

Singapore Management University

## Institutional Knowledge at Singapore Management University

---

Research Collection School Of Economics

School of Economics

---

11-2009

### Why Did Universities Precede Primary Schools? A Political Economy Model of Educational Change

Fali HUANG

*Singapore Management University, flhuang@smu.edu.sg*

Follow this and additional works at: [https://ink.library.smu.edu.sg/soe\\_research](https://ink.library.smu.edu.sg/soe_research)



Part of the [Behavioral Economics Commons](#), and the [Higher Education Commons](#)

---

#### Citation

HUANG, Fali. Why Did Universities Precede Primary Schools? A Political Economy Model of Educational Change. (2009). 1-36.

Available at: [https://ink.library.smu.edu.sg/soe\\_research/1306](https://ink.library.smu.edu.sg/soe_research/1306)

This Working Paper is brought to you for free and open access by the School of Economics at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection School Of Economics by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email [cherylds@smu.edu.sg](mailto:cherylds@smu.edu.sg).

# **Why Did Universities Precede Primary Schools? A Political Economy Model of Educational Change**

**Fali Huang**  
November 2009

Paper No. 13-2009

# Why Did Universities Precede Primary Schools? A Political Economy Model of Educational Change\*

Fali Huang

School of Economics

Singapore Management University

90 Stamford Road, Singapore 178903

Email: fhuang@smu.edu.sg

Tel.: 65-68280859

Fax: 65-68280833

November 6, 2009

## Abstract

Universities were first established in Europe around the twelfth century, while primary schools did not appear until the nineteenth. This paper accounts for this phenomenon using a political economy model of educational change on who are educated (the elite or the masses) and what is taught (general or specific/vocational education). A key assumption is that general education is more effective than specific education in enhancing one's skills in a broad range of tasks, including political rent-seeking. Its findings suggest that specific education for the masses is compatible with the elite rule, while mass general education is not, which refines the conventional association between education and democracy.

*JEL:* O10, O40, P16, N10.

*Key Words:* General Education, Specific Education, Elite Education, Mass Education, Long-Run Development.

## 1 Introduction

Modern school systems are usually operated on a well-structured, hierarchical ladder of grade levels starting from primary school, then moving to secondary school, and finally completed at the

---

\*I thank the editor, Vincenzo Quadrini, and an anonymous referee as well as participants at WEAI 2009 Kyoto conference for very helpful comments.

university of higher learning. This sequence has become so natural to us in modern times that it may be surprising to be reminded that this was not the case in most parts of human history. In fact, “as a formal structure, the university or highest level of education was invented before secondary schools, and in a sense, even before elementary schools.” (Collins 2000, P. 213). Universities were first established in Europe around the twelfth century, while primary schools did not appear until the nineteenth. Universities provided general education for a very small elite who would become the future leaders for State and Church, while primary schools as well as vocational secondary schools were initially created for the poor masses to teach them basic reading, writing, arithmetic and some specific vocational skills. There is some consensus that the modern education system for the masses arose around 1870 in western Europe and the U.S., while mass general education at the university level started after the second World War and is still an ongoing process (Trow 1967, Mueller, Ringer and Simon 1977, Meyer, Ramirez and Soysal 1992, Goldin 1999, Bowen 2003).

So during a long time in history, only a few ruling elite were educated and the elite education focused on general education, while mass schooling occurred much later, around the time of industrialization, and it started from specific or vocational education and gradually moved to general education. This seems to be a general trend of educational change at least in the history of western Europe. This paper uses a political economy model to account for such a historical trend in terms of who are educated (the elite or the masses) and what is taught (general or specific education). A main theoretical puzzle is why, on the one hand, specific education preceded general education for the masses, while on the other hand, the elite were trained exclusively in general education.<sup>1</sup>

The intuition begins with the conventional differences between general and specific education or human capital (Becker 1964, Bennett 1967).<sup>2</sup> General education increases one’s general ability in learning new knowledge and technologies, but it is more costly to acquire in terms of intellectual effort because it needs to build on a broad and abstract knowledge base. Skill-specific education, in contrast, teaches task-specific skills and is thus less costly to acquire, but it is narrow in knowledge

---

<sup>1</sup>This paper focuses on the evolution of organized education conducted in schools, while treating unorganized learning in daily life as the default starting point. See Section 6 for more discussions on this.

<sup>2</sup>The theoretical distinction between general and specific human capital has become widely adopted in economics; see, for example, Gould, Moav and Weinberg (2001), Lazear (2003), Bertocchi and Spagat (2004), Krueger and Kumar (2004a, b), Wasmer (2006), Iyigun and Owen (1999, 2006).

coverage and not deep enough to facilitate further learning and easy adaptation to new things.<sup>3</sup> An implication of the differences in these two kinds of education is that, when technological changes are slow, it is more beneficial to get specific skills than general education since there is less need to adapt to new technologies. This partially explains why specific education preceded general education for the masses.

Another important difference between general and specific education is from the political aspect. General education increases one's general productivity in solving problems, which implies that it also increases one's skills of extracting gains and defending one's interests when there are political conflicts. Franklin D. Roosevelt, for example, observed that "A man who has never gone to school may steal from a freight car, but if he has a university education he may steal the whole railroad." And in general there is evidence that better educated people are better organized and more active politically (Meyer and Rubinson 1975, Collins 1971). Specific skills, in contrast, increase one's task-specific productivity but not necessarily their political rent-seeking skills or organizing efficiency. So when the potential rent is large enough, the elite will focus only on general education even when the knowledge stock changes slowly, thus not following the natural precedence of specific education that arises from pure economical concerns.

The implication for mass education, however, is the opposite. On the one hand, the elite benefit from a better educated and hence more productive workforce through larger tax revenues, the more so when the capital-labor ratio is larger. But on the other hand, the elite's political privileges may be reduced or threatened when the masses acquire more general education, which makes them better equipped in defense of their own interests. This is consistent with Alexis de Tocqueville's famous proposition "To enlighten the people is to destroy kings."<sup>4</sup> So the ruling elite may find it

---

<sup>3</sup>In reality, an education program may contain a mixture of both specific and general education, which makes it less straightforward to classify it into either category. It is still possible, however, to ascertain the main component and purpose of an education program and categorize it accordingly. The primary school under the elite rule, for instance, by teaching pupils preliminary reading and number skills in preparation for later vocational training and by indoctrinating blind obedience, can be categorized as specific education. In contrast, the primary school that prepares pupils for future academic endeavors has to cultivate independent thinking besides teaching the basic skills, and so it belongs to an organic part of general education.

<sup>4</sup>The landed elite in England, for example, opposed mass schooling because they believed that "giving education to the labouring classes of the poor ... would render them factious and refractory, ... would render them insolent to their superiors" (quoted in Lindert 2004, pp. 100). And such attitudes "could be found on the part of propertied

beneficial to support mass specific education, but not mass general education, especially when the potential loss of political rent is large relative to the productivity gain. It suggests that under the elite rule mass general education is more likely to be repressed and delayed than specific education.

These results are applicable not only to the distinction between general and specific education, but also to different subjects in general education such as humanities versus natural sciences; if humanities are more likely to increase one's political rent-seeking skills than natural sciences, then an implication is that the ruling elite will invest earlier and more in humanities than in natural sciences, the more so when the political rent is larger. In stark contrast, the political elite would be more willing to teach the masses natural sciences than humanities.<sup>5</sup> These results bring new insights to our understanding of educational change over time and across countries.

This paper contributes to the literature in several ways. First, it proposes a simple political economy model to account for the major educational change in history, especially in terms of who are educated and what is taught. This is an important issue itself that warrants serious research effort, given the crucial role of knowledge and human capital in promoting economic growth and general welfare. Second, this model can be readily extended to address important questions such as why in history the progress of natural sciences lagged behind humanities, and the more so in societies with larger political rents or with an open-elite system such as that in pre-modern China. Third, this paper generates new insights in the relationship between education, political regimes and economic growth. For example, it shows that specific education for the masses can be supported by the elite rule or other forms of non-democratic political systems, since it is not necessarily conducive to democratization.

This new insight on compatibility between mass specific education with the elite rule refines the conventional association between education and democracy (Acemoglu et al. 2005). It suggests that the average schooling level may not be a sufficient statistic for education, and what children are taught in terms of general versus specific education is of great importance. Specifically, it may shed

---

conservatives in practically any country and century" (Lindert 2004, pp. 101).

<sup>5</sup>For similar reasons, the elite would be willing to teach the masses social, moral and civic values that help solidify their rule, which is one of the reasons for the public support of mass education (see, for example, Bowles and Gintis 1976, Lott 1990, Gradstein and Justman 2002). This is also consistent with the state building rationale of mass schooling (Green 1990, Lindert 2004).

light on the puzzling fact that Prussia (and then Germany) was the first to start mass education but a stable democracy was established much later than others (Lindert 2004); a possible explanation is that mass education in Germany was predominantly vocational with a heavy dose of obedient and patriotic conditioning (Green 1990), which could work well for the elite or authoritarian rule. In contrast, another leader in mass education, the U.S., emphasized general education and established democracy from early on (Goldin 1999, 2001).

This paper is related to a number of studies. The evidence for the initial increase and then decline of vocational education in the process of economic development is shown by Bertocchi and Spagat (2004), who propose that the elite restrict access to general education because it brings more social status compared with vocational education. The externalities of individual education on the distribution of human capital level and economic growth are studied by Galor and Tsiddon (1997), while Galor, Moav and Vollrath (2009) show that mass education may be delayed by land owners instead of industrialists. Brezis and Crouzet (2006) focus on the training of elites in history. Grossman and Kim (2003) examine the choice between an egalitarian and an elitist education policy where property of the rich may be threatened by the poor. None of these studies, however, analyzes the evolution of the entire education system that covers historical periods both before and after the Industrial Revolution in addition to major trends in both the elite and mass education.

The paper proceeds as follows. The next section briefly describes the relevant historical evidence of educational change in western Europe and the U.S. The basic model setup and the analysis of the benchmark model are provided in the following two sections, respectively. The main results of the paper are analyzed in Section 5, while some important extensions of the model are considered in Section 6. The final section concludes. All technical proofs are relegated to the Appendix.

