Endogenous Determination of Public Budget Allocation Across Education Stages

Xuejuan Su*

March, 2005

Abstract

This paper studies the endogenous determination of public budget allocation across hierarchical education stages. In less developed economies, the top class has dominant political power to implement its most preferred policy, which is characterized by exclusive participation and good schooling quality at higher education at the expense of basic education. In developed economies, the budget allocation is more balanced; under certain parameters, it leads to expanded participation of the middle class in higher education. The model offers an explanation to the observed cross-country policy difference, and is broadly consistent with historical evidence.

JEL classification: I28; H42; D78

Keywords: Education policy, Multi-stage education, Public budget allocation

*Author address: Economics, Finance and Legal Studies, University of Alabama, Tuscaloosa, AL 35487-0224. Email address: xsu@cba.ua.edu. Telephone: (205)348-7841. I thank Mark Gradstein, Michael Kaganovich, Tilman Klumpp, Paul Pecorino, the editor, two anonymous referees and seminar participants at the Midwest Economic Theory Conference and the Public Choice Society Meeting for their helpful comments. All remaining errors are my own.

1. Introduction

Countries differ drastically in their public budget allocation policies across education stages. Consequently the relative schooling quality at different stages of education exhibits a wide range of variation. Reported in Tables 1a and 1b are some 1996 cross-country data and summary statistics (UNESCO '99 Statistical Yearbook). There countries are divided into three groups: African countries, other non-OECD countries and OECD countries. For instance, in Malawi, the public current expenditure per pupil as GNP per capita is 9% at the primary stage, 27% at the secondary stage, and 1,580% at the tertiary stage. So the relative schooling quality between secondary and primary education is a ratio of 3, while that between tertiary and primary education is a stunning ratio of 176. The pattern is even clearer from the mean values (not weighted by population) of the three groups, that relative schooling quality ratios are significantly higher for non-OECD countries (especially African countries) than OECD countries.[1]

This policy difference across countries is astonishing, especially so when one considers the nature of education. Education is hierarchically organized, and finishing lower stages successfully is a prerequisite for attending higher stages. Consequently participation rates drop when a cohort of individuals moves up along the education ladder. Reported in Table 2 are the summary statistics of the survival rates, calculated as the ratio of the gross enrolment ratios across education stages.[2] Again the pattern is obvious: the survival rates are much lower for non-OECD countries (especially African countries) than OECD countries.

[1] Public current expenditure includes that for goods and services consumed within the current year, such as staff salaries and books and teaching materials. So the observed pattern can hardly be attributed to economies of scale in providing higher education, which mainly concerns public capital expenditure on construction, renovation and major repairs of buildings.

[2] Gross enrollment ratio may exceed 100% since it is the number of pupils enrolled at a given level of education, regardless of age, expressed as a percentage of the population in the relevant official age-group. Hence the calculated survival rate may exceed 100%.

The hierarchical nature of education implies fundamental asymmetry between basic and higher education. Funding basic education has both the direct and the indirect benefits: basic education produces human capital in its own name, and this human capital also acts as an intermediate input and improves learning efficiency if higher education is ever attended. On the other hand, funding higher education has mainly direct but little indirect benefits: human capital output from higher education can be viewed as a terminal output, mainly useful in goods production but playing little role in facilitating future learning. Intuitive cost-benefit analysis dictates that an efficient budget allocation should tilt towards basic education.

This intuition remains valid even when individuals differ in their initial qualifications (young children's pre-school preparations). With a two-stage education model that directly captures the hierarchical nature, Su (2004) shows that the lower is the average qualification in an economy, as is the case in many less developed countries (LDCs) where a big fraction of the population is poorly prepared, the more benefit arises from funding basic instead of higher education. Allocation policies focusing on basic education improve both efficiency (measured by either aggregate output or aggregate human capital) and income equality. Yet "what is" seems to contradict "what should be". As from Tables 1a and 1b, while developed countries adopt more or less balanced public budget allocation policies, it is the LDCs that adopt policies strongly favoring higher education. It is then of interest to know how these socially suboptimal policies may arise and even perpetuate, and what are their macroeconomic effects.

The social suboptimality of LDCs education policies are further corroborated by strong empirical evidence. A recent survey by Psacharopoulos and Patrinos (2004) (and the references therein) reports that while the returns to education are roughly the same across primary, secondary and tertiary stages for OECD countries (8.5, 9.4 and 8.5 percentage respectively), they are significantly higher at the primary stage, moderately higher at the secondary stage than the tertiary stage for non-OECD countries (for example, 25.4, 18.4 and 11.3 for sub-Sahara Africa and 17.4, 12.9 and 12.3 for Latin America/Caribbean). The failure to equate marginal returns across education stages suggests a socially inefficient public budget allocation policy. So the question of interest is, why in LDCs is the funding for basic education so low relative to the funding for higher education?

Following Su (2004), this paper develops a two-stage education model to study the endogenous determination of the public budget allocation policy across hierarchical education stages. The two-stage technology explicitly models the hierarchical nature of education, and proves extremely useful in analyzing stage-specific education policies. An individual's initial qualification (pre-school preparation) is determined by his family background such as wealth, income and/or parental human capital level. Through basic education, the initial qualification is augmented into qualification/preparation for higher education, and better qualification means higher learning efficiency and better potential to benefit. Consequently, individuals with different family background may have different optimal participation decisions and hence different policy preferences. The bottom class has extremely low initial qualification and never attends higher education, so it prefers to allocate all resources to basic education for maximum benefit. On the other hand, the top class has high initial qualification and attends higher education, so it prefers a balanced budget allocation across the two education stages.

Since public education is essentially a publicly provided private good, there is a congestion effect. With fixed budget allocation, schooling quality is inversely related to the enrollment at a given stage. Consequently, the top class' policy preference also depends on other groups' participation decisions, especially that of the middle class. The middle class may opt into higher education when finding it attractive. Yet the entry of the middle class lowers the schooling quality and makes it less attractive to the top class. Thus the top class adjusts its policy preference accordingly. When the middle class has relatively low initial qualification as in many LDCs, the top class benefits more from cutting funding for basic education and thus disqualifying the middle class for higher education. This prevents the congestion effect and guarantees exclusive benefit for the top class in higher education. On the other hand, when the middle class has relatively high initial qualification, the top class is better off not to attempt the exclusion. This is simply because the cost to disqualify the middle class would be too high for the top class to bear, whose qualification (learning efficiency) in higher education would also be adversely affected if the schooling quality at basic education were poor enough to disqualify the middle class. Overall the top class' policy preference depends on the distribution of initial qualifications (or equivalently, wealth, income and/or human capital of the parent generation) in the population, and hence the development stage of an economy.

