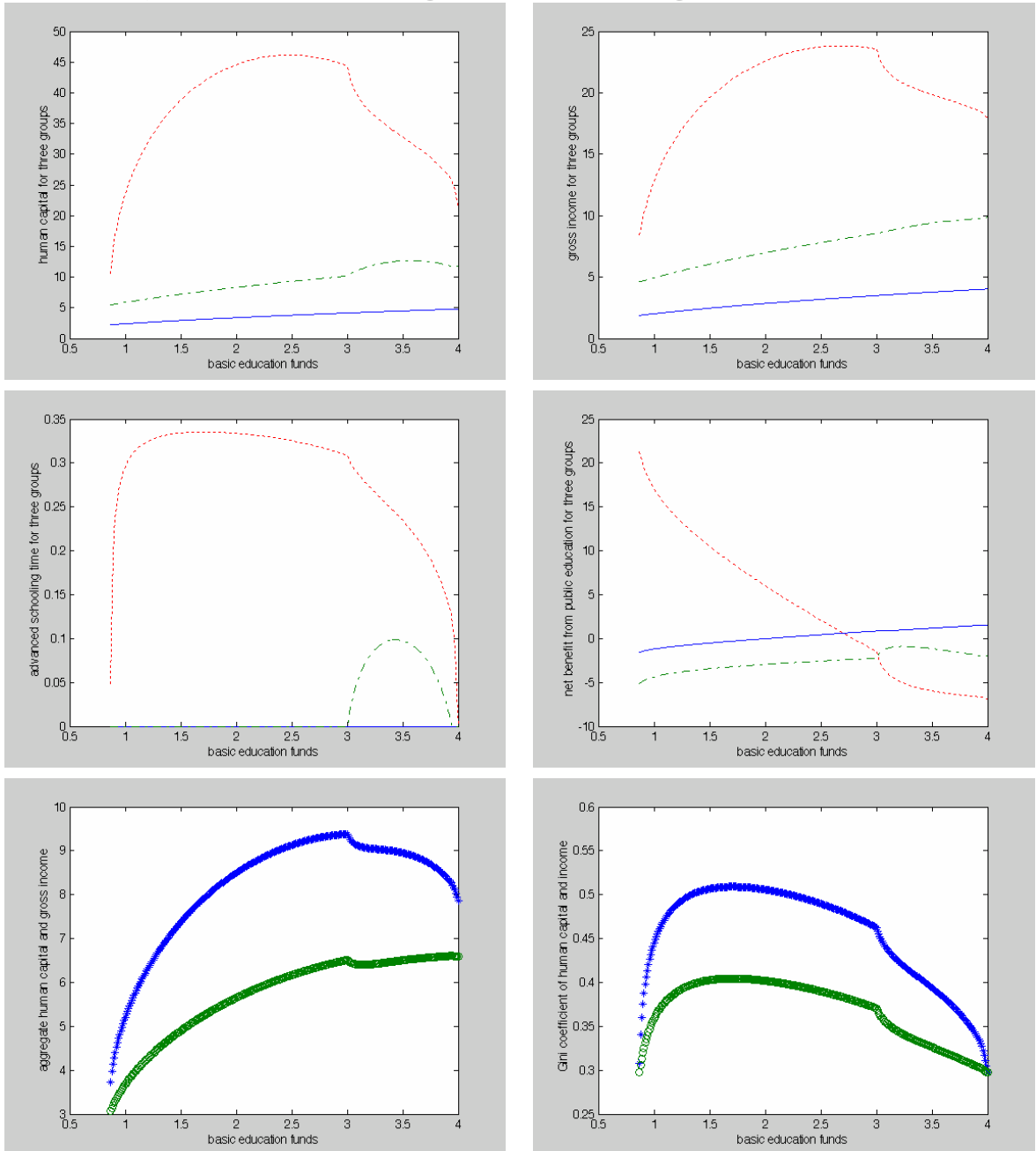
	% of Total Education Expenditure			Ratio of Exp. per Pupil	
	Prim. (7)	Second. (8)	Tert. (9)	(10)= $\frac{(8)/(3)}{(7)/(2)}$	(11)= $\frac{(9)/(4)}{(7)/(2)}$
Bangladesh **	38.1	43.0	10.1	2.3	3.4
Brazil **	34.0	35.7	21.1	1.4	3.5
Chad *	57.5	25.9	16.6	3.2	16.5
Japan **	35	40	15.2	1.1	0.9
Lesotho **	48.6	27.7	16.7	2.1	8.9
Niger **	49.3	28.6	19.9	3.5	12.1
United Kingdom *	24.4	46.7	20.1	2.0	1.4
United States *	31.4	35.6	26.2	1.2	1.1

Countries with 1999/2000 data are indicated with \* and those with 2000/2001 data are indicated with \*\*.