## 2 Historical Evidence

The main results of this paper suggest that the educational change can be categorized into three phases: (1) *Elite Education*. The political elite are educated first and in general education while the masses do not receive any formal education for a long time. (2) *Stratified Education*. When the capital-labor ratio becomes large enough, it may be beneficial for the elite to support specific

or technical education for the masses, while the elite still receive general education, and so the education system is stratified. (3) *Mass Education*. After the elite rule is replaced by democracy, which is endogenously determined by the general education level of the masses, mass education may still be dual-track initially, where both specific education and general education are pursued. Over time, however, mass education will eventually move towards general education. This is broadly consistent with the western history.

Before the nineteenth century, the masses hardly received any formal education at all, while the elites were educated in classical studies that were designed mainly to train free thinkers, rather than preparing for any specific or narrow vocations. Over the years, and especially during the seventeenth and eighteenth centuries, scientific knowledge as to the forces and laws of the physical world was gradually accumulated by scholars. Important development and great applications were made in the nineteenth century, which kicked in the era of modern economic growth (Mokyr 2002). In the second phase of the Industrial Revolution, the rise of science-based new industries such as chemicals and electricity made the grasp of systematic scientific knowledge essential for worker productivity, and the need to train the masses with relevant technical knowledge was strongly felt. As a consequence, vocational education for the masses started in the leading industrial countries around 1870. The elite education, however, was hardly changed from before and run separately from the mass schools. So the education system was highly stratified, especially in continental Europe. Though democracy was achieved in the leading industrial countries by the end of World War I, the dual-track education system continued for decades, and a unified general education system was not attempted or realized until after World War II (Collins 2000).

**Elite Education.** In the late eleventh and early twelfth centuries, a number of schools, in time known as universities, came into prominence. They represent a reaction to the changed conditions of the times such as the rise of the city and the revival of trade and industry. They offered studies in four traditional faculties, namely, arts, law, medicine, and theology, and trained future leaders for State and Church. During the Renaissance in the fifteenth and sixteenth centuries, language and literature of ancient Greece and Rome formed the main content of elite education and were known as humanities. The typical humanistic schools of Europe were the German Gymnasium



and the English public schools. The grammar school of the American colonies was a transplanted English public school. This humanistic education dominated the training of the elite classes of society for the next four hundred years. For example, the form of the English public schools was not drastically modified until the report of the royal commissioners of investigation in 1864.

**Stratified and Mass Education.** As commercial and industrial sectors gradually become dominant economic activities, the Industrial Revolution started first in England around the middle of the eighteenth century and later spread to other western countries. The industrialization process brought forth fundamental changes across Europe, transforming it from a primarily agrarian society to a mature industrial society by the end of the nineteenth century; it increased demand for skilled workers and set the stage for the beginning of mass education in western Europe and America.

**Germany.** Prussia was the first modern state to start mass education, and by 1825, it had taken over education from the Church and made it an instrument of the State. Its education system, however, was a highly stratified two-class school system, and it was continued after the formation of the German Empire in 1871. Children of the laboring classes attended the basic elementary schools (Volksschule: folk-schools), which were designed to create an intelligent but obedient and patriotic citizenship; a vast program of vocational education was, about a half century later, developed for the graduates of the Volksschule. For those who were to form the official and directing class of society, a classical training track composed of Gymnasium and university was provided. Up at least to 1914, schools for the masses educated about 92% children, while schools for the leaders educated 8% (Cubberley 1920, pp. 578). The basic pattern of this two-class education system has not been dramatically changed ever since.<sup>6</sup> The establishment of a stable liberal parliamentary republic was not achieved in West Germany until its major military defeat in the World War II and in East Germany until the reunification of Germany in 1990.

**France.** The Law of 1833, which established lower and higher primary schools, marked the first time in French history that an earnest effort was made to provide education for the great mass of the people. The basic form of French national education was achieved under the Third Republic (1870), which started a lasting democracy with universal male suffrage. The French state

---

<sup>6</sup>Currently in Germany about 25% secondary-school-age children enter the Gymnasium, and the majority rest attend some schools that eventually lead to vocational training (Encyclopedia Britannica).

school system was also of two-class, where schools for the masses were composed by the primary and vocational schools, and schools for the elite included the lycees and universities. No concerted drive was made to reform this system until World War I. Reforms eventually took place in 1960s when the access to secondary education was open up for all children, and since 1980s the lycees and then higher education were becoming accessible to the great majority of young people.

**England.** Starting from the Elementary Education Act of 1870, England started to establish compulsory elementary education for the masses. The curriculum of most elementary schools was limited to reading and writing. The Balfour Act of 1902 equalized and unified educational provisions and made the general provision of secondary education possible. New modern-type universities were created in various cities to extend the educational ladder upward. However, even after World War I only a small proportion of pupils, around 1920 about 10 per cent, transferred from the upper grades of the elementary into the secondary schools, and only “1 per cent or less of these former elementary-school pupils entered the university” (Ulich 1961, pp. 117-118). It was not until the Education Act of 1944 that the old class division of secondary schools for the privileged, on the one hand, and elementary and vocational schools for the masses, on the other hand, was broken down. Starting from 1960s the number of comprehensive schools has grown so that most state-maintained secondary schools are now comprehensive.

**United States.** The American education was initially similar to that in Europe, focused almost exclusively on training a small elite through the Latin grammar schools and universities. In the middle of the eighteenth century, the elementary or common school had been established in different states to provide the basic education to children of the masses. The full white manhood suffrage was largely complete by the 1840s in the U.S. After 1870, the mass secondary school started to grow, which was largely terminal, providing an increasingly vocational education for the new body of white-collar workers (Trow 1967). The rapid growth of higher education occurred after World War II, and as a consequence, prompted the transformation of the terminal secondary system into a mass preparatory system (Goldin 1999, 2001).

In summary, the three phases of educational change, namely, elite education, stratified education, and mass education (from dual track to mass general education) are broadly consistent with

historical evidence of western Europe and the U.S.. Similar school systems have been reproduced in many other countries, where mass general education links closely with democracy (Cubberley 1920).

### 3 The Basic Model

There are overlapping generations with a fixed population of measure  $N$ . Each individual lives for two periods, accumulating human capital in childhood and participating in production at adulthood.

**Technology and Endowment.** In every period the economy produces a single homogeneous good that can be used for consumption and investment. The production function at time  $t + 1$  is

$$Y_{t+1} = H_{t+1}^\alpha K_{t+1}^{1-\alpha},$$

where  $H_{t+1}$  is the aggregate stock of human capital and  $K_{t+1}$  is the aggregate physical capital. They depreciate fully after one period that is equivalent to an individual's adulthood.

The total amount of human capital  $H_{t+1}$  is composed of two distinct forms, general human capital  $H_{g,t+1}$  and task-specific human capital  $H_{s,t+1}$ , where

$$H_{t+1} \equiv A_{t+1}H_{g,t+1} + A_tH_{s,t+1}.$$

This suggests that workers with general human capital can access the current knowledge stock  $A_{t+1}$ , while those with specific human capital can only access  $A_t$ , the knowledge stock in the previous period. It captures an essential difference between these two kinds of human capital, where general human capital allows one to transform new knowledge into productivity, while specific human capital does not.<sup>7</sup>

The knowledge stock  $A_{t+1}$  grows at an endogenously determined speed  $\phi(H_{g,t})$  such that

$$A_{t+1} = A_t(1 + \phi(H_{g,t})), \tag{1}$$

where  $\phi' > 0$  and  $\phi(0)$  is a very small positive number. That is, the speed of knowledge accumulation is strictly increasing in general human capital  $H_{g,t}$ , but constant with regard to the stock of specific

---

<sup>7</sup>Similar results can be obtained if some degree of complementarity between  $H_{g,t+1}$  and  $H_{s,t+1}$  exists, for example, as in  $H_{t+1} = [(A_{t+1}H_{g,t+1})^\rho + (A_tH_{s,t+1})^\rho]^{1/\rho}$  where  $\rho \leq 1$ .

skills. This assumption reflects the fact that general human capital is more effective in generating new knowledge than specific skills. The initial stock of knowledge  $A_0 > 0$  is exogenously given.

**Preferences.** Individuals are identical in preferences, which are represented by a log-linear utility function<sup>8</sup>

$$u_{ti} = (1 - \beta) \log c_{ti} + \beta \log(z + b_{ti}),$$

where  $c_{ti}$  is the adulthood consumption of individual  $i$  in generation  $t$ ,  $b_{ti}$  is his bequest for offspring,  $\beta \in (0, 1)$  indicates the relative weight of bequest, and  $z > 0$  represents some threshold level of income. The budget constraint is  $c_{ti} + b_{ti} \leq I_{ti}$ , where  $I_{ti}$  is individual  $i$ 's income at adulthood. As a result of utility maximization, his optimal bequest is  $b_{ti} = \max\{\beta(I_{ti} - z(1 - \beta)/\beta), 0\}$ . That is, only when an individual's income is higher than a certain level indicated by  $Z \equiv z(1 - \beta)/\beta$ , would there be any resources left as bequest. The bequest  $b_{ti}$  can be invested in physical capital or human capital for the next generation. The total bequest in society is

$$B_t = \sum_i b_{ti} = \sum_i \max\{\beta(I_{ti} - Z), 0\}.$$

**Investment in Physical and Human Capital.** Let  $m_{kti}$ ,  $m_{gti}$ , and  $m_{sti}$  denote respectively the amount of material resources invested in the three capital forms, namely physical capital, general human capital, and specific human capital. The sum of them cannot exceed the total savings so that  $m_{kti} + m_{gti} + m_{sti} \leq b_{ti}$  must hold. The aggregate physical capital is thus

$$K_{t+1} = \sum_i m_{kti}.$$

Each worker is endowed with a basic unit of specific human capital, namely the raw labor, even without any education expenditure; to acquire human capital above the basic level, however, a positive amount of material is needed. The human capital production function  $h(\cdot)$  is the same for both general and specific human capital, where  $h(\cdot)$  is concave with  $h(0) = 0$  and  $h'(0) = \gamma < +\infty$ . An individual can invest in only one type of human capital, which is either

$$h_{g,t+1,i} = h(m_{gti}) \quad \text{or} \quad h_{s,t+1,i} = h(m_{sti}) + 1.$$

---

<sup>8</sup>As long as a person with a higher income saves more, the exact specifications of utility function are not essential for the qualitative results.

The aggregate general and specific human capital are thus  $H_{g,t+1} = \sum_i h_{g,t+1,i}$  and  $H_{s,t+1} = \sum_i h_{s,t+1,i}$ , respectively.