Individuals' policy preferences endogenously determine the actual public budget allocation policy. An essential assumption here is that an individual's political power may depend on his economic power, so that the top class may enjoy disproportionate political leverage in policy making. This is most likely in LDCs, where the combination of dominant economic power and less democracy gives the top class dominant political power. Then the actual budget allocation is the one most preferred by the top class, - strongly favoring higher education at the cost of basic education, and exclusive participation of the top class in higher education. On the other hand, in developed economies, the top class has less incentive and less means to exclude the middle class from higher education, so the actual budget allocation is more balanced and leads to expanded participation. Allowing asymmetric political power across the population, this model offers an explanation to the observed pattern in the data.

The rest of the paper is structured as follows. Section 2 discusses the relationship of this paper to the existing literature on education policies and education systems. Section 3 lays out the two-stage education model, and analyzes individuals' optimal participation decisions under a given allocation policy. Section 4 then characterizes the top class' policy preference over the entire range of feasible allocation policies. Section 5 discusses the robustness of main results with regard to several assumptions and draws the conclusion.

2. Related literature

The current paper, in general, is related to the literature on the political economy of education policies. On the issue of education finance, Glomm and Ravikumar (1992), Fernandez and Rogerson (1995), Epple and Romano (1996), Gradstein and Justman (1997) and Kaganovich and Zilcha (1999) focus on public versus private provision of education; Benabou (1996) and Fernandez and Rogerson (1997) focus on local versus state funding for education. This paper takes a centralized public education system as given, and studies the budget competition among hierarchically organized education stages.

Another strand of the literature studies the differential incidence of public education, namely rich individuals may benefit more from public education than poor individuals. 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) and Blanden and Machin (2004) documents similar pattern in the United Kingdom. Recently, Gradstein (2004a, 2004b) analyzes this issue in a rent-seeking setup, where rich individuals, by investing more in rent-seeking, secures a bigger fraction of benefit from public budget. Su (2004) analyzes this in a hierarchical education framework, where rich individuals, with their high qualifications, benefit more from publicly financed higher education due to exclusive participation. That study treats the budget allocation policy as exogenously given. This paper takes a step forward and studies the endogenous determination of the budget allocation policy, which offers an explanation to the puzzling pattern observed in cross-country policy data.

Also closely related is the growing literature on the evolution of education systems. Bourguignon and Verdier (2000) shows that when the positive externality of education is large enough, the self-interested top class may prefer to subsidize the education of the lower class, who cannot afford education on its own. Galor and Moav (2001) is based on the complementarity between physical and human capital in production. It shows that with physical capital accumulation, the top class may prefer to support public education for the lower class to sustain their profit rates. Galor and Moav (2004) is based on credit constraint on human capital accumulation, so as an economy develops, it undergoes the transition from physical capital accumulation to selective human capital accumulation to universal human capital accumulation. In these studies, education is modeled as a one-step process without hierarchical stages. Brunello and Giannini (2004) compares the efficiency of two school designs, the stratified system and the comprehensive system. It focuses on the distinction between vocational training and general education. Bertocchi and Spagat (2004) also distinguishes between general education and vocational training, and focuses on the social role of access to general education. Their vocational training and general education are modeled as two distinct and parallel tracks. The current paper complements the existing literature by adding the hierarchical structure to education, and focuses on public budget competition across hierarchical education stages. The evolution of the education system - the initial emergence of higher education and the later expanded participation of the middle class into higher education - is intrinsic to the hierarchical nature of education, and does not rely on technological advancement.

3. Two-stage education

This is a successive generations model with heterogeneous individuals. Within a generation (indexed by subscript t), individuals differ in their initial qualifications, which is assumed to be an increasing function of their parental human capital levels.[3] With little loss of generality, it is assumed that there are 3 groups of individuals with measures λ_i, $i = 1, 2, 3$. 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 $h_0^1 < h_0^2 < h_0^3$. For simplicity, it is assumed that the lifetime income of a generation 0 individual is the same as his human capital level, i.e., $y_0^i = h_0^i$, $i = 1, 2, 3$. Each individual lives for three periods, and is replaced by one offspring at the end of his life.

Individuals accumulate human capital through tax-financed public education.[4] The education system is hierarchical and consists of two stages. The lower stage is basic education, which occurs in the first period. The human capital output from basic education determines an individual's qualification for higher education in the second period. Unqualified individuals cannot benefit from higher education, so they forego that stage and work for the second and third periods. Qualified individuals weigh the benefit against the cost, and make their own participation decisions. If they decide to attend higher education in the second period, they work for the third period only.

[3] The initial qualification can be viewed as a child's pre-school preparation, which results from early childhood development. That a child's initial qualification depends on the parental human capital level is a commonly adopted specification in the literature, see Glomm and Ravikumar (1992), Eckstein and Zilcha (1994), Benabou (1996) and Kaganovich and Zilcha (1999). For more discussions, see Section 5.

[4] I deliberately abstract from private education expenditure to focus on the effect of public budget allocation. Surely 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 versus private school choice and/or education vouchers (Glomm and Ravikumar (1992), Epple and Romano (1998), Nechyba (1999), Caucutt (2002), Ladd(2002)). However, within the current framework of hierarchical education, adding private education expenditure renders the model analytically insolvable. The necessary simplification is dictated by the focus of the present study. There are more discussions on this issue in Section 5.

In this model no individual chooses not to attend basic education, since it is necessary for any human capital accumulation. Without basic education, an individual has no human capital and hence earns zero lifetime income, which obviously cannot be optimal. In the first period of life, an individual acquires human capital $h_t^{b,i}$ from basic education according to the following technology:

$$h_t^{b,i} = Bh_{t-1}^i G_t^b \tag{1}$$

This technology is fairly standard. There are two raw inputs: the initial qualification captured by the term h_{t-1}^i, and the schooling quality in basic education G_t^b. Here G_t^b is the aggregate public funds allocated to basic education. Since the population is normalized to 1 and all individuals attend basic education, G_t^b is also the average expenditure per student, i.e., the schooling quality. As can be seen, the distribution of $h_t^{b,i}$ is a linear transformation of the distribution of h_{t-1}^i. Namely, low initial qualification implies low qualification for higher education, and high initial qualification implies high qualification for higher education.[5]

At the beginning of the second period, an individual chooses the probability p_t^i with which he attends higher education. If he does attend, higher education produces human capital according to the following technology:

$$h_t^{a,i} = \begin{cases} 0 & \text{if } h_t^{b,i} \leqslant \widehat{h} \\ A(h_t^{b,i} - \widehat{h})G_t^a / P_t & \text{if } h_t^{b,i} > \widehat{h} \end{cases} \tag{2}$$

Again there are two inputs. The intermediate input is the individual's qualification for higher education, which depends on his performance in basic education, $h_t^{b,i}$,[6] with the threshold $\widehat{h}$ capturing the hierarchical nature: Unqualified individ-

[5]It is purely for algebraic simplicity that the technology exhibits constant instead of diminishing marginal returns on either input. A more general model that allows diminishing marginal returns is analyzed in Appendix B, and is shown that main results remain robust.

