The Technological Origins of the High School Movement

Kim, Se-Um

Brown University

20 October 2008

The Technological Origins of the High School Movement

Se-Um Kim[0]

October 20, 2008

Abstract

This paper argues that the emergence of knowledge hierarchies in the modern U.S. firms since the late 19th century, expedited by huge progress in communication technology, played a significant role in the expansion of mass secondary education called the high school movement in the U.S. in the early 20th century. To analyze the causal connections among these historical events, the paper presents a dynamic model in which the complementarity between individual skills is crucial to production. Middle-skilled individuals could help increase the payoff to the high-skilled by supervising low-skilled production workers as middle managers in firms, and so some of potential top managers with high skill actively supported the expansion of mass education to the secondary level some time after a sophisticated form of production organizations had started to emerge. This theoretical explanation is consistent with the existing historical evidence in the literature.

Keywords: High School Movement, Communication Technology, Skill Complementarity, Knowledge Hierarchies, Middle Managers, Public Secondary Education

JEL Classification Numbers: D20, I20, O10, O40

[0]Brown University (Se-Um_Kim@brown.edu). I am highly indebted to Oded Galor, my main advisor, for his cordial guidance throughout the project and his constructive criticisms that have helped improve the quality of the paper significantly. I especially thank him for his patience while I was stuck in the middle of the project. I am also highly obliged to Peter Howitt and David Weil for their valuable comments and guidance as my advisors. I also thank Eren Arbatli, Quamrul Ashraf, Daeho Kim, Hernando Zuleta and other seminar participants at Brown University for their helpful comments and genuine interest in my paper. All remaining errors are mine.

1 Introduction

In the early 20th century, more specifically from 1910 to 1940, the U.S. went through a great expansion in public secondary education called the *high school movement*, which had never been witnessed in any other countries before and would not be matched by the other parts of the world for several decades (Goldin, 1999).[1] This paper presents a formal dynamic model to analyze what economic incentive gave rise to this unprecedented and unparalleled expansion of secondary education.

To find an answer to why the high school movement in the U.S. began in the early 20th century, this paper focuses on the emergence of knowledge hierarchies in the modern U.S. firms in the late 19th century, as documented in detail by Chandler (1977). He argues that the huge progress in communication technology, which had originated from the introduction of the railroad, the telegraph and the telephone in the mid 19th century, contributed to the emergence of multi-layered firms where the role of the *middle management* was important in bridging the *top management* and production workers.[2] In addition, Chandler (1977) points out that the emergence of knowledge hierarchies in Europe and Japan was neither as extensive nor as early as that in the U.S.,[3] implying that the emergence of knowledge hierarchies was also

[1] Goldin (1998) also writes that "not only was the high school movement from 1910 to 1940 a uniquely-American phenomenon, the secondary school as we know it today was a uniquely-American invention" (p. 350). In addition, Table 1 in Goldin (1998) provides vivid statistical testimony of the drastic effect of the high school movement on the huge increase in the number of high school graduates. It shows that 61.4% of the cohort of men born in 1926-1930 completed secondary education, and so did 80.9% of the cohort of men born in 1946-1950. This is in stark constrast with 16.9% completion rate of the cohort born in 1886-1890.

[2] Garicano and Rossi-Hansberg (2006a) formalize the idea of Chandler (1977) with a simplified model based on Garicano and Rossi-Hansberg (2006b). Garicano and Rossi-Hansberg (2006a), however, presents a model in which there are only two skill levels and so the role of middle managers is not explicitly considered.

[3] "In Europe and Japan, however, the new institution appeared in smaller numbers and, at least until after World War II, spread more slowly than it did in the United States. Because it came slower and later, its builders and administrators have often looked to the American experience for models and precedents" (p. 498). As specific examples of the emergence of knowledge hierarchies in the U.S. much earlier than in Europe and Japan, he mentions the American Tobacco Company, meat-packing companies such as Armour and Swift, light-machinery makers such as Singer Sewing Machine and McCormick Harvester,

a phenomenon unique to the U.S. similarly to the high school movement.[4]

On the other hand, Goldin (1999) points out that the great expansion of public secondary education in the U.S. from 1910 to 1940 was adaptation "to the needs of the modern workplace of the early twentieth century" (p. S69) and that "the impetus to expand education to the secondary level was primarily a grassroots movement led by parents, employers, and even young people themselves" (p. S71). Furthermore, although Goldin and Katz (2000) stop short of saying that the economic incentive of potential top managers in multi-layered firms was the main driving force of the high school movement,[5] they point out that "the growth of big business, as described by Alfred Chandler, increased the demand for managers" and that "Compared with businesses in other industrial countries, American large-scale manufacturing was distinctive in its heavy reliance on skilled managers for decision making rather than shop-floor craftsmen" (pp. 791-792).[6]

Combining the above elements together, this paper argues that the emergence of knowledge hierarchies in the U.S. in the late 19th century played a significant role in promoting the remarkable high school movement in the U.S. in the early 20th century, as potential top managers of multi-layered organizations, who wanted to enjoy high economic payoffs with the formation of knowledge hierarchies, actively supported the expansion of mass education to the secondary level to have sufficient middle managers required for running optimal multi-layered organizations.[7] To make this argument more

and heavy-machinery nakers such as General Electric and Westinghouse.

[4] This observation gives strong support to this paper's attempt to connect by causality the emergence of knowledge hierarchies in the late 19th century and the expansion of mass secondary education in the early 20th century; if one historical event unique to the U.S. has caused the other event, then this event will remain unique to the U.S. for some time, too.

[5] They write, "We can infer, from the evidence just presented, that structural changes in the economy around 1900 increased the relative demand for workers with a high school education. Even though we do not have direct evidence that the high school movement began because of an increase in educational returns, we can estimate returns in 1914 at the start of the high school movement in an educational leader – Iowa" (p. 793).

[6] They do not present a theoretical model to connect this point with the high school movement, however.

[7] Evidence that jobs of *middle managers* really required high school education is provided by Goldin (1999) as follows: "Ordinary white-collar jobs customarily had an entry

convincing, inspired by Antras, Garicano and Rossi-Hansberg (2008) and Garicano and Rossi-Hansberg (2006b),[8] this paper formulates a dynamic model in which *middle managers* with the middle level of skill could help increase the earnings of the *top managers* with the high level of skill, by supervising low-skilled production workers and thereby allowing the top managers to avoid facing routine problems raised by the workers and specialize in the most difficult tasks.[9]

This paper is related mainly to three strands of the literature. First, this paper serves as a complement to Bourguignon and Verdier (2000), which focuses on positive externalities from human capital, and Galor and Moav (2006),[10] which emphasizes the complementarity between physical and human capital as incentives for rich capitalists to support the introduction of public education.[11] We present a formal dynamic model that focuses on the complementarity between individual skills made significant by huge progress in communication technology and the emergence of knowledge hierarchies, as

requirement of several years of high school or a high school diploma, whereas most ordinary production work had no such entry requirement in the 1900-1920s era" (p. S81).

[8] These papers formulate static models rather than dynamic ones, not considering an intergenerational connection.

[9] Although the model focuses only on the rise of *middle managers* as a cause of the high school movement, this paper should not be misunderstood as arguing that that was the only empirically significant cause of the event. Rather this paper should be understood as an attempt to present another possible candidate, which is shown to be significant in theory, for explanatory variables on the right hand side of a regression equation.

[10] Closely related to Galor and Moav (2006) is the literature of a unified growth theory, which tries to explain the very long-run movement of key economic variables throughout human history in a single framework. See Galor and Moav (2004), Galor, Moav and Vollrath (2008), and Murtin (2008) for studies focusing on a similar period to that considered in this paper. A thorough and much detailed review of the literature of a unified growth theory is presented in Galor (2005).

[11] As another complement to this literature, Galiani, Heymann, Dabus and Tohme (2008) argue that in the case of land-rich economies such as Latin American countries, the need to have skilled workers working for the human-captial-intensive service sector was the main motive for the rich elite to support public education. On the other hand, as an antithesis to the literature, Kochar (2008) shows that in the case of underdeveloped countries such as India, litte interest in sustaining decent public education is due to the fact that an increase in the poor's schooling causes a hike in their wages, thereby leading to a decline in the rich's profits. Another study with a slightly different focus is Su (2006), which analyzes whether basic or higher education is more emphasized in public budget allocation chosen exclusively by the privileged class depending on the development stage.

strong economic incentives for potential high-income earning top managers to support the expansion of mass education.[12] Second, this paper is hugely indebted to the literature that either empirically describes or theoretically explains the emergence of multi-layered knowledge hierarchies in the U.S. in the late 19th century. Chandler (1977) gives a very detailed and extensive description of the emergence of knowledge hierarchies in the U.S. in the late 19th century caused by the huge progress in communication technology, which had originated from the introduction of the railroad, the telegraph and the telephone in the mid-19th century. Garicano and Rossi-Hansberg (2006a) presents a formal static model consistent with Chandler's (1977) description of the emergence of knowledge hierarchies, based on another static model presented in Garicano and Rossi-Hansberg (2006b). Antras et al. (2008) provides another static model of knowledge hierarchies in the context of offshoring, which we extend to a dynamic overlapping-generations setting while also endogenizing the proportion of each skill level among the population in the equilibrium that is assumed to be exogenous in their model.[13] Third, as an empirical basis of our theoretical model we rely on the extensive empirical literature that testifies the historical significance of the high school movement in the U.S. in the early 20th century such as Goldin (1998,

[12] To our best knowledge, no theoretical model in the literature has considered the complementarity between individual skills as an incentive for (potentially) rich people to support mass education, although Goldin and Katz (2000) have already suggested that the emergence of knowledge hierarchies as documented by Chandler (1977) was connected to the increase in the demand for educated managers in the early 20th century. The closest theoretical work to this paper might be Doepke and Zilibotti (2005), which argues that the introduction of child labor regulation was the product of conflict between unskilled adult workers who would earn higher wages by restricting child labor and skilled adult workers whose wages would be decreased due to less supply of unskilled workers to work with if child labor regulation was introduced. However, they just assume that production is carried out by using both skilled and unskilled labor as inputs, not considering the formation of organizations. On the other hand, a study by Rajan and Zingales (2005) considers a model in which oligarchs own frims and need to hire managers and laborers who are complementary to each other, but they do not explicitly consider the formation of optimal production organizations either and reach the conclusion that reforms increasing access to education are very hard to implement unless the initial distribution of factor endowments is favorable.

[13] Knowledge hierarchies considered in Saint-Paul (2007) are distinct from those considered in the above-mentioned studies, in that communication in his model takes place in a top-down, not a bottom-up way as in other studies mentioned.

1999 & 2001) and Goldin and Katz (1997, 2000 & 2001).

This paper is structured as follows. The basic structure of our theoretical model based on the literature mentioned earlier is presented in Section 2. The evolution of the model economy depending on the exogenous improvements in communication technology is described in Sections 3, 4 and 5, in which main results of the paper are summarized in five propositions. The results in Sections 3 and 4 show that public secondary education is not implemented yet in the political equilibrium when communication technology has not been improved enough and so the most sophisticated form of knowledge hierarchies, for which a three-layered firm is the proxy in our model, does not emerge yet. Section 5 analyzes the evolution of the economy when communication technology has been so much improved that the most sophisticated form of knowledge hierarchies finally emerges. In Subsection 5.1 we show that public secondary education is still not implemented in the political equilibrium when there is no support yet from those who want to be top managers in three-layered firms and hire the optimal number of middle managers. In Subsection 5.2, however, we show that public secondary education starts to be implemented when there is enough extra support from those who want to be top managers in three-layered firms and employ the optimal number of middle managers, as a result of a sufficient increase in top managers' wages and so a strong incentive to have enough middle-skilled individuals in the economy. Section 6 concludes the paper.

2 The Basic Structure of the Model

Inspired by Galor and Moav (2006), this paper considers an overlapping-generations economy in which individuals live for two periods and are endowed with one unit of time per period. Each parent has one child and each child has only one parent, and so there is no change in the number of people per generation over time. In addition to agents' age, there are two more varieties of heterogeneity among individuals: innate ability and family wealth. We assume for simplicity that innate ability of a child is determined independently of her parent's ability and wealth.

2.1 Formation of Production Organizations

This subsection addresses the evolution of the prevailing form of production organizations depending on the exogenous progress in communication technology, extending the static setup in Antras et al. (2008) to a dynamic context.[14] Following Antras et al. (2008) and Garicano and Rossi-Hansberg (2006a, 2006b), production in each period is carried out using labor and knowledge in competitive environments. Agents optimally choose to produce together in a team or remain self-employed, and each production worker and self-employed agent draws one unit measure of problems per unit of production time, which coincides with the second period of her life in our model. Output y is normalized to the measure of problems solved. An individual i with skill $z_i \in (0,1)$ can solve a proportion z_i of the problems that she faces randomly. Knowledge is cumulative in the sense that an individual with higher skill is able to solve every problem that another individual with lower skill can solve.

While a self-employed agent i with skill $z_i \in (0,1)$ enjoys expected income $y_i = z_i$, agents in teams can communicate their knowledge and help others solve problems. First, each of production workers draws one unit measure of problems to be solved, and they ask middle managers[15] to help with problems that they cannot solve for themselves. After trying to solve those problems, middle managers ask the top manager to help with problems that even they cannot solve. Communication between different layers of a firm forces higher-ranked managers to incur costs in terms of time spent.

To incorporate into our model the progress in communication technology over time that made the emergence of knowledge hierarchies possible as documented by Chandler (1977), and also to serve our main purpose of figuring out the technological origins of the high school movement in a simple way, we make the following assumption: the time cost of communication

[14] The results in this subsection are not summarized in propositions or lemmas since most of them are rather straightforward extensions of the setup in Antras et al. (2008).

[15] Antras et al. (2008) use the term *middle managers* in the context of offshoring, as managers in production facilities in host countries who help minimize the cost of transmitting knowledge across borders. In contrast, this paper uses the term more generally for intermediate layers of managers in any knowledge hierarchies.

between ranks in a team in time t is $h_t \in (0,1)$, and there is an exogenous decrease in the communication costs over time given by

$$h_{t+1} = \frac{h_t}{1+g},\ t = 0, 1, 2, \cdots, \tag{1}$$

where a constant g satisfies $0 < g < 1$.[16][17]

For simplicity, assume that there are three fixed levels of problem-solving skill, the learning technology of which will be specified in the next subsection:

$$0 < z_l < z_m < z_h < 1. \tag{2}$$

With the three levels of skill specified above, it is straightforward to see that it cannot be optimal to have two agents with the same skill level working in different ranks of a team or to have subordinates with higher skill than their managers'. This is because managers would not help increase the output of subordinates, while only the production costs would rise by their wages. So there are only three forms of production possible: three-layered teams, two-layered teams, and self-employment, which are proxies for the most sophisticated form of knowledge hierarchies, the prototypical and less sophisticated one, and the production outside knowledge hierarchies in reality respectively.