Individuals have heterogenous innate abilities, which affect their costs of investing in general human capital. The ability distribution is identical across generations, so is the individual ability in the same family. Specifically, the ability of an individual  $i$  in any generation  $t$  is  $a_i$ , which has a distribution of  $F(\cdot)$  on the interval  $[a_l, a_h] \subset [0, +\infty)$ . The investment cost of general human capital is  $C(a_i, h_{gti}) > 0$ , where  $C_1 < 0$  and  $C_2 < 0$ . That is, it is less costly in effort to invest in general education if one has a higher ability or if one's parent has higher general human capital. The cost of investing in specific human capital is normalized to zero for all individuals. This captures another difference between general and specific human capital, that is, it is more costly to acquire general human capital since it needs to build on a much broader and more abstract knowledge base.

## 4 Educational Change: The Benchmark Case

We first consider the benchmark case where there is no political conflict and individuals are endowed with an identical amount of physical capital  $K_0/N$ . It is straightforward to show that as knowledge accumulates over time, the economy first invests in physical capital, then in specific human capital, and finally in general human capital, where individuals with higher abilities invest earlier.

**Production.** Production is operated in a perfectly competitive environment. Given the capital return rate  $r_{t+1}$ , the wage rate per unit of specific human capital  $w_{s,t+1}$  and the wage rate per unit of general human capital  $w_{g,t+1}$ , producers in period  $t + 1$  choose the amount of physical capital and two types of human capital to maximize profits. That is,

$$\{K_{t+1}^*, H_{s,t+1}^*, H_{g,t+1}^*\} = \arg \max (A_{t+1} H_{g,t+1} + A_t H_{s,t+1})^\alpha K_{t+1}^{1-\alpha} - r_{t+1} K_{t+1} - w_{s,t+1} H_{s,t+1} - w_{g,t+1} H_{g,t+1}.$$

The inverse demand functions for these production factors are

$$r_{t+1} = (1 - \alpha) k_{t+1}^{-\alpha}, \tag{2}$$

$$w_{s,t+1} = \alpha A_t k_{t+1}^{1-\alpha}, \tag{3}$$

$$w_{g,t+1} = \alpha A_{t+1} k_{t+1}^{1-\alpha}, \tag{4}$$

where  $k_{t+1} \equiv K_{t+1}/H_{t+1}$  is the ratio between the stock of physical capital and the aggregate human capital. Not surprisingly, the wage rate of general human capital is higher than that of specific skills, and both increase in the capital-labor ratio  $k_{t+1}$ .

**Investment in Physical and Human Capital.** The initial income of individuals is assumed to be larger than  $Z$  so that bequests are positive.<sup>9</sup> In any period  $t$ , the bequest  $b_t$  is allocated among  $m_{kt}$  and  $m_{st}$  or  $m_{gt}$  to maximize a child's adult income minus any relevant investing cost.<sup>10</sup> So the relevant objective function is  $\max\{I_{s,t+1}^*, I_{g,t+1}^* - C(a_i, h_{gti})\}$  where

$$\begin{aligned} I_{s,t+1}^* &\equiv \max_{m_{st}} r_{t+1}(b_t - m_{st}) + w_{s,t+1}h(m_{st}) + w_{s,t+1}, \\ I_{g,t+1}^* &\equiv \max_{m_{gt}} r_{t+1}(b_t - m_{gt}) + w_{g,t+1}h(m_{gt}) + w_{s,t+1}, \end{aligned}$$

taking as given  $r_{t+1}$ ,  $w_{s,t+1}$ , and  $w_{g,t+1}$ . Income  $I_{s,t+1}^*$  is obtained when a child invests in specific human capital, while  $I_{g,t+1}^*$  applies when he invests in general human capital.

Define  $\psi(k_{t+1})$  as the extra income of having general education than specific education:

$$\psi(k_{t+1}) \equiv I_{g,t+1}^* - I_{s,t+1}^*. \quad (5)$$

The following lemma shows that it is more beneficial to pursue general education when the capital-labor ratio  $k_{t+1}$  is higher.

**Lemma 1**  $\psi(k_{t+1}) > 0$  and  $\psi'(k_{t+1}) > 0$ .

An individual  $i$  will invest in general education if and only if  $I_{g,t+1}^* - C(a_i, h_{gti}) \geq I_{s,t+1}^*$ , which is equivalent to  $\psi(k_{t+1}) \geq C(a_i, h_{gti})$ . As individuals are identical except for their innate abilities, it is straightforward to establish the following lemma.

**Lemma 2** *There exists a unique threshold ability  $a_{t+1}^*$  in each generation  $t + 1$  where*

$$\psi(k_{t+1}) - C(a_{t+1}^*, 0) = 0 \quad (6)$$

---

<sup>9</sup>When the initial income is smaller than  $Z$ , there is no capital investment in the initial periods. Suppose the first period that investment starts is  $t_0$  so that  $I_{t_0} = Z$  holds, which uniquely determines  $t_0$ , since the income  $I_t$  strictly increases over time due to  $g > 0$ . Then we can use period  $t_0$  as our new initial period.

<sup>10</sup>The subscript  $i$  is suppressed whenever possible to simplify the notations. Here we implicitly assume that the child makes the allocation decision, possibly under parental guidance.

such that individuals with higher abilities  $a_i \geq a_{t+1}^*$  invest  $m_{gt}^*$  in general education, while others invest  $m_{st}^*$  in specific education. For interior solutions,  $m_{gt}^*$  and  $m_{st}^*$  are determined respectively by

$$\alpha A_{t+1} k_{t+1} h'(m_{gt}^*) = 1 - \alpha, \quad (7)$$

$$\alpha A_t k_{t+1} h'(m_{st}^*) = 1 - \alpha. \quad (8)$$

The threshold ability  $a_{t+1}^*$  declines over time as  $k_{t+1}$  increases.

This lemma suggests that in each period  $t + 1$  it is optimal for individuals with lower abilities than the threshold level  $a_{t+1}^*$  to invest in specific education  $m_{st}^*$  and others to invest in general education  $m_{gt}^*$ . And over time more and more people switch from specific to general education as the threshold ability  $a_{t+1}^*$  declines. Note that the optimal education investments  $m_{gt}^*$  and  $m_{st}^*$ , once positive, are independent of individual abilities and thus constant across individuals.

**Market Equilibrium.** In the market equilibrium, the demand and supply of physical capital as well as those of the two types of human capital are equal:

$$H_{s,t+1}^* = NF(a_{t+1}^*)h(m_{st}^*) + N,$$

$$H_{g,t+1}^* = N(1 - F(a_{t+1}^*))h(m_{gt}^*),$$

$$K_{t+1}^* = B_t - NF(a_{t+1}^*)m_{st}^* - N(1 - F(a_{t+1}^*))m_{gt}^*,$$

which determine their respective market prices in each period.

**Educational Change.** Let  $t_1$  denote the first time that specific human capital starts to accumulate. The comparative statics of  $a_{t+1}^*$  imply that when the capital-labor ratio  $k_{t+1}$  is too small,  $a_{t+1}^* > a_h$  is possible for some periods so that no one acquires general education. Let  $t_2$  denote the first time when (6) holds for individuals with the highest ability  $a_h$ , or equivalently,

$$\psi(k_{t_2}) - C(a_h, 0) = 0. \quad (9)$$

Then general human capital starts to accumulate from  $t_2$  onwards. The following proposition shows that when the exogenous knowledge accumulation speed  $\phi(0)$  is small enough we get  $t_1 < t_2$ , that is, specific education precedes general education.

**Proposition 1** *Before  $t_1$  all savings are invested in physical capital accumulation, where  $t_1$  is determined by*

$$A_{t_1-1}k_{t_1} = \frac{1-\alpha}{\alpha\gamma}. \quad (10)$$

*When  $\phi(0)$  is small enough, specific education starts earlier than general education. i.e.,  $0 < t_1 < t_2$ .*

The intuition is as follows. Before  $t_1$ , no one invests in human capital because investing in physical capital yields higher returns, as there is a natural endowment of specific human capital  $N$  even without investment. Only when the capital-labor ratio exceeds a certain threshold level, which happens after period  $t_1$ , does the investment in human capital start. During the early periods  $t \in [t_1, t_2]$ , only specific human capital is invested, as the knowledge gap across periods,  $\phi(0)A_t$ , is too small to justify the investing cost of general education.<sup>11</sup> After period  $t_2$ , however, the knowledge gap becomes large enough for the ablest individuals to acquire general education, and as a result, knowledge starts to accumulate at ever faster speeds  $\phi(H_{gt}) > \phi(0)$  than before, gradually making it beneficial for less able individuals to invest in general education, too. Eventually, when the knowledge stock increases fast enough, all individuals will invest in general education after a certain period  $t_3$ , where  $t_3$  is uniquely determined by

$$\psi(k_{t_3}) - C(a_l, 0) = 0. \quad (11)$$

It is obvious that  $t_2 < t_3$  must hold.

The evolution of these two types of human capital over time is illustrated in Figure 1. The average level of specific human capital remains at the endowment level 1 before period  $t_1$ , since then it gradually increases over time as all individuals start to acquire specific education from  $t_1$ ; it starts to decline some time after  $t_2$  when the ablest individuals switch from specific to general education, and eventually returns back to its initial level 1 at period  $t_3$  when all have switched. This inverted-U shape of specific education is observed in many countries during their economic development process (Bertocchi and Spagat 2004). The accumulation of general human capital, in contrast, starts later (at period  $t_2$ ) than specific education but its stock keeps increasing over time.

---

<sup>11</sup>In the alternative case with large  $\phi(0)$ , it is possible for general education to precede specific education. Then the political economy model can be used to explain why the opposite path occurs for mass education. See Section 6 for further discussions.