[6]This assumption follows directly from Lucas (1988) that existing human capital improves the learning efficiency in subsequent education. Su (2004) adopts a similar specification in a discrete stage education model.

uals simply cannot benefit from higher education. Evidence in support of this qualification threshold is almost self-evident. For example, it would be a waste of time for a student to sit in a class on dynamic programming, if he barely knows any algebra; or equally meaningless would it be for a student to attend a lecture on English literature, when his vocabulary is meager. It is worth pointing out that this threshold is intrinsic to the learning technology and not a policy choice.[7]

The other input is the schooling quality G_t^a/P_t. Here G_t^a is the aggregate public funds allocated to higher education, P_t is the fraction of population that attends higher education, so G_t^a/P_t measures the schooling quality in higher education.

After education, an individual joins the labor force and works. Labor is inelastically supplied, so the flat income tax rate τ_t used to finance public education for the next generation is not distortionary. If an individual has only basic education, he has two periods of working time. The discount factor is assumed to be 1 without loss of generality, and the lifetime income is given by:

$$y_t^{b,i} = 2h_t^{b,i} \tag{3}$$

If an individual attends higher education, he has only one period working time, and the lifetime income is given by:

$$y_t^{a,i} = (h_t^{b,i} + h_t^{a,i}) \tag{4}$$

Individuals choose their higher education participation decisions p_t^i optimally to maximize their lifetime income. The optimality dictates that $p_t^{i*} = 1$ iff $y_t^{a,i} > y_t^{b,i}$; $p_t^{i*} = 0$ iff $y_t^{a,i} < y_t^{b,i}$; and $p_t^{i*} \in [0,1]$ iff $y_t^{a,i} = y_t^{b,i}$. The corresponding maximum lifetime income is defined by $y_t^i = p_t^{i*} y_t^{a,i} + (1 - p_t^{i*})\, y_t^{b,i}$.

[7] 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 exist, the latter has to be higher than the former and hence more exclusive. More discussions on this issue are provided in Section 5.

A feasible allocation policy is one that satisfies government budget balance $G_t^b + G_t^a = R_t$, where R_t is the tax revenue given by $\tau \sum_{i=1}^{3} \lambda_i y_{t-1}^i$. Taking a feasible allocation policy G_t^b and G_t^a as given, an equilibrium consists of individual participation decisions $\{p_t^{i*}\}_{i=1}^3$, and the fraction of total population in higher education P_t, such that:

1. *Given G_t^b, G_t^a and P_t, $\{p_t^{i*}\}_{i=1}^3$ are individuals' optimal choices;*
2. *The consistency condition is satisfied: $P_t = \sum_{i=1}^{3} \lambda_i p_t^{i*}$.*

Proposition 1. *For generation t, given the distribution $h_{t-1}^1 < h_{t-1}^2 < h_{t-1}^3$, if the allocation policy G_t^b and G_t^a is such that $Bh_{t-1}^3 G_t^b > \widehat{h}$ and $G_t^a > 0$, then there exists a unique equilibrium with $P_t > 0$.*

This proposition establishes the existence and uniqueness of the equilibrium under a given allocation policy. The result is robust so long as there is negative externality in the education technology. In the public education system, the total funding for higher education is fixed, so there is negative externality due to congestion. The more individuals attend higher education, the lower is the schooling quality measured as expenditure per student, and the lower is the benefit from attending higher education.

Proposition 2. *With the equilibrium P_t, there exist $\widetilde{h} = \widehat{h}/(1 - \frac{P_t}{AG_t^a})$ such that $p_t^{i*} = 1$ iff $h_t^{b,i} > \widetilde{h}$; $p_t^{i*} = 0$ iff $h_t^{b,i} < \widetilde{h}$; and $p_t^{i*} \in [0,1]$ iff $h_t^{b,i} = \widetilde{h}$.*

This proposition characterizes how individuals base their participation decisions on their qualifications for higher education (and hence initial qualifications), taking the allocation policy as given. Given the hierarchical nature of education, after basic education, individuals endogenously stratify into two cohorts. Those with low qualifications ($h_t^{b,i} < \widetilde{h}$) forego higher education,[8] and those with

[8]Notice that $\widetilde{h} > \widehat{h}$, so that even some qualified individuals may choose to forego higher education. This is because their benefits from attending higher education are small compared to their opportunity costs of foregone wage income in the second period.

high qualifications ($h_t^{b,i} \geq \widetilde{h}$) attend higher education. The perfect stratification guarantees that the initial social ranking is preserved overtime. So in this model without any uncertainty, there is no social mobility.[9]

In this centralized public education system, any attending student enjoys the same schooling quality, so the share of public resource that contributes to his education depends only on his participation decision. All individuals attend basic education, so each enjoys his fair share of G_t^b; yet only individuals with high qualifications attend higher education but not those with low qualifications, so G_t^a only benefits the former but not the latter. Individuals' policy preferences thus depend on $\widetilde{h}$ and hence P_t, the participation decisions of the entire population.

4. Top class' policy preference

An individual's policy preference is given by his value function (the maximum lifetime income) over all feasible allocation policies. For generation t, taking the public budget for education R_t as given, feasible allocation policies are G_t^b and G_t^a such that $G_t^b + G_t^a = R_t$. It is assumed that $R_t > \frac{\widehat{h}}{Bh_{t-1}^3} \equiv c_3^l$ (the reason for this notation will be clear in the coming proposition), where budget allocation is a non-trivial issue. If $R_t \leq c_3^l$, no individual could ever qualify for higher education, and any resource allocated to higher education would be a waste. When $R_t > c_3^l$, the policy space $G_t^b \in [0, R_t]$ can be divided into nested subintervals according to different groups' participation decisions in higher education.

Proposition 3. *For given public budget R_t, there exists a minimal group i such that $p_t^{i*} > 0$ for some $G_t^b \in [0, R_t]$. More specifically, there exist $0 < \frac{\widehat{h}}{Bh_{t-1}^3} = c_3^l < b_3^l < ... < c_i^l < (b_i^l \leq b_i^r) < c_i^r < ... < b_3^r < c_3^r = R_t$, such that for any $j < i$, $p_t^{j*} = 0$ for all $G_t^b \in [0, R_t]$; for any $j \geqslant i$, $p_t^{j*} = 0$ when $G_t^b \in [0, R_t] \backslash (c_j^l, c_j^r)$; $p_t^{j*} \in (0,1)$ when $G_t^b \in (c_j^l, c_j^r) \backslash [b_j^l, b_j^r]$; and $p_t^{j*} = 1$ when $G_t^b \in [b_j^l, b_j^r]$. Overall p_t^{j*} changes*

[9] With regard to the effect of uncertainty on social mobility, see the discussions in Section 5.

continuously with G_t^b.

This Proposition divides the entire policy space into nested subintervals. The overall pattern is that a certain group in more likely to attend higher education when the allocation policy falls in the mid-range and balances funding basic and higher education. This mid-range of policies is group-specific, shrinks with low initial qualifications, and may be empty for those with extremely low initial qualifications. When the allocation policy changes from extremely unbalanced to balanced, the top class is always the first to attend higher education, followed sequentially by the middle class and then possibly the bottom class.