For example in a three-layered team, $n_l(1 - z_l)$ problems are asked to be solved by middle managers per unit of time, where n_l is the number of production workers in an organization and z_l is the skill level of production workers. Since managers optimally join organizations with the exact number of production workers that fully consumes all their time, it must be that $n_l h_t(1 - z_l) = n_m$, where n_m is the number of middle managers in an organization. After middle managers try to solve the problems referred to them by production workers, $n_l(1 - z_m)$ problems still remain unsolved

[16] Another way to interpret Equation 1 is to think of it as capturing the sequence of incremental secondary innovations caused by an initial drastic innovation (Helpman, 1998), in the forms of the railroad, the telegraph and the telephone.

[17] Antras et al. (2008) and Garicano and Rossi-Hansberg (2006a, 2006b) consider static models and conduct comparative static analysis with respect to communication costs, as opposed to explicitly considering a decrease in costs over time in a dynamic model.

and are asked to be solved by the top manager, where z_m is the skill level of middle managers. So it must be that $n_l h_t(1-z_m) = 1$ for a similar reason to the above, and the final output of the firm is given by $n_l z_h$, where z_h is the skill level of the top manager.[18] Note that the optimal numbers of middle managers and production workers for each top manager of a firm is given by $n_m = \frac{1-z_l}{1-z_m}$ and $n_l = \frac{1}{h_t(1-z_m)}$ apiece, and output of a three-layered firm, given by $\frac{z_h}{h_t(1-z_m)}$, increases over time as h_t decreases exogenously.[19]

Denoting the earnings of production workers, middle managers and the top manager by w_l, w_m and w_h respectively, the zero profit condition with perfect competition and the above two conditions for optimal team formation imply that the wages of the top manager in a three-layered organization is given by

$$\begin{aligned} w_h &= n_l z_h - n_l w_l - n_m w_m \\ &= \frac{z_h - z_l - z_m h_t(1-z_l)}{h_t(1-z_m)}, \end{aligned} \tag{3}$$

which increases over time as top managers can hire a greater number of production workers with communication costs declining. Note that the optimal ratio between ranks in an organization is fixed by parameters, and so if the proportion of z_h among the whole population is low enough, then $w_m = z_m$

[18] "Note the source of complementarity between skills in our model: An able top manager increases the productivity of all workers in the team. The more knowledgeable subordinates, the larger the team and the more can managers leverage their knowledge" (Garicano & Rossi-Hansberg, 2006b, p. 8).

[19] This increase in the number of production workers working in and output of a firm with h_t decreasing over time, is also consistent with the introduction of a wider variety of *general purpose technologies* (*GPTs*) than just communication technology as emphasized in this paper. GPTs can be widely used in various sectors, leading to productivity growth in the overall economy (Bresnahan & Trajtenberg, 1995). For example, Rosenberg (1963) emphasizes the role of the general-purpose machine tool industry as "a transmission center in the diffusion of the new technology" (p. 426) in American industrialization, while David (1990) focuses on the electric dynamo as a "general purpose engine" around which "the evolution of techno-economic regimes formed" (p. 355). We continue to focus on huge progress in communication technology as the main source, however, for the consistency of our analysis with Chandler (1977). See Helpman and Trajtenberg (1998), Aghion and Howitt (1998) and Howitt (1998) for formal models analyzing the macroeconomic effects of GPTs.

and $w_l = z_l$ as some individuals with z_m and z_l remain self-employed outside the organizations.[20]

Now consider a two-layered organization in which a high-skilled individual is the manager and low-skilled individuals work as production workers. By optimal team formation, it must be that $n_l h_t(1-z_l) = 1$, and if there is a sufficiently low proportion of z_h, $w_l = z_l$. Then by the zero profits condition,

$$\begin{aligned} w_h &= n_l z_h - n_l w_l \\ &= \frac{z_h - z_l}{h_t(1-z_l)}, \end{aligned} \tag{4}$$

which increases over time as high-skilled managers can hire a greater number of production workers with communication costs declining. Every middle-skilled individual is self-employed in this scenario and earns z_m. Note that output produced by this type of two-layered firm is given by $\frac{z_h}{h_t(1-z_l)}$, which is obviously less than $\frac{z_h}{h_t(1-z_m)}$, output of a three-layered firm considered above.[21]

In addition to the above form of two-layered organizations, it might be possible that a high-skilled individual forms a two-layered organization with middle-skilled individuals. In this case, we can derive the following:

$$\begin{aligned} w_h &= n_m z_h - n_m w_m \\ &= \frac{z_h - z_m}{h_t(1-z_m)}. \end{aligned} \tag{5}$$

However, note that $\frac{z_h - z_l}{h_t(1-z_l)} - \frac{z_h - z_m}{h_t(1-z_m)} = \frac{(z_m - z_l)(1-z_h)}{h_t(1-z_l)(1-z_m)} > 0$ and so the latter form of a two-layered organization cannot arise in an equilibrium. In addition, we ignore the case of a two-layered organization made up of a

[20] In fact we do not need to care about whether there are sufficient low-skilled individuals in the economy, because as will be discussed in detail in Section 4, we can assume that there is always a huge potential influx of immigrants who want to come to and work in the economy whenever excess demand for production workers arises.

[21] It is also obvious that output of a self-employed individual is always dominated by output of any type of a production organization.

middle-skilled individual and low-skilled individuals, which can be justified by assuming that higher education required to obtain high skill is essential for providing leadership skill to a leader of an organization.[22] As a result, we can focus only on three forms of production: self-employment, a two-layered organization in which a high-skilled individual is the manager and low-skilled individuals work as production workers, and a three-layered organization in which a high-skilled individual is the top manager, middle-skilled individuals are middle managers and low-skilled individuals are production workers.

Now we can figure out which form of production prevails in the economy with exogenous progress in communication technology over time. First note that regardless of the form of production, earnings of the middle-skilled and the low-skilled remain the same.[23] So only the preferences of the high-skilled determine which mode of production prevails in an equilibrium. For the payoff to the self-employed high-skilled to be lower than that to the manager of a two-layered firm, it must be that

$$z_h \leq \frac{z_h - z_l}{h_t(1 - z_l)}, \tag{6}$$

which implies that

$$h_t \leq \frac{z_h - z_l}{z_h(1 - z_l)} \equiv h^*. \tag{7}$$

On the other hand, for the payoff to the manager of a two-layered firm to be lower than that to the top manager of a three-layered firm, it must be that

$$\frac{z_h - z_l}{h_t(1 - z_l)} \leq \frac{z_h - z_l - z_m h_t(1 - z_l)}{h_t(1 - z_m)}, \tag{8}$$

[22] Or we might assume that high school graudates have hard time funding setup costs of an organization under severe capital market imperfections that we consider in this paper.

[23] It will be shown that the number of the middle-skilled is always enough to leave some of them self-employed without public secondary education. After public schooling is introduced, Kim (2008) shows, selective college admission can be used to make sure that there are enough middle-skilled individuals in the economy. As will be explained in more detail in Section 4, we can assume a constant influx of immigrants into the economy so that there are always enough low-skilled individuals.

which implies that

$$h_t \leq \frac{(z_h - z_l)(z_m - z_l)}{z_m(1 - z_l)^2} \equiv h^{**}. \tag{9}$$

We can also easily show that $h^{**} < h^*$[24]. Note that w_h keeps increasing over time with the sophistication of organizational forms and the exogenous decrease in h_t, while payoffs to the middle-skilled and the low-skilled remain constant regardless of time and occupational choices.[25] This implies that aggregate output in the economy keeps increasing as far as there is no big change in the proportion of each skill level among the population. So output growth in our model is rather trivial as it is mainly driven by the exogenous progress in communication technology over time.

In summary, the forms of production chosen by the high-skilled evolve with the decrease in communication costs over time as shown in Figure 1.

2.2 Individuals

We consider an OLG economy where individuals live for two periods. There is a constant number of individuals in each generation, as each family is made up of one parent and one child and so there is no population growth over time. In the first period of their lives, individuals devote their entire time to education.[26] In the second period of their lives, individuals are self-employed or work for a team as production workers, middle managers or the top manager, depending on their optimal career choices.

The skill level of an individual i depends on her innate ability and her

[24] $h^* - h^{**} = \frac{z_h - z_l}{z_h(1-z_l)} - \frac{(z_h - z_l)(z_m - z_l)}{z_m(1-z_l)^2} = \frac{(z_h - z_l)\{z_m(1-z_h) + z_l(z_h - z_m)\}}{z_h z_m(1-z_l)^2} > 0.$

[25] In reality, the rate of return to secondary education would increase as demand for the middle-skilled by firms increased. Goldin (1999) shows that education-skill premium for middle managers had been very high before World War I, although it decreased after 1920 as a result of an increase in the supply of high school graduates with the high school movement going on. While one might be able to come up with a different formal model from ours that captures that effect too, for simplicity we rather focus only on potential top managers' demand itself as the main cause of the introduction of public secondary education. Considering the rate-of-return-to-education channel additionaly, while making the analysis less tractable, might make it easier to derive the main results of our paper.

[26] For simplicity, first period consumption of an individual is abstracted from. It may be thought of as included in her parent's consumption.

level of formal education.[27] Assume that each individual i is endowed with one of two levels of ability: a_h and a_l, denoting high and low ability respectively. We focus on secondary education as the level of formal education with which one's own innate ability makes the most difference in her skill level after graduation,[28] and the term *high ability* in our model is used to indicate the level of competence required to take full advantage of secondary education, the meaning of which will be made clear below. We also assume for simplicity that innate ability of a child is determined independently of her parent's ability and wealth. Denote the proportion of individuals with high ability among the population by q and that of those with low ability by $1-q$, which stay constant over time.

Also denote three levels of formal education by e_h, e_m and e_l, which represent the completion of higher education, secondary education and elementary education, respectively.[29] Then z_i, each individual's level of problem-solving skill, is expressed as follows:

$$z_i = \begin{cases} z_h \text{ if } e_i = e_h, \\ z_m \text{ if } a_i = a_h \text{ and } e_i = e_m, \\ z_l \text{ if } e_i = e_l \text{ or if } a_i = a_l \text{ and } e_i = e_m. \end{cases} \tag{10}$$

Since we use the term *high ability* as meaning the level of competence necessary to take full advantage of secondary schooling, only individuals with

[27] Antras et al. (2008) and Garicano and Rossi-Hansberg (2006a) assume that the supply of skill is exogenously given, while Garicano and Rossi-Hansberg (2006b) consider endogenous learning of knowledge. However, what they consider is not so much formal education as on-the-job training, as they assume that the learning costs are paid to employees by the firm and that the learning costs are equivalent to the cost of processing information decided by factors such as the improvements in the speed of microprocessors. In contrast, this paper, as shown below, explicitly considers endogenous skill formation through formal education, thereby making it possible to consider the effect of the emergence of knowledge hierarchies on the high school movement in the early 20th century.

[28] Needless to say, one would be able to find examples in the real world implying that the innate ability of college graduates also makes a difference in their skill as top managers. The determination of skill levels presented in (10) below, however, is the simplest possible way to incorporate into our dynamic model the fact that a high school graduate has few ways to overcome the deficit in one's own ability, while a college graduate has many ways to do so.

[29] It is assumed that there is no dropout for simplicity.

high ability can acquire the middle level of skill with secondary education. This implies that individuals with $a_i = a_l$ can never be middle managers in a three-layered organization. Some of them who are lucky enough to go to college thanks to having wealthy parents become top managers, as they can take advantage of close connections with other talented college graduates fostered during and after their college years and so their skill levels can be thought of as high regardless of their own innate ability.[30] However, the others who are not as fortunate become production workers or self-employed with a low level of skill. This implication of (10) seems to be consistent with the reality despite its simplicity.[31] A key assumption regarding q, the proportion of high ability individuals, is such that with high school education only the minority of the population can master the skill level required of middle managers in a three-layered organization, but that a large enough proportion of people can enjoy the benefit of secondary education in that q is not too small, either.[32] This assumption will be

[30] These connections between the alumni of a college can be either those between private business executives and government officials, or those among business executives in trades complementary to one another. Another potential way for untalented college graduates to enjoy high payoffs is by *dynastic management*, studied by Caselli and Gennaioli (2005a, 2005b). Since descendants of rich families are overrepresented among college graduates as will be shown later, many of them have parents who own production facilities and hand them down to their offsprings. In this process know-how to run a firm effectively can also be handed down, implying that many college graduates with wealthy parents have significant advantages over the others. In addition they might be able to take advantage of their parents' connections with other people, too.

[31] The reason that any connections among high school alumni are not enough to overcome the deficit in one's innate ability is as follows: while top managers focus on general management problems that do not necessarily hinge on their own specific working skills, the tasks for white-collar middle managers usually require specific individual working skills. For example, a middle manager with a problem typewriting documents fast enough, would not be able to rely on any connections with her high school alumni to fix the problem and so an individual with low ability has few ways to overcome her handicap to become a middle manager.

[32] Assuming that the proportion q can be arbitrarily large would entail a criticism that the introduction of public secondary education in our model is rather trivial, as the benefit of the high school movement would be just assumed to spread among arbitrarily many people. To avoid this potential criticism we impose the limit that the proportion q is less than $\frac{1}{2}$. On the other hand, while avoiding the above-mentioned criticism is important, we need to have a lower limit of q so that our model can predict the introduction of public secondary education, which is of course consistent with the historical facts. Since *high ability* in this paper refers to the level of competence to take full advantage of secondary

formalized in (A1) later.

It is assumed that elementary schooling was already universal at $t = 0$ and that elementary schooling costs zero for simplicity.[33] However, secondary schooling costs K_1 per individual, and that cost can be publicly funded or paid by the student as tuition. Every individual who optimally chooses to get secondary education is given a chance to go to high school under public education or if they pay full tuition without public education. When public secondary education is implemented, each individual i should pay $\tau_t b_t^i$ in the first period of her life, where τ_t is the common tax rate for funding public education and b_t^i is intergenerational transfer from her parent.[34][35] So the decision to implement public secondary education is equivalent to the decision to raise positive tax by a majority vote in which every individual of the old generation participates in each period.[36] For

education and not that to become professional economists or brain surgeons, assuming that high ability individuals are not too few totally makes sense.