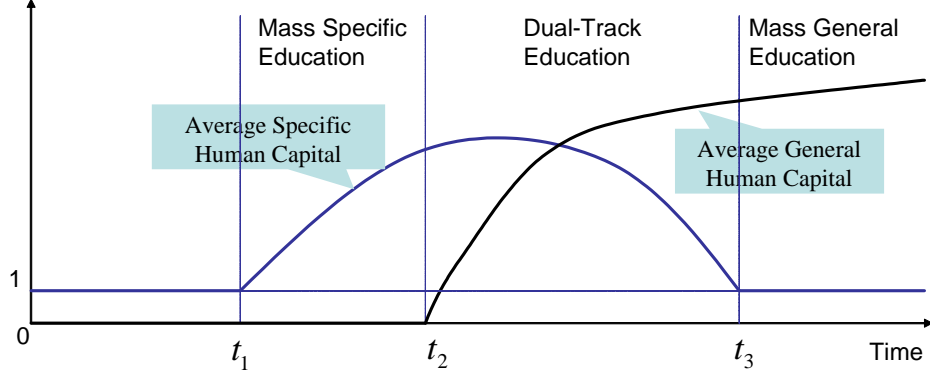


Figure 1: Educational Change in the Basic Model: From Specific to General Education

This benchmark model demonstrates the implications on the educational change by different technical features of general and specific education. They are consistent with the change of mass education, but not with elite education. Indeed, to account for *differences* between education of the elite and the masses, a political economy model is needed. Furthermore, note that investment in general human capital is inefficient due to its positive externality on the knowledge stock and hence on productivity. That is, compared to the social optimal result, individuals tend to invest too late and too little in general human capital. Such inefficiencies may be partially mitigated in a political economy model where rent-seeking opportunities are available.

## 5 Educational Change: A Political Economy Model

In the beginning of the political economy model, individuals are of two types, the ruling elite and the masses. Each elite member is endowed with an identical amount of physical capital  $K_0/N_e$ , where  $N_e$  denotes the measure of the elite, while the masses have no capital endowment. The elite have the political power and thus can impose tax and other public policies on the masses.

The tax-extracting ability of an elite member is increasing and concave in his general human capital  $h_{gt}^e$ , but decreases in the average general human capital of the masses denoted by  $\bar{h}_{gt}^p$ . Specifically, an elite individual gets a tax revenue

$$\tau(h_{gt}^e)(1 - d(\bar{h}_{gt}^p))I_t^p,$$

where  $\tau' > 0$ ,  $\tau'' < 0$ , and  $\tau(\cdot) \leq 1/N_e$ .  $I_t^p$  denotes the aggregate income of the masses;  $d(\bar{h}_{gt}^p)$  denotes the ability of the masses to counter-act the elite's political ruling and tax extraction, where  $d' > 0$  and  $d(0) = 0$  for normalization.

Let  $\hat{h}$  denote the threshold level of workers' average general human capital such that

$$1 - d(\hat{h}) \equiv 0.$$

Note that when the masses acquire enough general human capital such that  $\bar{h}_{gt}^p \geq \hat{h}$ , their abilities to defend their own interests become so high that the tax revenue of the elite becomes zero or negative, since  $1 - d(\bar{h}_{gt}^p) \leq 1 - d(\hat{h}) = 0$ . In other words, the cost of ruling exceeds the benefit for the elite when  $\bar{h}_{gt}^p \geq \hat{h}$ , and so they are better off by extending political power to the masses. That is, the political transition from the elite rule to democracy occurs when the masses become too difficult to rule.<sup>12</sup> In this context, *democracy* is defined as the political system where all individuals share the political power so that there is no expropriating tax, and public policies are chosen through majority voting by all individuals. The other elements of the model are the same as before.

In this political economy model, general human capital enhances an individual's political bargaining ability as represented by the tax-generating skill  $\tau(h_{gt}^e)$  of the elite and the tax-evading skill  $d(\bar{h}_{gt}^p)$  of the masses, while specific human capital does not or is at least less effective in doing so. This is a natural implication of the essence of general human capital, which is the ability to transform knowledge into productivity in whatever task at hand; the specific human capital, in contrast, only increases an individual's productivity in the specific task that he is trained.

## 5.1 Elite Education

An individual elite's objective function is  $\max\{I_{s,t+1}^{e*}, I_{g,t+1}^{e*} - C(a_i, h_{gti}^e)\}$  where

$$\begin{aligned} I_{s,t+1}^{e*} &\equiv \max_{m_{st}^e} r_{t+1}(b_t^e - m_{st}^e) + w_{s,t+1}h(m_{st}^e) + w_{s,t+1} + \tau(0)(1 - d(\bar{h}_{g,t+1}^p))I_{t+1}^p, \\ I_{g,t+1}^{e*} &\equiv \max_{m_{gt}^e} r_{t+1}(b_t^e - m_{gt}^e) + w_{g,t+1}h(m_{gt}^e) + w_{s,t+1} + \tau(h(m_{gt}^e))(1 - d(\bar{h}_{g,t+1}^p))I_{t+1}^p. \end{aligned}$$

---

<sup>12</sup>Other models of political transition, either forced or voluntary (see Huang 2008 for more discussions), would not change the main results on educational change in this paper.

The two new terms that associate with tax revenues are the only structural difference from a typical individual's incomes in the benchmark case.

The net benefit of having general education over specific education for an elite

$$\psi^e(k_{t+1}) \equiv I_{g,t+1}^{e*} - I_{s,t+1}^{e*}$$

is higher than that in the benchmark case, since

$$\psi^e(k_{t+1}) = \psi(k_{t+1}) + \Delta\tau_{t+1},$$

where  $\Delta\tau_{t+1} \equiv (\tau(h(m_{gt}^e)) - \tau(0))(1 - d(\bar{h}_{g,t+1}^p))I_{t+1}^p > 0$  is the extra tax revenue obtained by an elite member with general human capital  $h(m_{gt}^e)$ . This means that, *ceteris paribus*, being in the ruling group makes one more likely to invest in general human capital than in specific skills, and the more so when the extra tax revenue  $\Delta\tau_{t+1}$  is larger.

Let  $a_{t+1}^{e*}$  denote the threshold ability among the elite to invest in general education. Similar as (6), it is uniquely determined by

$$\psi^e(k_{t+1}) - C(a_{t+1}^{e*}, 0) = 0.$$

Let  $t_{he}$  denote the first time the elite start to invest in general education and  $t_{le}$  denote the period after which all elite members acquire general education; then following similar arguments as in (9) and (11) of the basic model,  $t_{he}$  and  $t_{le}$  are determined respectively by

$$\psi^e(k_{t_{he}}) - C(a_h, 0) = 0,$$

$$\psi^e(k_{t_{le}}) - C(a_l, 0) = 0.$$

**Proposition 2** *The elite education starts earlier and invests more compared with the benchmark case; when the potential tax revenue is large enough, the elite invest only in general education. Specifically,  $t_{he} < t_2$  and  $t_{le} < t_3$ .*

In this political economy model, concentration of capital endowment in the few elite and the extra tax revenue they can extract are two channels for the society to have a higher total bequest, and hence a larger aggregate stock of physical capital in each period. As a result, if the elite ever

invest in specific or general human capital, they will do so at a much earlier time and invest more in it than in the benchmark model. Furthermore, this proposition suggests that when the potential tax revenue is large enough, the elite may skip the stage of specific education to start directly with general education, which is in sharp contrast with the benchmark case where specific education precedes general education. Such a reverse of order in educational change arises because general education, by increasing one's skills in political rent seeking, brings high returns to individuals over and above its role in production.<sup>13</sup>

## 5.2 Stratified Education under Elite Rule

Consistent with broad historical trends, we assume that, even after period  $t_{le}$ , the after-tax income of workers is still too low for them to have bequest. The elite, however, may find it beneficial for their own sake to provide free public education, since it increases workers' productivity and hence the tax revenue. As specific education is less harmful than general education in raising workers' political threat to the elite, it is the natural focus of public education under the elite rule.

Let  $m_{gt}^p$  and  $m_{st}^p$  denote per student public education investment in general and specific education, respectively. The elite choose the optimal general education level  $m_{gt}^{p*}$  to maximize their joint pool of tax revenue  $T_{g,t+1}$  taking into consideration the public education expenditure  $(N - N_e)m_{gt}^p$ , where

$$T_{g,t+1}^* \equiv \max_{m_{gt}^p} (1 - d(\bar{h}(m_{gt}^p)))(I_{t+1}^p - (N - N_e)m_{gt}^p).$$

The optimal specific education level  $m_{st}^{p*}$  is similarly chosen to get

$$T_{s,t+1}^* \equiv \max_{m_{st}^p} (1 - d(\bar{h}_{g,t+1}^p))(I_{t+1}^p - (N - N_e)m_{st}^p),$$

taking as given the average general human capital  $\bar{h}_{g,t+1}^p$  of the masses, which stays at the basic level 0 for most times under the elite rule and starts to increase only after period  $t_{2e}$  determined in (13). The optimal choices are summarized in the following proposition.

---

<sup>13</sup>The propertied elite, for example, could have taken specific education in order to enhance the management and values of their properties rather than going through classical studies and assuming positions in the state and church. The fact that the opposite did happen in history is what this proposition tries to account for.

**Proposition 3** *Under the elite rule, free mass education in specific skills starts in some period  $t_{1e}$ , which is determined by*

$$A_{t_{1e}} k_{t_{1e}+1}^{1-\alpha} = 1/\alpha\gamma, \quad (12)$$

and  $t_{1e} > t_{le}$ . *General education for the masses will never be supported when  $d'(0) \geq 1$  or as long as  $h_{gt}^{e*} \leq \tilde{h}$ , where  $\phi(N_e \tilde{h}) = d'(0)/(1 - d'(0))$ ; and if ever supported, it will never go beyond  $\hat{h}$ .*

This proposition suggests that when the capital-labor ratio  $k_{t+1}$  becomes large enough, investing in physical capital becomes less profitable than investing in public education to increase the human capital of the masses and thus their productivity. Then the decision on which type of education, general or specific education, to provide for the masses depends on the trade-off between their different effects on increasing the tax base versus reducing the tax rate; the more effective general education is in lowering the tax rate through  $1 - d(\bar{h}_{g,t+1}^p)$  than enlarging the tax base  $I_{t+1}^p$ , the less likely it is to be supported by the elite. In fact, it will never be supported under the elite rule when  $d'(0) \geq 1$  holds, in which case general education is too powerful in enabling the masses to resist the elite rule and evade taxes. Though it is possible for the elite to provide general education for the masses when it is less harmful politically ( $d'(0) < 1$ ) and when the elite are skillful enough in political ruling ( $h_{gt}^{e*} > \tilde{h}$ ), the investment of mass general education is capped above because the elite rule cannot be sustained once the masses' average general human capital reaches the threshold level  $\hat{h}$ .