A direct implication of this proposition is that we should never observe a policy with $G_t^b \leqslant c_I^l$, since it is Pareto dominated. When $G_t^b \leqslant c_I^l$, all individuals attend basic education only, and they would strictly prefer the policy $G_t^b = R_t$.[10] So there is consensus among all individuals that basic education should be funded at least to a reasonable level, simply due to the hierarchical nature of education.

On the other hand, not all individuals may agree that higher education should be funded at least to a reasonable level. Or quite contrary, it is possible that no individuals would like to fund higher education. When the public budget R_t is so low that there exists no b_3^l or b_3^r, partial participation of the top class in higher education implies that the lifetime income with higher education is always the same as that without, with the latter linearly increasing in G_t^b. So all individuals prefer the policy $G_t^b = R_t$. It is easy to see that higher education is desirable (at least to some individuals in the economy) only when R_t is sufficiently high.

The asymmetry in individual preference for funding basic and higher education arises solely from the education hierarchy. As the foundation, basic education is always necessary and desirable to all individuals. Yet higher education only becomes desirable when there is sufficiently high level of education resources

[10]Notice that this lower bound on funding basic education is inversely related to the initial qualification of the top class.

available. Crudely speaking, higher education is a luxury good for a society. Only when the economy develops to a certain degree, will higher education become socially feasible. From now on it is assumed that R_t is sufficiently high so that higher education is desirable to the top class. It is also assumed that R_t is not too high to induce the bottom class to attend higher education.

Holding the top class' status h^3_{t-1} fixed, there are three possibilities associated with the middle class' status: (1) h^2_{t-1} is low and there exists no c^l_2 or c^r_2; (2) h^2_{t-1} is moderate so that there exist c^l_2 and c^r_2 but not b^l_2 and b^r_2; and (3) h^2_{t-1} is high so that there exist b^l_2 and b^r_2. As can be seen soon, the top class' policy preference depends critically on the participation decision of the middle class, and hence its initial qualification level.

Proposition 4. *When R_t is sufficiently high and h^2_{t-1} is low, the top class' most preferred allocation policy is $G^3 = \frac{R_t}{2} + \frac{\widehat{h}}{2Bh^3_{t-1}} + \frac{\lambda_3}{2A}$, with $\frac{\partial G^3}{\partial \widehat{h}} > 0$, $\frac{\partial G^3}{\partial h^3_{t-1}} < 0$, and $\frac{\partial G^3}{\partial \lambda_3} > 0$.*

With low initial qualification, the middle class never attends higher education and is non-distinguishable from the bottom class. The most preferred allocation policy of the top class then has three terms. The first term $\frac{R_t}{2}$ is just the familiar notion of balanced allocation to equalize marginal returns across education stages. The second term $\frac{\widehat{h}}{2Bh^3_{t-1}}$ adjusts for the qualification threshold in higher education. The higher is the qualification threshold, the better is the schooling quality needed at basic education to meet it, so G^3 increases with $\widehat{h}$. On the other hand, for a given threshold, the better is the top class' initial qualification, the lower is the schooling quality needed, so G^3 decreases with h^3_{t-1}. The third term $\frac{\lambda_3}{2A}$ adjusts schooling quality for the participation rate in higher education. Here λ_3 measures the exclusiveness of higher education, compared to the uniform participation in basic education. For any given $G^a_t = R_t - G^b_t$, the smaller is λ_3, the better is the schooling quality in higher education, and hence the more attractive is higher education to the top class. Hence G^a_t decreases with λ_3, or equivalently

G^3 increases in λ_3. In contrast to Bertocchi and Spagat (2004), in this model the top class does not value the exclusive access to higher education per se; it values the exclusiveness only because of the congestion effect in higher education - the inverse relationship between participation and schooling quality.

Proposition 5. *When R_t is sufficiently high and h_{t-1}^2 is moderate, the top class' most preferred allocation policy is G^3 if $G^3 \notin (c_2^l, c_2^r)$; c_2^l if $G^3 \in (c_2^l, c_2^r)$ and $\frac{\widehat{h}}{2Bh_{t-1}^2} - \frac{\widehat{h}}{2Bh_{t-1}^3} > \frac{\lambda_3}{A}$; and c_2^r if $G^3 \in (c_2^l, c_2^r)$ and $\frac{\widehat{h}}{2Bh_{t-1}^2} - \frac{\widehat{h}}{2Bh_{t-1}^3} < \frac{\lambda_3}{A}$.*

With moderate initial qualification, the middle class may attend higher education. Its entry lowers the schooling quality and hence the top class' benefit from higher education. If this entry effect is confined, namely $G^3 \notin (c_2^l, c_2^r)$, it affects the top class' policy preference in a small subinterval but not the most preferred policy. If this entry effect is broad enough, it affects the top class' most preferred policy. More specifically, when the initial qualification of the middle class is moderately low, i.e., $\frac{\widehat{h}}{2Bh_{t-1}^2} - \frac{\widehat{h}}{2Bh_{t-1}^3} > \frac{\lambda_3}{A}$, the top class prefers to cut funding on basic education to disqualify the middle class for higher education.[11] Of course the top class also suffers from poor schooling quality at basic education, but this cost can be more than compensated by the ensured exclusive benefit at higher education. On the other hand, when the initial qualification of the middle class is moderately high, i.e., $\frac{\widehat{h}}{2Bh_{t-1}^2} - \frac{\widehat{h}}{2Bh_{t-1}^3} < \frac{\lambda_3}{A}$, the cost to disqualify the middle class by cutting funding for basic education is too big for the top class to bear, so the top class prefers alternatively to cut funding for higher education and make it unattractive to the middle class.

Notice that this proposition implies a critical level for h_{t-1}^2. When the initial qualification of the middle class is slightly below this critical level, the top class prefers a policy that focuses more on higher education. When the initial qualification of the middle class is slightly above this critical level, the top class prefers

[11] For other mechanisms of exclusion, see the discussions in Section 5.

a policy that focuses more on basic education. In either case, if the top class gets its way, the current middle class will not attend higher education. But through the inter-generational link, these two cases have quite different long term consequences. In the former, poor schooling quality in basic education means lack of development opportunities for the middle and bottom classes, so their descendants may get stuck in the underdevelopment trap. In the latter, good schooling quality in basic education means good development opportunities for the middle and bottom classes, so their descendants are more likely to move above the threshold and attend higher education. With this threshold effect, two similar economies may follow drastically distinct development paths.[12]

Proposition 6. *When R_t is sufficiently high and h^2_{t-1} is high, the top class' most preferred allocation policy is $\widetilde{G}^3 = \frac{R_t}{2} + \frac{\widehat{h}}{2Bh^3_{t-1}} + \frac{\lambda_3+\lambda_2}{2A}$.*

With high initial qualification, the middle class attends higher education. The top class finds its exclusion too costly, either by cutting funding for basic education to disqualify the middle class, or by cutting funding for higher education to make it less attractive to the middle class. So the top class again prefers a balanced allocation policy, now with expanded participation at higher education.