[33] The first transformation of American education took place in the mid 19th century in the form of the expansion of elementary schooling (Goldin, 1998).

[34] We assume for simplicy that tax for funding public education is levied on intergenerational transfer, not the second period income of each parent specified below. Since parental transfer is shown below to be a fixed proportion of the second period income, introducing more general income tax would only make the analysis more complicated, without leading to any changes in the qualitative results of this paper.

[35] This paper regards the public provision of education as the only way of redistributing resources in the economy. See Levy (2005) for the implications of a political economic model, in which income redistribution is another redistributive policy tool available in addition to the public provision of education, for the formation of political coalitions and the resultant policies.

[36] Why does every individual's preference count when it is decided whether to introduce public secondary education? The answer is that the decision on public secondary education, as will be made clear below, has potential to affect the educational and occupational choices of everyone regardless of her ability and the amount of bequests received from her parent, in addition to the fact that everyone in the economy needs to share the costs of secondary education once public education starts to be implemented. This assumption that a majority vote in which every individual, not just some previleged ones, participates decides whether to introduce public secondary education, is consistent with the quotation from Goldin (1999) in the Introduction that "the impetus to expand education to the secondary level was primarily a grassroots movement led by parents, employers, and even young people themselves" (P. S71). Assuming that some group, who economically benefits from public secondary education, can have a decisive influence by lobbying, would not change the qualitative results of this paper and might make the analysis much easier, while causing controversy over not considering the democratic nature of American society. It is assumed that in each period, each old individual votes on the issue of funding pub-

simplicity, it is assumed that there are only two alternatives: no funding with zero tax rate and full funding with $\tau_t = \frac{K_1}{\overline{b_t}}$, where $\overline{b_t}$ is the average level of parental transfer at time t.[37]

On the other hand, an individual should pay K_2 ($> K_1$) privately as tuition for higher education.[38] We assume that college admission is non-selective and every individual who optimally demands college education and can pay tuition is admitted.[39] In addition, a rather extreme form of capital market imperfections is assumed: for an individual i to get higher education, she needs to have $(1 - \tau_t)b_t^i \geq K_2$ with public secondary education implemented, and $b_t^i \geq K_1$ to get secondary education without public secondary schooling.[40] In other words, there is no borrowing for education, which might not be very implausible for the periods on which the model focuses.

---

lic education perfectly reflecting her young child's preference, being aware of her child's innate ability. This assumption is made to reconcile the majority vote with the utility function presented below. It will get clear soon that, as for the vote on the issue of public funding of schooling, there is no conflict between a parent's own consumption and the transfer to her child and so the above assumption is not problematic.

[37] It should be noted that $\overline{b_t}$ is the arithmetic mean, not the median of parental transfer to avoid any confusion in later sections.

[38] K_2 is in general different from expenditure per student as there are other sources of revenue for colleges such as alumni's donations and income from their financial endowments. We assume for simplicity that there is no financial aid, and so every student pays the same amount of tuition. K_2, the real value of college tuition, is also assumed to be constant over time, which is equivalent to assuming that the increase in nominal college tuition over time has been in line with the overall inflation.

[39] This assumption is consistent with the historical fact that almost every applicant was admitted even to highly prestigious colleges before the introduction of selective admission criteria in the early 20th century (Karabel, 2005). As implied by the fixed amount of tuition, the supply of education is assumed to be perfectly elastic and so every individual who optimally chooses to apply to college can be admitted paying K_2 unless colleges choose to adopt selective admission criteria. In other words, only the demand for education determines the equilibrium education level in the economy unless selective college admission is implemented, and this simplification allows us to focus on the incentive of the general public to vote for public education. In addition, in an OLG model in which one generation lasts for 30 years or more, the elasticity of supply of education must be quite high, since the supply side has enough time to adjust for the time trend of demand.

[40] Galor and Zeira (1993) consider this type of non-convexity or indivisibility in human capital investment under a less extreme form of capital market imperfections than ours, focusing on the effect of inequality under capital market imperfections on human capital investment and economic growth.

In the second period of her life, each individual is self-employed or works for a team as a production worker, a middle manager or a top manager, depending on her optimal career choice. Denote second period income of an individual i by I^i_{t+1}. For simplicity, it is assumed that the internationally set interest rate is equal to zero for every period.[41]

Preferences of an individual i of a generation t, defined over second period consumption c^i_{t+1} and the transfer to her child b^i_{t+1}, are represented by the following homothetic, log-linear utility function:[42]

$$u^i_t = (1-\beta)\log c^i_{t+1} + \beta \log b^i_{t+1}, \tag{11}$$

where $\beta \in (0,1)$.

Then an individual optimally chooses her consumption and intergenerational transfer under the following budget constraint:

$$c^i_{t+1} + b^i_{t+1} \leq I^i_{t+1}. \tag{12}$$

Solving the problem of maximizing (11) subject to (12) gives the optimal consumption and transfer as follows:

$$c^i_{t+1} = (1-\beta)I^i_{t+1},\ b^i_{t+1} = \beta I^i_{t+1}. \tag{13}$$

Then the indirect utility of the individual is given by

$$v^i_t = \log(1-\beta)^{1-\beta}\beta^\beta I^i_{t+1}. \tag{14}$$

Therefore, individuals care only about I^i_{t+1}, and choose e_i given a_i and b^i_t to maximize I^i_{t+1}, under the borrowing constraint in the first period.[43]

[41] We assume that individuals can lend any amount of money at this rate in the international capital market, although they cannot borrow anything under the extreme form of capital market imperfections.

[42] This type of *joy-of-giving* preferences has been widely adopted in the recent literature closely related to this paper, the representative one of which is Galor and Moav (2006).

[43] Although an individual cares only about maximizing second period income and then just allocates the income for consumption and the transfer to her child in a fixed proportion, she could be thought of as fully representing her child's interest in the majority vote on public education because the vote involves no conflict between her own consumption

3 The Era of Everybody being Self-employed ($h^* < h_{t+1} \leq h^0$)

With the basic structure of the model presented in Section 2, now we can proceed to see whether public secondary education arises in each stage of the progress in communication technology. In the first stage in which communication costs are so high that production in organizations is not carried out yet, everybody is self-employed as shown in Figure 1 and so $y_i = z_i$, $i = h, m, l$.

Assume that $0 < K_1 < K_2 < 1$ and define $\widehat{b_t} \equiv \frac{K_2}{1-\tau_t} = \frac{K_2}{1-\frac{K_1}{\overline{b_t}}}$, which denotes a threshold level of bequests over which one can afford to get higher education with public secondary education implemented. Also assume that $\overline{b_t} > K_1$, $t = 0, 1, 2, \cdots$, which implies that capital market imperfections are mainly binding on college education. Then we obtain the following lemma:

Lemma 1 *$K_1 < \overline{b_t} < K_1 + K_2 < \widehat{b_t}$ if and only if $\overline{b_t} < K_1 + K_2$, and $K_1 < \widehat{b_t} < K_1 + K_2 < \overline{b_t}$ if and only if $\overline{b_t} > K_1 + K_2$.*

Proof. i) $(K_1 + K_2) - \widehat{b_t} = (K_1 + K_2) - \frac{K_2}{1-\tau_t} = (K_1 + K_2) - \frac{\overline{b_t}K_2}{\overline{b_t}-K_1} = \frac{K_1(\overline{b_t}-K_1-K_2)}{\overline{b_t}-K_1} \begin{cases} < 0 \text{ iff } \overline{b_t} < K_1 + K_2, \\ > 0 \text{ iff } \overline{b_t} > K_1 + K_2. \end{cases}$

ii) $\overline{b_t} - \widehat{b_t} = \overline{b_t} - \frac{K_2}{1-\tau_t} = \overline{b_t} - \frac{\overline{b_t}K_2}{\overline{b_t}-K_1} = \overline{b_t}\left\{1 - \frac{K_2}{\overline{b_t}-K_1}\right\} \begin{cases} < 0 \text{ iff } \overline{b_t} < K_1 + K_2, \\ > 0 \text{ iff } \overline{b_t} > K_1 + K_2. \end{cases}$

iii) $\widehat{b_t} \equiv \frac{K_2}{1-\tau_t} > K_1$ is obvious. Q.E.D. ■

Lemma 1 shows that as the average level of bequests in the economy increases, the tax rate for funding public secondary education decreases and so individuals with less bequests can afford to get college education than before.

Now assume that $\overline{b_t} < K_1 + K_2$ during this period, which implies that an average individual suffers from capital market imperfections.[44] The second period income of an individual i with $a_i = a_h$, depending on her education

and her child's welfare.

[44] Note that the *average* individual is an individual with the arithmetic mean level of bequests, not one with the median level.

level chosen and with the borrowing constraint ignored for the moment, is given as follows:

$$\begin{aligned}
\text{i) } e_i &= e_h \to I^i_{t+1} = \begin{cases} z_h + b^i_t - K_1 - K_2 \\ \text{without public education,} \\ z_h + (1-\tau_t)b^i_t - K_2 \\ \text{with public education.} \end{cases} \qquad (15) \\
\text{ii) } e_i &= e_m \to I^i_{t+1} = \begin{cases} z_m + b^i_t - K_1 \text{ without public education,} \\ z_m + (1-\tau_t)b^i_t \text{ with public education.} \end{cases} \\
\text{iii) } e_i &= e_l \to I^i_{t+1} = \begin{cases} z_l + b^i_t \text{ without public education,} \\ z_l + (1-\tau_t)b^i_t \text{ with public education,} \end{cases}
\end{aligned}$$

where $(1-\tau_t)b^i_t = \left(1 - \frac{K_1}{\overline{b_t}}\right) b^i_t \begin{cases} > b^i_t - K_1 \text{ if } b^i_t < \overline{b_t}, \\ < b^i_t - K_1 \text{ if } b^i_t > \overline{b_t}. \end{cases}$

Note that no individual with high ability would be willing to get only elementary schooling with public secondary education implemented, since individuals need to pay the tax regardless of whether they attend high school.[45]

Before figuring out the preferences of individuals for public education depending on their wealth inherited from their parents, we make the following assumption, which implies that acquiring one-step higher skill is worthwhile if one has high ability, and those with high ability are neither the majority nor too few among the population:

$$\begin{aligned}
z_h &> z_m + K_2, \qquad (A1) \\
z_m &> \max\left\{z_l + K_1,\ (4-2\sqrt{2})z_l + (2\sqrt{2}-3)\right\}, \text{ and} \\
\frac{2-\sqrt{2}}{2} &< q < \overline{q},
\end{aligned}$$

[45] This results from the fact that the first period of one's life is entirely devoted to education in the model. If some fraction of the first period can be spent supplying labor, then some individuals might optimally choose to participate in the labor market as soon as they graduate from elementary school.

where $\overline{q} \equiv \dfrac{-(4z_l - z_m - 3) - \sqrt{(4z_l - z_m - 3)^2 - 8(1-z_l)^2}}{4(1-z_l)}$.[46]

The following lemma shows that $\frac{2-\sqrt{2}}{2} < \overline{q} < \frac{1}{2}$ is always satisfied given (A1), which means that there is no contradiction in the inequality $\frac{2-\sqrt{2}}{2} < q < \overline{q}$ and people with high ability must be the minority. The proof is relegated to the Appendix.

Lemma 2 *Given (A1),*

$$\frac{2-\sqrt{2}}{2} < \overline{q} \equiv \frac{-(4z_l - z_m - 3) - \sqrt{(4z_l - z_m - 3)^2 - 8(1-z_l)^2}}{4(1-z_l)} < \frac{1}{2}.$$

With $z_h > z_m + K_2$ and $z_m > z_l + K_1$ as assumed in (A1) and given Lemma 1 and capital market imperfections, we obtain Figure 2 based on (15), in which solid lines represent second period income of high ability individuals for each level of bequests and education under the borrowing constraint. Note that we consider both a case in which public education is implemented and the other case in which there is no public education. Lines representing the two cases for getting only elementary education are omitted, as it is obvious from (15) that a rational individual with $a_i = a_h$ never prefers those options regardless of her wealth level.

For a given level of b_t^i, one needs to consider only the best scenarios in terms of second period income when she decides whether to vote for public education. Therefore, we obtain the following preferences of high ability agents on public education depending on b_t^i:[47]

i) People with $b_t^i < \overline{b_t}$ vote for public education, hoping to get e_m.

ii) People with $b_t^i \in [\overline{b_t}, K_1 + K_2)$ vote against public education, hoping to get e_m.

iii) People with $b_t^i \geq K_1 + K_2$ vote against public education, hoping to get e_h.

[46] When we take one set of plausible values of z_l and z_m such as $z_l = 0.2$ and $z_m = 0.5$, $(4 - 2\sqrt{2})z_l + (2\sqrt{2} - 3) \simeq 0.0627$, which must be less than $z_l + K_1$.

[47] It does not matter for individuals' voting preferences regarding public education, whether the solid part of the "e_m without public education" line intersects that of the "e_h with public education" line in Figure 2.

In the case of people with high ability, therefore, only relatively poor individuals support public secondary education since they would enjoy the benefit of becoming middle-skilled while paying less than the actual costs of secondary education. On the other hand, relatively rich ones have no incentive to vote for public schooling since they can pay less to get the same levels of education without public education.

To determine whether public secondary education is implemented in this period, we need to look at the preferences of people with low ability, too. The second period income of an individual i with $a_i = a_l$, depending on her education level chosen and with the borrowing constraint ignored for the moment, is given as follows:

$$\begin{aligned}
\text{i) } e_i &= e_h \to I^i_{t+1} = \begin{cases} z_h + b^i_t - K_1 - K_2 \\ \text{without public education,} \\ z_h + (1-\tau_t)b^i_t - K_2 \\ \text{with public education.} \end{cases} \qquad (16) \\
\text{ii) } e_i &= e_m \to I^i_{t+1} = \begin{cases} z_l + b^i_t - K_1 \text{ without public education,} \\ z_l + (1-\tau_t)b^i_t \text{ with public education.} \end{cases} \\
\text{iii) } e_i &= e_l \to I^i_{t+1} = \begin{cases} z_l + b^i_t \text{ without public education,} \\ z_l + (1-\tau_t)b^i_t \text{ with public education.} \end{cases}
\end{aligned}$$

With $z_h > z_m + K_2$ and $z_m > z_l + K_1$ as assumed in (A1), and given Lemma 1 and capital market imperfections, we obtain Figure 3 based on (16), a similar graph to Figure 2 but now for the preferences of low ability individuals. Lines representing the other three cases are omitted as it is obvious from (16) that a rational individual never prefers those options regardless of her wealth level.