The following analysis focuses on the case with  $d'(0) \geq 1$  where the elite won't support mass general education. So from  $t_{1e}$  onwards, workers begin to acquire specific human capital through public education. This enables the income of workers to grow at a faster speed than before, and it will eventually reach the threshold  $Z$  at some period  $t_Z$ .<sup>14</sup> Then from  $t_Z$  onwards workers start to have bequest and may choose to invest in general education  $h_{gt}^{p*}$  by themselves.

The net benefit of investing in general education over receiving free specific education is

$$\psi^p(k_{t+1}) \equiv [1 - N_e \tau(h_{gt+1}^{e*})(1 - d(\bar{h}_{g,t+1}^p))][w_{g,t+1} h(m_{gt}^{p*}) - w_{s,t+1} h(m_{st}^{p*}) - r_{t+1} m_{gt}^{p*}]$$

---

<sup>14</sup>Note that  $t_Z > t_{1e}$  is guaranteed when  $1/\gamma < Z$ , since  $w_{s,t_{1e}+1} = \alpha A_{t_{1e}} k_{t_{1e}+1}^{1-\alpha} = 1/\gamma$  by conditions (3) and (12).

for a worker, which is smaller than  $\psi(k_{t+1})$  in the benchmark case not only because of the confiscatory tax, but also because of the free specific education provided by the elite; both elements discourage the poor's investment in general education and thus help preserve the elite rule.

When the knowledge stock becomes large enough, however, workers with high abilities may find it beneficial to invest in general education  $h_{gt}^{p*}$ , instead of receiving the free specific education. Let  $t_{2e}$  denote the first period when the ablest workers start to acquire general education. Then following similar arguments as before it is uniquely determined by

$$\psi^P(k_{t_{2e}+1}) - C(a_h, 0) = 0. \tag{13}$$

The level of general education  $h_{gt}^{p*}$  among the masses, however, is still lower than that of the elite, since the masses lack the rent-seeking opportunity to generate tax revenues.

### 5.3 Mass Education under Democracy: From Dual-Track to General Education

Let  $t_m$  denote the first period when the threshold level of workers' average general human capital  $\widehat{h}$  is reached; that is, from period  $t_m$  onwards,  $d(\overline{h}_{g,t}^p) \geq 1$  holds so that the ruling cost becomes too high for the elite to obtain any tax revenues. This either forces or induces the elite to extend political power to the masses.<sup>15</sup> Without loss of generality, we assume that at period  $t_m$  democracy replaces the elite rule so that confiscatory tax is abolished.

The education policy under democracy is characterized by  $(m_{gt}^{d*}, m_{st}^{d*})$ , where  $m_{gt}^{d*}$  and  $m_{st}^{d*}$  denote each person's contribution to the public investment in general and specific education, respectively. It is determined by majority voting, where each individual votes for the policy that is closest to his favorite choice of public education investment, which maximizes the future income of his child. Opting out of public specific education to pursue privately financed general education is allowed, in which case an individual still needs to pay his share of the education tax.<sup>16</sup> If mass

---

<sup>15</sup>The elite may try to control how the masses are educated in order to prevent  $\overline{h}_{g,t}^p$  from being too high, but their success is not always guaranteed. The declining cost of print, for example, may have contributed to the failure of such elite control.

<sup>16</sup>If individuals opting out of public specific education do not need to pay their share of the education tax, it essentially leads to a similar situation as if the education choice is purely left to individuals; in this case their choice problems become exactly the same as in the benchmark case, where mass education also evolves from dual-track to general education.

general education is publicly provided, very few, if any, would choose to opt out of it, and so to simplify exposition, we assume that opting out of public general education is not allowed by law.

If a child is going to take general education, his future income will be

$$I_{g,t+1}^{d*} \equiv \max_{m_{gt}^d, m_{st}^d} r_{t+1}(b_{ti} - m_{st}^d - m_{gt}^d) + w_{g,t+1}h(m_{gt}^d) + w_{s,t+1} - C(a_i, h_{gti}).$$

The FOC for the interior solution  $m_{gt}^{d*} > 0$  is  $\alpha A_t k_{t+1} h'(m_{gt}^{d*}) = 1 - \alpha$ , the same as (7) in the benchmark case. So for all individuals who take general education,  $m_{gt}^{d*}$  will be identical, independent of a child's ability and parental bequest or education level. It is straightforward to see that one's preferred public investment in specific education is zero if his child takes general education.

Similarly, if the child is going to take specific education, his future income will be

$$I_{s,t+1}^{d*} \equiv \max_{m_{gt}^d, m_{st}^d} r_{t+1}(b_{ti} - m_{st}^d - m_{gt}^d) + w_{s,t+1}h(m_{st}^d + \varepsilon^*) + w_{s,t+1},$$

where  $\varepsilon^*$  denotes the extra educational investment that arises if some individuals opt out of public specific education but still pay the educational tax  $m_{st}^{d*}$ . The FOC for the interior solution  $m_{st}^{d*} > 0$  is  $\alpha A_t k_{t+1} h'(m_{st}^{d*} + \varepsilon^*) = 1 - \alpha$ , which implies that  $m_{st}^{d*}$  will be the same for all individuals. And one's preferred public investment in general education is zero if his child takes specific education.

These results imply that, conditional on the type of education, there is a consensus over the amount of investment. The conflict of interest thus lies in the different choices of education type across individuals. If a simple majority of people choose to attend public specific education, then the dual-track education policy will be the equilibrium choice under majority voting; if the opposite is true, then general education will instead be publicly financed.

If free public specific education is provided, every individual has to pay the education tax  $m_{st}^{d*}$ , and those who choose to pursue general education instead need to pay the extra investment  $m_{gt}^{d*}$ ; so the net benefit of having general education over specific education is

$$\psi^d(k_{t+1}) \equiv I_{s,t+1}^{d*} - I_{g,t+1}^{d*} = w_{g,t+1}h(m_{gt}^{d*}) - w_{s,t+1}h(m_{st}^{d*} + \varepsilon^*) - r_{t+1}m_{gt}^{d*},$$

which has to be at least as large as the cost  $C(a_i, h_{gti})$  for an individual to choose general education.

Since the benefit of general education  $\psi^d(k_{t+1})$  under democracy is larger than  $\psi^p(k_{t+1})$  under the elite rule, workers whose parents acquired general education under the elite rule will continue

to do so, and so will workers who are slightly less smart. Let  $a_t^w$  denote the ability of the threshold worker that invests in general education. It is thus uniquely determined by

$$\psi^d(k_{t+1}) = C(a_t^w, 0).$$

It is less clear whether children of the old elite will still acquire general education, since several forces affect their choices in the opposite directions. On the one hand, the knowledge stock is larger than before and their parents' general education levels are higher than before and than others, both of which would make them more likely to invest in general education; on the other hand, the disappearing of the political rent would reduce the benefit of having general education. What happened in history seems to suggest that the old elite children will continue to acquire general education, which is assumed here. So if specific education is to be publicly supported, individuals from the old elite families and workers with  $a_i \geq a_t^w$  will opt out of public specific education and pursue privately financed general education, while the rest with measure  $(N - N_e)F(a_t^w)$  will attend it.<sup>17</sup>

Let  $a^{median}$  denote the threshold ability level that is determined by

$$N_e/N + (1 - F(a^{median}))(1 - N_e/N) = \frac{1}{2}.$$

When  $a^{median} < a_t^w$  holds, less than half of the population choose to invest in general education, so free specific education will be provided under majority voting. So the lower  $a^{median}$  is, the more likely the specific education is publicly provided and the later the free general education policy is to be adopted. Note that, based on the above identity,

$$\frac{\partial a^{median}}{\partial (N_e/N)} = -\frac{F(a^{median})}{-F'(a^{median})(1 - N_e/N)} > 0,$$

which implies that a smaller number of old elite leads to a lower  $a^{median}$  and thus delays the arrival of mass general education under democracy. As the knowledge stock keeps increasing over time,  $a_t^w$  decreases so that more and more workers will choose general education. In a certain period  $t_{3m}$  such that

$$a_{t_{3m}}^w \geq a^{median}$$

---

<sup>17</sup>This means  $\varepsilon^* = \left(\frac{N}{(N - N_e)F(a_t^w)} - 1\right)m_{st}^{d*}$  must hold for a balanced budget for public specific education, and it is uniquely determined.



holds for the first time, the education policy will switch from the dual-track system to free general education for all. Thus we have proved the following proposition.

**Proposition 4** *Under democracy, the dual-track education system of free specific education plus private general education is more likely to continue and to last longer when  $N_e/N$  is smaller (or when the initial income inequality is higher); it will eventually be replaced by free general education for all.*

It is useful to note that the existence of dual-track education system in this proposition is not due to the usual suspect that insiders who receive general education protect their own privileges by restricting access to general education (Collins 1971, Bertocchi and Spagat 2004). The driving force is the heterogenous costs among individuals: Those with higher abilities and with parents who already acquired general education (the old elite belong to this group) have lower costs  $C(a_i, h_{gti})$  than others. When the initial income inequality is high, a large proportion of the population have high costs of investing in general education and thus prefer to continue with free specific education. In other words, the dual-track education system under democracy is the rational choice of the majority, not the conspiracy of the privileged. The majority support for mass general education, however, will eventually arise when the knowledge stock changes fast enough to warrant so. That said, one should be aware of the positive externalities of general education over knowledge accumulation, and hence an earlier adoption of mass general education than mandated by majority voting can be welfare improving.