Propositions 4, 5 and 6 together demonstrates how the top class' most preferred policy changes with the initial qualification of the middle class. The change is non-monotonic and possibly discontinuous. When the initial qualification of the middle class is low, the top class prefers a balanced allocation policy, since the congestion effect at higher education is absent. A small increase in h^2_{t-1} leads the top class to prefer cutting funding on basic education to disqualify the middle class for higher education. A bigger increase in h^2_{t-1}, however, leads the top class to increase funding for basic education, and the middle class is either better prepared for higher education in the future, or attends higher education in the current

[12] See Azariadis and Drazen (1990) for their discussion on the threshold effects.

period. Intended to explain cross-country data pattern, this model is also broadly consistent with historical evidence. As first in high school, and currently in higher education, the expanded participation of the middle class occurs swiftly.[13]

5. Discussions and conclusion

The above analysis focuses on the top class' policy preference. There is reason to believe that in many LDCs, the top class enjoys disproportionate political leverage in policy making, and the actual policy is (at least close to) the one most preferred by the top class. Empirical evidence suggests that an individual's political power may depend on his economic power, and Brad et al. (1995) argues that even within fully democratic political systems, political participation can be de facto denied to the bottom class in the population.[14] The top class is most likely to have dominant political power when it also has dominant economic power and the society is less democratic, as in many LDCs.

It is essential in this model that the top class enjoys dominant political power in many LDCs. Otherwise it is impossible that the endogenously determined budget allocation policy should be so adverse to the middle and bottom classes, which is the dominant majority of the population. Only with dominant political power, either by law or de facto, can the top class implement its most preferred policy. In many LDCs where the initial qualification of the middle class is relatively low, the most preferred policy strongly favors higher education at the cost of basic education, and ensures exclusive participation of the top class in higher education.

---

[13] As examples, see Goldin and Katz (1998) for the "high school movement" in the US between 1910 and 1940, and Blanden and Machin (2004) for the expansion of higher education in the UK since 1960 till now.

[14] Engerman and Sokoloff (2001) documents widespread of wealth and/or education requirement for suffrage till the beginning of the twentieth century both in the Old World and the New World. For the interaction of such franchise requirement and the education system, see Saint-Paul and Verdier (1993), Gradstein and Justman (1997) and Bertocchi and Spagat (2004).

The lack of development opportunities for the middle and bottom classes may keep them in the underdevelopment trap, which reinforces the economic and political dominance of the top class. On the other hand, in many developed countries where the initial qualification of the middle class is relatively high, the top class has less incentive and less means to exclude the middle class from higher education. So the actual policy, regardless of whether it is most preferred by the top class or some compromise outcome among the classes, is more balanced, and leads to expanded participation of the middle class in higher education. Thus the model offers an explanation to the puzzling pattern observed in cross-country policy data.

In the analysis of endogenous policy determination, it is assumed that the parent generation takes the size of public budget R_t as given, and chooses only its composition across education stages. The income tax rate τ acts only passively to balance the government budget. This assumption allows us to focuses on the intra-generational conflict of interests across different groups, and simplifies from the confounding inter-generational conflict of interests where parental altruism has to be explicitly modeled as relative weights on self and child. Numerical analysis in Appendix C suggests that various levels of R_t (corresponding to various levels of τ) has mainly quantitative but little qualitative impact on the main results.

Besides the political side story, this model also hinges on the hierarchical education technology. The remaining part of this section discusses the robustness of the main results with regard to several simplifying assumptions on the education technology.

In the two-stage education, one assumption is that there is no private education expenditure - either in the form of private versus public school choice, or as a direct complement to public education expenditure. These choices may be important, especially in LDCs for the current issue analysis. After all it is hard to imagine that the top class families send their kids to the same poor quality public school as the bottom class families in basic education. Yet if private education expenditure is allowed in the model, and if rich families invest proportionately more in their children's education than poor families, then it will magnify the qualification difference after basic education across groups. Namely the top class children will be even more advantaged compared to their poorer counterparts, and they are more likely to prefer an allocation policy favoring higher education. Thus our assumption of no private education expenditure is on the conservative side, and requires more stringent conditions for such policies favoring higher education to arise endogenously.

A related issue is on the effect of possibly random ability. There is no uncertainty in this paper, and a child's initial qualification is directly linked to his parent's human capital level. If ability is randomly assigned, social immobility as in the current model may no longer hold. The imperfect correlation across generations implies that there is always a chance that a top class parent has a child with low qualification. So either for insurance ex ante, or for optimum ex post, the introduction of random ability will have dampening effect on the main result. In the extreme case when ability is independently distributed from parental income/education, when ability alone determines children's initial qualifications (only nature but not nurture), and parents choose the policy before the uncertainty is resolved, all individuals would prefer the same allocation policy ex ante, and there will be no conflict of interests across groups.

In this model, there is no exclusion mechanism in higher education other than that imposed by the technology. Anyone that is qualified and wants to attend higher education can get it for free. One may argue that in reality, there are extra policy-chosen requirements. For example, admission to higher education may be based on entrance exam scores, which can be set above the technology-intrinsic threshold level.[15] Another possible exclusion mechanism is through tuition, so that if credit constrained, qualified poor individuals may not be able to afford higher

[15] If set below, this policy-chosen requirement is not binding, since no individuals who pass the exam without meeting the technology-intrinsic threshold level will choose to attend higher education, where their benefits are zero.

education. Yet another more subtle exclusion mechanism is through different academic tracks, so that only individuals on certain track can choose to attend higher education, while others are automatically excluded from this opportunity. Adding extra exclusion mechanism will give the top class more manipulative power to prevent the congestion effect in higher education. In this model since there is non-monotonicity in the top class' most preferred policy, adding extra exclusion mechanism will also have non-monotonic effects. When the middle class is of relatively low qualification, the extra exclusion mechanism saves the top class from resorting to lower basic education quality to exclude the middle class, then the policy will favor higher education less. When the middle class is of relatively high qualification, the extra exclusion mechanism helps the top class to keep the middle class out, then the policy will favor higher education more. Overall it will have quantitative but not qualitative impact on the endogenous determination of allocation policies favoring higher education.

Without denying the importance of technological advancement in the evolution of education systems, this paper draws attention to the hierarchical nature of education systems. Budget competition between hierarchically organized education stages is an important issue and has profound macroeconomic impact.

References

Azariadis, C. and Drazen, A. 1990, “Threshold Externalities in Economic Development”, Quarterly journal of Economics, 105(2), pp. 501-526.

Benabou, R. 1996, "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance." American Economic Review, 86(3), pp. 584-609.

Bertocchi, G. and Spagat, M. 2004, "The Evolution of Modern Educational Systems: Technical vs. General Education, Distributional Conflict, and Growth." Journal of Development Economics, 73(2), 559-582.