From Figure 3, we infer the following preferences of low ability agents on public education depending on b^i_t:[48]

[48] It does not matter for individuals' voting preferences regarding public education, whether the solid part of the "e_l without public education" line intersects that of the "e_h with public education" line in Figure 3.

i) People with $b_t^i < K_1 + K_2$ vote against public education, hoping to get e_l.

ii) People with $b_t^i \geq K_1 + K_2$ vote against public education, hoping to get e_h.

Therefore, people with low ability do not support public secondary education regardless of their wealth level. Non-rich ones vote against public schooling because they cannot afford to get higher education anyway and cannot become middle-skilled with secondary education, either. On the other hand, rich individuals do not support public education because they would pay less to get the same level of education without public schooling.

Before proceeding to see if public education is implemented during this period, we first consider the evolution of bequests over time when public education has never been implemented before. For this purpose, We define the proportions of poor and rich people respectively at time 0 as follows:

$$\begin{aligned} \lambda_P &\equiv \Pr(b_0^i \in [0, K_1)), \\ \lambda_R &\equiv \Pr(b_0^i \in [K_1 + K_2, \infty)), \end{aligned} \tag{17}$$

where an individual is considered poor if she cannot afford even secondary schooling without public education implemented, while one is regarded as rich if she can pay for college education without public funding.

We also assume the following inequalities among parameters of the model so as to impose reasonable restrictions on the evolution of the level of bequests over time, and impose a restriction that the proportion of rich families at time 0 is reasonably low:

$$\begin{aligned} K_1 &> \beta(z_l + K_1),\ K_1 + K_2 > \beta(z_m + K_2) > \beta K_1 + K_2, \\ \beta z_h &> K_1 + K_2, \text{ and } 0 \leq \lambda_R < 1 - \frac{1}{2q(2-q)}. \end{aligned} \tag{A2}$$

Note that (A2) is definitely consistent with (A1). For future use, we present the following lemma regarding how small the upper limit of λ_R is.

The lemma also shows that the inequality $0 \leq \lambda_R < 1 - \frac{1}{2q(2-q)}$ in (A2) is not nonsense. The proof is relegated to the Appendix.

Lemma 3 *Given (A1),* $0 < 1 - \frac{1}{2q(2-q)} < \frac{(1-z_m)q}{(1-z_l)+(1-z_m)q} < \frac{1}{2}$.

Consider now the distribution of skill without public secondary education implemented based on Figures 2 and 3. People with $a_i = a_h$ and $b_t^i < K_1$ are forced to choose e_l as their education level and so acquire z_l as their skill level. Individuals with $a_i = a_h$ and $b_t^i \in [K_1, K_1 + K_2)$ choose e_m and so acquire z_m as their skill level. Agents with $a_i = a_h$ and $b_t^i \geq K_1 + K_2$ want to get e_h and as a result acquire a skill level z_h. On the other hand, people with $a_i = a_l$ and $b_t^i < K_1 + K_2$ choose e_l and acquire z_l as their skill level. Those with $a_i = a_l$ and $b_t^i \geq K_1 + K_2$ want to get e_h and as a result acquire a skill level z_h. Hence there exist all three levels of skill in the equilibrium, and self-employed individuals make earnings accordingly.

Given that $b_{t+1}^i = \beta I_{t+1}^i$ from Section 2, the equilibrium skill and payoff distributions discussed above, and $K_1 > \beta(z_l + K_1)$, $K_1 + K_2 > \beta(z_m + K_2)$ and $\beta z_h > K_1 + K_2$ as assumed in (A2), we obtain Figure 4, a phase diagram depicting the evolution of b_t^i over time with public education having never been implemented:[49]

Note that there are three intersections between the $b_{t+1}^i = \beta I_{t+1}^i$ graph and the 45 degree line denoted by A, B and C in Figure 4. The first three conditions in (A2), except that $\beta(z_m + K_2) > \beta K_1 + K_2$, make sure that the $\beta(z_l + b_t^i)$ line is located above the 45 degree line at $b_t^i = K_1$, that the $\beta(z_m + b_t^i - K_1)$ line is below the 45 degree line at $b_t^i = K_1 + K_2$, and that the $\beta(z_h + b_t^i - K_1 - K_2)$ line (and so the $\beta(w_h + b_t^i - K_1 - K_2)$ line in later periods) is above the 45 degree line at $b_t^i = K_1 + K_2$. This configuration implies that without public secondary education implemented, the long-run level of bequests for descendants of originally rich families (with proportion λ_R) keeps increasing with w_h increasing as a result of the decrease in communication costs over time, while the long-run level of bequests for

[49] Of course w_h in the $\beta(w_h + b_t^i - K_1 - K_2)$ needs to be replaced by z_h in the period considered in this section. We leave w_h in Figure 4 so that we can refer to it in the following sections in which high-skilled individuals work as top managers of firms.

descendants of originally non-rich families (with proportion $1-\lambda_R$) oscillates between points A and B in Figure 4. This prediction of high and increasing inequality and low intergenerational mobility well fits the reality without public education.

Based on the above discussions, we present the following proposition:

Proposition 1 *Given (A1) and (A2), in the period when self-employment prevails due to high communication costs, public secondary education is not implemented.*

Proof. From the above analysis, the only group of people who vote for the introduction of public secondary education in this period are those who have high ability but are relatively poor ($a_i = a_h$ and $b_t^i < \overline{b_t}$). The proportion of those people, $\Pr(a_i = a_h \text{ and } b_t^i < \overline{b_t}) = q \Pr(b_t^i < \overline{b_t})$ by the independence between child's ability and parental ability or wealth. Now suppose that $q \Pr(b_t^i < \overline{b_t}) \geq \frac{1}{2}$, so as to show that this leads to a contradiction. From Figure 4, it is clear that $\Pr(b_t^i < \overline{b_t}) \leq 1 - \lambda_R$ as $\overline{b_t} < K_1 + K_2$ during this period. Then $q \Pr(b_t^i < \overline{b_t}) \leq q(1 - \lambda_R)$. As a result, it follows that $\frac{1}{2} \leq q \Pr(b_t^i < \overline{b_t}) \leq q(1 - \lambda_R)$ must hold. However, Since $q < \overline{q} < \frac{1}{2}$ from (A1) and Lemma 2 proven based on (A1), and $0 \leq \lambda_R < \frac{1}{2}$ from (A2) and Lemma 3 proven based on (A1), $q(1 - \lambda_R) < \frac{1}{2}$ and so the inequality $\frac{1}{2} \leq q \Pr(b_t^i < \overline{b_t}) \leq q(1 - \lambda_R)$ cannot hold true. Q.E.D. ■

Proposition 1 shows that in the period when there is no demand for middle managers working in a multi-layered organization, public secondary education cannot be implemented. As the outcome of the majority vote favors no public secondary education, the distribution of skill in the equilibrium shown above continues to hold. There exist all three levels of skill in the equilibrium, and self-employed individuals make earnings accordingly.

4 The Era of Two-Layered Organizations ($h^{**} < h_{t+1} \leq h^*$)

Before starting the main analysis of this section, we make a crucial simplifying assumption that is consistent with long history of mass immigration into

the U.S.: There is always a huge potential influx of immigrants who want to come to and work in the economy whenever excess demand for production workers arises, but who are willing to be excluded in the political process at both the national and the local level. This assumption is consistent with the history of immigration into the U.S. from the mid-19th century to the mid-20th century, the period that we focus on. These immigrant workers, regardless of the level of education that they had acquired in their native countries, usually started working in blue-collar jobs when they first landed in America while having few ways to make their voices heard in the political process (Brogan, 2001).[50] This assumption implies that we do not need to care about whether there is a sufficient number of low-skilled individuals required for forming optimal multi-layered organizations in the economy.

Based on the above assumption, in the period in which communication costs have decreased enough that production in two-layered organizations starts to emerge but three-layered organizations do not emerge yet, every individual with $z_i = z_h$ becomes the manager of a two-layered firm and employs individuals with $z_i = z_l$ (including immigrants) as production workers. On the other hand, every individual with $z_i = z_m$ in the economy becomes self-employed and so do any individuals with $z_i = z_l$ who are left out of organizations. Then following the results from Section 2, earnings to each skill level is given by

$$w_h = \frac{z_h - z_l}{h_t(1 - z_l)},\ y_m = z_m,\ w_l = z_l. \tag{18}$$

We still assume that $\overline{b_t} < K_1 + K_2$ during this period, which implies that an average individual still suffers from capital market imperfections. Then

[50] In fact this type of immigration persists even into the 21st century, in the form of hotly debated Hispanic illegal immigration. While most of the new immigrants continued to reside in the U.S. over the subsequent generations and so the demographic and political dynamics have been significantly affected in a long haul, we choose not to focus on the issue and assume that the total population in the economy remains unchanged. Introducing exogenous population growth would not affect the qualitative results of the political economy model that we consider. We also implicitly assume that the descendants of the immigrants are fully integrated into the mainstream society and their preferences on public education are identical to those of the original mainstream citizens.

the second period income of an individual i with $a_i = a_h$, depending on her education level chosen and with the borrowing constraint ignored for the moment, is given as follows:

$$\begin{aligned}
\text{i) } e_i &= e_h \rightarrow I^i_{t+1} = \begin{cases} \frac{z_h - z_l}{h_{t+1}(1-z_l)} + b^i_t - K_1 - K_2 \\ \text{without public education,} \\ \frac{z_h - z_l}{h_{t+1}(1-z_l)} + (1-\tau_t) b^i_t - K_2 \\ \text{with public education.} \end{cases} \qquad (19) \\
\text{ii) } e_i &= e_m \rightarrow I^i_{t+1} = \begin{cases} z_m + b^i_t - K_1 \text{ without public education,} \\ z_m + (1-\tau_t) b^i_t \text{ with public education.} \end{cases} \\
\text{iii) } e_i &= e_l \rightarrow I^i_{t+1} = \begin{cases} z_l + b^i_t \text{ without public education,} \\ z_l + (1-\tau_t) b^i_t \text{ with public education.} \end{cases}
\end{aligned}$$

From (6) and (7), $\frac{z_h - z_l}{h_{t+1}(1-z_l)} \geq z_h$ during this period and so we obtain exactly the same result regarding high ability individuals' voting preferences as we saw in Section 3:

i) People with $b^i_t < \overline{b_t}$ vote for public education, hoping to get e_m.

ii) People with $b^i_t \in [\overline{b_t}, K_1 + K_2)$ vote against public education, hoping to get e_m.

iii) People with $b^i_t \geq K_1 + K_2$ vote against public education, hoping to get e_h.

The intuition for the above voting preferences depending on the level of bequests is the same as in Section 3. Now consider the voting preferences of people with low ability. The second period income of an individual i with $a_i = a_l$, depending on her education level chosen and with the borrowing constraint ignored for the time being, is given as follows:

$$
\begin{aligned}
\text{i) } e_i &= e_h \rightarrow I_{t+1}^i = \begin{cases} \frac{z_h - z_l}{h_{t+1}(1-z_l)} + b_t^i - K_1 - K_2 \\ \text{without public education,} \\ \frac{z_h - z_l}{h_{t+1}(1-z_l)} + (1-\tau_t) b_t^i - K_2 \\ \text{with public education.} \end{cases} \qquad (20) \\
\text{ii) } e_i &= e_m \rightarrow I_{t+1}^i = \begin{cases} z_l + b_t^i - K_1 \text{ without public education,} \\ z_l + (1-\tau_t) b_t^i \text{ with public education.} \end{cases} \\
\text{iii) } e_i &= e_l \rightarrow I_{t+1}^i = \begin{cases} z_l + b_t^i \text{ without public education,} \\ z_l + (1-\tau_t) b_t^i \text{ with public education.} \end{cases}
\end{aligned}
$$

Noting that $\frac{z_h - z_l}{h_{t+1}(1-z_l)} \geq z_h$ during this period, we obtain exactly the same result regarding low ability individuals' voting preferences as we saw in Section 3:

i) People with $b_t^i < K_1 + K_2$ vote against public education, hoping to get e_l.

ii) People with $b_t^i \geq K_1 + K_2$ vote against public education, hoping to get e_h.

The intuition for the above voting preferences depending on the level of bequests is the same as in Section 3. Note that the only group of people who vote for the introduction of public education are still those who have high ability but are relatively poor ($a_i = a_h$ and $b_t^i < \overline{b_t}$), with the results for both high and low ability individuals combined. Now we present the following proposition, the proof of which is identical to that of Proposition 1:

Proposition 2 *Given (A1) and (A2), in the period when two-layered production organizations emerge but three-layered firms do not due to an intermediate level of communication costs, public secondary education is still not implemented.*

Proposition 2 shows that when there exists only a prototype of production organizations in which middle managers are not important, public

secondary education is still not implemented. As the outcome of the majority vote favors no public education, we obtain exactly the same distribution of skill in the equilibrium as in Section 3: people with $a_i = a_h$ and $b_t^i < K_1$ are forced to choose e_l as their education level and so acquire z_l as their skill level. Individuals with $a_i = a_h$ and $b_t^i \in [K_1, K_1 + K_2)$ choose e_m and so acquire z_m as their skill level. Those with $a_i = a_h$ and $b_t^i \geq K_1 + K_2$ want to get e_h and as a result acquire a skill level z_h. On the other hand, people with $a_i = a_l$ and $b_t^i < K_1 + K_2$ choose e_l and acquire z_l as their skill level. Agents with $a_i = a_l$ and $b_t^i \geq K_1 + K_2$ want to get e_h and as a result acquire a skill level z_h. Hence there exist all three levels of skill in the equilibrium. It is ensured that every individual with $z_i = z_h$ forms a two-layered organization with the optimal number of production workers $n_l = \frac{1}{h_{t+1}(1-z_l)}$ (See Section 2), because of the assumption that we have made at the beginning of this section. Therefore, we do not need to care about whether the ratio of people with $a_i = a_h$ and $b_t^i < K_1$ or with $a_i = a_l$ and $b_t^i < K_1 + K_2$ to those with $b_t^i \geq K_1 + K_2$ is greater than $\frac{1}{h_{t+1}(1-z_l)}$, as immigrant workers who do not participate in the political process make up for any shortage of production workers demanded by the high-skilled in the economy.