## 5.4 Summary

The timing of educational change in this political economy model is illustrated in Figure 2. The elite start to invest in general education after  $t_{he}$ , before which all savings are invested in physical capital accumulation. As the aggregate stock of general human capital, though so far composed completely by that of the elite, keeps increasing, the knowledge stock expands at ever faster speeds than before, and it eventually induces the elite for their own interests to support public specific education for workers starting from  $t_{1e}$ . Though the elite may never want to support general education for the masses, the ablest workers invest in general education themselves from period  $t_{2e}$  onwards. When

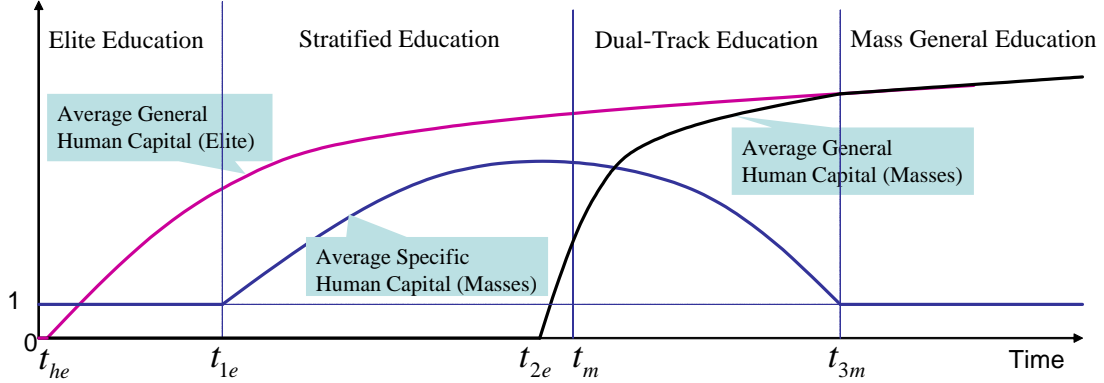


Figure 2: Educational Change in the Political Economy Model: From Elite to Mass General Education

the average general human capital of workers reaches the threshold level  $\hat{h}$  in period  $t_m$ , the elite rule becomes too costly to maintain and is thus replaced by democracy, under which there is no expropriating tax and public education is chosen by majority voting. The dual-track education system includes public specific education for majority and private general education for the rest, which will finally give way to public general education for all after period  $t_{3m}$ .

The existence of political rent-seeking opportunities facilitates investment in general human capital for the elite and encourages public education in specific skills for workers; it thus may increase productivity in general, especially when the distortion caused by such rent-seeking activities is relatively minor. In the later stage of development, however, it tends to prolong specific education and delay investment in general human capital for the masses. More importantly, this political economy model accounts for the sharp contrasts between the elite education and mass education in terms of timing and content: elite education starts much earlier and focuses on general education, while mass education starts much later and initially focuses on specific education; though mass specific education is compatible with and thus can be well supported by the elite rule, mass general education is much less so, and hence most likely to occur under democracy.

## 6 Extensions and Discussions

**Humanities versus Natural Sciences.** In the model, the contents of general education are assumed homogenous in terms of their contributions to productivity and rent-seeking effectiveness. They are not necessarily so. For example, humanities seem to be more effective in increasing one's rent-seeking skills than natural sciences, while the opposite tends to be true for improving production productivity in general (Murphy, Shleifer, Vishny 1991).

The model can be readily extended to accommodate such heterogeneity in general education. For example, the production function can be adjusted by replacing  $H_{g,t+1}$  with  $\sigma_1 H_{g,t+1}^R + \sigma_2 H_{g,t+1}^N$ , where  $H_{g,t+1}^R$  denotes the aggregate human capital in humanities,  $H_{g,t+1}^N$  denotes the aggregate human capital in natural sciences, and  $\sigma_1 \leq \sigma_2$ ; the rent-seeking skill  $\tau(h_{g,t+1})$  can be changed to  $\tau(\sigma_3 h_{g,t+1}^R + \sigma_4 h_{g,t+1}^N)$ , where  $\sigma_3 \geq \sigma_4$ ; and finally an individual can invest simultaneously in these two types of general human capital.

Following similar arguments as in the model, it is easy to get the following results. In the basic model without rent-seeking opportunities, individuals will invest more in natural sciences than in humanities, which is also true under democracy in the political economy model. Under the elite rule, in contrast, the elite will invest more in humanities than in natural sciences, the more so if the gains from rent-seeking are larger. And everything else equal, societies with relatively more gains from rent-seeking than others would invest more in humanities but less in natural sciences. It would be interesting to test these results empirically in future research.

**Closed-Elite versus Open-Elite.** The political economy model adopted in this paper is a model of closed-elite in that the elite group is fixed and determined by birth, and as a result only a fixed group in society is entitled to gains from rent-seeking. An alternative is an open-elite model, such as the bureaucratic system in feudal China, where the elite were recruited through standard exams that were open to all members in society (Collins 2000). In such a model, the access to rent-seeking gains is endogenous, since anyone can choose to invest in general education to become the elite, if the benefit of doing so outweighs the cost.

From a naive point of view, the open-elite model seems to be better since it gives everyone a more equal chance to join the elite than the closed-elite model. A quite surprising result emerging from

our analysis shows that the opposite is true, however, at least from the whole society’s perspective. Following similar arguments in the model, one can find that the open-elite system is more effective in stabilizing the elite rule in comparison to the closed-elite system, since the richest and smartest individuals will invest in general education and join the elite, which implies that the repressed masses get either no education at all or only specific education and thus have much limited ability to achieve political power and overthrow the elite rule. Furthermore, societies with the open-elite system tend to over-invest in humanities and under-invest in natural sciences than those with the closed-elite system. Both results seem consistent with the contrasting differences between feudal China and western Europe.

**Educational Differences Across Countries.** An implication of the model is that general education of the masses starts earlier and thus its knowledge stock grows faster in society with a more diverse talent pool. This may speak to the educational differences between the U.S. and Europe where vocational education is emphasized more in Europe than in the U.S. (Goldin 2001, Krueger and Kumar 2004b). A possible reason could be that the U.S. as an immigrant country has a relative more diverse distribution of talent and hence started mass general education earlier. Specifically, suppose that talent  $a_i$  is distributed on  $[a_l^U, a_h^U]$  in society U and on  $[a_l^E, a_h^E]$  in society E, where  $a_l^U < a_l^E$ ,  $a_h^U > a_h^E$ , and the median is kept the same. So abilities are more diversely distributed in society U than in society E. Since  $C_1 < 0$  and  $a_h^U > a_h^E$ , it follows from (13) that  $t_{2e}^U < t_{2e}^E$  must hold, which means that general education of the masses starts earlier in society U than in society E; and this in turn leads to earlier democratization ( $t_m^U < t_m^E$ ) and earlier beginning of mass general education ( $t_{3m}^U < t_{3m}^E$ ).

**Discussions on Modeling Choices.** As briefly mentioned in the Introduction, the model focuses on the changing level and content of *organized* learning conducted in schools. This is by no means intended to deny the ubiquitous presence of informal learning in daily life, which long preceded any form of organized learning in history and has always been playing an important role in knowledge accumulation. In fact, its existence is acknowledged in the assumption that each individual is endowed with a basic unit of specific human capital even without organized education, and its indispensable role is also reflected in the assumption  $\phi(0) > 0$ , which means

that the knowledge stock still grows, though slowly, even when there is no organized education. Furthermore, investment in organized education becomes worthwhile in the model only after the knowledge stock accumulated by such informal learning becomes large enough.

This paper does not explicitly model why learning becomes organized in schools at some point for a couple of reasons. First, the exact organization of learning presumably involves different economic forces than those considered in the model, and thus its analysis is beyond the scope of this paper. Second, considering it explicitly is not likely to change the main results of this paper. One may argue, for example, that if the economy of scale is more easily reached in general education than vocational education, this may explain why organized education began from universities among the elite.<sup>18</sup> This argument by itself, however, does not explain why, on the contrary, vocational education preceded general education for the masses. A plausible explanation is readily provided by the same political economy model as before: Even though it is best in this scenario to provide general education for the masses from pure economic considerations, the elite may choose the opposite because of the political concerns that general education makes the masses more difficult to rule. This suggests that the political economy model captures some fundamental forces underlying the different paths taken by elite and mass education, which are the main focus of this paper.

## 7 Concluding Remarks

For many years in history only the few ruling elite were educated and the elite education focused on general education rather than technical or specific skills, while the masses started to receive education in the nineteenth century in western Europe together with industrialization; and the mass education was initially focusing on specific vocational skills and then gradually moving to general education. Such educational change in terms of who are educated and what they learn is accounted for in this paper in a framework of long run economic and political development.

The main results of this paper suggest that the educational change can be categorized into three phases, namely, the elite education phase where only the elite are educated and it focuses on general education, the stratified education phase that includes both general education for the

---

<sup>18</sup>This may also be achieved by assuming large enough  $\phi(0)$ ; see Proposition 1.

elite as before and specific education for the masses, and finally, the mass education phase under democracy, which continues with the dual-track education system including general education for some and specific education for majority, and eventually converges to providing general education for all. These results are broadly consistent with the historical evidence.

More interestingly, the paper finds that, though mass specific education may be supported under the elite rule, this is very unlikely for mass general education, which occurs typically under democracy. In other words, the elite rule is often associated with a lack of general education for the masses. This is probably driven by an important difference between general and specific education, where general education is more likely to increase an individual's skills in whatever tasks at hand, which by definition include skills of extracting gains and defending one's own interests in political conflicts. The same insights suggest that if humanities and social sciences are more effective in improving one's rent-seeking skills than natural sciences, the elite will invest more in the former than in the latter, the more so in societies where the gains from rent-seeking are larger. As shown in the paper, the model can also be readily extended in other ways to yield intriguing results that may account for cross-country differences in education and its effects on economic and political development.

## References

- [1] Acemoglu, Daron, Simon Johnson, James A. Robinson and Pierre Yared. (2005). "From Education to Democracy?" *American Economic Review* 95(2), 44-49.
- [2] Becker, Gary. (1964). *Human Capital: A Theoretical and Empirical Analysis*. Chicago: University of Chicago Press.
- [3] Bennett, Jr., William S. (1967). "Educational Change and Economic Development," *Sociology of Education* 40(2), 101-114.
- [4] Bertocchi, Graziella, and Michael Spagat. (2004). "The Evolution of Modern Educational Systems: Technical vs. General Education, Distributional Conflict, and Growth," *Journal of Development Economics* 73, 559-582.