Betts, J. R. 1998, “The Impact of Educational Standards on the Level and Distribution of Earnings”, American Economic Review, 88(1), pp. 266-275.

Bishop, J. 1977, “The Effect of Public Policies on the Demand for Higher Education”, Journal of Human Resource, 12(3), pp. 285-307.

Blanden, Jo and Machin, Stephen. 2004, "Educational Inequality and he Expansion of UK Higher Education." Scottish Journal of Political Economy, 51(2), pp. 230-249.

Bourguignon, F. and Verdier, T. 2000, "Oligarchy, Democracy, Inequality and Growth." Journal of Development Economics, 62(2), pp. 285-313

Brady, H. E; Verba, S. and Schlozman, K. L. 1995, "Beyond SES: A Resource Model of Political Participation," American Political Science Review, 89(2), pp. 271-94.

Brunello, G. and Giannini, M. 2004, "Stratified or Comprehensive? The Economic Efficiency of School Design." Scottish Journal of Political Economy, 51(2), pp. 173-193.

Caucutt, E. M. 2002, "Educational Vouchers When There Are Peer Group Effects - Size Matters". International Economic Review, 43(1), pp. 195-222.

Costrell, R. 1994, “A Simple Model of Educational Standards”, American Economic Review, 84(4), pp. 956-971.

Eckstein, Z. and Zilcha, I. 1994, “The Effects of Compulsory Schooling on Growth, Income Distribution and Welfare”, Journal of Public Economics, 54(3), pp. 339-359.

Engerman, S. L. and Sokoloff, K. L. 2001, "The Evolution of Suffrage Institutions in the New World," NBER Working Papers 8512.

Epple, D. and Romano, R. E. 1996, "Ends against the Middle, Determining Public Service Provision When There Are Private Alternatives." Journal of Public Economics, 62(3), pp. 297-325.

_________. 1998, "Competition between Private and Public Schools, Vouchers, and Peer Group Effects." American Economic Review, 88(1), pp. 33-62.

Fernandez, R. and Rogerson, R. 1995, "On the Political Economy of Education Subsidies." Review of Economic Studies, 62(2), pp. 249-262.

_________. 1997, "Keeping People Out: Income Distribution, Zoning, and the Quality of Public Education." International Economic Review, 38(1), pp. 23-42.

Galor, O. and Moav, O. 2001, "Das Human Kapital." CEPR discussion paper 2701.

_________. 2004, "From Physical to Human Capital Accumulation: Inequality and the Process of Development." Review of Economic Studies, 71, pp. 1001-26.

Glomm, G. and Ravikumar, B. 1992, "Public versus Private Investment in Human Capital Endogenous Growth and Income Inequality." Journal of Political Economy, 100(4), pp. 813-834.

Goldin, C. and Katz, L. F. 1998, "Human Capital and Social Capital: The Rise of Secondary Schooling in America, 1910 to 1940." NBER working paper 6439.

Gradstein, M. and Justman, M. 1997, "Democratic Choice of an Education System: Implications for Growth and Income Distribution." Journal of Economic Growth, 2(2), pp. 169-183.

Gradstein, M. 2004a, "Voting on Meritocracy," European Economic Review, 48, pp. 797-803.

_________. 2004b, "The Differential Incidence of Public Education: Causes and Consequences," Department of Economics, Ben Gurion University, mimeo.

Hansen, W. L. 1970, "Income Distribution Effects of Higher Education", American Economic Review, 60(2), pp. 335-340.

Kaganovich, M. and Zilcha, I. 1999, "Education, Social Security, and Growth." Journal of Public Economics, 71(2), pp. 289-309.

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

Le Grand, J. 1982, "The Strategy of Equality: Redistribution and the Social Services", George Allen and Unwin (Publishers) Ltd.

Lucas, R. E. Jr. 1988, "On the Mechanics of Economic Development," Journal of Monetary Economics, 22(1), pp. 3-42.

Nechyba, T. 1999, "School Finance Induced Migration and Stratification Patterns: the Impact of Private School Vouchers." Journal of Public Economic Theory, 1(1), pp. 5-50.

Peltzman, S. 1973, "The Effect of Government Subsidies-in-kind on Private Expenditures: the Case of Higher Education", Journal of Political Economy, 81(1), pp. 1-27.

Psacharopoulos, G and Patrinos, H. A. 2004, "Returns to Investment in Education: a Further Update." Education Economics, 12(2), pp. 111-134.

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

Saint-Paul, G. and Verdier, T. 1993, "Education, Democracy and Growth," Journal of Development Economics, 42(2), pp. 399-407.

Su, X. 2004, "The Allocation of Public Funds in a Hierarchical Educational System." Journal of Economic Dynamics and Control, 28(12), pp. 2485-2510.

Appendix A

Proposition 1.

Proof. The equilibrium solution for P_t satisfies $P_t = \sum_{i=1}^{3} \lambda_i p_t^{i*}$. The left hand side is continuous and strictly increasing in P_t. The right hand side, denoted as $f(P_t)$, is a decreasing step function of P_t, with $\lim_{P_t \to 0+} f(P_t) > 0$, and $f(1) \leq 1$. The single crossing point of the two curves is the equilibrium solution for P_t, and $P_t > 0$. Subsequently $\{p_t^{i*}\}_{i=1}^{3}$ can be uniquely determined. ■

Proposition 2.

Proof. Individuals take the equilibrium solution for P_t parametrically, and solve for their own probabilities p_t^{i*} . Given $P_t < AG_t^a$, i.e., $1 - \frac{P_t}{AG_t^a} > 0$, it is obvious that $p_t^{i*} = 1$ iff $h_t^{a,i} > h_t^{b,i}$, or equivalently, $h_t^{b,i} > \widetilde{h}$; $p_t^{i*} = 0$ iff $h_t^{b,i} < \widetilde{h}$; and $p_t^{i*} \in [0, 1]$ iff $h_t^{b,i} = \widetilde{h}$. ■

Proposition 3.

Proof. This is proof by construction. For group 3, when $G_t^b \in (c_3^l, c_3^r)$, since $Bh_{t-1}^3 G_t^b > \widehat{h}$ and $G_t^a > 0$, by Proposition 1 there exists a unique economic equilibrium with $P_t > 0$. By Proposition 2, $p_t^{3*} > 0$. It is obvious that $p_t^{3*} = 0$ when $G_t^b \in [0, R_t] \backslash (c_3^l, c_3^r)$.

Now consider the equation $A(Bh_{t-1}^3 G_t^b - \widehat{h})(R_t - G_t^b)/\lambda_3 = Bh_{t-1}^3 G_t^b$. The left-hand side (LHS) is quadratic and concave in G_t^b, while the right-hand side (RHS) is linear in G_t^b. When $G_t^b \longrightarrow c_I^l+$, $LHS < RHS$; when $G_t^b \longrightarrow c_I^r-$, $LHS < RHS$. Then for $G_t^b \in (c_3^l, c_3^r)$, if the equation has 0 real solution, then $LHS < RHS$, and $p_t^{3*} < 1$. If the equation has 1 or 2 real solutions, denoted as b_3^l and b_3^r where $c_3^l < b_3^l \leq b_3^r < c_3^r$, then $LHS \geq RHS$ iff $G_t^b \in [b_3^l, b_3^r]$, and $p_t^{3*} = 1$.