Table 3 Parameters used for Numerical Examples

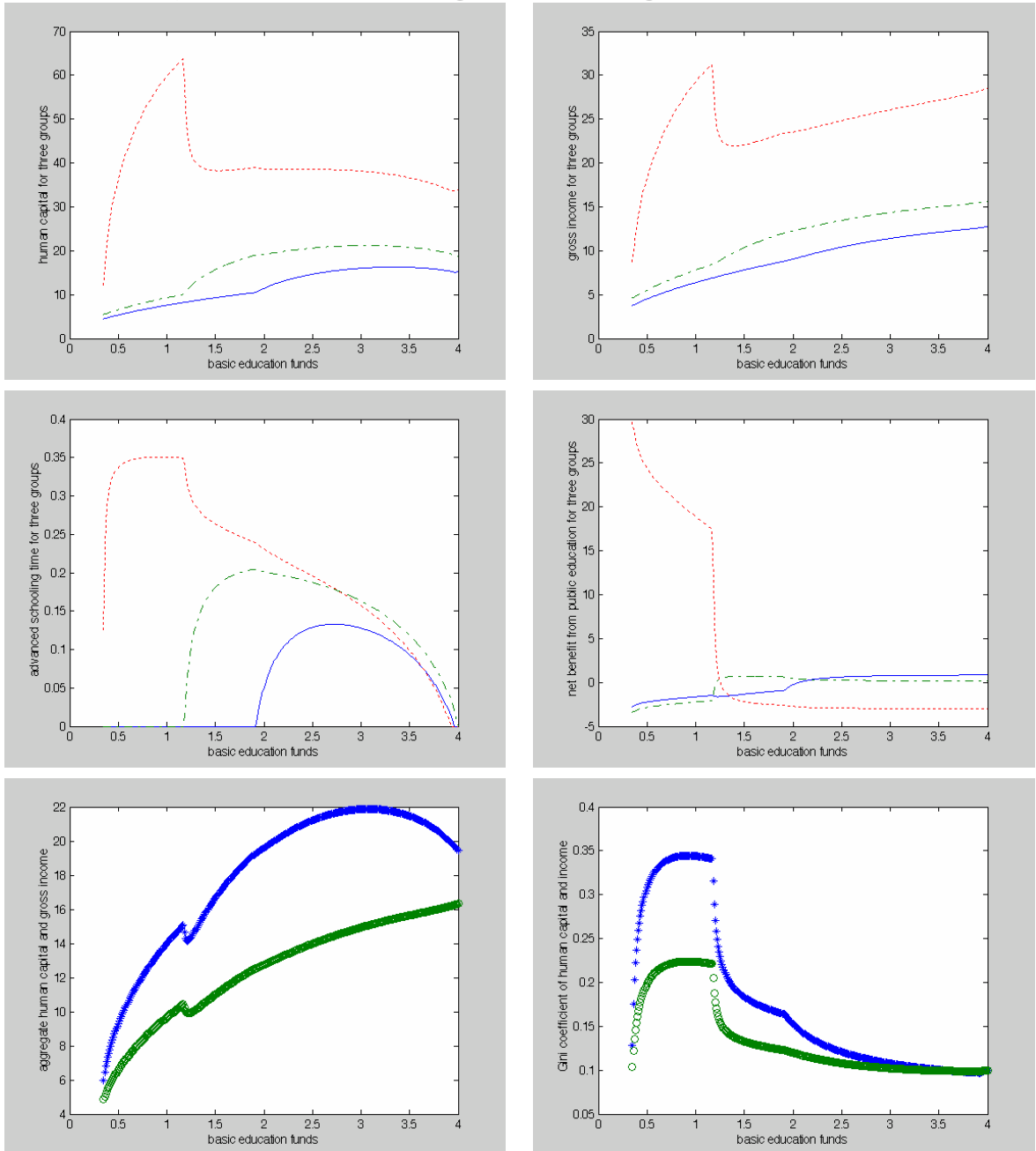
Basic learning technology coefficient	$B = 6$
Advanced learning technology coefficient	$A = 5$
Compulsory schooling duration	$l = 0.16$
Schooling quality elasticity	$\theta = 0.5$
Learning qualification elasticity	$\delta = 0.5$
Qualification threshold	$\hat{h} = 10$