5 The Era of Three-Layered Organizations ($h_{t+1} \leq h^{**}$)

In the period in which communication costs have finally decreased enough that production in three-layered organizations starts to emerge, every individual with $z_i = z_h$ becomes the top manager of a three-layered firm and employs individuals with $z_i = z_m$ as middle managers and individuals with $z_i = z_l$ as production workers.[51] On the other hand, some individuals with $z_i = z_m$ or z_l who are left out of organizations become self-employed. Then

[51] With public secondary education having never been implemented, Proposition 4 below shows that the ratio between the number of middle skilled individuals to that of high skilled ones is greater than the optimal ratio of middle managers to the top manager in a three-layered firm. Even after the implementation of public secondary education, Kim (2008) shows, selective college admission is sufficient for top managers to find enough middle managers required for running optimal three-layered firms.

following the results from Section 2, earnings to each skill level is given by

$$w_h = \frac{z_h - z_l - z_m h_t(1-z_l)}{h_t(1-z_m)},\ w_m = z_m,\ w_l = z_l. \tag{21}$$

5.1 Capital Market Imperfections Still Bind for an Average Individual

First, consider the period during which $\overline{b_t} < K_1 + K_2$ still holds, as in the previous two eras.[52] Then the second period income of an individual i with $a_i = a_h$, depending on her education level chosen and with the borrowing constraint ignored for the moment, is given as follows:

$$\begin{aligned}
\text{i) } e_i &= e_h \to I^i_{t+1} = \begin{cases} \frac{z_h - z_l - z_m h_{t+1}(1-z_l)}{h_{t+1}(1-z_m)} + b^i_t - K_1 - K_2 \\ \text{without public education,} \\ \frac{z_h - z_l - z_m h_{t+1}(1-z_l)}{h_{t+1}(1-z_m)} + (1-\tau_t) b^i_t - K_2 \\ \text{with public education.} \end{cases} \\
\text{ii) } e_i &= e_m \to I^i_{t+1} = \begin{cases} z_m + b^i_t - K_1 \text{ without public education,} \\ z_m + (1-\tau_t) b^i_t \text{ with public education.} \end{cases} \\
\text{iii) } e_i &= e_l \to I^i_{t+1} = \begin{cases} z_l + b^i_t \text{ without public education,} \\ z_l + (1-\tau_t) b^i_t \text{ with public education.} \end{cases}
\end{aligned} \tag{22}$$

From (8) and (9), $\frac{z_h - z_l - z_m h_{t+1}(1-z_l)}{h_{t+1}(1-z_m)} \geq \frac{z_h - z_l}{h_{t+1}(1-z_l)}$ during this period and so we obtain exactly the same result regarding high ability individuals' voting preferences as we saw in Sections 3 and 4:

i) People with $b^i_t < \overline{b_t}$ vote for public education, hoping to get e_m.

ii) People with $b^i_t \in [\overline{b_t}, K_1 + K_2)$ vote against public education, hoping to get e_m.

[52] Since $\overline{b_t}$ is the arithmetic mean level of bequests, not the median, one might think that it is possible that $\overline{b_t} < K_1 + K_2$ had already stopped being satisfied in the previous era. This would imply, however, that the introduction of public secondary education took place one generation before the emergence of a sophisticated form of production organizations: Seeing the emergence of a sophisticated form of firms next period, the majority vote would favor the introduction of public education in a similar way to Proposition 5 below. This is obviously not consistent with the historical facts.

iii) People with $b_t^i \geq K_1 + K_2$ vote against public education, hoping to get e_h.

The intuition for the above voting preferences depending on the level of bequests is the same as in Section 3. Now consider the voting preferences of people with low ability. The second period income of an individual i with $a_i = a_l$, depending on her education level chosen and with the borrowing constraint ignored for the time being, is given as follows:

$$\begin{aligned}
\text{i) } e_i &= e_h \rightarrow I_{t+1}^i = \begin{cases} \frac{z_h - z_l - z_m h_{t+1}(1-z_l)}{h_{t+1}(1-z_m)} + b_t^i - K_1 - K_2 \\ \text{without public education,} \\ \frac{z_h - z_l - z_m h_{t+1}(1-z_l)}{h_{t+1}(1-z_m)} + (1-\tau_t) b_t^i - K_2 \\ \text{with public education.} \end{cases} \qquad (23) \\
\text{ii) } e_i &= e_m \rightarrow I_{t+1}^i = \begin{cases} z_l + b_t^i - K_1 \text{ without public education,} \\ z_l + (1-\tau_t) b_t^i \text{ with public education.} \end{cases} \\
\text{iii) } e_i &= e_l \rightarrow I_{t+1}^i = \begin{cases} z_l + b_t^i \text{ without public education,} \\ z_l + (1-\tau_t) b_t^i \text{ with public education.} \end{cases}
\end{aligned}$$

Noting that $\frac{z_h - z_l - z_m h_{t+1}(1-z_l)}{h_{t+1}(1-z_m)} \geq \frac{z_h - z_l}{h_{t+1}(1-z_l)}$ during this period, we obtain exactly the same result regarding low ability individuals' voting preferences as we saw in Sections 3 and 4:

i) People with $b_t^i < K_1 + K_2$ vote against public education, hoping to get e_l.

ii) People with $b_t^i \geq K_1 + K_2$ vote against public education, hoping to get e_h.

The intuition for the above voting preferences depending on the level of bequests is the same as in Section 3. Note that the only group of people who vote for the introduction of public education are still those who have high ability but are relatively poor ($a_i = a_h$ and $b_t^i < \overline{b_t}$), with the results for both high and low ability individuals combined. Now we present the following proposition, the proof of which is identical to that of Proposition 1:

Proposition 3 *Given (A1) and (A2), in the period when three-layered firms finally emerge as a result of the lowering of communication costs but when* $\overline{b_t} < K_1 + K_2$ *still holds with an implication that an average individual still suffers from capital market imperfections, public secondary education is still not implemented.*

Proposition 3 implies that even if there emerges a sophisticated form of production organizations in which middle managers play a crucial role, public secondary education might not be introduced immediately if the attractiveness of being the top manager of a firm is not that high to the high-skilled and so they do not actively support public secondary education yet. To see why, note that in our model $\overline{b_t}$ increases over time because the payoff to top managers increases with the improvement in communication technology, while payoffs to the others remain unchanged. Therefore, $\overline{b_t} < K_1 + K_2$ is equivalent to saying that the payoff to top managers in the economy is not that high yet, and so the risk of being forced to remain self-employed and earning less than a top manager is not that high, either. This means that there is no sufficient incentive for the high-skilled to support public secondary education so that they can hire enough middle managers required for running an optimal firm. As a result, there is no support for public education from potential top managers yet[53] and this leads to public education not being implemented in this period, which is consistent with the time lag between the emergence of knowledge hierarchies in the late 19th century and the high school movement in the early 20th century.

As the outcome of the majority vote favors no public secondary education, we still obtain exactly the same distribution of skill in the equilibrium as in Sections 3 and 4: people with $a_i = a_h$ and $b_t^i < K_1$ are forced to choose e_l as their education level and so acquire z_l as their skill level. People with $a_i = a_h$ and $b_t^i \in [K_1, K_1 + K_2)$ choose e_m and so acquire z_m as their skill level. Individuals with $a_i = a_h$ and $b_t^i \geq K_1 + K_2$ want to get e_h and as a result acquire a skill level z_h. On the other hand, agents with

[53] Remember that the only group of people who vote for the introduction of public education in this period are those who have high ability but are relatively poor ($a_i = a_h$ and $b_t^i < \overline{b_t}$), and want to be middle-skilled with secondary education.

$a_i = a_l$ and $b_t^i < K_1 + K_2$ choose e_l and acquire z_l as their skill level. Those with $a_i = a_l$ and $b_t^i \geq K_1 + K_2$ want to get e_h and as a result acquire a skill level z_h. Hence there exist all three levels of skill in the equilibrium. As before we can count on the assumption of potential immigrants into the economy to avoid a complicated consideration of the ratio of the number of the low-skilled to that of the high-skilled in the economy. We cannot adopt the same simplifying strategy, however, regarding the ratio of the number of the middle-skilled to that of the high-skilled as new immigrants cannot be expected to have the level of skill required of middle managers due to either language barriers or cultural hardships.[54] To ensure that every individual with $z_i = z_h$ can form a three-layered organization with the optimal number of middle managers $n_m = \frac{1-z_l}{1-z_m}$ (See Section 2), the ratio of the number of people with $a_i = a_h$ and $b_t^i \in [K_1, K_1 + K_2)$ to that of people with $b_t^i \geq K_1 + K_2$ needs to be greater than $\frac{1-z_l}{1-z_m}$. Otherwise some fraction of individuals with $z_i = z_h$ cannot find enough middle managers that they want to hire and as a result are forced to remain self-employed, suffering the decline of payoffs from $\frac{z_h - z_l - z_m h_{t+1}(1-z_l)}{h_{t+1}(1-z_m)}$ to z_h.[55]

Now turn to the question of whether it is possible to have enough middle-skilled individuals in the economy for three-layered organizations to function optimally without public secondary education implemented. In fact we can

[54] Of course in reality we see some foreigners working as middle-managers or even in better jobs. However, it seems that these cases are rather exceptions than a rule. Assuming that new immigrants can be middle-skilled right after arriving in the economy would make our analysis much simpler, while causing the same controversy as we have already pointed out.

[55] To justify this assumption, we can think of the most efficient organizational form and the optimal ratios between ranks in organizations at the time as a social consensus reached by every agent in the economy depending on the communication costs, and so those specific form and ratios at the time need to be strictly abided by. This implies that those with z_h during this period have only two alternatives: either becoming a top manager in a three-layered organization with the optimal numbers of middle managers and production workers employed, or remaining self-employed. Neither forming a two-layered organization only with production workers nor running a three-layered organization with suboptimal ratios between ranks is not possible. Qualitative results in this paper would not change even if we allowed any form of suboptimal organizations to be formed by the high-skilled who failed to find the optimal number of middle managers, so that they could earn higher than z_h but still less than $\frac{z_h - z_l - z_m h_{t+1}(1-z_l)}{h_{t+1}(1-z_m)}$.

show that it must be the case for a reasonable range of the proportion of the people who are totally free of any borrowing constraints for education at time 0, as given in (A2).

Proposition 4 *Given (A1) and (A2), and considering the long run distribution of bequests when public education has never been implemented, the ratio between the number of middle-skilled individuals and that of high-skilled individuals in the economy is greater than the optimal ratio between middle managers and the top manager in a three-layered firm when* $\overline{b_t} < K_1 + K_2$ *holds.*

Proof. From Figure 4, it is clear that in the long run with public education having never been implemented, the $1 - \lambda_R$ proportion of people are located in the interval $b_t^i \in [K_1, K_1 + K_2)$, while the λ_R proportion of people are found in the interval $b_t^i \in [K_1 + K_2, \infty)$. So when both $h_{t+1} \leq h^{**}$ and $\overline{b_t} < K_1 + K_2$ hold, the ratio between the number of middle-skilled individuals and that of high skilled individuals is given by $\frac{\Pr(a_i = a_h \ \& \ b_t^i \in [K_1, K_1+K_2))}{\Pr(b_t^i \geq K_1+K_2)} = \frac{q(1-\lambda_R)}{\lambda_R}$, from the independence between child's innate ability and parental ability or bequest. Now suppose that $\frac{q(1-\lambda_R)}{\lambda_R} \leq \frac{1-z_l}{1-z_m}$, where the RHS is the optimal ratio. We are to show that this leads to a contradiction. It is straightforward to show that this inequality is equivalent to $\lambda_R \geq \frac{(1-z_m)q}{(1-z_l)+(1-z_m)q}$. From Lemma 3 proven based on (A1) $1 - \frac{1}{2q(2-q)} < \frac{(1-z_m)q}{(1-z_l)+(1-z_m)q}$, and so it must be that $1 - \frac{1}{2q(2-q)} < \frac{(1-z_m)q}{(1-z_l)+(1-z_m)q} \leq \lambda_R$. However, this contradicts $\lambda_R < 1 - \frac{1}{2q(2-q)}$ in (A2). Q.E.D. ■

Proposition 4 implies that if λ_R is small enough within a very reasonable range as given in (A2), then people with a skill level z_h have no problem finding as many people with a skill level z_m as they want to hire as middle managers in firms even without public secondary education. This implication strengthens the argument made in proposition 3 that in this period potential top managers have no incentive whatsoever to support public education for the purpose of having enough middle-skilled individuals whom they can hire as middle managers.

5.2 Capital Market Imperfections do not Bind Any More for an Average Individual

Next, consider the period during which $\overline{b_t} > K_1 + K_2$ holds finally, as an average individual escapes from the borrowing constraint for education.[56] Then from Lemma 1, we have $K_1 < \widehat{b_t} < K_1 + K_2 < \overline{b_t}$, where $\widehat{b_t} \equiv \frac{K_2}{1-\tau_t}$.

Then the second period income of an individual i with $a_i = a_h$, depending on her education level chosen and with the borrowing constraint ignored for the moment, is exactly the same as in (22), but now Figure 5, which depicts high ability individuals' optimal choices, looks different from Figure 2 due to the change in the order of K_1, $K_1 + K_2$, $\overline{b_t}$ and $\widehat{b_t}$. Note that in Figure 5 lines representing the two cases for elementary education only are omitted as it is obvious that a rational individual with $a_i = a_h$ never prefers those options regardless of her wealth level.

Based on Figure 5, we obtain the following preferences of high ability agents depending on b_t^i:

i) People with $b_t^i < \widehat{b_t}$ vote for public education, hoping to get e_m.

ii) People with $b_t^i \in [\widehat{b_t}, \overline{b_t})$ vote for public education, hoping to get e_h.

iii) People with $b_t^i \geq \overline{b_t}$ vote against public education, hoping to get e_h.

In the case of people with high ability, relatively poor individuals are still the only group who supports public secondary education since they would enjoy the benefit of becoming middle-skilled or high-skilled while paying less than the actual costs of secondary education. In contrast with what happened in the previous periods, however, as secondary schooling costs would decline with public secondary education implemented,[57] those who

[56] Note that from Figure 4, $\overline{b_t}$ keeps increasing as the payoff to top managers keeps increasing with the improvement in communication technology, while payoffs to others remain unchanged. As a result, it is clear that even when an *average* individual with the *arithmetic mean* level of bequests finally escapes from the borrowing constraint, one with the *median* level of bequests still suffers from that (In fact everyone except a few descendants of originally rich families does). Therefore, this paper does NOT argue that for public secondary education to be implemented, a *median* individual had to totally escape from the borrowing constraint for education.