Then go on to group 2 and consider the equation $A(Bh_{t-1}^2 G_t^b - \widehat{h})(R_t - G_t^b)/\lambda_3 = Bh_{t-1}^2 G_t^b$. When $G_t^b \longrightarrow b_3^l+$, $LHS < RHS$; when $G_t^b \longrightarrow b_3^r-$, $LHS < RHS$. Then for $G_t^b \in (b_3^l, b_3^r)$, if the equation has 0 or 1 real solution, then $LHS \leq RHS$, and $p_t^{2*} = 0$. If the equation has 2 distinct solutions, denoted as c_2^l and c_2^r, where $b_3^l < c_2^l < c_2^r < b_3^r$, then $LHS > RHS$ iff $G_t^b \in (c_2^l, c_2^r)$, and $p_t^{2*} > 0$.

Next switch to the equation $A(Bh_{t-1}^2 G_t^b - \widehat{h})(R_t - G_t^b)/(\lambda_3 + \lambda_2) = Bh_{t-1}^2 G_t^b$. For $G_t^b \in (c_2^l, c_2^r)$, if it has 0 real solution, then $LHS < RHS$, and $p_t^{2*} < 1$; if it has 1 or 2 real solutions, denoted as b_2^l and b_2^r with $c_2^l < b_2^l \leq b_2^r < c_2^r$, then $LHS \geq RHS$ iff $G_t^b \in [b_2^l, b_2^r]$, and $p_t^{2*} = 1$.

We follow similar steps for group 1. In total there are 3 groups and hence at most 6 steps to consider to find the minimal group i.

For any $j < i$, it is obvious that $p_t^{j*} = 0$, continuous (constant) for all $G_t^b \in [0, R_t]$. For any $j \geqslant i$, it is obvious that $p_t^{j*} = 0(1)$, continuous (constant) when $G_t^b \in [0, R_t] \backslash (c_j^l, c_j^r)(G_t^b \in [b_j^l, b_j^r])$. Continuity of p_t^{j*} when $G_t^b \in (c_j^l, c_j^r) \backslash [b_j^l, b_j^r]$ follows when applying the Implicit Function Theorem to the equation $A(Bh_{t-1}^j G_t^b - \widehat{h})(R_t - G_t^b)/(\sum_{s=j+1}^3 \lambda_s + p_t^{j*}\lambda_j) = Bh_{t-1}^j G_t^b$. Continuity of p_t^{j*} at the points c_j^l, c_j^r and b_j^l, b_j^r follows directly from the construction. ■

Proposition 4.

Proof. When h_{t-1}^2 is low, there exists no c_2^l or c_2^r. For the top class, $y_t^3 = 2Bh_{t-1}^3 G_t^b$ for $G_t^b = [0, R_t] \backslash [b_3^l, b_3^r]$, and $y_t^3 = Bh_{t-1}^3 G_t^b + A(Bh_{t-1}^3 G_t^b - \widehat{h})(R_t - G_t^b)/\lambda_3$ for $G_t^b = [b_3^l, b_3^r]$. With sufficiently high R_t, the top class' lifetime income is maximized at $G^3 = \frac{R_t}{2} + \frac{\widehat{h}}{2Bh_{t-1}^3} + \frac{\lambda_3}{2A}$, and hence $\frac{\partial G^3}{\partial \widehat{h}} > 0$, $\frac{\partial G^3}{\partial h_{t-1}^3} < 0$, and $\frac{\partial G^3}{\partial \lambda_3} > 0$. ■

Proposition 5.

Proof. When h_{t-1}^2 is moderate, there exist c_2^l and c_2^r but not b_2^l and b_2^r. For the top class, y_t^3 remains the same as that in Proposition 4 for $G_t^b = [0, R_t] \backslash (c_2^l, c_2^r)$. When $G_t^b \in (c_2^l, c_2^r)$, the middle class continuously adjusts its participation probability p_t^{2*} with respect to G_t^b, so the schooling quality at higher education cannot be directly determined. However, it can be indirectly determined using the equation $y_t^{a,2} = y_t^{b,2}$, namely $\frac{R_t - G_t^b}{\lambda_3 + p_t^{2*}\lambda_2} = \frac{Bh_{t-1}^2 G_t^b}{A(Bh_{t-1}^2 G_t^b - \widehat{h})}$. Hence $y_t^3 = 2Bh_{t-1}^3 G_t^b + (h_{t-1}^3/h_{t-1}^2 - 1)\widehat{h} + (h_{t-1}^3/h_{t-1}^2 - 1)\widehat{h}^2/(Bh_{t-1}^2 G_t^b - \widehat{h})$ for $G_t^b \in (c_2^l, c_2^r)$, and $\frac{\partial^2 y_t^3}{\partial G_t^b} > 0$. Convexity together with continuity of y_t^3 with respect to G_t^b implies that the candidates for local maxima are the end points c_2^l and c_2^r.

It is then obvious that the top class' most preferred policy is still G^3 if $G^3 \notin (c_2^l, c_2^r)$. When $G^3 \in (c_2^l, c_2^r)$, consider $c_2^m \equiv \frac{c_2^l + c_2^r}{2} = \frac{R_t}{2} + \frac{\widehat{h}}{2Bh_{t-1}^2} - \frac{\lambda_3}{2A}$. By the symmetry of the quadratic function that defines G^3, if $c_2^m > G^3$, namely $\frac{\widehat{h}}{2Bh_{t-1}^2} - \frac{\widehat{h}}{2Bh_{t-1}^3} > \frac{\lambda_3}{A}$, we have $y_t^3|_{G_t^b = c_2^l} > y_t^3|_{G_t^b = c_2^r}$, so the top class' most preferred policy is c_2^l. Similarly if $c_2^m < G^3$, the top class' most preferred policy is c_2^r. ■

Proposition 6.

Proof. When h_{t-1}^2 is high, there exist b_2^l and b_2^r. For the top class, y_t^3 remains the same as that in Proposition 5 for $G_t^b = [0, R_t] \backslash [b_2^l, b_2^r]$. When $G_t^b = [b_2^l, b_2^r]$, $y_t^3 = Bh_{t-1}^3 G_t^b + A(Bh_{t-1}^3 G_t^b - \widehat{h})(R_t - G_t^b)/(\lambda_3 + \lambda_2)$. With sufficiently high R_t, the top class' lifetime income is maximized at $\widetilde{G}^3 = \frac{R_t}{2} + \frac{\widehat{h}}{2Bh_{t-1}^3} + \frac{\lambda_3 + \lambda_2}{2A}$. ■

Appendix B

The Extended Model with Diminishing Marginal Returns

Now the original model is extended to allow diminishing marginal returns in the two-stage education technology:

$$h_t^{b,i} = B(h_{t-1}^i)^{\alpha_1} (G_t^b)^{\beta_1} \tag{5}$$

$$h_t^{a,i} = \begin{cases} 0 & \text{if } h_t^{b,i} \leqslant \widehat{h} \\ A(h_t^{b,i} - \widehat{h})^{\alpha_2} (G_t^a / P_t)^{\beta_2} & \text{if } h_t^{b,i} > \widehat{h} \end{cases} \tag{6}$$

All other aspects of the model remain the same, and so is the definition of the equilibrium.