- [5] Bowen, James. (2003). *A History of Western Education*. London : Routledge.
- [6] Bowles, Samuel, and Herbert Gintis. (1976). *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life*. New York: Basic Books.
- [7] Brezis, Elise S., and Francois Crouzet. (2006). "The Role of Higher Education Institutions: Recruitment of Elites and Economic Growth." In T. Eicher (eds.) *Institutions and Economic Growth*, MIT Press, 191-213.
- [8] Collins, Randall. (1971). "Functional and Conflict Theories of Educational Stratification," *American Sociological Review* 36 (6), 1002-1019.
- [9] Collins, Randall. (2000). "Comparative and Historical Patterns of Education." In *Handbook of the Sociology of Education*, Maureen T. Hallinan (eds) (2000). Springer. New York.
- [10] Cordasco, Francesco. (1976). *A Brief History of Education*. Roman Littlefield Publishers.
- [11] Cubberley, Ellwood. (1920). *The History of Education : Educational Practice and Progress Considered as a Phase of the Development and Spread of Western Civilization*. Houghton Mifflin, Boston.
- [12] Galor Oded, Omer Moav and Dietrich Vollrath. (2009). "Inequality in Land Ownership, the Emergence of Human Capital Promoting Institutions and the Great Divergence," *Review of Economic Studies* 76, 143-179.
- [13] Galor, Oded, and Daniel Tsiddon. (1997). "Technological Progress, Mobility, and Growth," *American Economic Review* 87, 363-382.
- [14] Goldin, Claudia. (1999). "A Brief History of Education in the United States," NBER Working Paper, Historical Series no. 119.
- [15] Goldin, Claudia. (2001). "The Human Capital Century and American Leadership: Virtues of the Past," *Journal of Economic History* 61, 263-292.

- [16] Gould, Eric, Omer Moav and Bruce Weinberg. (2001). "Precautionary Demand for Education, Inequality and Technological Progress," *Journal of Economic Growth* 6(4), 285-315.
- [17] Gradstein M., and M. Justman. (2002). "Education, Social Cohesion and Growth," *American Economic Review* 92, 1192-1204.
- [18] Green, Andy. (1990). *Education and State Formation: The Rise of Education Systems in England, France and the USA*. New York: St. Martin's Press.
- [19] Grossman, Herschel I., and Minseong Kim. (2003). "Educational Policy: Egalitarian or Elitist?" *Economics and Politics* 15(3), 225-246.
- [20] Huang F. (2008). "The Coevolution of Economic and Political Development from Monarchy to Democracy." Working Paper, Singapore Management University.
- [21] Iyigun, Murat, and Ann Owen. (1999). "Entrepreneurs, Professionals, and Growth," *Journal of Economic Growth* 4(2), 211-30.
- [22] Iyigun, Murat, and Ann Owen. (2006). "Adaptive Skills, Technological Progress, and Growth," *European Economic Review* 50(3), 565-79.
- [23] Krueger, Dirk, and Krishna Kumar. (2004a). "US-Europe Differences in Technology-Driven Growth: Quantifying the Role of Education," *Journal of Monetary Economics* 51(1), 161-190.
- [24] Krueger, Dirk, and Krishna Kumar. (2004b). "Skill-Specific rather than General Education: A Reason for US-Europe Growth Differences?" *Journal of Economic Growth* 9(2), 167-207.
- [25] Lazear, Edward P. (2003). "Firm-Specific Human Capital: A Skill-Weights Approach," NBER Working Papers 9679.
- [26] Lindert, Peter H. (2004). *Growing Public: Social Spending and Economic Growth since the Eighteenth Century*. Cambridge University Press
- [27] Lott, John R., Jr. (1990). "An Explanation for Public Provision of Schooling: The Importance of Indoctrination." *Journal of Law and Economics* 33(1), 199-231.



- [28] Meyer, John W., and Richard Rubinson. (1975). "Education and Political Development," *Review of Research in Education* 3, 134-162.
- [29] Meyer, John W., Francisco O. Ramirez, and Yasemin Soysal. (1992). "World Expansion of Mass Education, 1870-1980" *Sociology of Education* 65(2), 128-149.
- [30] Mokyr, Joel. (2002). *The Gifts of Athena: Historical Origins of the Knowledge Economy*. Princeton University Press, Princeton.
- [31] Mueller, D.K, Ringer, F. and Simon, B. (eds.). (1977). *The Rise of the Modern Educational System: Structural Change and Social Reproduction 1870-1920*, Cambridge University Press, Cambridge.
- [32] Murphy, Kevin, Andrei Shleifer, and Robert Vishny. (1991). "The Allocation of Talent: Implications for Growth," *Quarterly Journal of Economics* (106), 503-530.
- [33] Trow, M. (1967). "The Second Transformation of American Secondary Education," in *Class, Status, and Power*, Bendix, R. and Lipset, S.M.(eds.), Routledge & Kegan Paul, London.
- [34] Ulich, Robert. (1961). *Education of Nations : A Comparison in Historical Perspective*. Harvard University Press. Cambridge, Massachusetts.
- [35] Wasmer, Etienne. (2006). "General versus Specific Skills in Labor Markets with Search Frictions and Firing Costs," *American Economic Review* 96(3), 811-831.

## APPENDIX

### Proof of Lemma 1.

**Proof.** Note that the incomes  $I_{g,t+1}^*$  and  $I_{s,t+1}^*$  arise from a common function  $I_{t+1}^* \equiv r_{t+1}(b_t - m_t^*) + w_{t+1}h(m_t^*) + w_{s,t+1}$  where  $w_{t+1}$  can take two distinct values, either  $w_{s,t+1}$  or  $w_{g,t+1}$ . By the Envelope Theorem we know  $\partial I_{t+1}^*/\partial w_{t+1} = h(m_t^*) > 0$  if  $m_t^* > 0$ , that is, the income  $I_{t+1}^*$  strictly increases in wage rate  $w_{t+1}$ ; but then  $w_{g,t+1} > w_{s,t+1}$  implies  $I_{g,t+1}^* > I_{s,t+1}^*$ , which is equivalent to  $\psi(k_{t+1}) > 0$ .

Observe that

$$\begin{aligned} \psi'(k_{t+1}) &= \frac{\partial I_{g,t+1}^*}{\partial k_{t+1}} - \frac{\partial I_{s,t+1}^*}{\partial k_{t+1}} \\ &= w'_{g,t+1}h(m_{gt}^*) - w'_{s,t+1}h(m_{st}^*) - r'_{t+1}(m_{gt}^* - m_{st}^*) \\ &= \alpha(1 - \alpha)k_{t+1}^{-\alpha}[A_{t+1}h(m_{gt}^*) - A_t h(m_{st}^*)] + \alpha(1 - \alpha)k_{t+1}^{-\alpha-1}(m_{gt}^* - m_{st}^*), \end{aligned}$$

where the second equality follows from the Envelop Theorem. Since  $A_{t+1} \geq A_t$  and  $h' \geq 0$ ,  $\psi'(k_{t+1}) > 0$  holds if  $m_{gt}^* \geq m_{st}^*$ , which is indeed true.

Now we prove that  $m_{gt}^* > m_{st}^*$  holds; that is, when an individual is considering the optimal investment in these two types of human capital, he would invest more resources in acquiring general education than in specific education. For an individual that invests in  $h_{s,t+1}^*$ , his objective function is

$$\max_{m_{st}} r_{t+1}(b_t - m_{st}) + w_{s,t+1}h(m_{st}) + w_{s,t+1}.$$

The FOC after plugging in (2) and (3) becomes

$$A_t k_{t+1} h'(m_{st}^*) = \frac{1 - \alpha}{\alpha} \text{ for } m_{st}^* > 0, \quad (14)$$

$$A_t k_{t+1} \leq \frac{1 - \alpha}{\alpha \gamma} \text{ for } m_{st}^* = 0. \quad (15)$$

So the threshold level for  $A_t k_{t+1}$  is  $\frac{1 - \alpha}{\alpha \gamma}$ , below which no resources are spent in accumulating specific human capital. The optimal amount of specific education  $m_{st}^*$  is a function of  $A_t$  and  $k_{t+1}$  obtained from (14).

For an individual that invests in  $h_{g,t+1}^*$ , his objective function is

$$\max_{m_{gt}} r_{t+1}(b_t - m_{gt}) + w_{g,t+1}h(m_{gt}) + w_{s,t+1} - C(a_i, h_{gti}).$$

The FOC is

$$\begin{aligned} A_{t+1}k_{t+1}h'(m_{gt}) &= \frac{1-\alpha}{\alpha} \text{ for } m_{gt} > 0, \\ A_{t+1}k_{t+1} &\leq \frac{1-\alpha}{\alpha\gamma} \text{ for } m_{gt} = 0. \end{aligned} \tag{16}$$

For interior solutions, the optimal amount of general education  $m_{gt}^*$  is a function of  $A_{t+1}$  and  $k_{t+1}$ , while it is independent of  $a_i$ . This implies that  $h(m_{gt}^*)$  is the same for all who invest in it, regardless of their individual abilities. Note that  $m_{gt}^* > m_{st}^* > 0$  because  $h''(\cdot) < 0$  and  $h'(m_{gt}) = \frac{1-\alpha}{\alpha A_{t+1}k_{t+1}} < \frac{1-\alpha}{\alpha A_t k_{t+1}} = h'(m_{st})$  holds for interior solutions. Since  $A_{t+1} > A_t$ , the threshold condition  $A_{t+1}k_{t+1} = \frac{1-\alpha}{\alpha\gamma}$  for general education investment must hold at some time earlier than that for specific education. This means  $m_{gt}^* \geq m_{st}^*$  is always true. ■

### Proof of Lemma 2.

**Proof.** The investment cost  $C(a_i, h_{gti})$  is strictly decreasing in  $a_i$  as  $C_1 < 0$ ,  $C_2 < 0$  and  $h_{gti} = h(m_{g,t-1}^*)$  does not depend on individual  $a_i$ , which is shown in the above proof. As  $\psi(k_{t+1})$  is independent of  $a_i$  while  $C(a_i, h_{gti})$  strictly decreases in  $a_i$ , a threshold ability  $a_{t+1}^*$  must exist so that the marginal person with ability  $a_{t+1}^*$  is indifferent between investing in either general or specific human capital. Note that when it is the first time that an individual reaches the threshold of starting to invest in general human capital, his parent must have not invested in it so that  $h(m_{g,t-1}^*) = 0$ . The reason is straightforward. If  $h(m_{g,t-1}^*) > 0$ , then  $\psi(k_t) - C(a_{t+1}^*, 0) \geq 0$  must hold for the parent, which together with  $k_{t+1} > k_t$  then implies  $\psi(k_{t+1}) - C(a_{t+1}^*, 0) > 0$  is true for the child. But this contradicts with (6). Based on (6) we have

$$\frac{\partial a_{t+1}^*}{\partial k_{t+1}} = \frac{\psi'(k_{t+1})}{C_1(a_{t+1}^*, h(m_{g,t-1}^*))} < 0,$$

since  $C_1 < 0$ . Similarly we can show that  $\partial a_{t+1}^*/\partial A_{t+1} < 0$  and  $\partial a_{t+1}^*/\partial A_t < 0$ . So  $a_{t+1}^*$  strictly decreases over time and in capital-labor ratio  $k_{t+1}$ . ■

### Proof of Proposition 1.

**Proof.** The equality in (15) is achieved in period  $t_1$ , from then on people start to invest in specific human capital. When  $\psi(k_{t_1}) - C(a_h, 0) < 0$ , no one invests in general education before period  $t_1$ , which means  $t_1 < t_2$  will hold. As  $\psi(k_{t_1})$  strictly increases in  $\phi(0)$  and is equal to zero when  $\phi(0) = 0$ , there must exist a unique level of  $\phi(0)$ , denoted by  $\tilde{g}$ , such that  $\psi(k_{t_1}; \tilde{g}) = C(a_h, 0)$ . So a sufficient condition for  $\psi(k_{t_1}) - C(a_h, 0) < 0$  is  $\phi(0) < \tilde{g}$ .