[57] Note that $\widehat{b_t} \equiv \frac{K_2}{1-\tau_t}$ keeps decreasing over time as $\tau_t = \frac{K_1}{\overline{b_t}}$ keeps decreasing with the increase in $\overline{b_t}$. Therefore, once public secondary education is implemented, the costs of secondary schooling (the tax) keeps decreasing and so an increasing number of relatively

are moderately poor now do not want to stop at secondary schooling and hope to get higher education under public secondary education, in the hope of becoming top managers and hiring middle managers after graduation. On the other hand, relatively rich ones have no incentive to vote for public schooling since they can pay less to get the same level of education without public schooling.

On the other hand, the second period income of an individual i with $a_i = a_l$, depending on her education level chosen and with the borrowing constraint ignored for the time being, is also exactly the same as in (23), but now Figure 6, which depicts low ability individuals' optimal choices, looks different from Figure 3 for the same reason as mentioned above. Note that lines representing the options that will never be preferred by any individual with $a_i = a_l$ are omitted in Figure 6.

Based on Figure 6, we obtain the following preferences of low ability individuals depending on b_t^i:

i) People with $b_t^i < \widehat{b_t}$ vote against public education, hoping to get e_l.

ii) People with $b_t^i \in [\widehat{b_t}, \overline{b_t})$ vote for public education, hoping to get e_h.

iii) People with $b_t^i \geq \overline{b_t}$ vote against public education, hoping to get e_h.

Therefore, now some proportion of people with low ability support public secondary education in contrast with what happened in the previous periods. Those who are moderately poor are the additional group supporting public education: as secondary schooling costs would decline with public secondary education implemented, moderately poor ones do not want to stop at secondary schooling and hope to get higher education under public secondary education, in the hope of becoming top managers and hiring middle managers after graduation. On the other hand, highly poor ones vote against public schooling because they cannot afford to get higher education anyway and cannot become middle-skilled with secondary education, either. In addition, relatively rich individuals do not support public education because they would pay less to get the same level of education without public schooling.

Combining the above results for both high and low ability individuals,

poor people can afford to get higher education over time.

we see that public secondary education is supported by people with high ability who are also not very rich ($a_i = a_h$ and $b_t^i < \overline{b_t}$) as before. However, there arises additional support for public education in this period: now low ability individuals who are also neither very rich nor very poor ($a_i = a_l$ and $b_t^i \in [\widehat{b_t}, \overline{b_t})$) vote for public secondary education, too, as they look forward to becoming top managers of three-layered firms and hiring the optimal number of middle managers.[58] Note that those with $b_t^i \geq \overline{b_t}$ do not support public secondary education even in this period as rich people still prefer not to pay the proportional tax required for sustaining public schooling.[59] The only group, among those who are not very rich, that does not support public secondary education is those with $a_i = a_l$ and $b_t^i < \widehat{b_t}$. These people do not support public secondary education because they know that they cannot afford college tuition anyway and getting only secondary education does not give them a chance to find a better paying job. Furthermore, it is noteworthy that while in the previous periods every individual with $a_i = a_h$ and $b_t^i < \overline{b_t}$ supported public secondary education hoping to become middle-skilled with only secondary education, now some of them, specifically with $a_i = a_h$ and $b_t^i \in [\widehat{b_t}, \overline{b_t})$, support public secondary education in the hope of becoming top managers of three-layered firms with higher education and hiring the optimal number of middle managers. In summary, while people in support of public education used to pursue only up to high school diplomas before, now most supporters of public education pursue college diplomas hoping to have enough people with $z_i = z_m$ whom they can employ as their middle managers.

As a result, if the number of people with $a_i = a_l$ and $b_t^i \in [\widehat{b_t}, \overline{b_t})$, the

[58] The reason that this key additional support for public education is from those who have low ability is that high ability individuals with the same range of bequests have already supported public education since the previous periods. In fact noting that $\widehat{b_t} < K_1 + K_2 < \overline{b_t}$ now, we can also think of those with $a_i = a_h$ and $b_t^i \in [\widehat{b_t}, \overline{b_t})$ as an additional support group for public education, because now they hope to get higher education and then become top managers of firms rather than getting only secondary education. Therefore, our model does NOT imply that those who were pivotal in making successful the high school movement in the early 20th century were all of low ability.

[59] This is in contrast with the result in Galor and Moav (2006) that support for public education becomes unanimous even among rich capitalists in the end, and could be thought of as an improvement on their work in a little more realistic direction.

additional support group, is large enough, then public secondary education has a chance to be implemented in this period. Note that the interval $[\widehat{b_t}, \overline{b_t})$ expands over time, because as $\overline{b_t}$ increases, $\tau_t = \frac{K_1}{\overline{b_t}}$ decreases and so $\widehat{b_t} \equiv \frac{K_2}{1-\tau_t}$ also decreases. Therefore, it might be that this additional support from low ability individuals who are moderately poor is enough for the introduction of public education. However, this is possible only in the later stage of this period, as $\widehat{b_t} \fallingdotseq \overline{b_t}$ in the early stage.

If the outcome of the majority vote is the rejection of public secondary education, which is a probable outcome in the earlier stage of this period, we obtain exactly the same distribution of skill in the equilibrium as in the previous sections: people with $a_i = a_h$ and $b_t^i < K_1$ are forced to choose e_l as their education level and so acquire z_l as their skill level. People with $a_i = a_h$ and $b_t^i \in [K_1, K_1+K_2)$ choose e_m and so acquire z_m as their skill level. Those with $a_i = a_h$ and $b_t^i \geq K_1 + K_2$ choose e_h and so acquire z_h as their skill level. On the other hand, people with $a_i = a_l$ and $b_t^i < K_1 + K_2$ choose e_l and acquire z_l as their skill level. Individuals with $a_i = a_l$ and $b_t^i \geq K_1 + K_2$ want to get e_h and as a result acquire a skill level z_h. Hence there exist all three levels of skill in the equilibrium. To ensure that every individual with $z_i = z_h$ can form a three-layered organization with the optimal number of middle managers $n_m = \frac{1-z_l}{1-z_m}$ (See Section 2), the ratio of people with $a_i = a_h$ and $b_t^i \in [K_1, K_1 + K_2)$ to people with $b_t^i \geq K_1 + K_2$ needs to be greater than $\frac{1-z_l}{1-z_m}$. In fact it is straightforward to show that this is really the case if (A1) and (A2) hold. This result, which is similar to Proposition 4, is obtained because we can use exactly the same graph as Figure 4 except that now $K_1 < \widehat{b_t} < K_1 + K_2 < \overline{b_t}$, with only the positions of $\widehat{b_t}$ and $\overline{b_t}$ interchanged.

Based on these considerations, the next proposition shows that public education must start to be implemented during this period given the previous assumptions that we made.

Proposition 5 *Given (A2), there exists a period t^* when public secondary education starts to be implemented during the period in which both $h_{t+1} \leq h^{**}$ and $\overline{b_t} > K_1 + K_2$ hold.*

Proof. Suppose that until now public education has never been implemented before, and that both $h_{t+1} \leq h^{**}$ and $\overline{b_t} > K_1 + K_2$ hold. Then based on the above discussion, we obtain the following expression for the proportion of the population supporting public secondary education at time t: $\Pr(a_i = a_h \ \& \ b_t^i < \overline{b_t}) + \Pr(a_i = a_l \ \& \ b_t^i \in [\widehat{b_t}, \overline{b_t})) = q \Pr(b_t^i < \overline{b_t}) + (1-q) \Pr(b_t^i \in [\widehat{b_t}, \overline{b_t}))$ by the independence between child's ability and parent's ability or wealth.

Now note that $\overline{b_t}$ increases over time to infinity as $w_h = \frac{z_h - z_l - z_m h_{t+1}(1-z_l)}{h_{t+1}(1-z_m)}$ goes to infinity and so does the long run level of bequests for descendants of originally rich family given by $\frac{\beta}{1-\beta}(w_h - K_1 - K_2)$. Note also that the long run level of bequests for descendants of originally non-rich family oscillates between two exogenous values (See Figure 4 with the positions of $\widehat{b_t}$ and $\overline{b_t}$ reversed). As a result, $\widehat{b_t} \equiv \frac{K_2}{1-\tau_t} = \frac{K_2}{1-\frac{K_1}{\overline{b_t}}}$ is monotonically decreasing from $K_1 + K_2$ to K_2 during this period as $\widehat{b_t} = \overline{b_t} = K_1 + K_2$ at the beginning of this period and $\overline{b_t}$ tends to infinity as time goes on. This implies that $\Pr(b_t^i \in [\widehat{b_t}, \overline{b_t}))$ is non-decreasing over time, while $\Pr(b_t^i < \overline{b_t})$ remains equal to $1 - \lambda_R$ in the long run as $\overline{b_t} \in \left(K_1 + K_2, \frac{\beta}{1-\beta}(w_h - K_1 - K_2)\right)$. Therefore, the proportion of the population supporting public education at time t, $q \Pr(b_t^i < \overline{b_t}) + (1-q) \Pr(b_t^i \in [\widehat{b_t}, \overline{b_t})) = q(1-\lambda_R) + (1-q) \Pr(b_t^i \in [\widehat{b_t}, \overline{b_t}))$ is non-decreasing over time.

Also note that $\beta(z_l + K_2) < K_2$ holds, as $K_1 > \beta(z_l + K_1)$ in (A2) implies that $K_2 > \beta(z_l + K_1) + K_2 - K_1 = \beta(z_l + K_2) + (1-\beta)(K_2 - K_1) > \beta(z_l + K_2)$. This implies that in Figure 4 (with the positions of $\widehat{b_t}$ and $\overline{b_t}$ interchanged), the $\beta(z_l + b_t^i)$ curve lies below the 45 degree line at $b_t^i = K_2$. On the other hand, $\beta(z_m + K_2) > \beta K_1 + K_2$ in (A2) is equivalent to $\beta(z_m + K_2 - K_1) > K_2$, which implies that in Figure 4 (with the positions of $\widehat{b_t}$ and $\overline{b_t}$ interchanged), the $\beta(z_m + b_t^i - K_1)$ curve lies above the 45 degree line at $b_t^i = K_2$. Therefore, $\Pr(b_t^i \in [\widehat{b_t}, \overline{b_t}))$ (and so $q(1-\lambda_R) + (1-q) \Pr(b_t^i \in [\widehat{b_t}, \overline{b_t}))$) is continuously non-decreasing with $\widehat{b_t}$ close to K_2 past the point B in Figure 4 (with the positions of $\widehat{b_t}$ and $\overline{b_t}$ interchanged), where a jump of the probability can arise.

Based on these considerations, now suppose that public secondary edu-

cation can never be implemented, which is equivalent to $q\,(1-\lambda_R)+(1-q)\Pr(b^i_\infty \in [K_2, \overline{b_t})) \leq \frac{1}{2}$. A contradiction caused by this inequality would prove that support for public secondary education becomes the majority in finite time t^* by the continuity of $q(1-\lambda_R)+(1-q)\Pr(b^i_t \in [\widehat{b_t}, \overline{b_t}))$ with $\widehat{b_t}$ close enough to K_2. The LHS of the above inequality is equal to $q\,(1-\lambda_R)+(1-q)q\,(1-\lambda_R) = (2-q)q\,(1-\lambda_R)$ from Figure 4 (with the positions of $\widehat{b_t}$ and $\overline{b_t}$ reversed). Then $(2-q)q\,(1-\lambda_R) \leq \frac{1}{2}$ is equivalent to $\lambda_R \geq 1-\frac{1}{2q(2-q)}$, which contradicts the condition $\lambda_R < 1-\frac{1}{2q(2-q)}$ given in (A2). Q.E.D. ■

In Proposition 3, we have shown that even in the period when three-layered firms finally emerge, public secondary education is still not implemented when $\overline{b_t} < K_1 + K_2$ still holds. This is because public secondary education in this period is supported only by those who want to become only middle-skilled and there is no support whatsoever from those who want to become top managers of three-layered firms and hire middle-skilled individuals as their middle managers. Since the payoffs to low-skilled and middle-skilled individuals remain unchanged at z_l and z_m apiece regardless of time and occupational choices in our model, the long-run evolution of $\overline{b_t}$ depends only on that of the wages of high-skilled individuals as top managers (See Figure 4). Then $\overline{b_t} < K_1 + K_2$ implies that the wages of high-skilled individuals as top managers in three-layered firms are still low due to communication costs being not very low yet, and so those who demand higher education to become high-skilled have few incentives to support public secondary education to have enough middle managers required for running optimal three-layered organizations.

In contrast, Proposition 5 shows that public secondary education starts to be implemented when $\overline{b_t} > K_1 + K_2$ holds. This is because now there is enough extra support from those who want to be top managers in three-layered firms and hire the optimal number of middle managers. $\overline{b_t} > K_1 + K_2$ implies that the wages of high-skilled individuals as top managers are now high with communication costs having decreased enough, and so some of those who demand higher education to become high-skilled now have huge incentives to support public secondary education to have sufficient middle

managers required for running optimal three-layered organizations. Therefore, the results presented in Propositions 3 and 5 are consistent with the quotations from Goldin (1999) and Goldin and Katz (2000) in the Introduction.[60]

If the outcome of the majority vote in the later stage of this period is the implementation of public education,[61] as implied by Proposition 5, it is clear from Figure 5 that people with $a_i = a_h$ and $b_t^i < \widehat{b_t}$ are forced to choose e_m as their education level and so acquire z_m as their skill level. Individuals with $a_i = a_h$ and $b_t^i \geq \widehat{b_t}$ choose e_h and so acquire z_h as their skill level. On the other hand, from Figure 6 it is straightforward to see that those with $a_i = a_l$ and $b_t^i < \widehat{b_t}$ choose e_l or e_m and acquire z_l as their skill level, while agents with $a_i = a_l$ and $b_t^i \geq \widehat{b_t}$ want to get e_h and as a result acquire a skill level z_h. Hence, there exist all three levels of skill in the equilibrium.[62]

[60] The transition from the era in which there was no public secondary education, to the era in which it is implemented, can be thought of as a historical example of the transition from an *oligarchic* society to a *democratic* one presented in Acemoglu (2008). As a result of mass public schooling there are now more people who can afford to get college education and at least some of them, who did not have any chance to be top managers of firms before, now have the opportunity. This is similar to the lowering of entry barriers against new entreprenuers in transition from an oligarchic to a democratic society in Acemoglu (2008). An interesting question is whether this example could be thought of as supporting Glaeser, Ponzetto and Shleifer (2007) which argues that education causes democracy, while refuting Acemoglu, Johnson, Robinson and Yared (2005) which argues otherwise. This seeming contradition might be because the definition of *democracy* used by Acemoglu (2008) is opposite to *oligarchy*, while that in the debate between Glaeser et al. (2007) and Acemoglu et al. (2005) is the antonym of *dictatorship*.