Figure 1 Numerical Example 1 - a Less Developed Economy



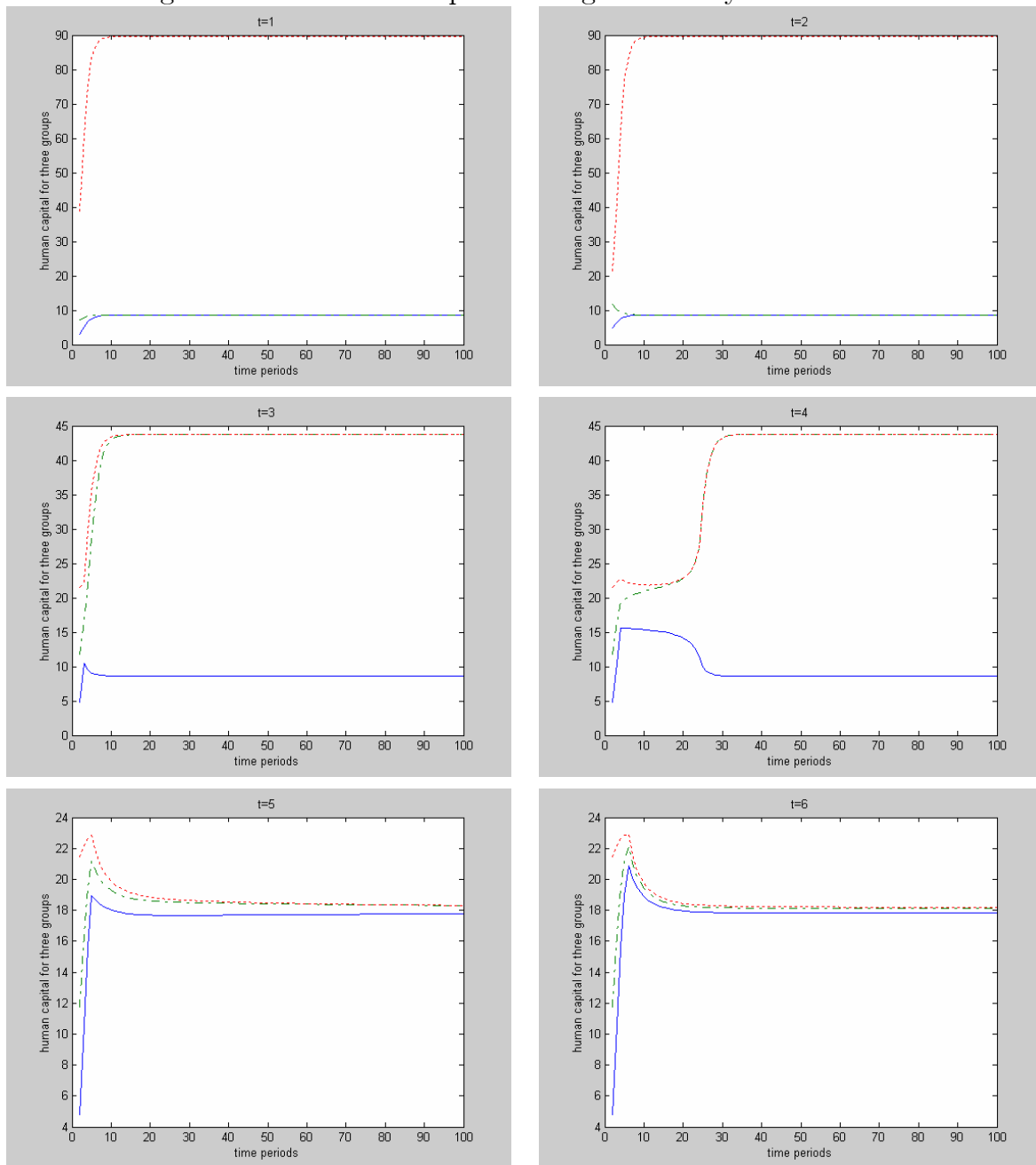
Note: solid line - group 1, dashed line - group 2, dotted line - group 3;  
 '\*' line - human capital, 'o' line - gross income.

Figure 2 Numerical Example 2 - a Developed Economy



Note: solid line - group 1, dashed line - group 2, dotted line - group 3;  
 '\*' line - human capital, 'o' line - gross income.

Figure 3 Numerical Example 3 - Long Run Policy Effects



Note: solid line - group 1, dashed line - group 2, dotted line - group 3.