Since there is no human capital investment before period  $t_1$ , we get  $H_{gt} = 0$ ,  $H_{st} = N$ , and  $A_t = A_{t-1}(1 + \phi(0))$  in any period  $t \leq t_1$ ; then  $A_{t+1}k_{t+1} = K_{t+1}(1 + \phi(0))/N$  by definition. So (10) can be rewritten as  $K_{t_1} = \frac{1-\alpha}{\alpha\gamma} \frac{N}{1+\phi(0)}$ . Thus  $K_1 < K_{t_1}$  is sufficient to insure that  $t_1 > 0$ , where  $K_1 = \beta(A_{-1}^\alpha K_0^{1-\alpha} - NZ)$ . This condition is equivalent to  $\phi(0) < \hat{g}$  where  $\hat{g}$  is determined by  $\hat{g} = \frac{(1-\alpha)N}{\alpha\gamma K_1} - 1$ . So any  $\phi(0) \leq \min\{\tilde{g}, \hat{g}\}$  is sufficient to establish the timing  $0 < t_1 < t_2$ . ■

### Proof of Proposition 2.

**Proof.** For an individual elite that invests in general human capital  $h_{g,t+1}^{e*}$ , his objective function is

$$\max_{m_{gt}^e} r_{t+1}(b_t^e - m_{gt}^e) + w_{g,t+1}h(m_{gt}^e) + w_{s,t+1} + \tau(h(m_{gt}^e))(1 - d(\bar{h}_{g,t+1}^p))I_{t+1}^p - C(a_i, h_{gt}^e).$$

The FOC is

$$\begin{aligned} [A_{t+1}k_{t+1} + k_{t+1}^\alpha I_{t+1}^p \tau'(h(m_{gt}^e))(1 - d(\bar{h}_{g,t+1}^p))/\alpha]h'(m_{gt}^e) &= \frac{1-\alpha}{\alpha} \text{ for } m_{gt}^e > 0, \\ A_{t+1}k_{t+1} + k_{t+1}^\alpha I_{t+1}^p \tau'(0)/\alpha &< \frac{1-\alpha}{\alpha\gamma} \text{ for } m_{gt}^e = 0. \end{aligned} \quad (17)$$

For interior solutions, the optimal general education for an elite  $m_{gt}^{e*}$  is higher than that in the benchmark case  $m_{gt}^*$ , since a higher general human capital enables him to get a larger tax revenue as reflected by  $k_{t+1}^\alpha I_{t+1}^p \tau'(h(m_{gt}^e))(1 - d(\bar{h}_{g,t+1}^p))/\alpha$  in the FOC. And the difference is larger when the tax base  $I_{t+1}^p$  is larger. The investment in general education among the elite will start earlier than in the benchmark case due to the positive term  $k_{t+1}^\alpha I_{t+1}^p \tau'(0)/\alpha$ , where  $1 - d(\bar{h}_{g,t+1}^p) = 1$  must hold due to  $\bar{h}_{g,t+1}^p = 0$ , since the education of the poor masses cannot start before the rich elite.

As in the benchmark case, the threshold ability  $a_{t+1}^{e*}$  strictly decreases over time and in capital-labor ratio  $k_{t+1}$  as long as  $\bar{h}_{g,t+1}^p = 0$ , which is true well after  $t_{le}$  as confirmed in the next proposition.

Observe that

$$\psi^{e'}(k_{t+1}) = \psi'(k_{t+1}) + \frac{\partial \Delta \tau_{t+1}}{\partial k_{t+1}} > 0,$$

where  $\partial \Delta \tau_{t+1} / \partial k_{t+1} = \partial(\tau(h(m_{gt}^{e*})) - \tau(0)) I_{t+1}^p / \partial k_{t+1} > 0$  due to  $\partial m_{gt}^{e*} / \partial k_{t+1} > 0$ , which can be obtained by the Implicit Function Theorem from equation (17). Note that  $a_{t+1}^{e*} < a_{t+1}^*$  holds because  $\psi^e(k_{t+1}) > \psi(k_{t+1})$  is true due to the extra positive term  $\Delta \tau_{t+1}$ . Given that  $a_{t+1}^{e*} < a_{t+1}^*$  and both decrease over time, we get immediately  $t_{he} < t_2$  and  $t_{le} < t_3$ .

If one invests in specific human capital, the FOC is the same as (14) and (15) so that  $m_{st}^{e*}$  has the same functional form as  $m_{st}^*$  in the benchmark case. However, their optimal levels at each period are different. Concentration of capital endowment in the few hands of the elite and the extra tax revenue allow the society to have a higher total bequest, and hence a larger aggregate stock of physical capital in each period. So if the elite ever invest in specific human capital, they do so at a much earlier time and invest more in it, that is,  $t_{se} < t_1$  and  $m_{st}^{e*} > m_{st}^*$ , where  $t_{se}$  is the first time the elite start to invest in specific education such that  $A_{t_{se}-1} k_{t_{se}} = \frac{1-\alpha}{\alpha \gamma}$ .

When the tax revenue is large enough, the elite will invest only in general education and never in specific skills. This is true when all elite individuals have invested in general education by period  $t_{se}$ , that is, when  $t_{le} \leq t_{se}$  or equivalently  $\psi^e(k_{t_{se}}) > C(a_l, 0)$  holds, which is true when the extra tax revenue  $\Delta \tau_{t_{se}}$  is large enough. Since  $\Delta \tau_{t_{se}}$  strictly decreases in  $N_e/N$ , the ratio of the elite versus the masses, there exists a unique level  $\hat{n}$  such that  $\Delta \tau_{t_{se}}(\hat{n}) = C(a_l, 0)$ . This implies that  $N_e/N \leq \hat{n}$  is a sufficient condition for  $\psi^e(k_{t_{se}}) > C(a_l, 0)$ . ■

### Proof of Proposition 3.

**Proof.** The optimal specific education level  $m_{st}^{p*}$  is chosen to get

$$\begin{aligned} T_{s,t+1}^* &\equiv (1 - d(\bar{h}_{g,t+1}^p))(I_{t+1}^{p*} - (N - N_e)m_{st}^{p*}) \\ &= (N - N_e)(1 - d(\bar{h}_{g,t+1}^p)) \max_{m_{st}^p} (w_{s,t+1} h(m_{st}^p) - m_{st}^p). \end{aligned}$$

The FOC for the specific education level  $m_{st}^{p*}$  is

$$\begin{aligned} \alpha h'(m_{st}^p) A_t k_{t+1}^{1-\alpha} &= 1 \text{ for } m_{st}^p > 0, \\ \alpha \gamma A_t k_{t+1}^{1-\alpha} &< 1 \text{ for } m_{st}^p = 0, \end{aligned} \tag{18}$$

The LHS of (18) is increasing over time, so there must exist some period  $t_{1e}$  such that the equality holds. Since  $A_{t_{se}}k_{t_{se}+1} = \frac{1-\alpha}{\alpha\gamma}$ , we know that  $\alpha\gamma A_{t_{se}}k_{t_{se}+1} = (1-\alpha) < 1$ , which means  $t_{1e} > t_{se}$ . This, combined with  $t_{se} > t_{le}$ , indeed confirms our assumption that  $t_{1e} > t_{le}$ .

The optimal general education level  $m_{st}^{p*}$  is chosen to get

$$\begin{aligned} T_{g,t+1}^* &\equiv (1 - d(\bar{h}(m_{gt}^{p*}))(I_{t+1}^{p*} - (N - N_e)m_{gt}^{p*})) \\ &= (N - N_e) \max_{m_{gt}^p} (1 - d(h(m_{gt}^p)))(w_{g,t+1}h(m_{gt}^p) - m_{gt}^p). \end{aligned}$$

The FOC for the general education level  $m_{gt}^{p*}$  is

$$\begin{aligned} \alpha A_{t+1}k_{t+1}^{1-\alpha}[1 - d(h) - d'(h)h]h' + d(h) - d'(h)h'm_{gt}^p &= 1 \text{ for } m_{gt}^p > 0, \\ \alpha\gamma A_{t+1}k_{t+1}^{1-\alpha}(1 - d'(0)) &< 1 \text{ for } m_{gt}^p = 0. \end{aligned} \tag{19}$$

where  $h$  is the shorthand for  $h(m_{gt}^{p*})$ . If  $d'(0) \geq 1$ , then (19) will always hold, which means the ruling elite will never support mass general education. If  $d'(0) < 1$ , the LHS of (19) is smaller than that in (18) if

$$\phi(N_e h_{gt}^{e*}) < \frac{d'(0)}{1 - d'(0)} \tag{20}$$

given that  $A_{t+1} = A_t(1 + \phi(N_e h_{gt}^{e*}))$ . The LHS of (20) is strictly increasing in  $h_{gt}^{e*}$ , which implies that there exists a unique threshold  $\tilde{h}$  such that (20) holds with equality. In other words, before their general education reaches  $\tilde{h}$ , the ruling elite will only support mass education in specific skills. Even after  $h_{gt}^{e*} > \tilde{h}$  is achieved, the elite may still choose not to support mass general education if  $T_{g,t+1}^* < T_{s,t+1}^*$ . Specifically, the investment in mass general education, if any, will never go beyond  $\hat{h}$  under the elite rule, otherwise the tax revenue will become zero or negative. ■