[61] The *later stage* of this period could actually arrive rather quickly, as K_1, the gap between $K_1 + K_2$ and K_2, could be small and the majority of support for public secondary education could start to appear well before $\widehat{b_t}$ reaches K_2. As a result, it might take only a couple of generations for public secondary education to be implemented after knowledge hierarchies of the most sophisticated form start to emerge, which is consistent with the historical facts of the U.S.

[62] To ensure that every individual with $z_i = z_h$ forms a three-layered organization with the optimal number of middle managers $n_m = \frac{1-z_l}{1-z_m}$ (See Section 2), the ratio of the number of people with $a_i = a_h$ and $b_t^i < \widehat{b_t}$ to that of people with $b_t^i \geq \widehat{b_t}$ needs to be greater than $\frac{1-z_l}{1-z_m}$. Now we face an important question: Will there continue to be enough middle-skilled individuals in the equilibrium so that three-layered firms can be run optimally, without any changes in the structure of the educational system under public secondary schooling? Kim (2008) shows that the answer to the above question is "No," and argues that selective college admission using both academic and non-academic criteria, as widely observed in the U.S., can be thought of as a solution to the problem. It also shows that once public secondary education is implemented, those groups who originally

6 Concluding Remarks

In this paper, we have tried to theoretically connect two historical events in the U.S. by causality: the emergence of knowledge hierarchies in the modern U.S. firms in the late 19th century with huge progress in communication technology, and the unprecedented and unparalleled expansion of public secondary education called the high school movement in the early 20th century. Based on a number of theoretical and empirical studies regarding at least one of the above two historical events, we have combined those rather fragmented elements into one consistent dynamic theory as presented in this paper.

Based on an OLG model, the basic structure of which is presented in Section 2 and in which the complementarity between individual skills is crucial to production, we have first shown in Sections 3 and 4 that when there has not yet emerged a sophisticated form of production organizations in which middle managers play a very crucial role in bridging the top manager and production workers to boost productivity significantly, public secondary education cannot be implemented as there is no demand for the middle-skilled required for forming sophisticated knowledge hierarchies. The same result applies, Section 5.1 shows, even when there has emerged a sophisticated form of production organizations that hire middle managers but the attractiveness of being top managers is not very high yet compared to remaining self-employed for the high-skilled individuals. This is because the extra support necessary for public secondary education to be implemented, from the potential top managers for the purpose of having enough middle-skilled individuals required for optimally forming a sophisticated form of knowledge hierarchies, is still not strong enough.

We have shown in Section 5.2, however, that public secondary education can be finally implemented when the payoff to top managers gets high enough compared to that of being self-employed for the high-skilled, as some

supported public secondary education continue to do so in subsequent generations and as a result public secondary education continues to be implemented in the future.

of potential top managers now have a strong incentive to support public secondary education as a way to have enough middle-skilled individuals in the economy whom they can hire as middle managers in firms. This prediction of our model is consistent with the historical fact that the high school movement took place in the U.S. in the early 20th century after a sophisticated form of knowledge hierarchies had started to emerge in the late 19th century.[63]

To briefly mention a possible direction of future research, a theoretical model answering the following questions would be a good complement to this paper: Will there continue to be enough middle-skilled individuals in the economy required for optimally running a sophisticated form of production organizations with public secondary education implemented? And if not, which change in the basic structure of the economy can correct the problem? Is this theoretically identified solution consistent with historical facts of the U.S.? Trying to find answers to these questions might reveal further causal connections among key historical events in the U.S. since the late 19th century, which would not have seemed related so much to one another.[64]

[63] It should be emphasized that this paper does NOT argue that economic development rather automatically leads to the adoption of better institutions in the end. On the contrary, we have focused on the case of the U.S., where democratic politics was already deep-rooted in the periods covered in this paper. Our assumption that everyone participates in a majority vote taking place every period regarding the introdcution of public education, is crucial to deriving the results. For example, if the few descendants of the originally rich families could unilaterally decide on the agenda, there would be no chance for public education to get implemented (Remember that those with $b_t^i \geq \overline{b_t}$ never support public schooling in any period in our model). Therefore, the democratic characterstic of the U.S. is essential to our analysis and the results of this paper cannot be applied to the case of countries with weak democratic foundations without explicitly considering the political charateristics of those countries. In this sense, this paper is consistent with Gradstein's (2008) view that the difference in the original political bias in countries can lead to multiple equilibria in terms of institutional quality and economic development.

[64] See Kim (2008) for the first step in this direction.

Appendix: Proofs of Lemmas 2 and 3

Proof of Lemma 2. i) Suppose that

$$\overline{q} \equiv \frac{-(4z_l - z_m - 3) - \sqrt{(4z_l - z_m - 3)^2 - 8(1-z_l)^2}}{4(1-z_l)} \geq \frac{1}{2}.$$

We are to show that this leads to a contradiction. This inequality is equivalent to $\sqrt{(4z_l - z_m - 3)^2 - 8(1-z_l)^2} \leq 1 + z_m - 2z_l$. Now focus on the term inside the square root on the LHS. $(4z_l - z_m - 3)^2 - 8(1-z_l)^2 = z_m^2 + 2(3-4z_l)z_m + 8z_l^2 - 8z_l + 1$ is positive if $z_m > (4-2\sqrt{2})z_l + (2\sqrt{2}-3)$ as given in (A1), because the determinant of the above quadratic form $(3-4z_l)^2 - (8z_l^2 - 8z_l + 1) = 8(1-z_l)^2 > 0$ and $z_m = (4-2\sqrt{2})z_l + (2\sqrt{2}-3)$ is the greater root of the quadratic equation $z_m^2 + 2(3-4z_l)z_m + 8z_l^2 - 8z_l + 1 = 0$. Since $1 + z_m - 2z_l = (1-z_l) + (z_m - z_l) > 0$ from (2), $\sqrt{(4z_l - z_m - 3)^2 - 8(1-z_l)^2} \leq 1 + z_m - 2z_l$ if and only if $(4z_l - z_m - 3)^2 - 8(1-z_l)^2 \leq (1 + z_m - 2z_l)^2$. It is straightforward to show that this inequality is equivalent to $(1-z_l)(z_m - z_l) \leq 0$, which is a contradiction to (2).

ii) Suppose that

$$\frac{2-\sqrt{2}}{2} \geq \overline{q} \equiv \frac{-(4z_l - z_m - 3) - \sqrt{(4z_l - z_m - 3)^2 - 8(1-z_l)^2}}{4(1-z_l)},$$

so as to show that this leads to a contradiction. It is straightforward to show that this inequality is equivalent to

$$2\sqrt{2}(1-z_l) - (1-z_m) \leq \sqrt{(4z_l - z_m - 3)^2 - 8(1-z_l)^2}.$$

It is easy to see that the LHS is positive as $1 - z_l > 1 - z_m$, and the term in the square root on the RHS has already been shown to be positive with (A1). So the above inequality holds if and only if $\left\{2\sqrt{2}(1-z_l) - (1-z_m)\right\}^2 \leq (4z_l - z_m - 3)^2 - 8(1-z_l)^2$ holds, which after some algebraic manipulations can be shown to be equivalent to $(1-z_m)(1-z_l) \leq 0$. This inequality is a contradiction to (2), which completes the proof. Q.E.D. ■

Proof of Lemma 3. i) Suppose that $0 \geq 1 - \frac{1}{2q(2-q)}$. We are to show that this leads to a contradiction. This inequality is equivalent to $2q^2 - 4q + 1 \geq 0$,

which holds true if and only if $q \leq \frac{2-\sqrt{2}}{2}$ or $q \geq \frac{2+\sqrt{2}}{2}$. However, this contradicts $\frac{2-\sqrt{2}}{2} < q < \overline{q} < \frac{1}{2}$ from (A1) and Lemma 2 proven based on (A1).

ii) Supposed that $1 - \frac{1}{2q(2-q)} \geq \frac{(1-z_m)q}{(1-z_l)+(1-z_m)q}$. We are to show that this gives rise to a contradiction. It is straightforward to show that this inequality is equivalent to $2(1-z_l)q^2 + (4z_l - z_m - 3)q + (1-z_l) \leq 0$. This inequality holds if and only if both $(4z_l - z_m - 3)^2 - 8(1-z_l)^2 = z_m^2 + 2(3-4z_l)z_m + 8z_l^2 - 8z_l + 1 \geq 0$ and $\frac{-(4z_l - z_m - 3) - \sqrt{(4z_l - z_m - 3)^2 - 8(1-z_l)^2}}{4(1-z_l)} \leq q \leq \frac{-(4z_l - z_m - 3) + \sqrt{(4z_l - z_m - 3)^2 - 8(1-z_l)^2}}{4(1-z_l)}$ hold. However,

$$\frac{-(4z_l - z_m - 3) - \sqrt{(4z_l - z_m - 3)^2 - 8(1-z_l)^2}}{4(1-z_l)} \leq q$$

is a contradiction to

$$q < \overline{q} \equiv \frac{-(4z_l - z_m - 3) - \sqrt{(4z_l - z_m - 3)^2 - 8(1-z_l)^2}}{4(1-z_l)}$$

in (A1).

iii) It is straightforward to see that $\frac{(1-z_m)q}{(1-z_l)+(1-z_m)q} < \frac{1}{2}$, as $1 - z_l > (1-z_m)q$. Q.E.D. ■

References

[1] Acemoglu, Daron (2008), "Oligarchic versus Democratic Societies," *Journal of the European Economic Association*, 6(1), 1-44.

[2] Acemoglu, Daron, Simon Johnson, James A. Robinson, and Pierre Yared (2005), "From Education to Democracy?" *American Economic Association Papers and Proceedings*, 95(2), 44-49.

[3] Aghion, Philippe and Peter Howitt (1998), "On the Macroeconomic Effects of Major Technological Change," in Helpman, E. (Ed.), *General Purpose Technologies and Economic Growth* (pp. 121-144), Cambridge, MA: MIT Press.

[4] Antras, Pol, Luis Garicano, and Esteban Rossi-Hansberg (2008), "Organizing Offshoring: Middle Managers and Communication Costs," in Helpman, E., D. Marin, and T. Verdier (Eds.), *The Organization of Firms in a Global Economy*, Harvard University Press.

[5] Bourguignon, Francois and Thierry Verdier (2000), "Oligarchy, Democracy, Inequality and Growth," *Journal of Development Economics*, 62(2), 285-313.

[6] Bresnahan, Timothy F. and Manuel Trajtenberg (1995), "General Purpose Technologies: Engines of Growth?" *Journal of Econometrics*, 65(1), 83-108.

[7] Brogan, Hugh (2001), *The Penguin History of the USA*, New York, NY: Penguin Books.

[8] Caselli, Francesco and Nicola Gennaioli (2005a), "Credit Constraints, Competition, and Meritocracy," *Journal of the European Economic Association*, 3(2-3), 679-689.

[9] Caselli, Francesco and Nicola Gennaioli (2005b), "Dynamic Management," mimeo, LSE.

[10] Chandler, Alfred D. (1977), *The Visible Hand: The Managerial Revolution in American Business*, Cambridge, MA: Belknap Press.

[11] David, Paul A. (1990), "The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox," *American Economic Association Papers and Proceedings*, 80(2), 355-361.

[12] Doepke, Matthias and Fabrizio Zilibotti (2005), "The Macroeco-

nomics of Child Labor Regulation," *American Economic Review*, 95(5), 1492-1524.

[13] Galor, Oded (2005), "From Stagnation to Growth: Unified Growth Theory," in P. Aghion and S. Durlauf (Eds.), *Handbook of Economic Growth* (pp. 171-294), North-Holland.

[14] Galor, Oded and Omer Moav (2004), "From Physical to Human Capital Accumulation: Inequality and the Process of Development," *Review of Economic Studies*, 71(4), 1001-1026.

[15] Galor, Oded and Omer Moav (2006), "Das Human Kapital: A Theory of the Demise of the Class Structure," *Review of Economic Studies*, 73(1), 85-117.

[16] Galor, Oded, Omer Moav and Dietrich Vollrath (2009), "Inequality in Land Ownership, the Emergence of Human Capital Promoting Institutions, and the Great Divergence," forthcoming *Review of Economic Studies*, 76(1).

[17] Galor, Oded and Joseph Zeira (1993), "Income Distribution and Macroeconomics," *Review of Economic Studies*, 60(1), 35-52.

[18] Garicano, Luis and Esteban Rossi-Hansberg (2006a), "The Knowledge Economy at the Turn of the Twentieth Century: The Emergence of Hierarchies," *Journal of the European Economic Association Papers and Proceedings*, 4(2-3), 396-403.

[19] Garicano, Luis and Esteban Rossi-Hansberg (2006b), "Organization and Inequality in a Knowledge Economy," *Quarterly Journal of Economics*, 121(4), 1383-1435.

[20] Galiani, Sebastian, Daniel Heymann, Carlos Dabus and Fernando Tohme (2008), "On the Emergence of Public Education in Land-rich Economies," *Journal of Development Economics*, 86(2), 434-446.

[21] Glaeser, Edward L., Giacomo Ponzetto and Andrei Shleifer (2007), "Why Does Democracy Need Education?" *Journal of Economic Growth*, 12(2), 77-99.

[22] Goldin, Claudia (1998), "America's Graduation from High School: The Evolution and Spread of Secondary Schooling in the Twentieth Century," *Journal of Economic History*, 58(2), 345-374.

[23] Goldin, Claudia (1999), "Egalitarianism and the Returns to Education During the Great Transformation of American Education," *Journal of Political Economy*, 107(6), S65-S94.

[24] Goldin, Claudia (2001), "The Human-Capital Century and American Leadership: Virtues of the Past," *Journal of Economic History*, 61(2), 263-292.

[25] Goldin, Claudia and Lawrence F. Katz (1997), "Why the United States Led in Education: Lessons from Secondary School Expansion, 1910 to 1940," NBER Working Paper 6144.

[26] Goldin, Claudia and Lawrence F. Katz (2000), "Education and Income in the Early Twentieth Century: Evidence from the Prairies," *Journal of Economic History*, 60(3), 782-818.

[27] Goldin, Claudia and Lawrence F. Katz (2001), "The Legacy of U.S. Educational Leadership: Notes on Distribution and Economic Growth in the 20th Century," *American Economic Association Papers and Proceedings*, 91(2), 18-23.