First, note that the parameter α_1 is simply an increasing transformation of the exogenously given parental human capital level. With our focus on the intra-generational instead of inter-generational analysis, the parameter α_1 has neither qualitative nor quantitative impact on all results.

Proposition 1 is robust to this more general specification, since the critical argument relies only on the congestion effect in higher education such that p_t^{i*} decreases with P_t.

Proposition 2 now depends on the parameter α_2. Taking the equilibrium solution P_t parametrically, when $\alpha_2 = 1$, there is only a lower threshold on qualification $h_t^{b,i}$ for attending higher education; while when $\alpha_2 < 1$, there is also an upper threshold on qualification $h_t^{b,i}$ for attending higher education. With diminishing marginal returns on qualification, if some individuals accumulate sufficiently high human capital at basic education, it may be optimal for them to opt out of higher education since the benefit can be outweighed by the opportunity cost of foregone wage income. Whether highly qualified individuals actually opt out of higher education or not depends on the distribution of initial qualifications and the public budget allocation policy. Since strong empirical evidence suggests that students from good family background on average benefit more from higher education, it is with little loss of generality by focusing on the lower threshold for attending higher education.

Propositions 3, 4, 5 and 6 are more sensitive to the combination of the parameters α_2, β_1 and β_2. Only with special parameter values can the analysis be simplified to a negative quadratic function of G_t^b to determine the sign of $h_t^{a,i} - h_t^{b,i}$, whose upper contour set (where $h_t^{a,i} - h_t^{b,i} \geqslant 0$, or equivalently the range of allocation policies under which a group i individual finds it optimal to attend higher education) is conveniently convex. Otherwise the analysis will involve a general polynomial function (of some increasing function of G_t^b), whose upper contour set may be the union of finitely many disjoint intervals. Furthermore, the intervals for different groups need not occur in the nested fashion as in Proposition 3, if highly qualified individuals opt out of higher education. Then the extended model requires a much heavier computational cost to determine the global optimum among many local optima, and there is a bigger likelihood that individuals' policy preferences exhibit discontinuous jumps.

Appendix C

Here numerical analysis is conducted to illustrate how the top class' most preferred policy (1) changes with the initial qualification of the middle class; (2) changes with the size of total public budget; and (3) may be drastically different from the socially efficient policy. The parameters are set at the following values: $\lambda_1 = 0.7$, $\lambda_2 = 0.2$ and $\lambda_3 = 0.1$; $h^1_{t-1} = 1$, $h^2_{t-1} \in \{2, 3, 4, 5, 6, 7\}$, and $h^3_{t-1} = 10$; $A = 1$, $B = 1$, and $\widehat{h} = 3$.

		h^2_{t-1}					
		2	3	4	5	6	7
$R_t = 2$	most preferred	1.20	1.12	**0.82**	**0.64**	2	2
	socially efficient	1.31	1.12	**2**	**2**	2	2
$R_t = 2.5$	most preferred	1.45	1.07	**0.79**	**0.63**	**1.55**	**1.55**
	socially efficient	1.56	1.07	**2.5**	**2.5**	**2.5**	**2.5**
$R_t = 3$	most preferred	1.61	1.05	**0.78**	**0.62**	1.80	1.80
	socially efficient	1.61	1.05	**2.02**	**1.98**	1.95	1.93

(1) Reading by the rows, the top class' most preferred policy firsts decreases and then jumps upward when the initial qualification of the middle class increases; (2) Reading by the columns, the top class' most preferred policy depends quantitatively but not qualitatively on the size of total public budget; and (3) Reading the bold numbers, the top class' most preferred policy differs most from the socially efficient policy when the middle class' initial qualification is neither too low nor too high, and/or when the size of total public budget is neither too low nor too high, so that it is most beneficial for the top class to manipulate the allocation policy to exclude the middle class from higher education. And a difference exists generally. Other parameter values are used as robustness check, and the patterns remain the same (results not reported here).

Table 1a Relative Schooling Quality across Education Stages

	Public Current Expenditure per Pupil (% of GNP per Capita)				
	Prim. (1)	Second. (2)	Tert. (3)	(4)=$\frac{(2)}{(1)}$	(5)=$\frac{(3)}{(1)}$
Malawi	9	27	1,580	3	176
Comoros	9	42	1253	4.67	139
Guinea	8	32	483	4	60
Lesotho	14	54	780	3.86	56
China	6	12	67	2	11
India	11	18	100	1.64	9
Brazil	10	30	98	3	10
Pakistan	9	15	94	1.67	10
U.S.	17	24	25	1.41	1.47
Japan	17	19	14	1.12	0.82
France	16	27	28	1.69	1.75
U.K.	17	20	40	1.18	2.35

Table 1b Summary Statistics for OECD and non-OECD countries

	Africa - 53 countries		Other - 107 countries		OECD - 29 countries	
	(4)=$\frac{\text{Second.}}{\text{Prim.}}$	(5)=$\frac{\text{Tert.}}{\text{Prim.}}$	(4)=$\frac{\text{Second.}}{\text{Prim.}}$	(5)=$\frac{\text{Tert.}}{\text{Prim.}}$	(4)=$\frac{\text{Second.}}{\text{Prim.}}$	(5)=$\frac{\text{Tert.}}{\text{Prim.}}$
obs.	27	28	59	62	22	29
mean	3.10	35.10	2.26	5.84	1.28	2.04
min	0.53	3.60	0.53	0.85	0.67	0.35
max	7.15	176	10	28.64	1.79	4

Table 2 Survival Rates across Education Stages

	Africa - 53 countries		Other - 107 countries		OECD - 29 countries	
	(6)=$\frac{\text{Second.}}{\text{Prim.}}$	(7)=$\frac{\text{Tert.}}{\text{Prim.}}$	(6)=$\frac{\text{Second.}}{\text{Prim.}}$	(7)=$\frac{\text{Tert.}}{\text{Prim.}}$	(6)=$\frac{\text{Second.}}{\text{Prim.}}$	(7)=$\frac{\text{Tert.}}{\text{Prim.}}$
obs.	23	32	60	65	19	28
mean	0.37	0.05	0.68	0.20	1.06	0.47
min	0.08	0.004	0.23	0.01	0.54	0.10
max	0.74	0.20	1.11	0.48	1.47	0.86