[28] Gradstein, Mark (2008), "Institutional Traps and Economic Growth," *International Economic Review*, 49(3), 1043-1066.

[29] Helpman, Elhanan (1998), "Introduction," in Helpman, E. (Ed.), *General Purpose Technologies and Economic Growth* (pp. 1-14), Cambridge, MA: MIT Press.

[30] Helpman, Elhanan and Manuel Trajtenberg (1998), "A Time to Sow and a Time to Reap: Growth Based on General Purpose Technologies," in Helpman, E. (Ed.), *General Purpose Technologies and Economic Growth* (pp. 55-84), Cambridge, MA: MIT Press.

[31] Howitt, Peter (1998), "Measurement, Obsolescence, and General Purpose Technologies," in Helpman, E. (Ed.), *General Purpose Technologies and Economic Growth* (pp. 219-252), Cambridge, MA: MIT Press.

[32] Karabel, Jerome (2005), *The Chosen: The Hidden History of Admission and Exclusion at Harvard, Yale, and Princeton*, Boston, MA: Houghton Mifflin.

[33] Kim, Se-Um (2008), "A Theory of Selective College Admission in the U.S.," mimeo, Brown University.

[34] Kochar, Anjini (2008), "Schooling, Wages, and Profits: Negative Pecuniary Externalities from Schooling and their Consequences for Schooling Investments," *Journal of Development Economics*, 86(1), 76-95.

[35] Levy, Gilat (2005), "The Politics of Public Provision of Education," *Quarterly Journal of Economics*, 120(4), 1507-1534.

[36] Murtin, Fabrice (2008), "American Economic Development or the Virtues of Education," mimeo, Stanford University.

[37] Rajan, Raghuram G. and Luigi Zingales (2005), "The Persistence of Underdevelopment: Constituencies and Competitive Rent Preservation," mimeo, University of Chicago.

[38] Rosenberg, Nathan (1963), "Technological Change in the Machine Tool Industry, 1840-1910," *Journal of Economic History*, 23(4), 414-443.

[39] Saint-Paul, Gilles (2007), "Knowledge Hierarchies in the Labor Market," *Journal of Economic Theory*, 137(1), 104-126.

[40] Su, Xuejuan (2006), "Endogenous Determination of Public Budget Allocation across Education Stages," *Journal of Development Economics*, 81(2), 438-456.

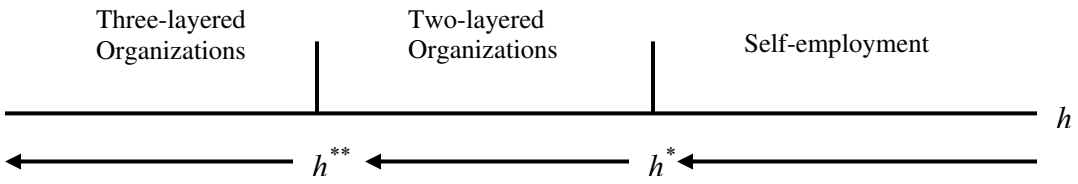

Figure 1: The evolution of production forms with the decline in communication costs

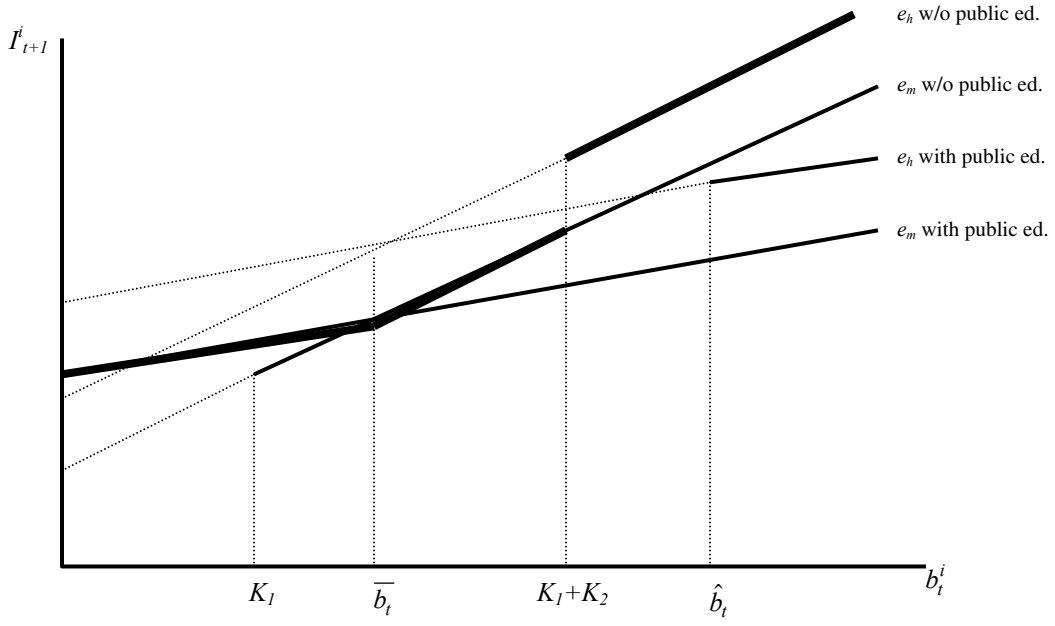

Figure 2: High ability individuals' preferences regarding public secondary education depending on the level of bequests, when $\overline{b_t} < K_1 + K_2$. The bold segments represent the maximum second period income for each level of bequests.

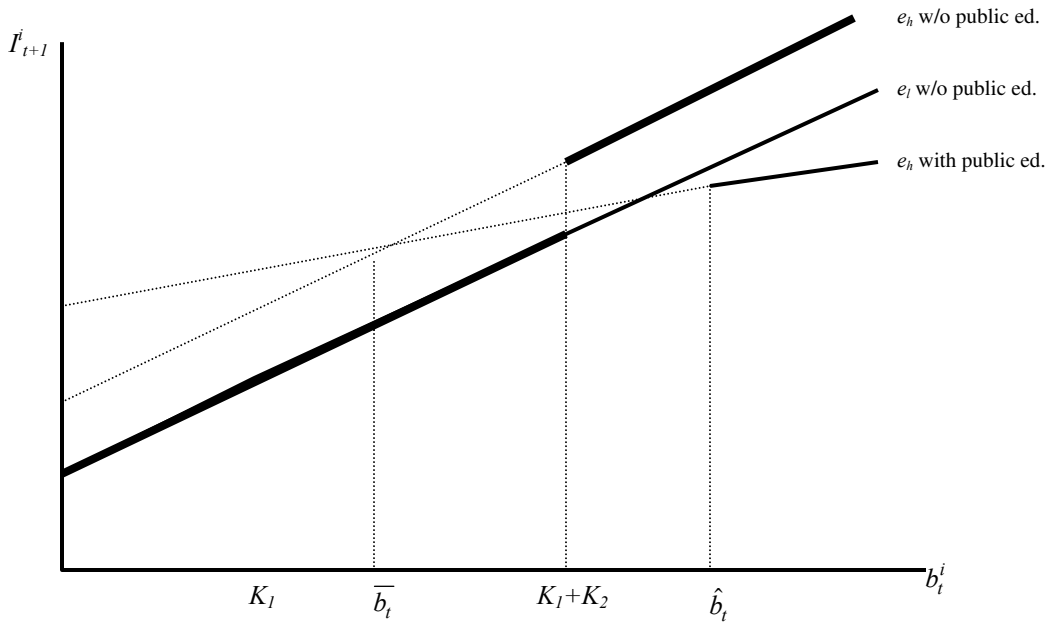

Figure 3: Low ability individuals' preferences regarding public secondary education depending on the level of bequests, when $\overline{b_t} < K_1 + K_2$. The bold segments represent the maximum second period income for each level of bequests.

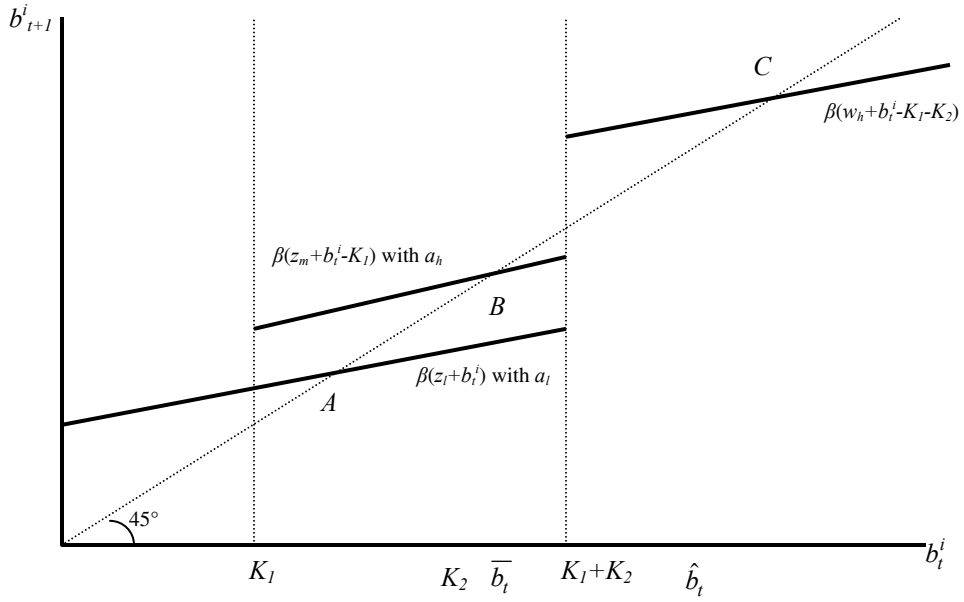

Figure 4: The evolution of bequests without public secondary education implemented. The long-run level of bequests for descendants of originally rich families keeps increasing, while that for descendants of originally non-rich families oscillates between points A and B.

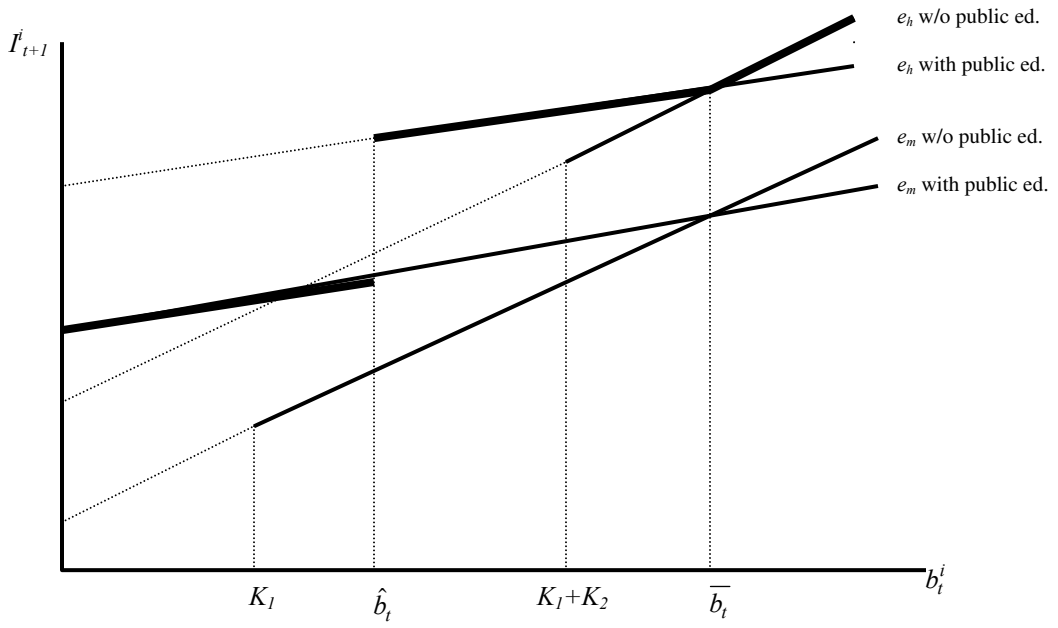

Figure 5: High ability individuals' preferences regarding public secondary education depending on the level of bequests, when $\overline{b_t} > K_1 + K_2$. The bold segments represent the maximum second period income for each level of bequests.

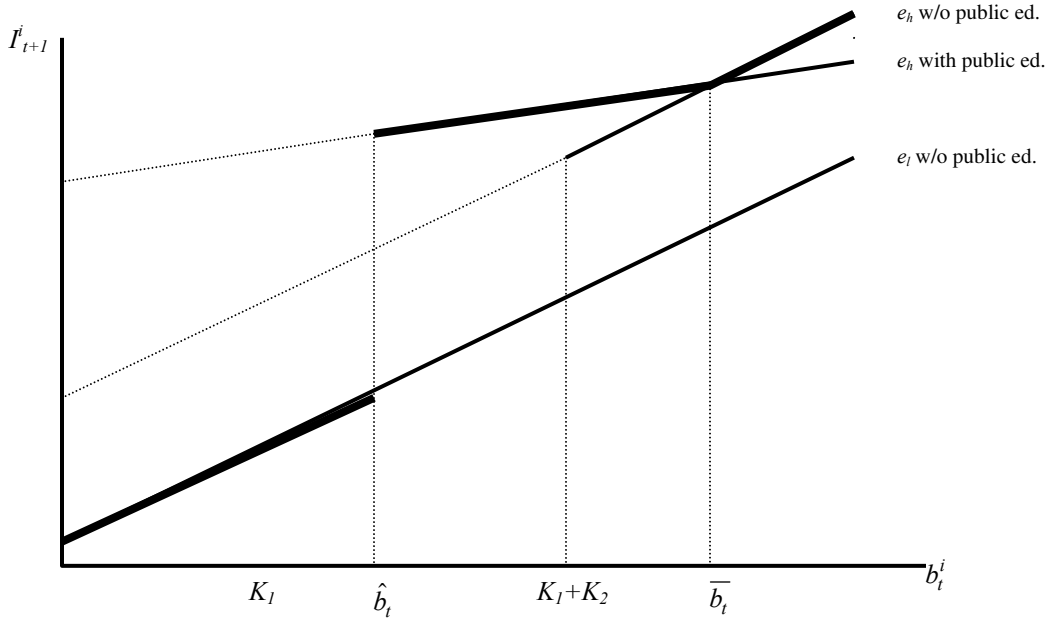

Figure 6: Low ability individuals' preferences regarding public secondary education depending on the level of bequests, when $\overline{b_t} > K_1 + K_2$. The bold segments represent the maximum second period income for each level of bequests.